Evolutionary Dynamics of Mitochondrial Mutations in the Origin and Development of Eukaryotic Sex

Arunas L Radzvilavicius

A thesis submitted for the degree of

Doctor of Philosophy

CoMPLEX
University College London
January 2017

I, Arunas L Radzvilavicius, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Principal supervisor: Prof. Andrew Pomiankowski, UCL

Subsidiary supervisor: Dr. Nick Lane, UCL

Abstract

Sexual reproduction is virtually universal among eukaryotes, suggesting that the last eukaryotic common ancestor was already sexual. It is very likely that the first sexual lineage already contained mitochondrial endosymbionts, each with its own genome of bacterial origin. In this thesis I develop a set of theoretical models that together form a framework for understanding the evolution of eukaryotic sex and further sexual traits—mating types, uniparental inheritance, sexual dimorphism and the early sequestration of a protected germline in higher metazoans—as a consequence of mitochondrial endosymbiosis.

First, I review currently dominating views on the origin of eukaryotes and selective forces that led to the evolution of meiotic sex early in the prokaryote-eukaryote transition. Sex likely emerged as a direct consequence of the mitochondrial endosymbiosis, and was essential for the further evolution of eukaryotic genome complexity. In Chapter 2, I show that the evolution of sexual cell fusion in the nascent eukaryotic lineage might have been driven by cytoplasmic mixing, temporarily masking the detrimental effects of defective organelles. The model introduced in Chapter 3 shows that self-incompatible mating types can evolve to ensure the efficient removal of mitochondrial mutations through asymmetric organelle transmission.

Frequent observations of paternal leakage and heteroplasmy pose a substantial challenge to the current understanding of uniparental organelle inheritance. In Chapter 4 I show that the evolutionarily stable pattern of cytoplasmic inheritance depends on which sex—male or female—governs the destruction of paternal organelles. Maternal regulation favours complete elimination of sperm mitochondria, while paternal control supports paternal leakage and heteroplasmy. Intersexual competition over the control of cytoplasmic inheritance may have driven the repeated evolution of mechanisms enforcing uniparental inheritance. Finally, I analyse the dynamics of mitochondrial mutation segregation in the evolution of the metazoan

germline. High mitochondrial DNA replication error rates in bilaterians favour early germline sequestration, while in basal metazoans gamete quality is maximized through repeated cell divisions in non-sequestered germline stem cell lineages.

Impact statement

The evolution of eukaryotic sexuality is a fundamental unsolved question. In this thesis, I present a set of mathematical models that together form a novel framework for understanding the evolution of eukaryotic sexuality as a consequence of mitochondrial endosymbiosis, with purifying organism-level selection against deleterious mitochondrial mutations playing a central role. This unified approach offers new solutions to unsolved or disputed questions in theoretical biology: the evolution of sexual cell fusion, the origin of self-incompatible mating types together with the asymmetric inheritance of cytoplasmic genes, the extraordinary diversity of mechanisms regulating mitochondrial inheritance, the persistence of heteroplasmy and paternal leakage and the evolution of early germline sequestration in metazoans with two sexes. This work therefore makes a major and potentially paradigm shifting contribution to our understanding of the evolution of eukaryotic sexual life cycles and developmental programmes in relation to mitochondrial genetics, laying the foundations for future theoretical and empirical research.

Table of Contents

List of Figures

CHAPTER 1. EVOLUTION OF EUKARYOTIC COMPLEXITY VIA ENDOSYMBIOSIS AND SEX

1.1 Summary

Sexual reproduction with nuclear fusion and reciprocal recombination is nearly universal among complex eukaryotes, but its early evolution remains shrouded in mystery. Myriad theories on the evolution of sex exist, but most of them focus on the advantages of sex in contemporary complex organisms and do not explain how, when and why sex first arose. It is becoming increasingly clear, however, that sex evolved as part of the evolutionary transition from prokaryotes to eukaryotes, in a pre-eukaryotic cell that already possessed mitochondrial endosymbionts. A new view is emerging in which sex is both a direct consequence of the bacterial-archaeal symbiosis, and a *sine qua non* for the further evolution of the genomic complexity of eukaryotes. Selective forces induced by the presence of at least two genomes of distinct origin within the same cell further shaped the evolution of eukaryotic reproductive strategies, from the evolution of two mating types with uniparental inheritance in protists, to sexes with extreme gamete-size dimorphism and sex-specific germline structure in multicellular eukaryotes.

1.2 Introduction

At its core, sex is a process of genetic mixing: it creates sets of alleles and trait combinations different from either of the two parents. In eukaryotes, two reproductive cells—the gametes—fuse, homologous chromosomes are paired and recombined into novel permutations of maternal and paternal pieces that are then passed on to the offspring (Page and Howley, 2003). Although parts of the molecular machinery used in meiotic recombination are derived directly from their prokaryotic precursors (Marcon and Moens, 2005; Goodenough and Heitman, 2014), sex is a trait unique to eukaryotes. Bacteria and archaea acquire new genes through the mechanisms of horizontal gene transfer (HGT), sometimes referred to as “bacterial sex” (Narra and Ochman, 2006), but these mechanisms always operate unidirectionally, produce pangenomes with different sets of genes in closely related lineages, and do not involve cell fusion nor recombination across the full length of the chromosomes. In contrast, sexual gene exchange in eukaryotes is always reciprocal and operates only among the representatives of the same species, producing vertical lineages with allele combinations that belong to the same general gene set.

Reproduction by the means of sex is strikingly ubiquitous among extant eukaryotes (Goodenough and Heitman, 2014; Speijer et al., 2015), in spite of numerous apparent advantages of clonal reproduction (Otto, 2009; Lehtonen et al., 2012). A lineage of asexual females, for example, would produce twice as many offspring as the sexual one, avoiding the burden of bearing males that do not directly invest in producing offspring (Lehtonen et al., 2012), or finding a compatible and healthy mating partner. And yet very few eukaryotic species revert to obligate asexuality, and those that do, are usually relatively short-lived on the evolutionary timescale (Vrijenhoek, 1998; Simon et al., 2003). Sex is costly, but it also appears to be vital for the long-term survival of complex life.

The most prevalent view on the evolutionary advantages of sex is that meiotic recombination breaks up the associations between mutations at different loci, allowing selection to act on individual genes independent of their backgrounds within finite populations (Otto, 2009). Reciprocal recombination could improve the response to fluctuating selection, e.g. in individuals infected by rapidly evolving parasites (Jaenike, 1978; Hamilton, 1980), or bring several beneficial mutations into the same lineage (Fisher, 1930). More notably, recombination increases the efficacy of selection against mildly deleterious mutations, what would otherwise irreversibly accumulate over time (Muller, 1964; Keightley and Otto, 2006). These canonical views were developed to account for the advantages of sex in modern eukaryotic populations, but do not tell us much about when, how and why sex first arose. Selective forces favouring sex at its origin could have been very different from the evolutionary pressures maintaining amphimixis in present-day populations of complex eukaryotes.

Recent phylogenetic analyses have shown that the last common ancestor of all eukaryotes (LECA) was already capable of full meiotic sex, as the genes underpinning nuclear fusion and meiotic recombination are present in virtually all extant eukaryotic clades (Ramesh et al., 2005; Schurko and Logsdon, 2008; Speijer et al., 2015). Although the prokaryote-to-eukaryote transition left no evolutionary intermediates, there is little doubt that early eukaryotic evolution was largely shaped by the ancient endosymbiotic association between the archaeal host and bacterial ancestors of mitochondria (Martin and Müller, 1998; Lane and Martin, 2010). It is therefore very likely that the cell that first became sexual, although not necessarily a fully-featured eukaryote, already possessed mitochondrial endosymbionts; sex was among the myriad eukaryote-specific traits such as internal membrane systems, active transport networks and phagocytosis that evolved within the context of this unique archaeal-bacterial partnership.

While canonical hypotheses for the evolution of sex are limited to population-genetic effects of reciprocal recombination among nuclear genes (Kondrashov, 1993;

Otto, 2009), shifting the focus from meiosis in modern complex organisms to its very origin in the nascent eukaryotic cell points to the central role of energetic and genetic revolution following the acquisition of mitochondria (Lane, 2011; Garg and Martin, 2016). Housing two genomes of distinct origin within the same cell had immense repercussions for the further evolution of complex life. If meiotic recombination is strictly reciprocal, the inheritance of mitochondrial genes is almost always biased towards one of the gamete classes or mating types, breaking the primordial symmetry of sex (Billiard et al., 2011; Birky, 2001). While this asymmetry is almost certainly related to the fundamental differences in genome structure, ploidy levels or the intricate interactions between the two genomes, the exact reasons behind it remain elusive (current explanations reviewed in Greiner et al., 2015). The sexual asymmetry is even more pronounced in higher eukaryotes characterized by true sexes with oogamy, where females produce few large oocytes and males specialize in mass-production of small and often motile sperm (Billiard et al., 2011). Once again, in an overwhelming majority of cases, mitochondria are transmitted through only one of sexes.

Oogamy appears to have evolved multiple times, and is virtually universal among complex multicellular organisms such as algae, plants and animals, where the unit of selection is not an individual cell, but a large community of clonal cooperating cells all descending from a single zygote (Kirk, 2006). A vast majority of clonal cells in these organisms do not survive past a single generation, and specialize in providing support for germline—the only population of cells contributing genetic material, both mitochondrial and nuclear, to the future generations (Buss, 1987; Michod and Roze, 2010). Often omitted from theoretical analyses of sex-role evolution, germline structure represents one of the most conspicuous sexually asymmetric traits of higher metazoans, with female germlines characterised by relatively low numbers of stem cell divisions, mitochondrial bottlenecks and atretic germ cell death (Krakauer and Mira, 1999), and males producing gametes continuously throughout adulthood without any apparent constraints on germline cell division (Spradling et al., 2011).

The origin of eukaryotic sex, mating types and oogamous sexes, germline-soma differentiation and traits involved in sexual selection have long dominated the list of the most puzzling questions in evolutionary biology. A recent surge of mitochondria-centric research suggests that shifting our focus to the effects of the ancient mitochondrial endosymbiosis has the potential of eliminating much of the mystery that surrounded these problems for more than a century. In this introduction to the rest of the thesis, I discuss recent theoretical developments that establish links between the mitochondrial genetics, energetics and the evolution of eukaryotic sexuality, and how these views fit into the general framework of eukaryotic evolution. The emerging consensus is that the mitochondrial endosymbiosis was a chief driving force behind the evolution of sexual life cycles and traits, while sex itself was largely responsible for much of the genomic complexity characteristic of the eukaryotic domain.

1.3 Endosymbiosis at the origin of eukaryotes

For over a century the concept of endosymbiosis—cells living within cells—has figured prominently in evolutionary hypotheses for eukaryote origin (Mereschkowsky 1905, 1910; Wallin, 1927; Sagan, 1967). The genomic data that became available within the last few decades confirmed the chimeric nature of eukaryotic genomes (Rivera et al, 1998): in addition to genes that are unique to the eukaryotic domain, genes related to information processing and storage (translation, transcription, splicing, replication) are mostly related to archaea (Yutin et al., 2008; Koonin, 2010), while genes for metabolic processes are mostly of bacterial origin. This suggests that the symbiosis that gave rise to a complex eukaryotic cell involved an archaeal partner, as well as a bacterial partner or partners (Koonin, 2010; Guy et al., 2014; Martin et al., 2015), one of which gave rise to mitochondria. Critically, while the nature of the archaeal partner is becoming increasingly well understood (Guy et al., 2014; Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017), eukaryotic genes of bacterial origin cluster with several

groups of present-day bacteria (Koonin, 2010), complicating the analysis and hindering the reconstruction of early eukaryogenic events. Not surprisingly, at least 20 different versions of the endosymbiotic theory have been proposed (Martin et al., 2015), and current data from gene phylogenies is not sufficient to dispel the ambiguity.

Several evolutionary hypotheses were historically centred around the assumption that the host cell that acquired mitochondria was already eukaryotic, and could have formed via an earlier fusion between members of bacterial and archaeal domains ("mitochondria late", reviewed in Guy et al., 2014; Martin et al., 2015). More recently, many of these ideas were criticized in favour of an alternative theory—that the last eukaryotic common ancestor already possessed mitochondria—dedicated energy-producing organelles of bacterial origin (Müller et al., 2012). In this "mitochondria-early" view, the origin of mitochondria and the eukaryotic cell was the same event, and amitochondriate eukaryotes have never existed.

Present-day eukaryotes that lack bona fide mitochondria bear highly simplified mitochondrion-related organelles (MROs)—hydrogenosomes and mitosomes (Tielens et al., 2002; Embley et al., 2003; Henze and Martin, 2003; Tovar et al., 2003; Stechmann et al., 2008; but see Karnkowska et al. (2016) for a recently documented loss of the organelle), found in anaerobes, some of which were once mistakenly thought to have been derived before the acquisition of mitochondria—the so-called Archezoa (Cavalier-Smith, 1983), including metamonads, parabasalids and microsporidia (van der Giezen and Tovar, 2005). Hydrogenosomes produce ATP by substrate-level phosphorylation, oxidizing pyruvate to form hydrogen, CO2 and acetate (Müller, 1988). ATP is then exported into cytosol via the standard mitochondrial ATP/ADP carrier (AAC). Furthermore, mitochondria and hydrogenosomes use protein import pathways of shared origin, and have conserved mechanisms of iron-sulphur cluster assembly. While mitosomes do not have a role in ATP production, they nevertheless retain the critical pathways of iron-sulphur cluster assembly (Tovar et al., 2003).

The realization that mitochondria and hydrogenosomes share common ancestry has led Martin and Müller (1998) to propose one of the earliest “mitochondria-early” theories—the so called *hydrogen hypothesis*. Based on molecular phylogeny, the endosymbiont that gave rise to mitochondria and MROs was of bacterial origin (Horner et al., 1996; Andersson et al., 1998), while the host was an archaeal cell (Cox et al., 2008; Williams et al., 2012; Spang et al., 2015). The hypothesis predicts that anaerobic metabolism of the bacterial ancestor of MROs produced molecular hydrogen as a waste product, which was then being used as an electron source for the autotrophic metabolism of the archaeon. Similar inter-domain symbiotic relationships with hydrogen as an electron transport intermediate are not rare (Stams and Plugge, 2009) and occur abundantly in Earth's crust and marine sediments.

The “mitochondria-early” theories imply further large-scale endosymbiotic gene transfer from bacterial symbionts to the archaeal host chromosomes, replacing the host’s archaeal pathways and membranes with bacterial counterparts and “transforming the archaeon from within” (Timmis et al., 2004; Martin et al., 2015), even though the mechanism of the archaeal lipid membrane replacement is not clear (Gould et al., 2016). The endosymbiont was gradually transformed into an organelle—the process which involved the reductive evolution of the symbiont genome (Gray et al., 1999; Timmis et al., 2004), evolution of protein transporters of inner and outer mitochondrial membranes (TIMs and TOMs, Dolezal et al., 2006; Kulawiak et al., 2013), and, most notably, the ATP/ADP carrier (AAC) capable—at least in modern eukaryotes—of exporting mitochondrial ATP in exchange for cytosolic ADP (Klingenberg, 2008). The hydrogen hypothesis itself does not predict when or how the AAC was established, only that ATP export was not an initial benefit of the symbiosis (Martin and Müller, 1998). Indeed, given the complete absence of sequences homologous to AAC proteins among present-day prokaryotes, it is unlikely that the mitochondrial carrier proteins were present from the very beginning of the endosymbiosis.

More recently, a competing view has been presented in the *phagocytic archaeon* theory (Poole and Neumann, 2011; Martijn and Ettema, 2013)—a prokaryote possessing actin cytoskeleton, membrane vesicle trafficking machinery and capable of primitive phagocytosis, with bacterial genes acquired via multiple pre-endosymbiotic HGT events. In this view, the precursor of the nuclear membrane might have formed as a strategy to protect the host genome from frequent and destabilizing horizontal gene transfer events. The hypothesis draws support from multiple archaeal homologues of proteins involved in eukaryotic membrane trafficking systems, such as ESCRT-III complex (Koonin, 2015; Spang et al., 2015; Klinger et al., 2016; Zaremba-Niedzwiedzka et al., 2017), and a recent attempt to resolve the order of eukaryogenic events via molecular phylogeny (Pittis and Gabaldon, 2016).

While it is not known how the mitochondrial endosymbiont was acquired (Poole and Gribaldo, 2014), *bona fide* eukaryotic phagocytosis might not have been necessary given that cases of bacteria invading bacterial hosts—although not archaeal—are known (Guerrero, 1986; von Dohlen et al., 2001). Archaeal sequences bearing similarity to the eukaryotic signature proteins do not all by themselves indicate the presence of phagotrophy either (Samson and Bell, 2009; Koonin, 2009; Makarova et al., 2010; Dey et al., 2016). Similarly, multiple independent HGT events into the archaeal host genome are not necessary to explain the apparent branching of eukaryotic sequences with multiple contemporary bacterial lineages—the same pattern is easily explained by a single endosymbiotic acquisition and continuous lateral gene transfer among the free living prokaryotes (Ku et al., 2015).

As correctly noted by Poole and Gribaldo (2014) the current evidence—most notably, the lack of evolutionary intermediates in the prokaryote-eukaryote transition—is in principle compatible with multiple hypotheses of eukaryote origin. Since key eukaryotic features are ancestral to the group, establishing the relative timing of their origins is difficult, if not impossible. Most attempts to resolve the order of eukaryogenic events, and to determine the nature of the symbiotic relationship, remain speculative.

1.4 Energetics of the prokaryote-eukaryote transition

Morphological complexity of the eukaryotic cell is unparalleled in the bacterial world. The full set of hallmark eukaryotic traits was present in an already complex LECA—the nucleus housing a relatively large genome organized into linear chromosomes, internal membrane systems such as the endoplasmic reticulum and Golgi, dynamic cytoskeleton, intracellular transport, peroxisomes, mitosis and meiotic sex, among many others (Koonin, 2010; Koumandou et al., 2013). The origin of this striking structural and genomic complexity in a pre-eukaryotic lineage poses a considerable and polarizing challenge to the current evolutionary theory: there are no true evolutionary intermediates between prokaryotes and eukaryotes, and phylogenetic analyses are limited by the chimeric nature of the eukaryotic genome and continuous horizontal gene transfer among prokaryotes of past and present (Koumandou et al., 2013; Poole and Gribaldo, 2014). The timing and mechanism of mitochondrial acquisition occupies a central position in these debates.

Lane and Martin (2010) approach the problem from a unique perspective of genome organization in relation to cellular bioenergetics. Large eukaryotic genomes with high protein expression levels are energetically costly, and this cost has to be met by the ATP produced mostly by respiratory chains embedded within bioenergetic membranes. In prokaryotes these membranes surround the cell, constraining the size and complexity of the individual due to shrinking surface-area to volume ratio. The only thermodynamically feasible way to make the transition, Lane and Martin (2010) argue, is through the endosymbiosis in which energy-generating membranes are internalized, while the size of local genomes needed to support the bioenergetic function of oxidative phosphorylation is reduced to a bare minimum. The acquisition of a mitochondrial organelle through endosymbiosis was therefore a critical event in

overcoming the bioenergetic constraints on prokaryotic genome size, and a prerequisite for the subsequent evolution of eukaryotic complexity.

1.5 Transforming endosymbiont into an organelle

Endosymbiotic hypotheses explain how the evolutionary transition from prokaryotes to eukaryotes was initiated. Lane and Martin's (2010) energetic considerations suggest that the acquisition of mitochondria—or, to be more precise, conversion of the bacterial symbiont into an ATP-exporting organelle with a small genome—permitted the expansion of the host cell genome to the size typical for modern eukaryotes. While bacterial-archaeal endosymbioses with large-scale horizontal gene transfer events were not necessarily rare (Nelson-Sathi et al., 2012, 2015), the mitochondrion arose only once. Even in the light of these seminal hypotheses, it remains unclear how and under what selective forces a bacterium within the host's cytoplasm acquired the protein import machinery of TIMs (protein translocators of the inner membrane) and TOMs (translocators of the outer membrane), and mitochondrial carrier proteins ultimately transforming it into an energy-producing organelle. Was the organellogenesis driven by the archaeal host cell or the endosymbiont?

Gross and Bhattacharya (2009) argue that the symbiont-to-organelle transformation was driven by the host, endorsing the widespread view that the prokaryotic endosymbionts were essentially "enslaved" by the pre-eukaryotic host cell (Cavalier-Smith, 2006). The argument is based on the assumption of constant selective pressure on the host cell to optimize the ATP production of the nascent organelle for the benefit of the group, i.e. the host and its endosymbiont population. In this model, the evolution of mitochondrial protein importers started at the outer endosymbiont membrane with the establishment of nuclear-coded β-barrel TOM and SAM (outer-membrane sorting and assembly machinery) complexes. The host proteins then gained access to the intermembrane space and the inner membrane of

the bacterium, establishing the TIM translocase complexes and mitochondrial carriers such as the AAC. In the view of Gross and Bhattacharya (2009), the organellogenesis was already underway before the initiation of the endosymbiotic ATP export, and the endosymbiont genome was already reduced and unable to encode its own protein import complexes, hence the selective pressure for the host protein insertion into the bacterial endosymbiont membranes.

An alternative view is that mitochondrial protein-import complexes and the mitochondrial solute carriers originated within the genome of the bacterial endosymbiont (Alcock et al., 2010). This view is supported by the apparent sequence homology and structural similarity of the main TIM and TOM complexes to contemporary bacterial proteins (Dolezal et al., 2006; Alcock et al., 2010), the curious observation that the import and assembly of TIM22—the complex responsible for the assembly of mitochondrial inner membrane carriers and TIM's—is dependent on its own presence within the inner membrane (Neupert and Herrmann, 2007) and possibly their tight coevolution with the respiratory chain complexes (Kutik et al., 2007; Kulawiak et al., 2013), suggesting their origin within the same genome. The endosymbiont's protein import machinery allowed the bacterium to reduce its genome size further, which quite possibly entailed a replicative advantage to individual endosymbionts.

If this view is correct, carrier proteins of the mitochondrial carrier family (MCF) evolved independently of the mitochondrial protein import machinery of TIMs and TOMs, and before the relocation of key bacterial genes into the host's genome, as their insertion into the inner membrane depended only on the pre-existing bacterial membrane protein insertion and assembly complexes, such as SecYEG and YidC (Driessen and Nouwen, 2008). The MCF proteins in organelles of modern eukaryotes transport a range of substrates, including carboxylates (malate, succinate, citrate), amino acids, nucleotides and dinucleotides (ATP, GTP, NAD^+), and protons (Kunji, 2004; Palmieri and Monné, 2016), but its original role in the nascent organelle remains unknown. One possibility is that it was used to import the cytosolic ATP to maintain

membrane potential at times of substrate shortage (Radzvilavicius and Blackstone, 2015), as has been demonstrated in ischemic liver cells (Belous et al., 2003) and cells lacking mitochondrial DNA (Giraud and Velours, 1997; Buchet and Godinot, 1998), or to import partially-oxidized organics.

1.6 Sex in the emerging complex cell

Proteins and pathways of the complex LECA descend from the ancestral sequences of symbiotic partners, modified through evolutionary tinkering, protein re-targeting and *de novo* evolution of new sequences (Kurland and Andersson, 2000; Koonin et al., 2004). LECA was already a complex cell with a nucleus, linear chromosome organization, internal membrane systems, motor proteins, mitosis and meiotic sex (Koonin et al., 2010). Some of these characteristic eukaryote features had to be established relatively early into the prokaryote-eukaryote transition, as the subsequent progress of the transition depended on their existence. The AAC is a prime example of such trait, as the massive energetically-constrained expansion of the proto-eukaryotic genome size (Lane and Martin, 2010) hinges upon the presence of the mitochondrial ATP/ADP exchanger. Whether originating in the endosymbiont or the host genome, the hallmark eukaryotic traits were all brought together into the lineage leading to LECA, with no surviving intermediates or early branching semi-eukaryotic species. Lane (2011) convincingly argued, that this pattern of eukaryotic trait evolution can only be consistent with the early evolution of sex—cell fusion and frequent recombination, laying the groundwork for the evolution of vertical mode of inheritance.

Genes required for nuclear fusion and meiosis permeate all eukaryotic groups, indicating that LECA was indeed a cell capable of eukaryotic sex (Ramesh et al., 2005; Schurko and Logsdon, 2008). This supports the view that the eukaryotic cell coevolved with sexual reproduction. What drove the evolution of sex in eukaryogenesis, however, is a matter of an ongoing debate. Alluding to the benefit of meiotic sex slowing down

the mutational meltdown, one group of theorists argues that sex evolved in response to the host genome damage generated by reactive oxygen species (ROS) produced within the mitochondrial respiratory chains (Gross and Bhattacharya, 2010; Hörandl and Hadacek, 2013; Speijer et al., 2015). Mitochondria of modern aerobic eukaryotes constitute a major source of intracellular ROS, where reduced complexes of the electron transport chain can readily donate electrons to molecular oxygen forming superoxide, hydrogen peroxide and, through Fenton chemistry, highly reactive hydroxyl radicals (Hallivell, 2006; Murphy, 2009). Gross and Bhattacharya (2010) argue that a single selective pressure—the rise in atmospheric oxygen levels and the mutagenicity of either environmental or internal ROS—can account for both the acquisition of mitochondria and the origin of sex. Aerobic bacterial symbionts in this model cleared the local environment of oxygen, which would have been toxic to the anaerobic archaeal host, although it is not clear why internalizing the source of reactive oxygen species would be selected for. Sex with reciprocal recombination among the host's chromosomes in this model evolves from archaeal conjugation in response to the frequent ROS-induced damage to host's genome. Likewise, in the model envisaged by Speijer et al. (2015), the pre-eukaryotic endosymbiosis occurs in an aerobic environment, where the phagocytic uptake of the bacterium results in high internal concentrations of reactive oxygen radicals within the cell, prompting the evolution of sex, peroxisomes, internal membranes and the protective nuclear compartment.

Internalization of the bacterial symbiont provided the host with a source of not only ATP, but also new genes that were migrating into the host's genome (Timmis, 2004) and appear to be largely responsible for the initial host genome expansion. Koonin (2006, 2009) argues, that this unidirectional flow of DNA from decaying endosymbionts included bacterial self-splicing group II introns that induced structural transformations of the host's genome. Proliferation and recombination of introns

scattered across the host's genome was highly mutagenic, and caused disintegration of the circular prokaryotic chromosomes of the host into linear pieces.

Lane (2011) reasons that small effective size of the nascent eukaryotic population—which was needed for the introns to fix in the first place (Koonin, 2009)—and high mutation rate together with general genome instability strongly favoured the evolution of sex with cell fusion and frequent reciprocal recombination. This is in stark contrast to prokaryotes, which seem to be evading the effects of mutational meltdown through HGT and unidirectional homologous recombination within a population of large effective size. Variable numbers of newly formed straight chromosomes, imprecise chromosome segregation at cell division and variable sets of genes of endosymbiont origin created a strong selective pressure for the evolution of cell fusions and genome duplications, restoring fully viable chromosome sets and masking detrimental mutations (Garg and Martin, 2016). In this view, the emerging eukaryotic cell was rescued by the early evolution of sex via cell fusion and recombination, which stabilized the host genome allowing it to expand further, e.g. through the acquisition of endosymbiont DNA and gene duplications (Makarova et al., 2005).

Cell fusion with cytoplasmic mixing curtails variance between host cells, reducing the efficacy of selection against detrimental symbionts on the level of the host (Chapter 2). Intermixing between the large endosymbiont populations of unrelated cells would therefore promote competition and facilitate the spread of selfish endosymbionts, to the detriment of the host cell (Radzvilavicius and Blackstone, 2015; Blackstone, 2016). On the other hand, with synergistic interactions between deleterious endosymbiont mutations, cell fusion-fission cycles could in fact mask the deleterious effects of mitochondrial mutations (Chapter 2). Recent theoretical analysis, presented also in this thesis, suggests that this effect could easily drive the fixation of host-cell alleles inducing cell fusion within the population of clonal proto-eukaryotes (Radzvilavicius, 2016a). Strikingly, the analysis predicts that when the cell fusion first arose in a nascent eukaryotic cell, it had to be frequent and clonal reproduction

relatively rare. This prediction is in agreement with the view that clonal cell divisions might have produced inviable combinations of the newly formed linear chromosomes, requiring frequent cell fusions, or even the syncytial state, to maintain viability (Garg and Martin, 2016).

1.7 Fundamental asymmetry of eukaryotic sex

In contrast to the fully symmetric recombination among nuclear chromosomes, mitochondrial inheritance in modern eukaryotes is virtually always biased towards one of the gametes (Birky, 2001). Typically, the gametes themselves belong to two or more self-incompatible mating types, with one of them contributing most of the mitochondrial DNA to the zygote, the organelle genomes of the other being discarded. While phylogenetic analyses trace the origin of sex before the last common ancestor of all eukaryotes, the mechanisms responsible for asymmetric inheritance of mitochondria are not conserved (Birky, 1995; Xu, 2005; Sato and Sato, 2013; Greiner et al., 2015), making the inference of the evolutionary timing of its origin virtually impossible. There is no reason to believe that the initial form of sex originating before LECA was already asymmetric in cytoplasmic transmission, but the ubiquity of the uniparental inheritance (UPI) among extant eukaryotes is nevertheless indicative of a strong and shared selective pressure favouring the asymmetry in transmission of mitochondrial genes.

The earliest explanations for the prevalence of UPI invoked inter-genomic conflicts and envisaged the evolution of uniparental inheritance together with mating types or sexes (Cosmides and Tooby, 1981; Hastings, 1992; Hurst and Hamilton, 1992; Hutson and Law, 1993), although others argue that mating types evolved for unrelated reasons (reviewed in Billiard et al., 2011). More recently, Lane (2011b) and Hadjivasiliou et al. (2012) suggested that uniparental inheritance evolved to improve the co-adaptation between nuclear and mitochondrial genomes, Christie et al. (2015) invoked simple selection against heteroplasmy, while Bendich (2013) argued that UPI

evolved as a mechanism of “DNA abandonment”, in which organelles carrying damaged DNA are discarded to prevent mutation accumulation.

Greiner et al. (2015) maintains the view that uniparental inheritance suppresses the spread of selfish cytoplasmic elements, but leads to the accumulation of deleterious mutations in the long term, and that periodic reversals to the biparental mode of organelle inheritance are needed to halt Muller’s ratchet in mtDNA. While there was some early evidence of excess deleterious polymorphisms in non-recombining mitochondrial genomes (Rand and Kann, 1996; Weinreich and Rand, 2000), a recent analysis of larger gene sets in humans and flies showed that in spite of smaller effective population size, mitochondrial loci experience similar efficacy of purifying selection as loci in the recombining nuclear genome (Cooper et al. 2015). In contrast to Greiner’s predictions, uniparental inheritance could be the mechanism maintaining strong purifying selection against detrimental mitochondrial mutations. Theoretical analyses indeed show that UPI increases variance in the mutational load between individuals, facilitating purifying selection against detrimental mitochondrial mutations, while symmetric organelle inheritance promotes the accumulation of mitochondrial defects (Bergstrom and Pritchard, 1998; Hadjivasiliou et al., 2013; Radzvilavicius, 2016a).

The fundamental asymmetry of sex is perhaps best reflected in complex multicellular organisms with two sexes—males, producing small and often motile sperm, and females producing large oocytes. Extreme gamete-size dimorphism appears to have evolved multiple times in organisms that already had mating types, and is virtually universal among complex multicellular organisms. For many, the satisfactory explanation for the evolution of oogamy has been provided by Parker et al. (1972) in what is known as the “disruptive selection” or the PBS model (Bell, 1978; Parker, 1978; Bulmer and Parker, 2002). Large immotile gametes in this model maximize their viability and survival, while numerous small and motile gametes evolve to increase the chance of successful fertilization, which under certain conditions leads

to divergence into females and males. Since multicellular development requires substantial initial investment, production of large eggs was predicted to correlate with organism size (Parker et al., 1972; Knowlton, 1974; Randerson and Hurst, 2001).

On the other hand, gamete size dimorphism has been interpreted as a means to regulate organelle or endosymbiont inheritance (Cosmides and Tooby, 1981; Hurst, 1990; Hastings, 1992). Even though UPI was already present in ancestral isogamous populations, given its apparent evolutionary instability it is still plausible that oogamy represents one of the mechanisms to reduce the amount of paternal organelles transmitted to the zygote. This view is supported by empirical observations of males actively discarding their own mitochondrial DNA in spermatogenesis (Nishimura et al., 2006; DeLuca and O'Farrell, 2012; Luo et al., 2013), even though the classical mathematical models predict that such paternally-controlled organelle destruction might not be possible to evolve (Randerson and Hurst, 1999; Hadjivasiliou, 2013; but see Chapter 4).

Mitochondrial inheritance in a vast majority of complex multicellular eukaryotes is matrilineal, exerting disparate selective forces on female and male germline development. Allen (1996) proposed that female germline cells act to preserve "template" mitochondrial DNA by repressing oxidative phosphorylation and reducing the production of damaging ROS (reactive oxygen species), while male germ cells specialize in short-term energy production required for motility, actively transcribe their mitochondrial DNA and, since they normally do not pass their mitochondria to the zygote, do not employ the same defence mechanisms against mitochondrial damage. The hypothesis was supported by mitochondrial activity studies in jellyfish, flies and zebrafish (dePaula et al., 2013a, 2013b), although a different pattern with active mitochondria in gametes of both sexes has been observed in a bivalve mollusc with doubly-uniparental inheritance of mtDNA (Milani and Ghiselli, 2015).

Consistent with the protective role of female germline is the observation that the number of germline cell divisions in females is typically lower than in males (Drost

and Lee, 1995; Crow, 2000), but only in higher metazoans with high mitochondrial copying error rates. In early-branching animals (e.g. *Porifera*, *Anthozoa* and *Placozoa*) with extremely low mtDNA evolution rates, gametes are generated from the multipotent stem cells that also give rise to somatic tissues. Mitochondrial quality in these organisms could instead be maximized through high numbers of cell divisions generating intra-cellular variance in mitochondrial mutation load and facilitating purifying selection (Radzvilavicius et al., 2015; Radzvilavicius, 2016a), or efficient molecular repair mechanisms (Hellberg, 2006). Likewise, the mitochondrial bottleneck in the female germline is thought to redistribute mtDNA variation from within the cell to between the gametes and organisms, facilitating purifying selection against defective mitochondria (Bergstrom and Pritchard, 1998; Roze et al., 2015; Haig, 2016), even though there is still no consensus on what causes the increase in variance, and whether it involves a significant reduction in mtDNA copy number beyond the normal level found in cells (Wai et al., 2008; Cao et al., 2007, 2009; Cree et al., 2009). Krakauer and Mira (1999) interpreted the massive atretic germ cell death in females, in which a vast majority of germ cells are eliminated, as an additional mechanism of mitochondrial quality control through the intra-organismal purifying selection. Recent experimental evidence in mice provided some evidence for the germline selection against severe mitochondrial mutations, although the exact mechanisms, again, are not known (Fan et al., 2008). While the results reported so far seem promising, further empirical studies with a broader variety of non-model organisms are clearly needed to elucidate the role of female germline organization in mtDNA quality preservation.

1.8 Concluding remarks and structure of the thesis

Unlike horizontal gene transfer in bacteria, eukaryotic sex involves reciprocal recombination across the full length of the chromosomes, resulting in vertical lineages that contain sets of alleles belonging to one particular, species-defining gene set. The

symmetry of nuclear gene exchange is in stark contrast to the high overall asymmetry of eukaryotic sex. Nowhere is the asymmetry of sexes more evident than in higher animals, where it manifests itself through the extreme divergence in sex roles: secondary sex characters such as ornaments or weapons in males, competing for mating opportunities with the choosier females. Strikingly, the sex-role divergence can be traced back to the gamete-size and number dimorphism—the core distinction between the two sexes (Trivers, 1972; Lehtonen et al., 2016), which itself is most often attributed to disruptive selection (Parker et al., 1972; Bulmer and Parker, 2002).

More recent work, however, points to the key role of organelle genomes in the evolution of sexual asymmetry. Multiple traits of female germline development, such as the low number of germline cell divisions, metabolic and transcriptional quiescence, mitochondrial bottlenecks and atretic germ cell death can all be attributed to the fundamental need to preserve the quality of maternally inherited mitochondrial genes. Male germline in higher metazoans, on the other hand, can be seen as specializing in mass production of motile reproductive cells with no specific adaptations related to the mitochondrial quality, as mitochondrial genes are inherited mostly through the maternal germline. While not strictly associated with gamete size, uniparental inheritance of mitochondria can be viewed as one of the selective forces driving the evolution of oogamy, and, likely, the mating types. Sexual reproduction itself can be attributed to purifying selection against mitochondrial mutations in early eukaryotic lineages (Radzvilavicius, 2016a).

In this thesis I take the extra-nuclear perspective to the evolution of eukaryotic sexual traits. In Chapter 2, I show that the evolution of sexual cell fusion in the nascent eukaryotic lineage might have been driven by cytoplasmic mixing, temporarily masking the detrimental effects of defective organelles. The model introduced in Chapter 3 shows that self-incompatible mating types can evolve to ensure the efficient removal of mitochondrial mutations through the asymmetric organelle transmission.

Frequent observations of paternal leakage and heteroplasmy pose a substantial challenge to the current understanding of uniparental inheritance. In Chapter 4 I show that the evolutionarily stable pattern of cytoplasmic inheritance depends on which sex—male or female—governs the destruction of paternal organelles. Maternal regulation favours complete elimination of sperm organelles, while males could favour paternal leakage and heteroplasmy. Finally, in Chapter 5 I propose a new hypothesis on the origin of germline sequestration in metazoans in which purifying selection against mitochondrial mutations plays a central role. I develop a simple mathematical model that shows that high mtDNA copying error rates in bilaterians favour reduced numbers of germline cell divisions, while in basal metazoans with slow mtDNA evolution, mitochondrial quality is maximized through segregational drift in multipotent stem cell lineages with unconstrained division.

1.9 Publications

Several articles based on the material presented in this thesis have been published in peer-reviewed journals or submitted for publication. Chapter 2, with little modification, has been published in the *Journal of Theoretical Biology* (Radzvilavicius 2016a), while the stochastic model supporting some of the results presented within the same chapter has been published in *Journal of the Royal Society Interface* (Radzvilavicius and Blackstone, 2015). Chapter 4 is based on the article by Radzvilavicius, Lane and Pomiankowski recently submitted to *Nature Ecology and Evolution*, while the manuscript elaborating on the hypothesis of Chapter 5 has been published in *PLOS Biology* and is also stored in *bioRxiv* preprint server (Radzvilavicius et al., 2015). A critique on the recently proposed mitochondrial-erosion hypothesis for the evolution of eukaryotic sex (Havird et al., 2015) has been published in *BioEssays* (Radzvilavicius, 2016b).

CHAPTER 2. EVOLUTIONARY DYNAMICS OF CYTOPLASMIC SEGREGATION AND FUSION: MITOCHONDRIAL MIXING FACILITATED THE EVOLUTION OF SEX AT THE ORIGIN OF EUKARYOTES

2.1 Summary

Sexual reproduction is a trait shared by nearly all complex life, but explaining its origin and evolution remains a major theoretical challenge. Virtually all theoretical work on the evolution of sex has focused on the benefits of reciprocal recombination among nuclear genes, paying little attention to the dynamics of mutation in the mitochondrial genome. Here I develop a mathematical model to study the evolution of alleles inducing cell fusion in an ancestral population of clonal proto-eukaryotes. Mitochondrial mixing masks the detrimental effects of faulty organelles and drives the evolution of sexual cell fusion despite the declining long-term population fitness. Cell-fusion alleles fix under negative epistatic interactions between mitochondrial mutations and strong purifying selection, low mutation load and weak mitochondrial-nuclear associations. I argue that similar conditions could have been maintained throughout eukaryogenesis, favouring the evolution of sexual cell fusion and meiotic recombination without compromising the stability of the emerging complex cell.

2.2 Introduction

Sexual reproduction with gamete fusion and reciprocal recombination is among the traits shared by virtually all eukaryotes (Ramesh et al., 2005; Goodenough and Heitman, 2014; Speijer et al., 2015). Current views on the evolutionary advantage of sex have it that recombination among nuclear genes exposes the hidden genetic variation in finite populations, breaks up unfavourable allele combinations under fluctuating selection, or rescues the genome from the mutational meltdown (Otto, 2009; Hartfield and Keightley, 2012). These views, however, are based on the long-term effects of recombination, and do not directly explain when or how these traits first arose. It is becoming increasingly clear that sex first appeared as a part of the evolutionary transition from prokaryotes to eukaryotes, most likely after the endosymbiotic acquisition of mitochondria (Gross and Bhattacharya, 2010; Lane and Martin, 2010; Speijer et al., 2015). There is therefore more than just recombination to the origin of sex, and a full account of the evolution of sexual reproduction has to account for the complex relationship between mitochondrial symbionts and the host.

The evolution of complex life can be conceptualized as a sequence of major evolutionary transitions (Buss, 1987; Maynard Smith and Szathmary, 1995). With each transition conflicts between the levels of individuality arise and have to be mediated for a stable higher-level unit to be established (Michod, 1997; Michod and Nedelcu, 2003; West et al., 2015). Conflict resolution often involves mechanisms reducing genetic variance within groups of lower-level units, eliminating the scope for defection and detrimental competition. In contrast, cell fusion allows for cytoplasmic mixing and the horizontal spread of detrimental mutants, facilitating evolutionary conflict and reducing the efficacy of purifying selection (Hastings, 1992; Bergstrom and Pritchard, 1998; Randerson and Hurst, 1999). The origin of cytoplasmic mixing at the early stages of eukaryogenesis therefore could have hindered the evolution of a stable higher-level unit—the eukaryotic cell (Radzvilavicius and Blackstone, 2015). While two mating

types and uniparental inheritance (UPI) might eliminate these issues in modern eukaryotes (Birky, 1995; Hadjivasiliou et al., 2013; Sato and Sato, 2013; Greiner et al., 2015), the mechanism of asymmetric inheritance presumably would not have been present during the early evolution of sex.

In their recent article, Havird et al. (2015) suggest a novel hypothesis for the evolution of eukaryotic sex, in which mitochondrial mutations play a central role. Owing to its high mutation rate, mitochondrial DNA (mtDNA) can quickly accumulate mutations compromising the function of the respiratory chain and diminishing a cell's viability, which prompts the evolution of compensatory nuclear modifications that could restore the cell's fitness. Recombination among the nuclear genes would potentially increase the rate at which new compensatory combinations of nuclear alleles are introduced, rapidly improving the match between the two genomes. While important in some ways, the hypothesis has several shortcomings (Radzvilavicius, 2016b) and does not account for one of the hallmark features of mitochondrial genetics—cytoplasmic segregation (Rand, 2008, 2011)—which facilitates the elimination of mitochondrial mutations via purifying selection at the level of an organism.

Another recent idea highlighting the role of mitochondria in the evolution of sex stems from the evolutionary history of endosymbionts and the biochemistry of cellular respiration (Blackstone and Green, 1999). Faced with stressful conditions constraining their growth and proliferation, proto-mitochondrial symbionts could have systematically manipulated the host cell phenotype using the by-products of oxidative phosphorylation. High emissions of reactive oxygen species (ROS), for example, could have served as a trigger for the host cell fusion and recombination, restoring favourable conditions for the endosymbiotic growth and proliferation. Similarly, for Speijer et al. (2015), mitochondrial acquisition gave rise to sex due to the ROS-induced genome damage and the need for frequent recombinational repair (see also Gross and Bhattacharya, 2010; Horandl and Hadacek, 2013).

While it is very likely that sex evolved in a cell that already possessed

mitochondria (Lane and Martin, 2010; Lane, 2014; Speijer et al., 2015), the conditions favouring the emergence of cell fusion and cytoplasmic mixing under these circumstances have not received substantial attention. Multiple factors are likely to affect the evolutionary dynamics of mitochondrial mixing and segregation, including the intensity of purifying selection, mutation rate, epistatic interactions, intracellular competition and the properties of early cell cycles, but the relative importance of these effects in the early evolution of meiotic sex is not known. Here I introduce an infinite-population model to study cytoplasmic segregation at the origin of sexual cell fusion. I analyse the spread of proto-nuclear alleles inducing cell fusion with cytoplasmic mixing under the effect of purifying selection against mitochondrial mutations. The model suggests a set of conditions promoting the emergence of sex in the form of eukaryotic cell fusion, and strongly supports the view that mitochondria could have represented one of the driving forces behind the origin of sexual life cycles.

2.3 Mathematical model for cytoplasmic segregation and the evolution of sexual mixing

2.3.1 Neutral segregational drift with clonal host reproduction

The eukaryotic cell can be modelled as a collective of mitochondria within a cytosol that also contains the host genome (I assume here that similar conditions also applied early in eukaryogenesis, i.e. after the acquisition of mitochondria). Consider an infinite population of cells, containing M mitochondria each and reproducing clonally. Mitochondria are found in one of two possible states, wild-type or mutant. Clonal reproduction is modelled by first duplicating the mitochondrial population of the cell and then randomly partitioning organelles to the two daughter cells through random sampling without replacement. A cell containing m mitochondrial mutants will give birth to a daughter with q mutations with probability

$$r(q|M,m) = \binom{2m}{q}\binom{2M-2m}{M-q}\binom{2M}{M}^{-1}. \quad (2.1)$$

The frequency distribution for the number of mutants per cell $p(x)$ after one round of clonal reproduction will therefore change to $p(q) = \sum_{x=0}^{M} r(q|M,x)p(x)$. The process of cytoplasmic segregation modelled in this manner represents a type of neutral genetic drift, conceptually similar to the Wright-Fisher process (Ewens, 2004), but assuming a finite group size at all stages of the life cycle.

For the initial frequency of mitochondrial mutations within the cell $f_0 = m_0/M$, variance after n clonal divisions is (see Appendix A for the details)

$$\mathrm{Var}(F_n) = f_0(1-f_0)\left[1-\left(1-\frac{1}{2M-1}\right)^n\right]. \quad (2.2)$$

Segregational drift therefore increases variance between host cells, which after a large number of clonal reproduction cycles converges towards $f_0(1-f_0)$, where all mutants have either reached fixation within the cell or were replaced by the wild type mitochondria. It follows then that the probability of reaching fixation is equal to the initial mutant frequency f_0.

It can be similarly shown (see Appendix A) that neutral segregation increases the homogeneity within the cell. The probability for two segregating units within the cell to be identical by descent can be expressed as

$$\varphi_n = \frac{\mathrm{Var}(F_n)}{f_0(1-f_0)} = 1-\left(1-\frac{1}{2M-1}\right)^n. \quad (2.3)$$

After a large number of clonal divisions, the identity-by-descent probability approaches one, $\varphi_\infty \to 1$, at which point the mitochondrial populations are fully clonal and no further change is possible.

2.3.2 Evolution of cytoplasmic mixing

Now consider a full population life cycle with mitochondrial mutation, selection and reproduction (Fig. 2.1). I assume an ancestral state without sex or cell-cell fusion,

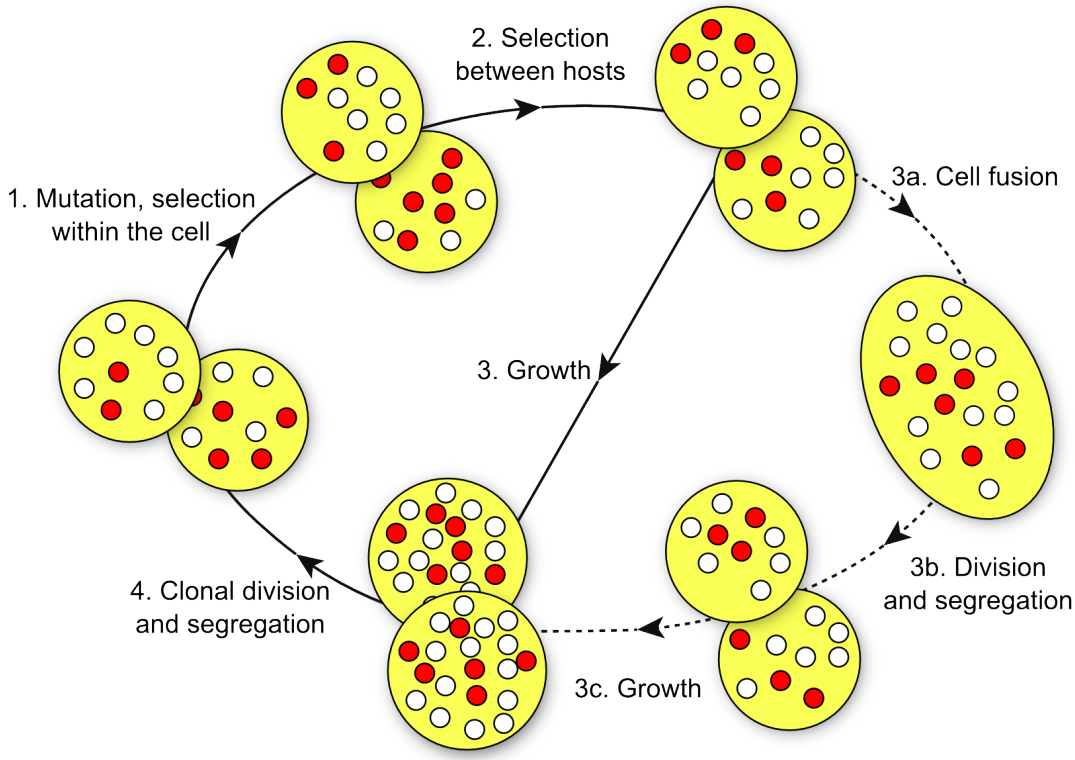

Figure 2.1. Population life cycle with asexual and sexual modes of reproduction. Red circles represent deleterious mitochondrial mutants within the host cell, wild-type organelles are left blank. Steps 1–4 (solid arrows) represent the life cycle of clonally reproducing individuals. Steps 3a–3c (dashed arrows) occur only if one of the cells meeting at random is a carrier of the cell fusion allele *H*.

where haploid hosts reproduce clonally as described above. The mode of reproduction is controlled by a single locus in the host's haploid genome, *h*/*H*. Only mutants carrying a copy of the allele *H* can initiate a temporary cell fusion with a randomly chosen partner (steps 3a-3c in Fig. 2.1) before proceeding to standard clonal reproduction.

The population state at generation t is represented by a $(M+1)\times 2$ matrix $\mathbf{P}^{(t)}$ with the matrix element $P_{m,n}^{(t)}$ depicting the frequency of cells with m mitochondrial mutants and the nuclear state n $(n = 0, 1)$. The column vector $\mathbf{P}_{\bullet,0}^{(t)}$ therefore corresponds to the wild type population with allele *h* and column $\mathbf{P}_{\bullet,1}^{(t)}$ contains entries pertaining to the cells with the nuclear allele *H*.

Mutation. The genotype-changing events in the population life cycle can be represented as matrix operations changing the population state $\mathbf{P}^{(t)}$. The population state after the mutation is therefore given by $\mathbf{P}^{(t,1)} = \mathbf{U}\mathbf{P}^{(t)}$, where $\mathbf{U}$ is

$(M+1)\times(M+1)$ transition matrix, with the element $U_{i,j}$ defined as the probability that a cell with j mutant mitochondria will contain i mutants after the transition. Mitochondrial mutation at the rate μ is modeled as a binomial event, giving the transition probabilities

$$U_{i,j} = \binom{M-j}{i-j} \mu^{i-j} (1-\mu)^{M-i}. \quad (2.4)$$

Selection on the lower level. In the case where mutant mitochondria have a competitive advantage within the cell ("selfish" mutants), the mutation step is followed by selection on the lower level. Selection among mitochondria of the same cell is modelled as a random sampling with replacement, with the probability to select a selfish mutant proportional to its replicative advantage $1+\kappa$. The population state after selection on the lower level is therefore $\mathbf{P}^{(t,2)} = \mathbf{W}\mathbf{P}^{(t,1)}$, where the binomial transition probabilities of matrix $\mathbf{W}$ are

$$W_{i,j} = \binom{M}{i} \left[\frac{j(1+\kappa)}{M+j\kappa}\right]^{i} \left(\frac{M-j}{M+j\kappa}\right)^{M-i}. \quad (2.5)$$

Selection between eukaryotic hosts. Selection at the higher level changes the relative genotype frequencies according to the host cell fitness. In matrix notation, the population state after selection is

$$\mathbf{P}^{(t,3)} = \frac{(\mathbf{I}\mathbf{w})\mathbf{P}^{(t,2)}}{\mathbf{w}^{\mathrm{T}}\mathbf{P}^{(t,2)}\overline{\mathbf{u}}_2}, \quad (2.6)$$

where $\mathbf{I}$ is the identity matrix, $\overline{\mathbf{u}}_2$ is a column vector of ones $(1,1)^{\mathrm{T}}$, and $\mathbf{w}$ is a column vector with the m-th element $w_m = \omega(m)$ corresponding to the fitness of a cell containing m mutants. Following the models by Hadjivasiliou et al. (2013) and Kuijper et al. (2015), I assume that the relative fitness of the cell depends on the number of mitochondrial mutants m and can be expressed as $\omega(m) = 1 - s(m/M)^{\xi}$. Parameter s here represents the strength of selection and ξ determines the strength of epistatic interactions between mitochondrial mutations. Empirical studies suggest that in modern eukaryotes $\xi > 1$, leading to so-called phenotypic threshold effects (Rossignol

et al., 2003). The relative effect of each new mutation therefore increases with the overall mutation load.

Reproduction and mitochondrial segregation. Cells carrying the nuclear allele *H* are capable of cytoplasmic fusion with the other *H*-type individuals as well as randomly chosen wild-type hosts *h*. Wild-type individuals do not initiate cell fusion, and mix their cytoplasmic contents only if randomly chosen by an individual carrying the allele *H*. The process of cell fusion in our model is represented by the convolution of corresponding genotype-frequency vectors, forming a temporary subpopulation of diploid zygotes each containing $2M$ mitochondria. Fusion is immediately followed by cell division with random partitioning of mitochondria between the two daughter cells. The population state after the sexual stage of the life cycle can then be expressed as

$$\begin{aligned}\mathbf{P}_{\bullet,0}^{(t,4)} &= \mathbf{P}_{\bullet,0}^{(t,3)}\overline{\mathbf{u}}_{M+1}^{\mathrm{T}}\mathbf{P}_{\bullet,0}^{(t,3)} + \mathbf{K}\left(\mathbf{P}_{\bullet,0}^{(t,3)} * \mathbf{P}_{\bullet,1}^{(t,3)}\right),\\ \mathbf{P}_{\bullet,1}^{(t,4)} &= \mathbf{K}\left(\mathbf{P}_{\bullet,1}^{(t,3)} * \mathbf{P}_{\bullet,1}^{(t,3)}\right) + \mathbf{K}\left(\mathbf{P}_{\bullet,0}^{(t,3)} * \mathbf{P}_{\bullet,1}^{(t,3)}\right).\end{aligned} \tag{2.7}$$

Asterisks here denote vector convolution, and $\overline{\mathbf{u}}_{M+1}^{\mathrm{T}}$ is a row vector of $M+1$ ones, so that $\overline{\mathbf{u}}_{M+1}^{T}\mathbf{P}_{\bullet,0}^{(t,3)}$ is the total frequency of the allele *h*. $\mathbf{K}$ is the transition matrix for the reductive cell division without the prior replication of mitochondria, implemented as selection without replacement with transition probabilities (Eq. 2.1) $K_{i,j} = r(i|M, j/2)$, where $i \in [0, M], j \in [0,2M]$.

The life cycle ends with the standard clonal replication, first duplicating the mitochondrial population within each cell and then partitioning the organelles between the two daughter cells. This gives the updated population state at the start of the next generation $\mathbf{P}^{(t+1)} = \mathbf{S}\mathbf{P}^{(t,4)}$, with transition probabilities (Eq. 2.1) $S_{i,j} = r(i|M, j)$, where $i, j \in [0, M]$.

The model is initialized in a random mitochondrial state $\mathbf{P}^{(0)}$ so that the whole population initially consists only of the wild-type individuals, i.e. $\overline{\mathbf{u}}_{M+1}^{T}\mathbf{P}_{\bullet,0}^{(0)} = 1$ and $\mathbf{P}_{\bullet,1}^{(0)} = \mathbf{0}$. After the equilibrium is reached at time t_E, the allele *H* is inserted at a small

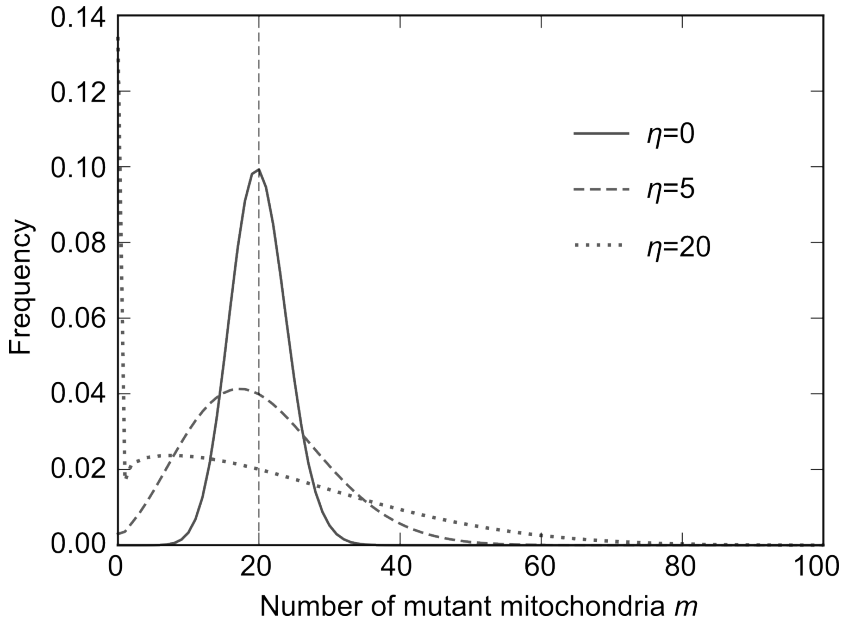

Figure 2.2. Cytoplasmic fusion opposes the effect of mitochondrial segregation of neutral mutations. Variance in the number of mutant mitochondria per cell increases in multiple rounds of clonal reproduction, but is reduced by cell fusion, resulting in an equilibrium when the two modes of reproduction alternate in time. η is the number of consecutive clonal generations without cytoplasmic mixing. The initial number of mutants per cell is $m_0 = 20$ (vertical line).

frequency $\chi = 0.001$, so that $\mathbf{P}_{\bullet,1}^{(t_E+1)} = \chi\, \mathbf{P}_{\bullet,0}^{(t_E)}$ and $\mathbf{P}_{\bullet,0}^{(t_E+1)} = (1-\chi)\, \mathbf{P}_{\bullet,0}^{(t_E)}$. In the following I present the results based on the numerical solution of the above system of equations for both equilibrium and transient states, obtained for multiple values of μ, κ, s, M and ξ.

2.4 Mitochondrial variation and the fitness cost of cytoplasmic mixing

2.4.1 Mitochondrial variation at the segregation–fusion equilibrium

Let us first look into the effect of recurrent cell fusion on mitochondrial variance between cells generated by cytoplasmic segregation. Starting with a fixed number of mutants per cell m_0 I allow for η clonal generations without cell fusion, followed by a single round of sexual reproduction. No mutation or selection occurs at this stage.

The results show that the effect of cytoplasmic mixing opposes the constant increase of mitochondrial variance between cells generated by segregational drift (Eq.

2.2), establishing an intermediate equilibrium (Fig. 2.2). While segregational drift alone results in diverging cell lineages with clonal intracellular populations of mitochondria, increasing the frequency of sexual cell fusion relative to the number of clonal generations η reduces the mitochondrial variation between cells, at the same time reducing homogeneity within the cell. Given the importance of heritable variance for the efficacy of selection on the higher level, frequent cell fusion could result in diminished population fitness, which I investigate further.

2.4.2 Mitochondrial mutation pressure

Here I return to the full population life cycle and analyse the effects of cellular fusion on the long-term population fitness. Mitochondrial mutants arise at a constant rate μ, but do not have an intra-cellular replication advantage over the cooperative organelles, i.e. $\kappa = 0$. Cell fusion rate is controlled by keeping the frequency of nuclear allele *H* at a constant level p_H, while allowing the mitochondrial population to evolve freely. Given that a cell carrying the *H* allele fuses with a randomly selected partner, the overall rate of sexual reproduction can be expressed as $R = p_H^2 + 2\,p_H(1 - p_H) = p_H(2 - p_H)$.

Owing to the effect of reduced variance in the number of mitochondrial mutants between cells, the long-term population fitness is reduced by increasing frequency of *H* (Fig. 2.3). In agreement with previous studies (Hadjivasiliou et al. 2013, Radzvilavicius et al., 2015), the detrimental effect of lower mitochondrial variance is more prominent with higher numbers of mitochondria per cell. Larger populations of segregating lower-level units dampen the effect of segregational drift (Eq. 2.2), reducing the efficacy of selection at the higher level.

2.4.3 Fast replicating “selfish” mutants

The fitness costs of cytoplasmic mixing can be exacerbated in the presence of so called “selfish” mitochondria—organelles that have gained mutations leading to the

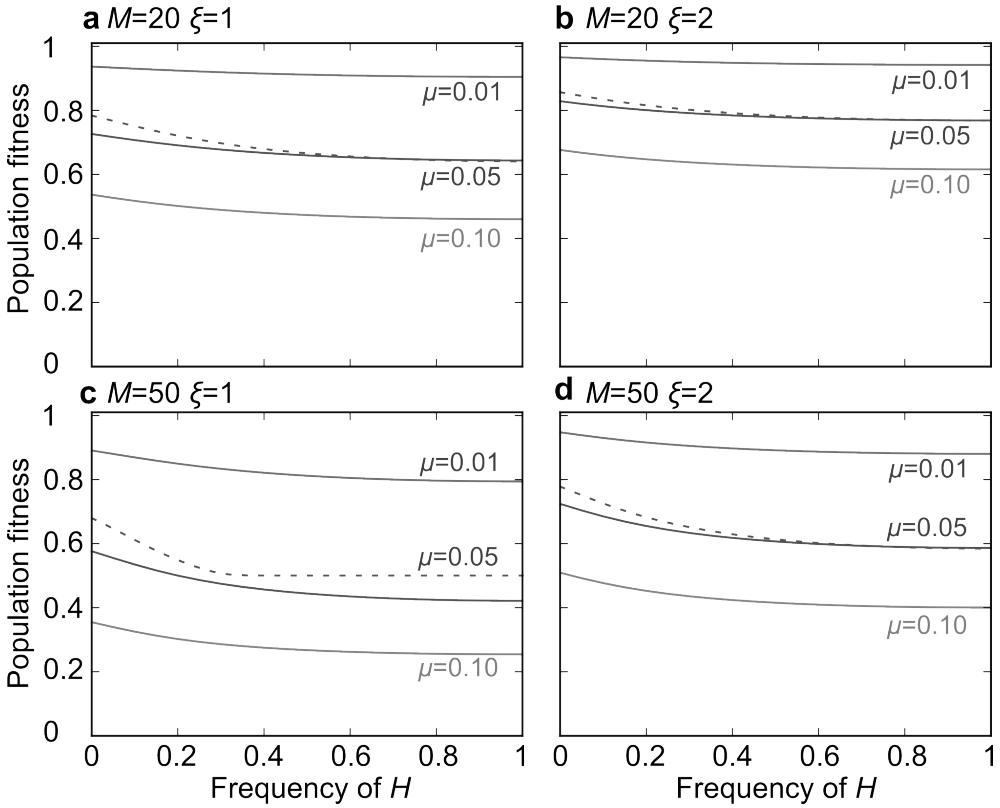

Figure 2.3. Frequent cytoplasmic mixing reduces the mean population fitness under mitochondrial mutation pressure. *H* is the cell fusion allele. M is the total number of mitochondria per cell, μ is the mitochondrial mutation rate and ξ is the strength of epistatic interactions. Selection strength is set to $s = 1$, except for $\mu = 0.05$, where dotted lines show the effect of weaker selection with $s = 0.5$.

faster reproduction rate but at the same time reducing their ability to participate in cooperative interactions. Selection at the lower level therefore increases the frequency of non-cooperative mitochondria, but cells with a significant proportion of selfish lower-level units suffer a fitness cost and replicate slower. This is an example of the evolutionary conflict between levels of selection, endangering the stability of the higher-level unit—the nascent eukaryotic cell (Schable and Wise, 1988; Taylor et al., 2002; Clark et al., 2012; Bastiaans et al., 2014). This time, I assume that mutants arise at a low rate ($\mu = 0.0001$) and proliferate mostly due to their ability to outcompete the cooperative mitochondria within the same cell.

The results confirm that clonal reproduction (low p_H) coupled with purifying selection at the higher level suppresses selfish mitochondrial competition within the cell and maintains high population fitness (Fig. 2.4). Unable to spread horizontally in

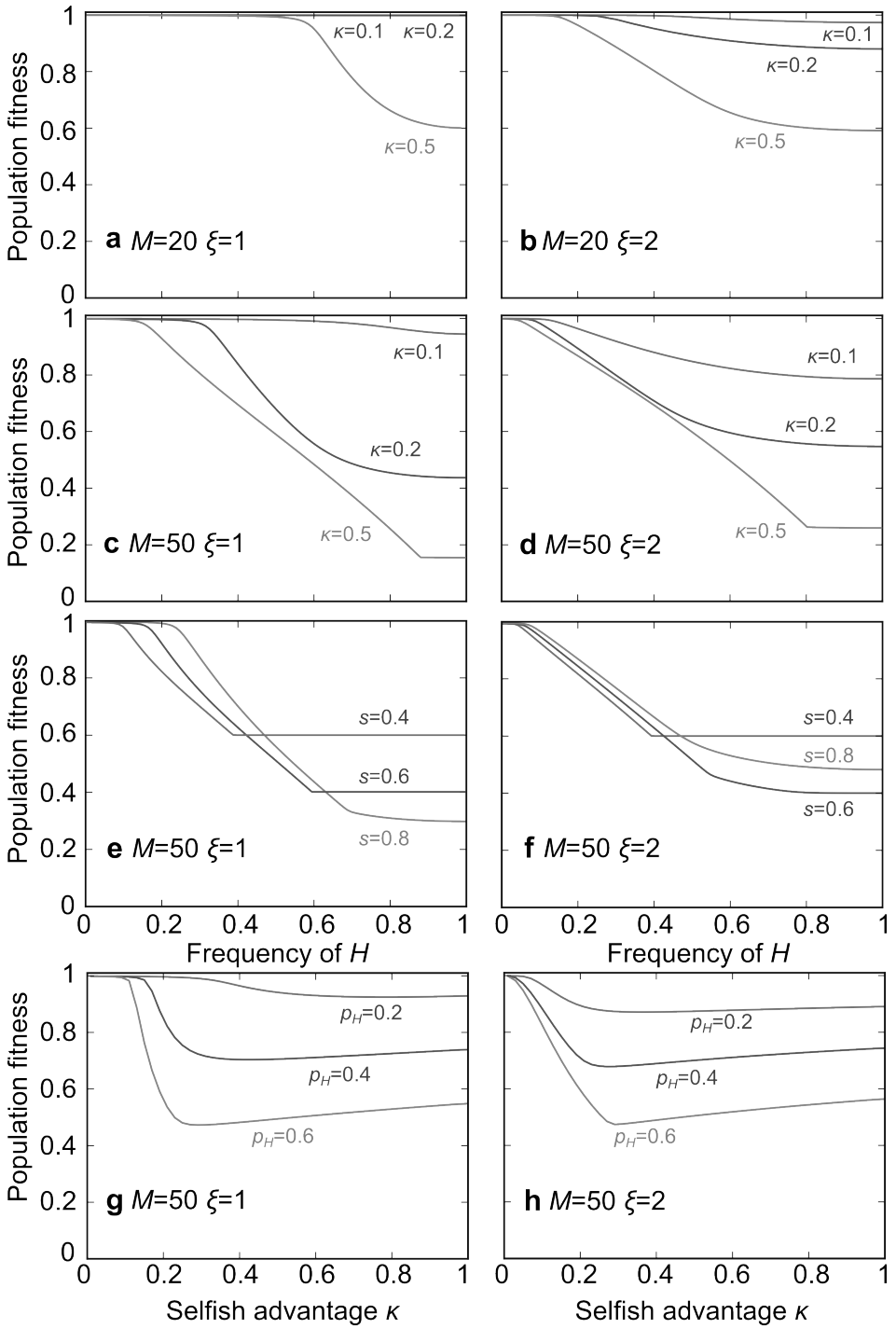

Figure 2.4. Fitness costs of cytoplasmic mixing can be exacerbated in the presence of selfish mitochondrial mutants. Cooperative interactions within groups of mitochondria break down more easily in larger groups ((**a**, **b**) vs. (**c**, **d**)) with strong epistasis and weak selection s (**e**-**f**). Excessive competition among mitochondria can become costly to selfish organelles, increasing the mean population fitness, if fast replicating deleterious mitochondria overtake the cell before being able to spread (**g**-**h**). Mutation rate is set to $\mu = 0.0001$, selections strength $s = 1$ unless indicated otherwise. Selfish advantage is set to $\kappa = 0.2$ in (**e**) and (**f**).

the absence of cell fusion, selfish mutants affect only those cell lineages in which rare mutations occur, allowing selection at the higher level to rapidly eliminate the affected

individuals. There is, however, a critical frequency p_H at which the non-cooperative mitochondria start proliferating faster (Fig. 2.4a-f). Frequent cytoplasmic mixing reduces the mitochondrial variance between cells and therefore lowers the efficacy at which selection can eliminate the affected cells. The conditions are more permissive for selfish proliferation in populations with high numbers of mitochondria per cell (Fig. 2.4a-d), weak selection and strong epistasis ξ (Fig. 2.4e, f). With increasing ξ the fitness function becomes increasingly flat at low m, allowing the non-cooperative mitochondria to reach high per-cell frequencies before they are eliminated by selection.

Further increase in p_H pushes the equilibrium towards the lower population fitness by increasing both the number of mutants per cell and the amount of affected lineages. Eventually all cell lineages contain some non-cooperative mitochondria at which point the population fitness becomes nearly independent of the frequency of *H* (Fig. 2.4a-f).

Interestingly, the equilibrium frequency of selfish mitochondrial mutants is not always a monotonic function of their relative replicative advantage κ (Fig. 2.4g-h). There is a critical value of κ correpsonding to the highest mutant load and the minimal population fitness. With the replicative advantage lower than the critical value of κ, the mutant spread through the population is limited by their replication rate; for higher κ, selfish mutants overtake their host cells too rapidly, allowing the purifying selection to efficiently suppress their further spread.

2.5 Invasion of *H* allele

In this section I consider an evolutionary scenario where alleles *H* are introduced into a population at a low frequency and evolve freely. The allele *H* changes the mode of reproduction by inducing temporary cell fusion with a randomly selected partner, mixing the mitochondrial populations of the two cells (Fig. 2.1).

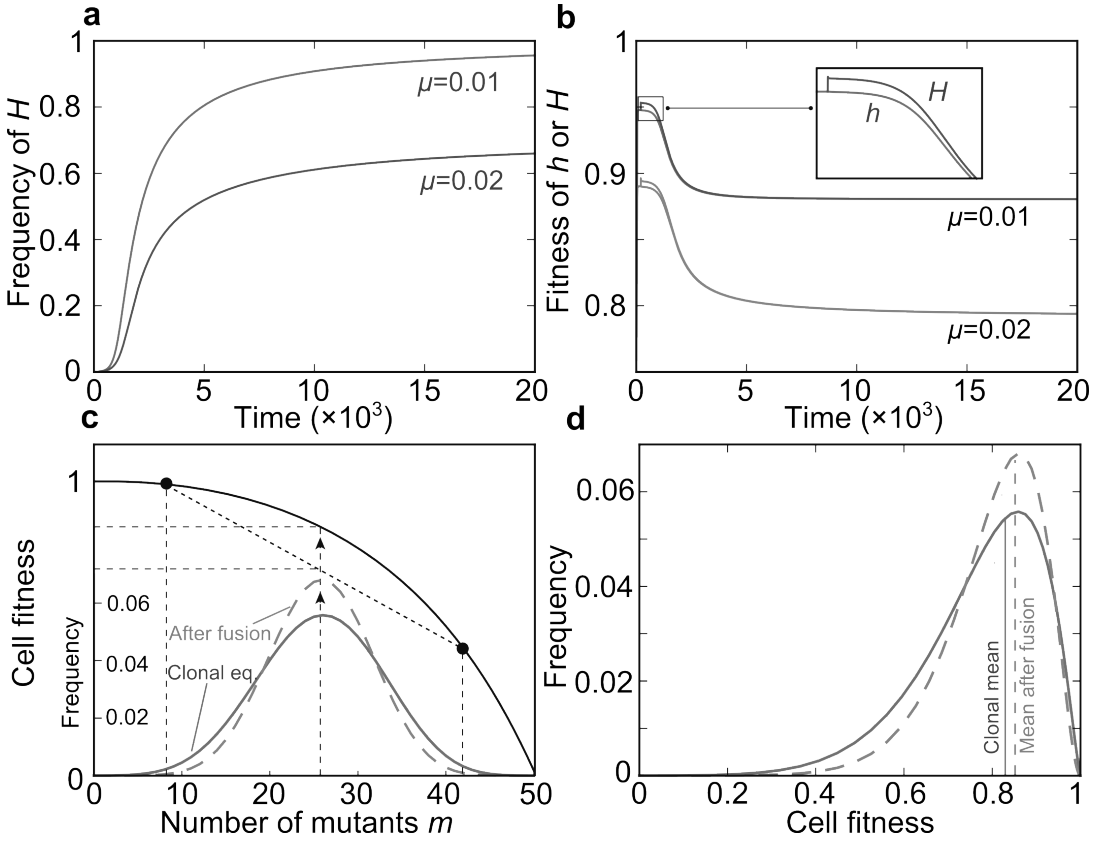

Figure 2.5. Invasion of the cell-fusion allele *H* into an ancestrally clonal population. A nuclear allele inducing cytoplasmic fusion invades and can reach fixation (**a**), but its spread reduces the long-term population fitness (**b**). A single cell fusion-division event increases the frequency of intermediate cytotypes, and reduces the frequency of cells with extreme mutant numbers (**c**). The loss of mitochondrial variance reduces the efficacy of selection and is detrimental in a long term. However, the reduced frequency of extreme cytotypes can be beneficial in a short term (**d**), if the intermediate cytoplasmic states have a higher fitness than expected from the additive interactions, i.e. with negative epistasis ($\xi > 1$) (**c**). The fitness advantage can be maintained if the mito-nuclear linkage is weak, e.g. if half of the mitochondria are inherited from a randomly selected partner which would otherwise reproduce clonally. $\mu = 0.04$, $\xi = 3$ in (**c**) and (**d**), $\xi = 2$ in (**a**) and (**b**). The number of mitochondria per cell is $M = 50$, $\kappa = 0$.

2.5.1 Epistasis between mitochondrial mutations

The results show that the cell-fusion allele *H* is able to invade, spread to an equilibrium frequency of $p_H < 1$ or reach fixation ($p_H = 1$) (Fig. 2.5a, 2.6). The invasion occurs in spite of the curtailed long-term population fitness due to the lower variance in the number of mitochondrial mutants among invaders and weaker selection (Fig. 2.5b). The necessary condition for successful invasion is $\xi > 1$, i.e. negative epistasis between deleterious mitochondrial mutations. The detrimental effect of every new mutation has to increase with the total mutational load, as indeed is the case in modern

eukaryotes.

Cytoplasmic mixing increases the frequency of intermediate cytoplasmic states, reducing the frequency of cells with extreme mutant numbers (both high and low, Fig. 2.5c). This reduced variance has a long-term fitness disadvantage due to the weakened response to selection (Fig. 2.5b). However, with negative epistasis ($\xi > 1$), the intermediate cytotypes have a higher fitness than expected from the linear combination of the extremes, which gives the invading allele *H* a short-term advantage (Fig. 2.5d). Invaders choose their mating partners randomly and therefore the mitochondrial-nuclear associations remain weak. This allows the allele *H* to acquire a long-lasting advantage over the clonally reproducing subpopulation *h* (Fig. 2.5b), even though its spread inevitably curtails the population fitness in the long term. The advantage is lost with positive epistasis ($\xi < 1$), in which case both short- and long-term effects of reduced mitochondrial variance become detrimental.

2.5.2 Further conditions favouring the spread of *H*

Under the pressure of deleterious mitochondrial mutations, the allele *H* spreads to high frequencies and fixes more readily with low mutation rates (Fig. 2.6a, b). Cytoplasmic fusion has a stronger evolutionary advantage with small mitochondrial population sizes, as segregational drift is more efficient in generating mitochondrial variance with small M (Eq. 2.2). A similar trend is observed with selfish mitochondria having a replicative advantage over their wild-type counterparts, where fast replication of mutants, i.e. large κ, diminishes the evolutionary advantage of the cell-fusion allele *H* (Fig. 2.6c, d). This time, there is a critical value of κ corresponding to a distinctive drop of equilibrium allele frequency p_H to zero. This fast transition occurs once the replicative advantage of selfish organelles becomes large enough to rapidly reduce the fitness of fusing hosts, whereas the high-fitness asexual lineages remain resistant to their spread. As the allele *H* spreads due to its short-term fitness advantage, its

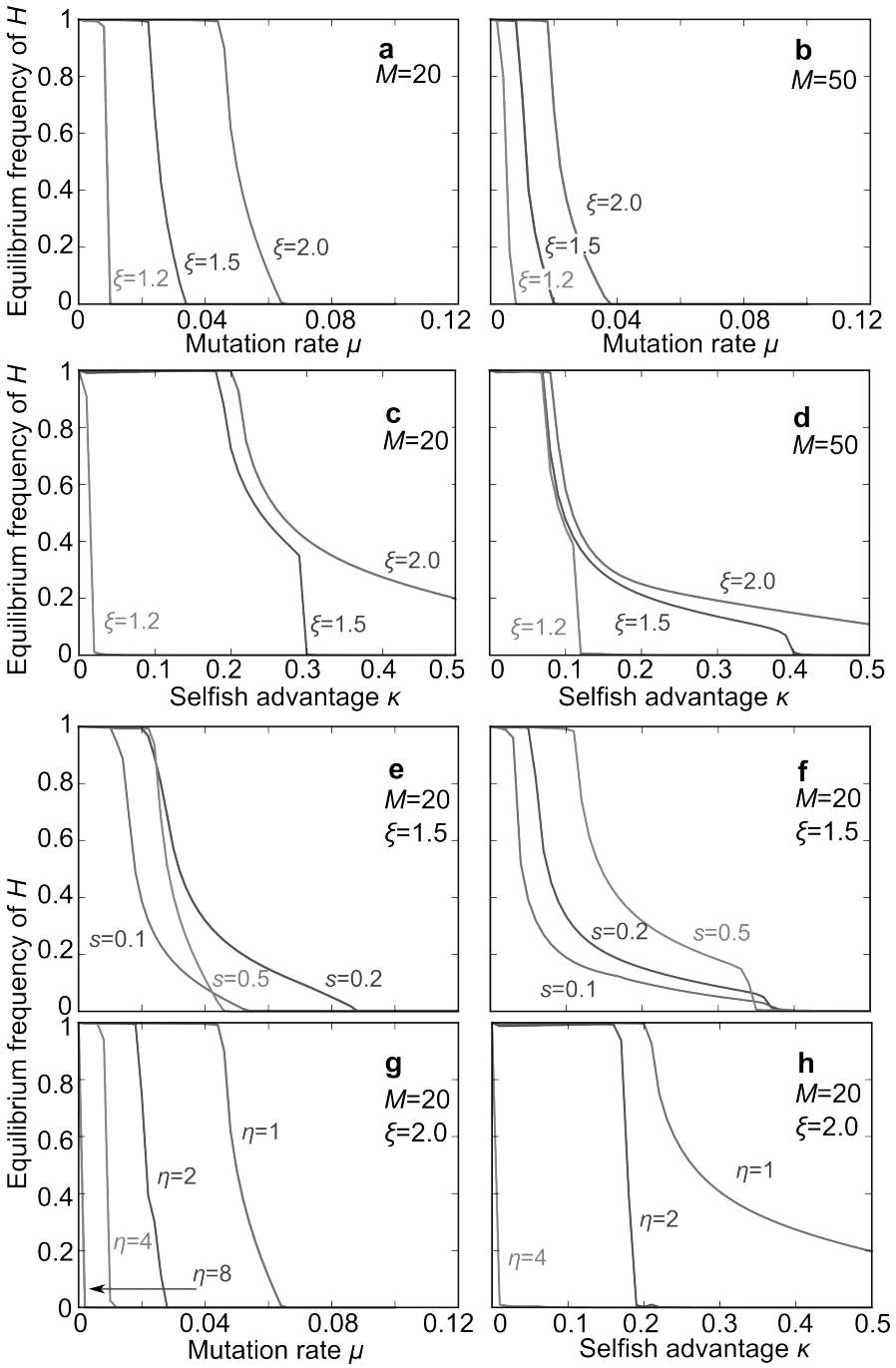

Figure 2.6. Conditions favouring the evolution of sexual cell fusion. Reduced mutation rates, small mitochondrial populations and negative epistatic interactions all promote the evolution of sexual cell fusion under mitochondrial mutation pressure (**a**-**b**). In the presence of selfish mitochondrial mutants, competition at the lower level has to be suppressed before cell fusion can be established (**c**-**d**). Cellular fusion evolves more easily under strong purifying selection at the higher level (conditions favouring the fixation of *H* are less strict under large s, **e**-**f**). With alternating clonal and sexual life cycle stages, the number of consecutive clonal divisions η must remain low (**g**-**h**). Mutation rate in (**c**), (**d**), (**f**) and (**h**) is set to $\mu = 0.0001$, s is the strength of selection, *H* is the allele inducing cell fusion with a randomly selected partner.

fixation is also facilitated by strong purifying selection at the higher level s (Fig. 2.6e, f).

2.5.3 Alternating life cycles and mitochondrial-nuclear associations

Consider now the case where mutants carrying the allele *H* are capable of inducing the cytoplasmic fusion only every η generations. It is indeed often the case in protists that individuals engage in sexual reproduction only occasionally, e.g. under stressful environmental conditions or starvation (Dacks and Roger, 1999; Goodenough et al., 2007). Since the clonal stage of the life cycle would result in higher variance due to segregational drift, η is likely to affect the evolutionary success of the cell fusion allele *H*. Indeed, the model shows that increasing the number of consecutive clonal divisions in-between the sexual fusion events has a very strong effect *opposing* the spread of the fusion allele *H* (Fig. 2.6g, h). As few as 4-8 clonal cell divisions could be enough to prevent the invasion of the cell fusion allele *H* under all reasonable conditions investigated.

The principal reason allowing the allele *H* to invade lies in the way that reduced mitochondrial variance affects mean fitness after every cellular fusion event. Reduced variance in the number of mutants gives a fitness advantage due to negative epistasis, but only if the association between the nuclear allele and the mitochondrial population of the same cell is temporary. This is most easily achieved through frequent fusions with randomly chosen partners that might otherwise reproduce clonally. With $\eta = 1$ mitonuclear associations are weakest, but become stronger when the same mitochondrial population persists within the lineage for several generations, i.e. $\eta > 1$. The detrimental long-term effects of reduced mitochondrial variation between the higher-level units become increasingly important as η grows. Fixation of the cellular fusion allele *H* therefore requires frequent cytoplasmic mixing, maintaining weak mitochondrial-nuclear associations.

2.6 Conclusions and discussion

It is very likely, that sex appeared in a cell that already possessed mitochondria, but not mating types or mechanisms constraining the cytoplasmic inheritance. The model presented here shows that sexual cell fusion can evolve simply to promote mitochondrial mixing, which temporarily masks the detrimental effects of faulty or selfish endosymbionts or organelles.

Low mitochondrial mutation rates allow cell fusion alleles to fix more readily (Fig. 2.6a, b). Admittedly, substantial variation in mitochondrial mutation rates among the extant eukaryotes makes the inference of the ancestral pace of mutation accumulation rather complicated. On the one hand, we know that evolution rates (a proxy for the true mutation rate) in intracellular symbiont genomes are typically elevated (Itoh et al., 2002; Marais et al., 2008). Mitochondrial evolution rates in higher animals, fungi and some plants can also be substantially higher than in their nuclear genomes (Lynch et al., 2006, 2008; Sloan et al., 2009). On the other hand, the mutation rate appears to be extremely low in most plants, early branching metazoans (Palmer and Herbon, 1988; Shearer et al., 2002; Huang et al., 2008) and many unicellular eukaryotes (Burger et al., 1995, 2013; Smith and Keeling, 2015). It is therefore not impossible that the initially high evolution rate at the beginning of the endosymbiotic association slowed down as the evolutionary transition progressed, facilitating the evolution of sex. High mutation rates in some present-day eukaryotes are then secondarily derived, perhaps owing to their high metabolic rates and active lifestyles.

While the initial symbiotic association remains shrouded in mystery (Martin and Muller, 1998; Embley and Martin, 2006; Martin et al., 2015), with mitochondrial endosymbiosis entering an obligatory phase, selection against mitochondrial mutations, e.g. the ones affecting the respiratory function of the cell, likely increased in strength (higher s) providing conditions more permissive for the evolution of cytoplasmic mixing (Fig. 2.6e, f). Indeed, empirical data reveal substantial purifying

selection acting on mitochondrial populations in modern animals (Elson et al., 2004; Stewart et al., 2006; Castellana et al., 2011, Cooper et al., 2015). Similarly, as the evolutionary transition progressed, selfish mitochondrial competition also had to become suppressed (reduced κ), through the proto-eukaryotic mechanisms of conflict mediation, e.g. honest signalling or reduced mitochondrial genomes (Radzvilavicius and Blackstone, 2015), thus promoting the evolution of cell fusion and sex (Fig. 2.6c, d).

Extrapolating from fitness interactions in modern eukaryotes (Rossignol et al., 2003), it is not unreasonable to assume negative epistatic interactions ($\xi > 1$) between detrimental mitochondrial mutations throughout eukaryogenesis, favouring the invasion of cell-fusion alleles (Fig. 2.6a). Indeed, with multiple mitochondria per cell and several copies of mtDNA per organelle, a critical number of deleterious mutations has to accumulate before cellular respiration is significantly impaired (Mazat et al., 2001; Rossignol et al., 2003). Mitochondrial endosymbiosis therefore created a unique genetic system in which strong synergistic interactions between deleterious mutations favour the evolution of sexual cell fusion. This is in stark contrast to deleterious mutations in the nucleus (or prokaryotic genomes), where negative epistatic interactions seem to be relatively uncommon (Kouyos et al., 2007).

Modern eukaryotes are capable of reproducing clonally in numerous consecutive generations, punctuated by occasional sex (Dacks and Roger, 1999; Goodenough et al., 2007). The model analysed here predicts, however, that when host-regulated cell fusion first arose, it had to be frequent—and clonal reproduction rare—in order to maintain the weak mitochondrial-nuclear associations responsible for the evolutionary advantage of cytoplasmic mixing, i.e. low values of η (Fig. 2.6g, h). Frequent cell fusion events might have been vital early in eukaryogenesis, since without the precisely coordinated chromosome and cell division machinery, consecutive reproduction cycles without cell fusion could have produced non-functional chromosome numbers or gene combinations, rendering the emerging

eukaryotic cell unviable. Garg and Martin (2016) even proposed that sex could have arisen in a syncytial eukaryote ancestor, in which multiple early nuclei of variable gene content complemented each other, even though the life-cycle stage with a single nucleus per cell would have still been needed to maintain efficient selection against new deleterious mutations. Likewise, frequent cell fusions would have masked mutations and stabilized the emerging eukaryotic genome in the presence of intron bombardment and endosymbiotic gene transfer (Lane, 2011).

One way to interpret the main result of this work is that the initial selective pressure driving the evolution of cell-cell fusion could have been mitochondrial, in which case the routine recombination among nuclear genes came as a fortunate side effect, maintaining the evolutionary advantages of sex past the evolution of the uniparental inheritance and until present day. Indeed, the molecular machinery for meiotic, reciprocal recombination had to evolve in the routine presence of cell fusion events, and the barrier separating eukaryotic sex from prokaryotic recombination might have never been crossed without mitochondria. The present work accounts only for the accumulation of deleterious mitochondrial mutations, but cell fusion could also benefit hosts in which the number of mitochondria fluctuates through random segregation and drift, as well as in the plausible case in which distinct mitochondrial haplotypes have complementary functions. Additionally, cell fusion could have been induced directly by protomitochondrial endosymbionts, through various manipulations of the host cell's life cycle under stressful conditions (Blackstone and Green, 1999). It is therefore very likely, that multiple mechanisms promoting cell fusion were in place, with mitochondrial selection pressure contributing to the ease with which sexual reproduction combining cytoplasmic fusion and reciprocal recombination came into the widespread existence.

2.7 Appendix A. Derivations

2.7.1 Variance and identity-by-descent relations

Let X_n be a random variable denoting the number of mutants within a cell after n rounds of clonal cell division, sampling without replacement M mitochondria from the doubled population of $2M$. The population mean is then simply equal to the initial number of mutants within the cell, $\mathrm{E}(X_n) = x_0$. Variance in the number of mutants can be expressed as

$$\mathrm{Var}(X_n) = \mathrm{E}[\mathrm{Var}(X_n|X_{n-1})] + \mathrm{Var}[\mathrm{E}(X_n|X_{n-1})]. \tag{2.8}$$

Given that the variance of the hypergeometric probability distribution used in sampling without replacement is

$$\mathrm{Var}(X_n|x_{n-1}) = \frac{x_{n-1}(M - x_{n-1})}{2M - 1}, \tag{2.9}$$

we can further write

$$\begin{aligned}
\mathrm{Var}(X_n) &= \mathrm{E}\left(\frac{X_{n-1}(M - X_{n-1})}{2M - 1}\right) + \mathrm{Var}(X_{n-1}) \\
&= \mathrm{E}\left(\frac{MX_{n-1}}{2M - 1}\right) - \mathrm{E}\left(\frac{X_{n-1}^2}{2M - 1}\right) + \mathrm{Var}(X_{n-1}) \\
&= \frac{M}{2M - 1}\mathrm{E}(X_{n-1}) - \frac{1}{2M - 1}\mathrm{E}(X_{n-1}^2) + \mathrm{Var}(X_{n-1}) \\
&= \frac{x_0 M}{2M - 1} - \frac{1}{2M - 1}[\mathrm{Var}(X_{n-1}) + x_0^2] + \mathrm{Var}(X_{n-1}) \\
&= \frac{x_0(M - x_0)}{2M - 1} + \left(1 - \frac{1}{2M - 1}\right)\mathrm{Var}(X_{n-1}).
\end{aligned} \tag{2.10}$$

Here x_0 is the initial number of mutants within a cell. With the boundary condition $\mathrm{Var}(X_0) = 0$ the solution is

$$\mathrm{Var}(X_n) = x_0(M - x_0)\left[1 - \left(1 - \frac{1}{2M - 1}\right)^n\right]. \tag{2.11}$$

Variance in the mutant frequency $P_n = \frac{X_n}{M}$ is then

$$\mathrm{Var}(P_n) = p_0(1 - p_0)\left[1 - \left(1 - \frac{1}{2M - 1}\right)^n\right]. \tag{2.12}$$

The genetic diversity (or lack of it) within a cell can be expressed as a

probability that two randomly selected mitochondria within the cell will be identical by descent, f_n. In our random segregation model two lower-level units are considered identical if they are either descendants of the same parent, or different parents that are identical by descent themselves due to associations in previous generations. After n generations we can then write

$$f_n = \frac{1}{2M-1} + \left(1 - \frac{1}{2M-1}\right) f_{n-1}. \quad (2.13)$$

The above recursion is straightforward to solve for the parameter of non-identity $h_n = 1 - f_n$. As initially all mitochondria within the cell are assumed to be unrelated, the boundary condition is $h_0 = 1$. It is then easy to show that

$$h_n = \left(1 - \frac{1}{2M-1}\right)^n, \quad (2.14)$$

and

$$f_n = 1 - \left(1 - \frac{1}{2M-1}\right)^n. \quad (2.15)$$

Comparing this result to the expression for the variance after n generations (Eq. 2.2), we notice that clonality within the cell is just a normalized mitochondrial variance between host cells,

$$f_n = \frac{\mathrm{Var}(p_n)}{p_0\,(1-p_0)}. \quad (2.16)$$

2.8. Appendix B. Stochastic dynamics of selfish mitochondrial mutations and evolution of cytoplasmic mixing

2.8.1 Acknowledgement

The work presented in this chapter stemmed in part from the collaboration with professor Neil W. Blackstone of Northern Illinois University, in which we made an attempt to integrate evolution of eukaryotic sex into the general framework of

eukaryogenesis as a levels-of-selection transition (Buss, 1987; Maynard Smith and Szathmáry, 1995). Trained under Leo Buss, Blackstone is one of the earliest proponents of the levels-of-individuality approach to understanding the evolution of biological complexity in general, and mitochondrial endosymbiosis in particular (Blackstone 1995; 2013; 2016). The individual-based model (Radzvilavicius and Blackstone, 2015), which I developed to evaluate the implications of selfish endosymbiont replication for the evolution of cell fusion within a finite population, fully supports and complements the results presented in this chapter. Here I only briefly introduce the model and the main results.

2.8.1 Individual-based simulation model

The individual-based simulation in its concept is identical to the deterministic infinite-population model introduced earlier in this chapter and was designed to simulate the life cycle depicted in Figure 2.1. Each haploid organism is represented as an independent object in the C++ implementation of the model[1], and the life-cycle events (symbiont mutation, intracellular selection within the group of M symbionts, selection between cells, cell fusion and division, Fig. 2.1) are simulated explicitly within a population of 10,000 individuals. Selfish endosymbionts arise at a rate μ and have a reproductive advantage $1 + k$, which essentially represents their selection coefficient within the group. Initially, cells reproduce clonally by first duplicating their cytoplasmic contents and then splitting into two. A rare mutation in the host nuclear locus *H*/*h*, however, can cause the host to temporarily fuse with a randomly chosen cell, mixing the endosymbiont populations before splitting into two daughter cells again. The probability of the mutant inducing a successful fusion within a single generation r can be adjusted, to account for facultative sexuality ($r < 1$). We analysed both the

[1] https://github.com/ArunasRadzvilavicius/ConflictAndSex

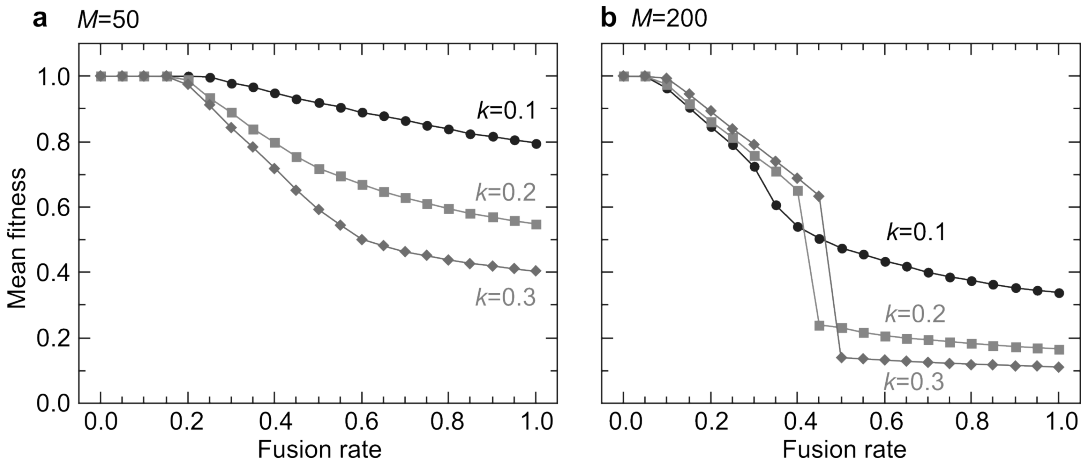

Figure 2.7. Mean population fitness at equilibrium as a function of cell fusion rate. Three curves correspond to three values of mutant reproductive advantage *k*. Owing to the horizontal spread of selfish mutants and reduced variance between the fusing hosts, there is a significant fitness cost associated with sexual mixing. With rare fusions, selfish proto-mitochondria arise and stay confined within distinct lineages, and are easily eliminated by selection, which results in high mean fitness. Increasing fusion rate allows for a limited spread of deleterious symbionts, leaving the rest of the population mutant-free. A fast transition occurs after the critical rate of cytoplasmic mixing is reached, at which point most of the population is overtaken by the selfish proto-mitochondria. Endosymbiont mutation rate is set to 10^{-5}.

population state at equilibrium with fixed fusion rates, and the evolution of the host cell-fusion allele *H* invading the clonal population after being introduced at a low frequency.

2.8.2 Cytoplasmic fusion facilitates the spread of selfish symbionts

First, we simulated a finite population at the dynamic equilibrium with fixed cell fusion rates. Mutant endosymbionts arise at a constant rate $\mu = 10^{-5}$ and proliferate faster than the wild-type endosymbionts within the cell, but purifying group-level selection favours cells with low numbers of selfish symbionts m according to the concave fitness function $\omega(m) = 1 - (m/M)^2$. Overall, the frequency of selfish mutants increases with the rising fusion rate in a pattern remarkably similar to the trends observed in the deterministic infinite-population model (Fig. 2.7). In populations dominated by clonally reproducing hosts, selection on the higher level is efficient enough at eliminating selfish symbionts, ensuring high levels of cooperation and high population fitness. Selfish endosymbionts that arise at a low mutation rate are mostly confined within isolated

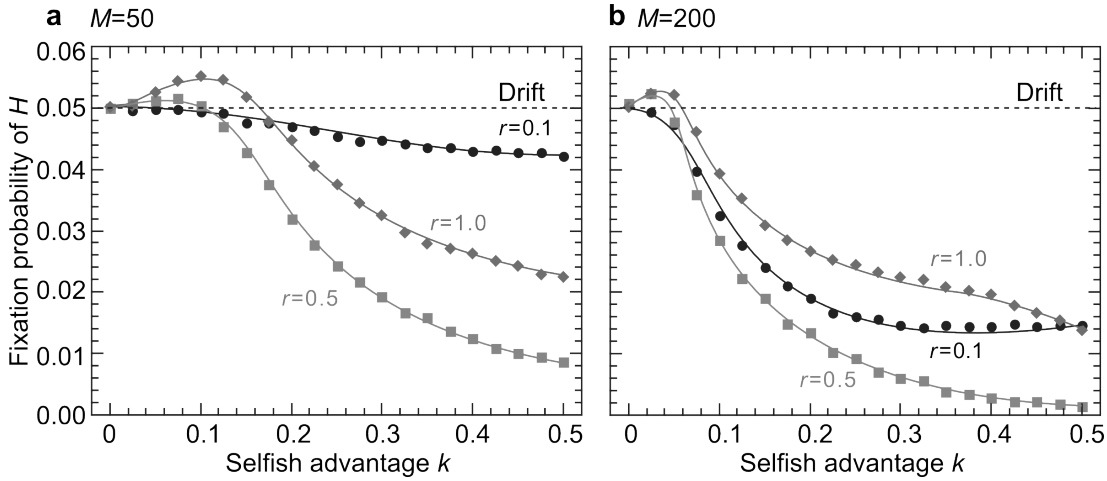

Figure 2.8. Fixation probability of cell-fusion allele *H* in a finite population of clonally reproducing individuals. The number of endosymbionts per host cell is set to $M = 50$ (**a**) or $M = 200$ (**b**). Neutral mutants fix with probability 0.05. Selection favours the invasion of *H* under low values of selfish replication rate k and high rate of induced fusions r. Endosymbiont mutation rate is set to 10^{-4}.

vertical host-cell lineages, maintaining the rest of the population in a virtually mutant-free state. This ensures high variance between higher-level units and facilitates purifying selection against fast replicating but deleterious endosymbionts.

In agreement with the deterministic dynamics in an infinite population (Fig. 2.4) increasing fusion frequency allows selfish endosymbionts to spread horizontally faster, reduces variance in fitness among the symbiont groups and hinders selection between host cells. As a consequence, the number of selfish endosymbionts grows rapidly as host cell fusions become more frequent (Fig. 2.7). The spread of selfish mutants is also favoured by large endosymbiont population sizes within the cytosol, reducing both the effect of segregational drift and the efficacy of group-level selection (Fig. 2.7b).

2.8.3 Invasion of the cell-fusion allele

Now, consider invasion of the cell-fusion allele *H* under the effect of purifying selection against selfish endosymbionts. Due to strong stochastic drift inherent to finite populations, measuring the allele frequency at equilibrium is impractical. To determine whether the cell-fusion mutant is favoured or opposed by selection, we instead calculate fixation probability of the allele *H*, by repeatedly running the simulation until

either fixation or extinction of the invader. For a neutral mutant evolving via drift alone, the fixation probability equals its initial frequency p_0, since after a large number of generations the whole population must consist solely of the descendants of a single individual at generation 0. The fixation probability exceeds p_0 if *H* is favoured by selection, and is lower if selection opposes its spread.

Despite the long-term fitness cost brought into the population by the cell-fusion allele, the spread of *H* can be favoured by selection, but only with low values of selfish advantage k and high invader cell-fusion rate r (Fig. 2.8a, b). The invasion of *H* is more likely with smaller symbiont population sizes, and is only possible with negative epistatic interactions between endosymbiont mutations, i.e. a concave fitness function. As the reproductive advantage of selfish endosymbionts increases, however, the fixation probability of *H* drops, suggesting that the cell-fusion alleles would be strongly selected against. Easily explained in terms of long- and short-term effects of reduced between-group variance (see Section 2.5), these trends support the infinite-population results depicted in Figures 2.6c and 2.6d, and show that the general outcome of the model and the predictions of the main hypothesis, are independent of the assumptions on the population size. At the same time, the results suggest that there are conditions under which the fitness cost of cytoplasmic mixing can prevent sex from evolving; and any advantages of sex or recombination would have to outweigh these costs.

CHAPTER 3. EVOLUTION OF MATING TYPES DRIVEN BY PURIFYING SELECTION AGAINST MITOCHONDRIAL MUTATIONS

3.1 Summary

Sexual cell fusion combines genetic material of two gametes, but why the two reproductive cells have to belong to distinct self-incompatible gamete classes is not known. In a vast majority of sexual eukaryotes, mitochondria are inherited uniparentally from only one of two mating types, which is thought to facilitate purifying selection against deleterious mitochondrial mutations and limit the inter-genomic conflicts. It remains unclear, however, whether uniparental inheritance can drive the evolution of self-incompatible gamete classes under the dynamics of cytoplasmic segregation and mixing. Here I show that two mating types in eukaryotes could have evolved together with the asymmetric transmission of mitochondrial genes as a mechanism of mitochondrial quality control. I develop a mathematical model to explicitly study the evolution of two self-incompatibility alleles linked to the nuclear locus controlling the pattern of organelle inheritance. The invasion of mating types is opposed by the short-term fitness benefit of mitochondrial mixing under negative epistasis and the lower chance of encountering a compatible mating partner. Nevertheless, under high mitochondrial mutation rates and low gamete mortality, purifying selection against defective mitochondria can drive two mating types to fixation. The invasion is further facilitated by the paternal leakage of mitochondria under paternal control of cytoplasmic inheritance. In contrast to previous studies, the model does not rely on the presence of selfish cytoplasmic elements, providing a more universal solution to the long-standing evolutionary puzzle of two sexes.

3.2 Introduction

Sex is among the traits universal to all complex life and was therefore already present in the last eukaryotic common ancestor. By combining the genetic material of two gametes, sexual cell fusion and meiotic recombination exposes the hidden genetic variation in finite populations, breaks up the unfavourable allelic combinations under fluctuating selection, and mitigates mutational meltdown (Otto, 2009; Hartfield and Keightley, 2012). In a vast majority of cases, the two mating partners belong to distinct self-incompatible gamete classes, i.e. mating types in isogamous protists or true sexes with anisogametes, but the selective forces behind this fundamental asymmetry remain elusive (Billiard et al., 2010). The existence of two gamete types in most eukaryotes is often regarded as an evolutionary conundrum, as it reduces the number of potential mating partners, which should be detrimental if the cost of finding a compatible gamete is high and the mating opportunities are limited.

Several non-exclusive explanations for the emergence of self-incompatible gamete types in a unisexual population have been proposed (Billiard et al., 2010). Two mating types might have appeared together with bipolar gamete-recognition systems ensuring efficient inter-cellular signalling (Hoekstra 1982; Hadjivasiliou et al., 2015), to promote outbreeding (Charlesworth and Charlesworth, 1979, Uyenoyama, 1988) or to improve mitochondrial-nuclear coadaptation (Hadjivasiliou et al., 2012). Particularly appealing has been the idea that mating types emerged to ensure the asymmetric inheritance of cytoplasmic genetic elements. Indeed, in a vast majority of eukaryotes, only one gamete class transmits its organelles—mitochondria and chloroplasts—to the zygote, although uniparental inheritance is not always complete, with paternal leakage and heteroplasmy being relatively common (Breton and Stewart, 2015). Early theoretical studies supported the hypothesis (Hastings, 1992; Hurst and Hamilton, 1992; Hutson and Law, 1993), but relied on the presence of the so-called “selfish” or

parasitic cytoplasmic elements, simplistic assumptions, and lacked generality or empirical support.

A more general view of the evolutionary advantage of uniparental inheritance (UPI) is that it improves the efficacy of purifying selection against mitochondrial mutations (Bergstrom and Pritchard, 1998; Hadjivasiliou, 2013), and therefore confers a long-term fitness advantage. Mitochondrial mixing under biparental inheritance, on the other hand, limits the strength of selection at the level of a cell and is costly in a long term (Chapter 2). Asymmetric transmission of mitochondria therefore counters the mutational meltdown in mitochondrial genomes, and this may account for the highly efficient purifying selection operating even in the absence of recombination (Cooper et al., 2015).

Of particular importance, a recent study explicitly accounting for the segregational drift of detrimental mitochondrial mutations found that the fitness benefit of UPI decreases in a frequency-dependent manner (Hadjivasiliou et al., 2013). As the frequency of the nuclear UPI allele increases, the fitness benefits increasingly spread to the biparental part of the population, limiting the invasion of UPI alleles. Due to the limited advantage of UPI and further fitness costs of reduced mating rate of self-incompatible gametes, it has been concluded that UPI alone is unlikely to drive the evolution of two self-incompatible mating types in an ancestral unisexual population (Hadjivasiliou et al., 2013). The robustness and generality of these results, however, might be limited by the assumptions of the study; in particular, it was assumed that gametes unable to find a suitable mating partner in a single mating attempt incur a severe fitness cost and do not contribute to the next generation, putting the self-incompatible gamete types at a strong disadvantage. Whether the asymmetric inheritance of mitochondria can facilitate the evolution of binary mating types under the mitochondrial mutation pressure and biologically realistic mating kinetics remains unclear.

Here I develop a mathematical model to investigate whether selection against deleterious mitochondrial mutations can drive the evolution of two self-incompatible mating types by establishing asymmetric transmission of mitochondrial genes. Mitochondria continuously accumulate mutations which are purged by selection on the level of the cell, with random mitochondrial drift generating additional variance at every cell division. Nuclear inheritance-restriction alleles regulate the selective destruction of mitochondria inherited either from the cell carrying the allele, or its mating partner with the opposite mating type. Mating kinetics play a central role in the model, with long mating periods and low gamete mortality rates favouring the invasion of self-incompatible gamete classes. Strikingly, paternal leakage of mitochondria—an evolutionarily stable state under paternal control of cytoplasmic inheritance—can facilitate the evolution of binary mating types under negative epistasis between mitochondrial mutations.

3.3 Modelling the evolution of uniparental inheritance

3.3.1 Two modes of UPI regulation

Mechanisms of uniparental inheritance vary substantially across eukaryotic species (Sato and Sato, 2013), but one general theme seems to be rather common: one of the mating partners tags its mitochondria with the mating-type specific marker protein, which is recognized by the partner's molecular factors after the gamete union, leading to the eventual degradation of the marked organelles. In mammals, for example, sperm mitochondria are tagged by the recycling marker protein ubiquitin and destroyed by the egg's cytoplasmic destruction machinery after the gamete union (Sutovsky et al., 1999). Similarly, in basidiomycete yeast *Cryptococcus neoformans,* genes *SXI1a* and *SXI2α* located in opposite mating types are responsible for tagging and recognition of paternal mitochondria (Yan et al., 2007).

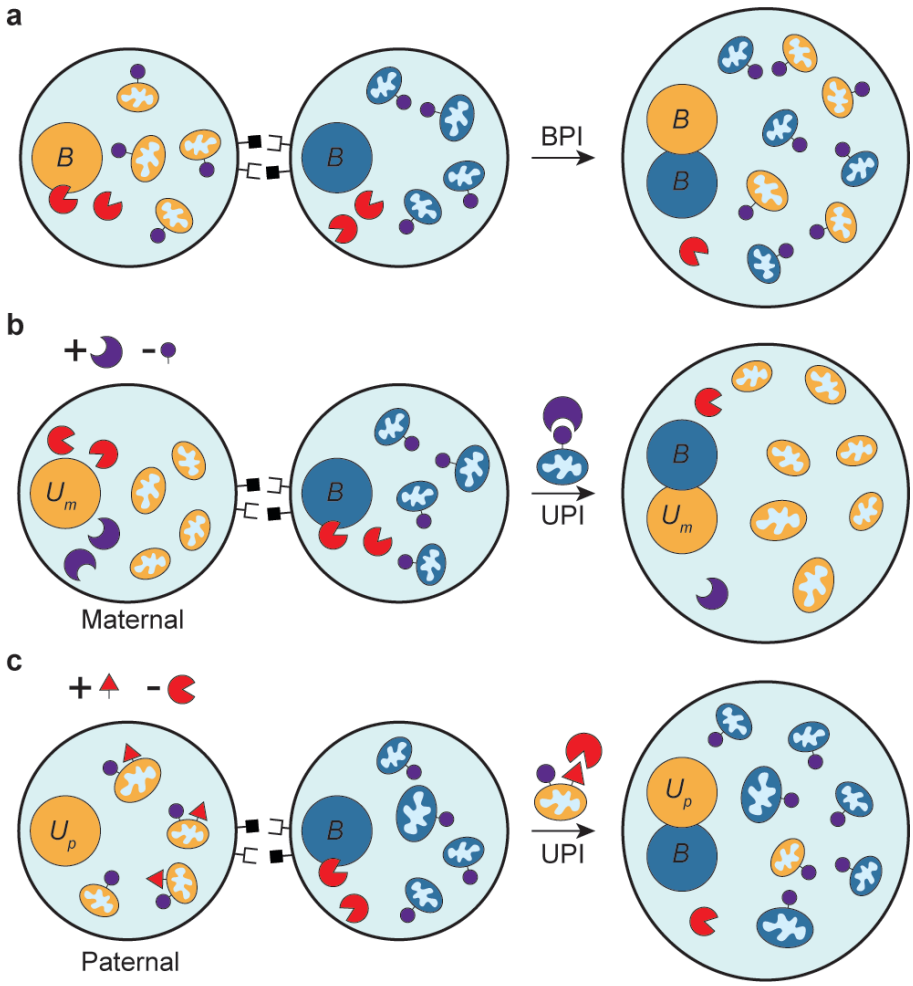

Figure 3.1. Two modes of uniparental inheritance differ in the way organelle inheritance is controlled. In an ancestral state, mitochondrial surface proteins do not have corresponding nuclear-coded molecular factors targeting the organelles for destruction (**a**). Maternal mode of UPI arises with one of the gametes developing the ability to recognize the mitochondrial marker protein universal to the whole population, while protecting its own organelles e.g. by removing the tag (**b**). Similarly, in paternal mode of UPI, one of the gametes marks its mitochondria with a new universally recognizable marker, but loses the ability to target it for destruction in its own cytoplasm (**c**). Alternatively, the paternal gamete can simply remove a part of its own organelle population before fertilization.

In the present work I assume that uniparental inheritance can be established in two ways, based on how the two gametes control organelle transmission (Fig. 3.1). First, a cell (wild-type nuclear allele *B* coding for biparental inheritance) can develop the ability to recognize and target for destruction a universal mitochondrial marker protein, at the same time protecting its own organelles from degradation, e.g. by ceasing the expression of the marker in its own mitochondria (allele U_m, Fig. 3.1b). I term this mechanism the “maternal” mode of UPI, and the gamete destroying its

partner's mtDNA the "maternal gamete"—the definition which does not rely on the prior existence of mating types or sexes and merely reflects the fact that the gamete controlling the cytoplasmic inheritance eliminates its partner's mtDNA, but not its own (Fig. 3.1b). Alternatively, a cell can start producing a new nuclear-coded and universally recognized mitochondrial marker, but lose the ability to recognize the tag in its own cytoplasm (allele U_p). In this case the gamete essentially controls the inheritance of its own organelles, as stopping the expression of the mitochondrial tag protein would make them unrecognizable in the zygote. I term this mechanism the "paternal" mode of UPI (Fig. 3.1c). In gamete unions with identical alleles at the mitochondrial-inheritance locus, both mating partners lack either the mitochondrial marker or the corresponding molecular destruction machinery. The inheritance of mitochondria is therefore biparental in $U_m \times U_m$ and $U_p \times U_p$ gamete unions.

3.3.2 Mating kinetics

The rules governing the frequency of gamete unions of distinct mating types are expected to play a critical role in the origin and evolutionary stability of mating types (Iwasa and Sasaki, 1987). Rare additional mating types, for instance, are favoured in models where mating opportunities are limited and only a short period of time is available to locate a suitable mating partner, but the same models penalize the newly emerging self-incompatible gamete classes in populations where all gametes are initially compatible. Perhaps a more realistic assumption is that the mating period can be significantly longer than the duration of a single cell-fusion attempt, and that gametes can survive for a sufficiently long time until another suitable mating partner arrives before they are eliminated from the population.

In this work I adopt the mating kinetics first developed in the models of Iwasa and Sasaki (1987), and Hutson and Law (1993). I assume a gamete pool in which the influx of gametes matches their removal due to random death and zygote formation.

The gamete influx rates are proportional to the allele frequencies within the infinite population, while gamete death rate is kept at a constant level δ. Cells within the gamete pool form random pairwise associations at rate $\beta/2$, and are removed from the pool if they are compatible and able to fuse. The allele frequencies in the next generation are calculated from the steady-state gamete-class frequencies within the mating pool. These frequencies correspond to the equilibria of the following system of equations

$$\frac{\mathrm{d}F_i}{\mathrm{d}t} = \begin{cases} -\delta F_i - \beta F_i + \delta f_i + \beta f_i \left(1 - \sum_{k \in \text{s.i.}} F_k^2\right) \; if \; i \; is \; self-compatible \\ -\delta F_i - \beta F_i (1 - F_i) + \delta f_i + \beta f_i \left(1 - \sum_{k \in \text{s.i.}} F_k^2\right) \; if \; i \; is \; self-incompatible \end{cases} \quad (3.1)$$

Here F_i denotes the frequency of the genotype i in the gamete pool, while f_i is the corresponding frequency within the infinite population. The sums here are over the self-incompatible gamete classes. I fix the mating rate at $\beta = 1$ and vary only the value of the death rate δ, as the steady-state frequencies of gamete classes depend only on the ratio of these rates β/δ. Under high gamete mortality rates this model recreates the dynamics observed by Hadjivasiliou et al., 2013.

3.3.3 Population life cycle

The model assumes an infinite population of unicellular haploid organisms, each containing M mitochondria, and is in many ways similar to the model developed in Chapter 2. The population state can be represented by the $(M+1)\times n$ matrix $\mathbf{P}$, where the matrix element $P_{i,j}$ denotes the frequency of cells in a nuclear state $j \leq n$ and containing m mutant mitochondria. The horizontal index j enumerates all possible nuclear states including the mating type and the mode of mitochondrial transmission.

As in Chapter 2, mitochondrial mutation is represented by the transition $\mathbf{P}^{(t,1)} = \mathbf{U}\mathbf{P}^{(t)}$, where the matrix element $U_{i,j}$ represents the probability that a cell with j mutant mitochondria will have i mutants after the mutation event,

$$U_{i,j} = \binom{M-j}{i-j} \mu^{i-j} (1-\mu)^{M-i}. \tag{3.2}$$

Mutation occurs in individual mitochondria, but selection acts on the level of the cell, i.e. on groups of mitochondria. In our model, cell fitness *w* directly depends only on the mitochondrial genotype. The updated population state after selection is then

$$\mathbf{P}^{(t,2)} = \frac{(\mathbf{I}\mathbf{w})\mathbf{P}^{(t,1)}}{\mathbf{w}^{\mathrm{T}}\mathbf{P}^{(t,1)}\mathbf{u}_n}, \tag{3.3}$$

where $\mathbf{I}$ is the identity matrix. $\mathbf{u}_n$ is a column vector of ones, so that $\mathbf{P}^{(t,1)}\mathbf{u}_n$ is the row-wise sum of $\mathbf{P}^{(t,1)}$ and $\mathbf{w}$ is a column vector containing all possible values of mitochondrial fitness $w_i = 1 - \left(\frac{i}{M}\right)^{\xi}$. Parameter ξ determines the magnitude of epistatic interactions between mitochondrial mutations. With $\xi > 1$ the fitness cost of every new mutation increases with the overall mutation load (negative epistasis). Negative epistasis between deleterious mitochondrial mutations leads to mitochondrial threshold effects, well known from empirical observations in eukaryotes (Rossignol et al., 2003).

Gametes in the gamete pool fuse at random, according to their equilibrium frequencies F_i. For asymmetric gamete unions, assuming that the gamete *k* is of maternal type, we have

$$\mathbf{z}_{kl} \sim \begin{cases} (\mathbf{\Phi}^{(\pi)}\mathbf{P}_{\bullet,k}^{(t,2)}) * (\mathbf{\Psi}^{(\pi)}\mathbf{P}_{\bullet,l}^{(t,2)}) \dfrac{2^{\delta_{k\neq l}} F_k F_l}{f_k f_l}, & \text{if } k \text{ and } l \text{ are compatible} \\ \mathbf{0}, & \text{otherwise} \end{cases} \tag{3.4}$$

An asterisk here denotes vector convolution, f_k is the frequency of cells in a nuclear state k, i.e. $f_k = \mathbf{u}_{\mathrm{M}+1}^{T}\mathbf{P}_{\bullet,k}^{(t,2)}$ and F_k is the corresponding equilibrium frequency in the gamete pool. The delta symbol $\delta_{k\neq l} = 1$ if $k \neq l$ and is 0 otherwise. The zygote-state vectors $\mathbf{z}_{kl}$ are scaled linearly to sum up to one. The two transition matrices $\mathbf{\Phi}^{(\pi)}$ and

$\boldsymbol{\Psi}^{(\pi)}$ are included to implement the mitochondrial inheritance bias where one of the gametes transmits more mitochondria than the other. We assume that the paternal gamete contributes πM mitochondria through sampling without replacement, and that $(2-\pi)M$ mitochondria come from the maternal gamete ($\pi = 0, \frac{1}{M}, \ldots, \frac{1-M}{M}, 1$). The two transition matrices therefore have elements

$$\Phi_{i,j}^{(\pi)} = \binom{[2-\pi]M}{i}\left(\frac{j}{M}\right)^{i}\left(1-\frac{j}{M}\right)^{(2-\pi)M-i}, \tag{3.5}$$

and

$$\Psi_{i,j}^{(\pi)} = \binom{j}{i}\binom{M-j}{\pi M-i}\binom{M}{\pi M}^{-1}. \tag{3.6}$$

The life cycle is completed by two meiotic cell divisions restoring the haploid state. At the start of the next generation the genotype frequencies then are

$$\mathbf{P}_{\bullet,k}^{(t+1,0)} = \mathbf{F}_2\mathbf{F}_1\left(\mathbf{z}_{kk} + \frac{1}{2}\sum_{l\neq k}\mathbf{z}_{kl}\right). \tag{3.7}$$

I do not differentiate between $\mathbf{z}_{kl}$ and $\mathbf{z}_{lk}$, i.e. both state vectors indicate the same zygote type. $\mathbf{F}_1$ and $\mathbf{F}_2$ are transition matrices for the two meiotic divisions implemented as mitochondrial sampling without replacement. Their corresponding elements are

$$F_{(1)i,j} = \binom{2j}{i}\binom{4M-2j}{2M-i}\binom{4M}{2M}^{-1}, \tag{3.8}$$

and

$$F_{(2)i,j} = \binom{j}{i}\binom{2M-j}{M-i}\binom{2M}{M}^{-1}. \tag{3.9}$$

The following results are based on the numerical solution of the above system of equations.

3.4 Asymmetric sex before mating types: UPI facilitates purifying selection against deleterious mitochondrial mutations

Eukaryotic sex depends on molecular mechanisms for self/non-self-recognition and gamete attraction (Goodenough and Heitman, 2014); similar mechanisms likely existed in both the bacterial ancestor of mitochondria and the archaeal host in the form of quorum-sensing and biofilm-formation systems triggered by external conditions (Ng and Bassler, 2009; Fröls, 2013). With the prokaryotic ancestry of eukaryotic gamete attraction machineries, the initial proto-eukaryotic self-recognition system was likely symmetrical, i.e. initial forms of sexual reproduction did not involve differentiation into distinct gamete classes (Goodenough and Heitman, 2014; Heitman, 2015). Self-incompatible mating types and sexes then could have been a later addition, driven, as I shall argue here, by purifying selection against mitochondrial mutations.

Assuming that the initial form of sexual reproduction was unisexual, I first model the evolution of uniparental inheritance in an ancestral population without mating types. While all pairs of gametes are capable of mating, only the unions between U_m and B (or U_p and B) involve asymmetric transmission of mitochondria (Fig. 3.1). As I show in Fig. 3.2, both nuclear alleles U_m and U_p invade and can reach the frequency of $e_2 = 0.5$, but the invasion dynamics differ substantially between the two modes of UPI control.

Starting at low allele frequencies, the invader U_m attains one of two distinct equilibria, $e_1 < 0.5$ under low mutation rates and strong negative epistasis, or $e_2 = 0.5$ under higher mutation rates (Fig. 3.2a)—the finding analogous to the modelling results of Hadjivasiliou et al. (2013). The analogous set of equilibria exists for the paternal invader U_p, but this time $e_2 = 0.5$ is the only asymmetric equilibrium which can be reached starting from low initial mutant frequencies (Fig. 3.2a). In this case the combination of the mutation rate and the initial frequency must exceed the unstable equilibrium $0 < e_1' < 0.5$, the characteristic frequency of which approaches zero at high mutation rates. Paternal leakage of mitochondria relaxes the conditions for the

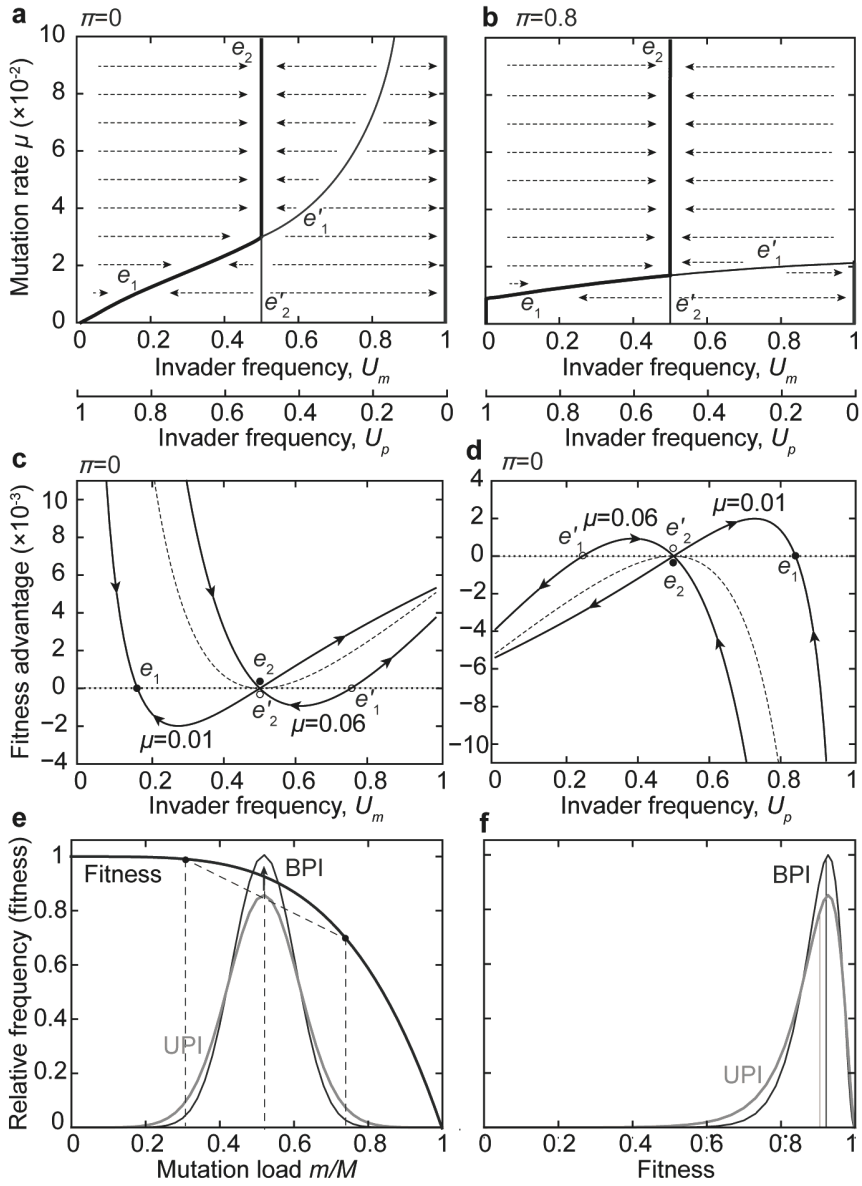

Figure 3.2. Invasion of the uniparental inheritance alleles in a population without mating types. Equilibria for the maternal (U_m) and paternal (U_p) UPI alleles for strict uniparental inheritance π=0 (**a**) and asymmetric inheritance with mitochondrial leakage π=0.8 (**b**). Equilibria e_1 and e_2 (thick lines) are stable, while e_1' and e_2' indicate unstable states. Fitness advantage of the UPI mutants U_m (**c**) and U_p (**d**) with π=0 depends on their frequency, explaining the stability of equilibria in (a). Dashed line indicates the fitness landscape at the saddle point $\mu = 0.03$, where equilibria e_1 and e_2 coincide. The strength of epistasis $\xi = 2$, $M = 50$. The asymmetric inheritance increases variance in the mutation load allowing for more efficient purifying selection and giving a long-term advantage, while mitochondrial mixing under the biparental inheritance (BPI) increases the frequency of intermediate cytoplasmic states (**e**). But with negative epistasis between mitochondrial mutations (**e**), mitochondrial mixing results in higher fitness than would be expected under linear fitness interactions (**f**), giving BPI a short-term advantage.

invasion of U_p to the equilibrium e_2 (Fig. 3.2b), but at the same time hinders the invasion of U_m under low mutation rates.

What determines the equilibrium frequencies of UPI alleles in a population without mating types? The model shows that the fitness advantage of U_m decreases as the allele invades from small frequencies (Fig. 3.2c), while the opposite is true for U_p (Fig. 3.2d), determining the locations of equilibria e_1 and e_2. This behaviour can be explained in terms of costs and benefits of asymmetric transmission of mitochondrial genes, and the statistical association of these benefits to the nuclear allele of mitochondrial inheritance control.

Uniparental inheritance of mitochondria increases the frequency of extreme cytotypes underrepresented at the mutation-selection equilibrium, reducing the abundance of the common intermediate cytoplasmic states (Fig. 3.2e). This increase in mitochondrial variance facilitates purifying selection against deleterious mutations and gives the UPI mutant a long-term advantage over the biparental population. Nevertheless, under negative epistasis (Fig. 3.2e) intermediate cytoplasmic states have higher fitness than expected from the linear combination of the extremes, penalizing the UPI invader and giving the biparental inheritance a short-term benefit (Fig. 3.2f). These short-term fitness effects, however, are relevant only if the association between the cytoplasmic state and the nuclear allele of inheritance control is weak (e.g. with paternal leakage, when part of the mitochondrial population is inherited from the unrelated gamete); otherwise the long-term variance-based effects dominate. The stable equilibria e_1 and e_2 are located where short-term effects match the long-term fitness advantage of the UPI invader.

Since the rate at which the resident *B* inherits the cytoplasm from the uniparental mutant rises with increasing frequency of U_m, the short-term fitness gains of *B* increase, halting the invasion at $e_1 < 0.5$ or $e_2 = 0.5$. Paternal leakage associated with U_m weakens the mito-nuclear associations and therefore reduces the strength of the long-term fitness effects. The opposite pattern of nuclear-cytoplasmic linkage applies to the invader with paternally-determined UPI: destroying part of one's own mitochondria increases the level of transmission asymmetry, but weakens the

statistical associations between the mitochondrial population and the allele U_p. As the allele U_p invades from small frequencies, the rate of biparental unions $U_p \times U_p$ increases, but so does the overall frequency of uniparental transmissions. The short-term fitness gains of invading U_p therefore increase with its overall frequency. As the invasion proceeds, the biparental unions between the identical gametes become more common, eventually reversing the trend of frequency-dependence (Fig. 3.2c, d).

3.5 Evolving paternal leakage in UPI without mating types

Suppose now that the level of paternal leakage π is itself an evolvable trait, controlled by a single nuclear allele U_p or U_m. The number of organelles transmitted to the progeny from each gamete can indeed be regulated genetically, e.g. by controlling the expression of gamete-specific protein markers tagging the organelles for selective degradation (e.g. ubiquitin). What pattern of organelle inheritance would we expect to evolve in the population without mating types? I performed an evolutionary invasion analysis to find the evolutionarily stable states (ESS; Eshel, 1983; Geritz et al., 1998) for paternal leakage π under both maternal and paternal modes of inheritance control. With the allele *B* fixed, I introduce the uniparental allele U_p or U_m corresponding to $\pi < 1$, and find its equilibrium frequency. A new uniparental-inheritance allele (an invader) is then inserted with a value of π different from the resident, and its spread is tracked until it either replaces the resident or is eliminated. The process is repeated for all values of π, finding the uninvadable states.

With maternal regulation of cytoplasmic inheritance, the sole non-invadable state with asymmetric transmission of mitochondria is the strict UPI, i.e. $\pi_{\mathrm{ESS}} = 0$ (Fig. 3.3a), at which the gamete U_m discards all mitochondria inherited from *B*. As paternal leakage π goes down, both the long-term fitness advantage of variance, and the nuclear-cytoplasmic associations become stronger. The equilibrium frequency of U_m

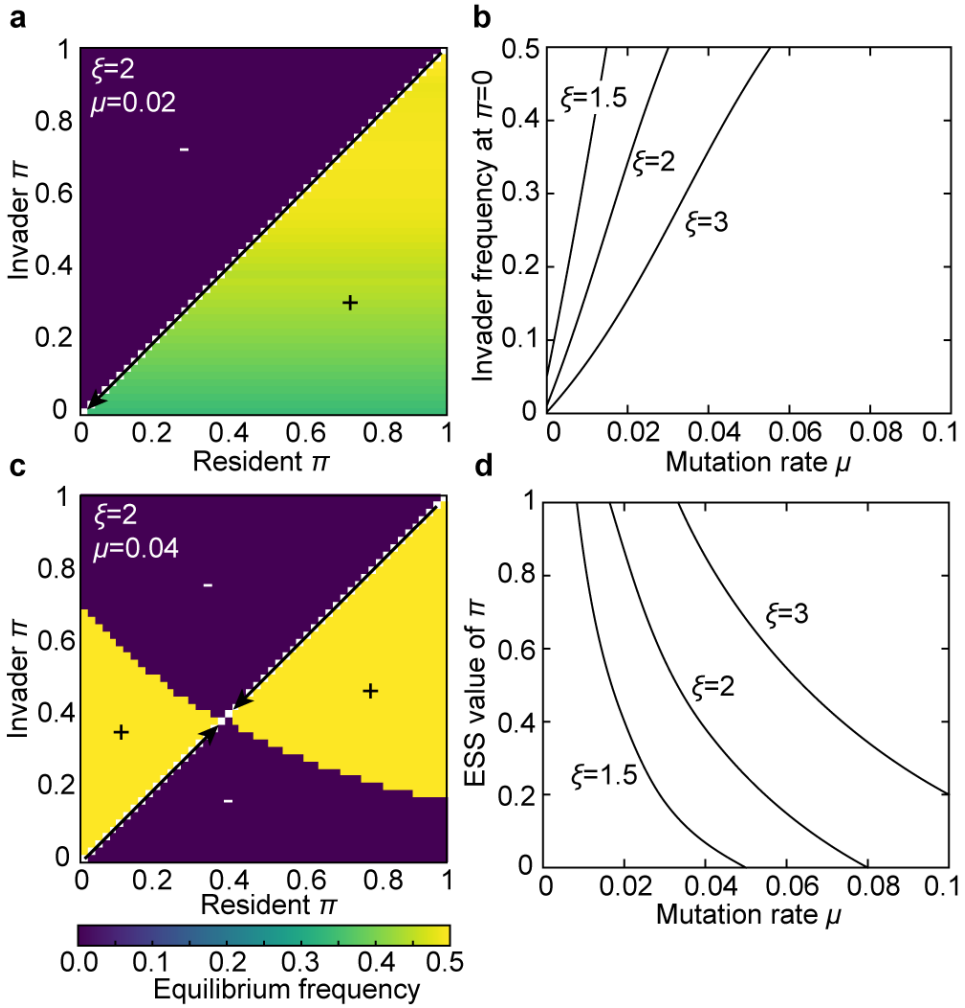

Figure 3.3. Evolutionary stability of asymmetric cytoplasmic inheritance in a population without mating types. Pairwise invasibility plot shows that with maternal control of cytoplasmic inheritance, the only asymmetric ESS is the strict UPI, i.e. π=0 (**a**), with the frequency of U_m at the ESS increasing with higher mutational load (**b**). Note that under different parameter values some resident states might not exist, as they do not invade the pure BPI state. In contrast, with paternal control the ESS can lie anywhere between the fully symmetric and strict uniparental inheritance (**c**), depending on mutation rates and epistatic interactions (**d**).

at π_{ESS} increases with mutation rate, consistent with the role of UPI increasing the efficacy of purifying selection against mitochondrial mutations (Fig. 3.3b). Likewise, weaker epistatic interactions ($\xi \to 1$) reduce the short-term fitness gains in symmetric gamete unions, and result in higher equilibrium frequencies of U_m (Fig. 3.3b).

Under paternal control of mitochondrial inheritance, the globally-attracting ESS lies between $0 \leq \pi_{\mathrm{ESS}} \leq 1$ (Fig. 3.3c), with the frequency of U_p allele always at $e_2 = 0.5$. The deviation of π to either side of its π_{ESS} is detrimental. With $\pi < \pi_{\mathrm{ESS}}$ the asymmetry of mitochondrial inheritance increases, but the strength of the long-term

mitochondrial-nuclear associations is diminished, whereas with $\pi > \pi_{\mathrm{ESS}}$, the strength of genetic linkage increases, but the long-term effects of partial UPI are reduced—in both cases to the detriment of the invader. The analysis further shows that the value of π_{ESS} goes down with increasing mutation rate μ and decreasing strength of epistasis ξ, as it reduces the short-term benefit of paternal leakage (Fig. 3.3d). The model therefore shows that limited mitochondrial mixing can be maintained in spite of its long-term fitness costs, but only if the inheritance-restriction allele causes the selective destruction of organelles inherited from the same mating type (paternal gamete by definition), and if the epistasis is negative. This observation will be further explored in Chapter 4; here it serves to illustrate that stable paternal leakage can be maintained and therefore must be accounted for in the further analysis of mating type evolution.

3.6 Invasion of self-incompatible mating-type alleles establishes the population-wide UPI

The above analysis shows that in a population without mating types the following equilibria with asymmetric inheritance of mitochondria are possible, depending on mutation rates and the mode of UPI nuclear regulation:

1) U_m at $e_1 < 0.5$ with strict UPI ($\pi = 0$);
2) U_m at $e_2 = 0.5$ with strict UPI ($\pi = 0$);
3) U_p at $e_2 = 0.5$ with paternal leakage ($0 \leq \pi \leq 1$).

The stable equilibrium for the allele U_p at $e_1 > 0.5$ (Fig. 3.2a, b) is not considered here, as it cannot be reached from low initial allele frequencies. These equilibria do not depend on the gamete mortality rate δ, as all gametes are universally compatible in the absence of mating types. While the frequency of uniparental mutants U_m or U_p can reach 0.5, the overall rate of asymmetric unions cannot exceed 0.5 as long as biparental matings between identical gametes (U_m×U_m, U_p×U_p, B×B) are allowed. The biparental gamete unions can be eliminated, if a mutation in the intracellular signalling

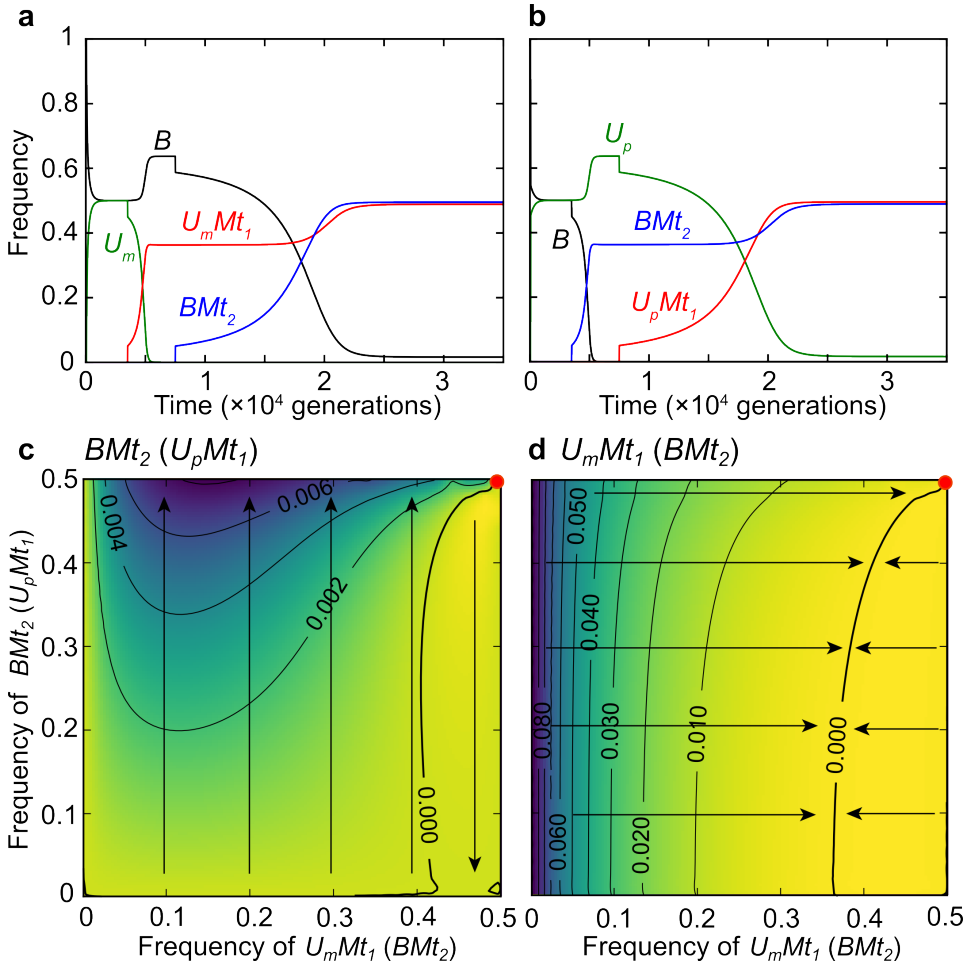

Figure 3.4. Invasion of self-incompatibility alleles *Mt₁* and *Mt₂* linked to the mitochondrial inheritance locus *U_m*/*B* (*U_p*/*B*). Due to the higher rate of uniparental gamete unions, U_mMt_1 (or BMt_2 under paternal control) replaces the universally compatible allele U_m (B), but to a lower frequency (**a**, **b**). Similarly, the self-incompatible BMt_2 (U_pMt_1 under paternal regulation) then increases in frequency at the expense of B (U_p), as it makes the asymmetric gamete unions more frequent. The fitness advantage of the invading BMt_2 (U_pMt_1) increases with its frequency, assuming fixed frequencies of other alleles (**c**). The fitness advantage of U_mMt_1 (BMt_2) over B (U_p) also increases as BMt_2 (U_pMt_1) invades, as less frequent biparental gamete unions reduce the short-term fitness advantage of the universal resident B (**d**). Arrows indicate the expected direction of evolution in the population consisting of U_mMt_1, BMt_2 and B. Parameter values are $\xi = 1.5$, $\mu = 0.03$, $\pi = 0.2$, $\delta = 10^{-5}$.

system leads to the incompatibility of cells carrying the same allele at the mitochondrial inheritance locus.

I next consider the invasion of self-incompatibility alleles Mt_1 and Mt_2, linked to either U_m/U_p or B. Due to the inherent symmetry of the system, the invasion of U_mMt_1 under maternal inheritance regulation is in fact equivalent to invasion BMt_2 under paternal control; similarly, the invasion of U_pMt_1 is equivalent to the invasion of BMt_2 under maternal regulation of cytoplasmic inheritance. It is therefore formally sufficient

to study the evolution under maternal control, as the case of paternal regulation can be obtained by simply reversing the order of allele invasion, given that the initial equilibrium state exists and is stable.

First, consider the maternal allele U_m at equilibrium and introduce the mutant form U_mMt_1. Self-incompatible gamete class U_mMt_1 invades and replaces U_m, but to a lower frequency than the ancestral U_m (Fig. 3.4a, b). Linkage to the mating type allele Mt_1 ensures that symmetric unions between gametes carrying U_m are forbidden, increasing the frequency of uniparental matings and therefore leading to the higher long-term fitness advantage of U_m associated with more efficient removal of deleterious mutations. In a population consisting solely of U_mMt_1 and B, all gamete unions involving U_mMt_1 are therefore uniparental. At the same time, however, the subpopulation of B enjoys higher short-term fitness benefits of mitochondrial mixing in symmetric $B\times B$ unions, limiting the spread of the first mating type allele.

Consider now the invasion of a second self-incompatibility allele Mt_2 linked to B. The gamete class BMt_2 spreads at the expense of B, as it reduces the frequency of fully symmetric gamete unions and ensures that BMt_2 inherits fit mitochondria from U_mMt_1 uniparentally more often than the wild-type B. With the constant frequency of U_mMt_1, the fitness advantage of BMt_2 over B increases with its frequency (Fig. 3.4c). The spread of BMt_2 also increases the long-term fitness advantage of the first invader U_mMt_1 (Fig. 3.4d), as it reduces the short-term advantage of B. The spread of a second mating-type allele therefore reinforces the fitness advantage of the first, allowing both to reach high frequencies and even fix at low gamete mortality rates δ (Fig. 3.4a, b). Similar dynamics are observed if the order of mating-type invasion is reversed, but the initial invasion of BMt_2 is now favoured less, as the allele's only advantage is the reduced frequency of symmetric gamete fusions. Under higher gamete mortality rates δ, BMt_2 does not invade unless U_mMt_1 is already present. Here I therefore focus on a more permissive case of U_mMt_1 invading first.

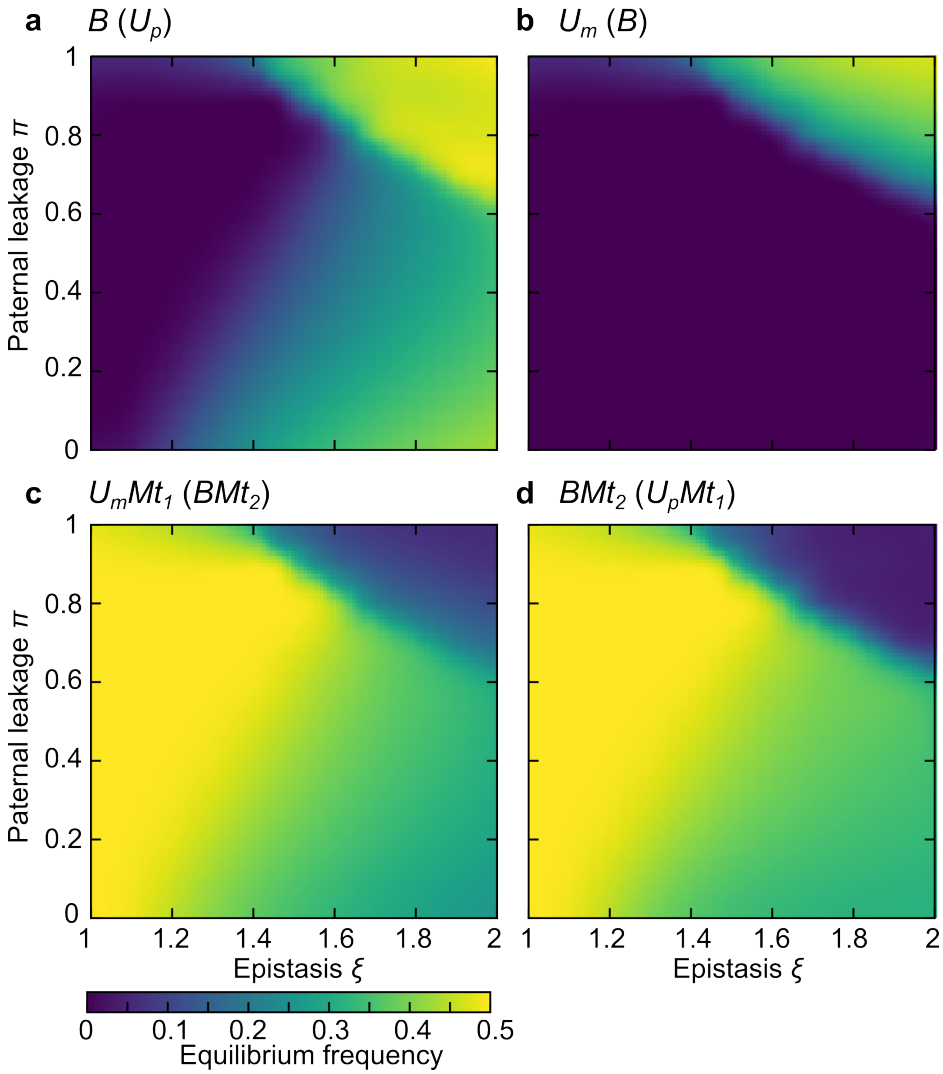

Figure 3.5. Invasion of mating-type alleles is favoured by weak epistasis between mitochondrial mutations. Self-incompatible mating types reach fixation under weak epistatic interactions, where the long-term fitness advantage of UPI is maximized. With stronger epistasis, the fixation of mating types is facilitated by intermediate levels of paternal leakage, indicating the importance of paternally regulated mitochondrial inheritance. Under strong negative epistasis and significant paternal leakage, mating types fail to replace the universally compatible gametes due to relatively weak long-term variance effects. Gamete death rate is set to $\delta = 0.001$, the mitochondrial mutation rate is $\mu = 0.03$. Genotypes in brackets correspond to paternal control of mitochondrial inheritance.

3.7 Conditions favouring fixation of mating types

The two mating type alleles fix only if they are capable of replacing the universally compatible forms U_m/U_p and B. This is opposed by two forces: 1) the short-term fitness advantage of mitochondrial mixing in subpopulations without mating types, and 2) the mating rate disadvantage of gametes that can mate only with a subset of the population. The spread and fixation of Mt_1 and Mt_2 is favoured by factors increasing

the long-term fitness effects of asymmetric transmission of mitochondria, and low gamete death rates δ.

First, regardless of the UPI regulation mode, the two self-incompatible mating types replace unisexual residents under weak epistatic interactions between mitochondrial mutations ($\xi \to 1$, Fig. 3.5). The short-term fitness advantage of mitochondrial mixing in biparental gamete unions applies only to the case of negative epistasis between deleterious mutations (Chapter 2); at $\xi = 1$, for example, the uniparental invaders become universally advantageous, with their equilibrium frequency being limited only by the frequency of compatible mating partners. As the strength of negative epistatic interactions increases ($\xi > 1$), the short-term fitness effects of mitochondrial mixing become increasingly more important, opposing the spread of the two mating-type alleles. This effect can be partially alleviated in the presence of paternal leakage, as it tends to reduce the short-term fitness advantage of universally compatible gametes, favouring the invasion of mating types (Fig. 3.5). The results therefore indicate the importance of paternal regulation of mitochondrial inheritance, as paternal leakage can be evolutionarily stable only under paternal control of cytoplasmic inheritance (note that the ESS level of paternal leakage depends on the mutation rate and epistasis (Fig. 3.3d), but also changes as the mating types invade). Under strong epistasis and high levels of paternal leakage π, however, the first mating type $U_m Mt_1$ fails to replace the universal form U_m, which subsequently prevents the invasion of Mt_2 (Fig. 3.5).

High mitochondrial mutation rates favour the spread of mating-type alleles linked to the mitochondrial-inheritance locus (Fig. 3.6). This is consistent with the long-term effect of asymmetric mitochondrial transmission enhancing the efficacy of purifying selection against mitochondrial mutations due to higher variance in mitochondrial mutation load (Hadjivasiliou, 2013; Chapter 2). As expected, the evolutionary success of self-incompatibility alleles depends on the cost associated with

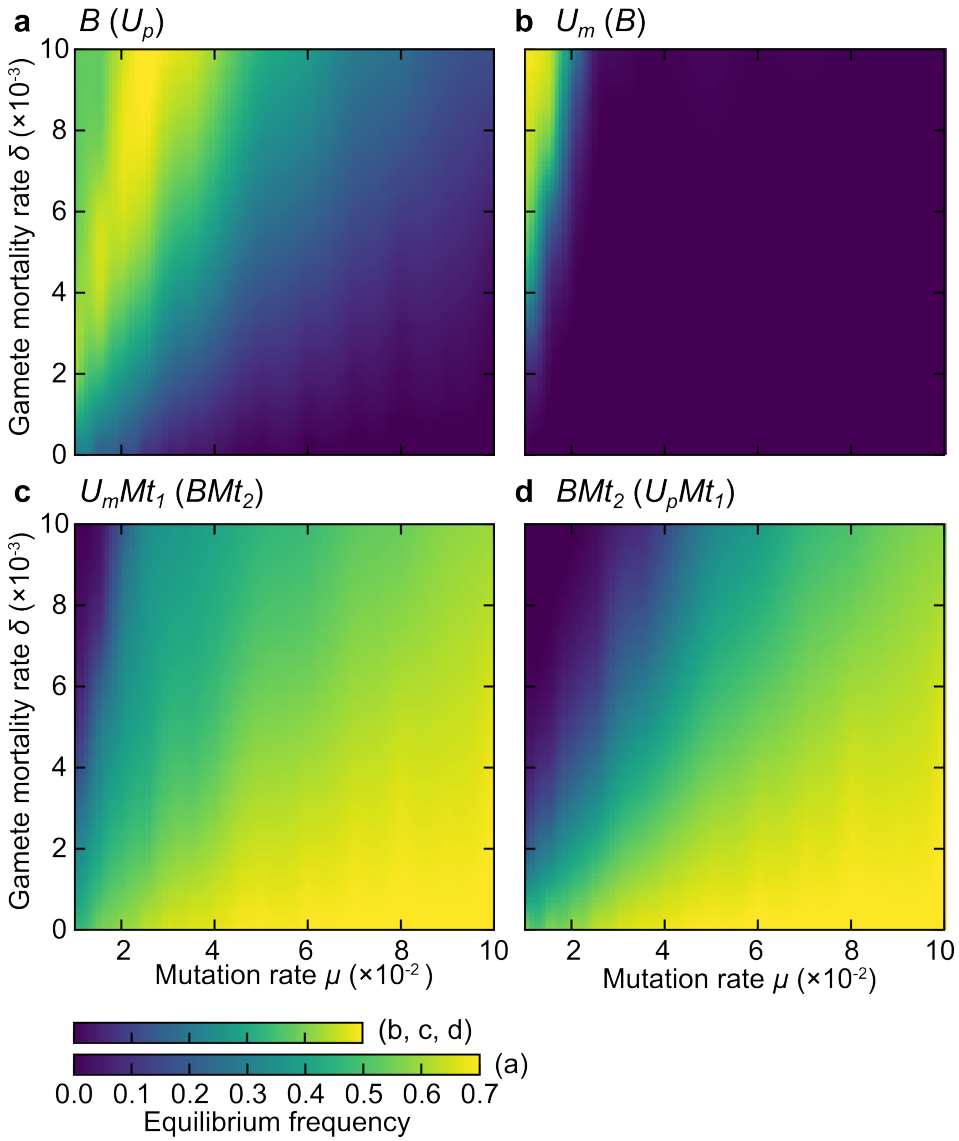

Figure 3.6. Invasion of self-incompatible mating type alleles is favoured by high mutation rates and low gamete mortality. Self-incompatible gamete classes U_mMt_1 (BMt_2 under paternal control) and BMt_2 (U_pMt_1) replace universal gamete types U_m (B) and B (U_p) under high mutation rates μ and low gamete mortality rates δ. The strength of epistasis is $\xi = 1.4$, the rate of paternal leakage of mitochondria is set to $\pi = 0.2$.

the lower frequency of potential mating partners within the population, which in the model is represented by the gamete mortality rate δ (Fig. 3.6). Under high mortality rates, gametes unable to find suitable mates are rapidly eliminated from the gamete pool; consequently, the universal gametes carrying B or U_m/U_p prevail. With low gamete death rates ($\delta \to 0$), on the other hand, self-incompatible mating types can persist longer, waiting for an encounter with a compatible mating partner, which favours their invasion and eventual fixation (Fig. 3.6). Indeed, while gamete self-

incompatibility reduces the overall frequency of potential mating partners within the population, with the evolution of mating types and mating-type associated pheromone-based intracellular attraction systems, gamete encounters are rarely random (Hadjivasiliou et al., 2015; Hadjivasiliou and Pomiankowski, 2016), and the part of the life cycle available for mate-finding could indeed be much longer than the time separating random encounters between mature reproductive cells.

3.8 Discussion

The existence of two sexes or mating types is often regarded as an evolutionary conundrum, as both males and females can mate only with one-half of the population. The fitness benefits of gamete differentiation must therefore exceed the costs associated with reduced mating opportunities. With the evolution of two mating types and associated asymmetric signalling systems (Hoekstra, 1982; Hadjivasiliou et al., 2015; Hadjivasiliou and Pomiankowski, 2016), the encounters between gametes are not random, and the cost of being able to mate with only a part of the population might not be as high as assumed in previous studies (Hadjivasiliou et al., 2013). I showed, that two mating-type alleles linked to the mitochondrial-inheritance locus can spread to fixation driven by the long-term advantage of asymmetric transmission of mitochondrial genes, which increases the efficacy of purifying selection against detrimental mitochondrial mutations, even though their invasion is opposed by the short-term fitness benefits of mitochondrial mixing in the biparental part of the population, and the limited availability of potential mating partners. Two mating types should therefore be expected to arise with fast accumulation of mitochondrial mutations, and high gamete-survival rates ensuring that fitness costs associated with the need to locate a suitable mating partner are low. Additionally, the results show that paternal leakage of mitochondria can favour the evolution of mating types under

negative epistasis, suggesting that paternal control of cytoplasmic inheritance could have played an important role in the evolution of self-incompatible gamete classes.

Can the model account for the evolution of more than two gamete classes? Multiple mating types with uniparental inheritance of mitochondria are indeed known. For example, in true slime mould *Physarum polycephalum* with several mating types, uniparental inheritance of mitochondria via selective organelle digestion follows a complex hierarchy according to the allele at one of the mating-type loci (Moriyama and Kawano, 2003). Consider the invasion of a third self-incompatible mating type Mt_3, linked to the mitochondrial inheritance locus with one of the existing alleles, say U_m. The invader can therefore mate with both initial gamete classes, but part of the gamete unions will remain biparental ($U_mMt_3 \times U_mMt_1$). The model shows that U_mMt_3 can invade only if it has a large initial mating-rate advantage, i.e. under gamete mortality rates δ high enough to compensate for the long-term fitness disadvantage due to less frequent uniparental unions. Interestingly, unions between certain pairs of gamete types in *Physarum polycephalum* are indeed biparental, indicating that some of the mating-type alleles might have arisen simply because of their mating-rate advantage (Moriyama and Kawano, 2003). Similar selective pressure has been suggested as a driving force for the evolution of multiple mating types in fungi without mobile gametes (Hurst, 1995).

The evolutionary stability of two mating types can therefore be explained by the high long-term fitness cost of biparental inheritance in the newly invading gamete class compared to the advantage of having more compatible mating partners. On the other hand, the model shows that a third mating type allele invades much more easily in tandem with a new allele at the mitochondrial-inheritance locus U_{m1} coding for a novel mitochondrial recognition and destruction machinery. In this case, unions between all three gamete types remain uniparental throughout the invasion. A stable population with three mating types at equal frequencies can therefore become established even at low gamete mortality rates δ; the same remains true for subsequently invading mating-type alleles. Under these assumptions the number of mating types in the

population would seem to be limited only by the amount of distinct molecular mechanisms controlling the transmission of organelle genomes. This scenario, however, is critically dependent on the simultaneous invasion of two novel mutations at linked loci, both controlling vastly complex processes of mate recognition and mitochondrial destruction, and is therefore highly unlikely.

How does the present analysis compare to the previous work suggesting that mating types could have evolved together with the uniparental inheritance of cytoplasmic genes? In the model of Hurst and Hamilton (1992), gamete classes evolve to eliminate the conflict between haploid organelle genomes inherited from distinct parents, where the selfish organelle destroys its opponent in the zygote (Hurst, 1995). Similarly, Hutson and Law (1993) considered the spread of “selfish” endosymbionts with fixed fitness costs. These authors have not considered intercellular dynamics nor mitochondrial segregation, both of which are now known to be of key importance in the evolution of uniparental inheritance (Hadjivasiliou, 2013; Chapter 2). Multiple cases of selfish organelles and genetic elements are known (Schable and Wise, 1998; Taylor et al., 2002; Clark et al., 2012), but it is unlikely that they occur at high enough frequencies to account for the striking universality of mating types with uniparental transmission of cytoplasm in eukaryotes. Selection against deleterious mitochondrial mutations, on the other hand, provides a more general explanation. Given the central role of mitochondria in eukaryotic metabolism, it is not surprising that mechanisms facilitating the removal of deleterious mitochondrial mutations are selected for. The evolution of two mating types with uniparental organelle transmission might therefore be a direct consequence of the requirement for high quality mitochondria in complex eukaryotes.

CHAPTER 4. SEXUAL CONFLICT EXPLAINS THE EXTRAORDINARY DIVERSITY OF MECHANISMS REGULATING MITOCHONDRIAL INHERITANCE

4.1 Summary

Uniparental inheritance of mitochondria is nearly universal among eukaryotes and is believed to facilitate selection against deleterious mitochondrial mutations and restrict inter-genomic conflicts. But this understanding of cytoplasmic inheritance is challenged by frequently detected paternal leakage of mitochondria and persistent heteroplasmy. It is typically assumed that paternal leakage and heteroplasmy are just episodic deviations from the general rule, indicating the breakdown of the uniparental inheritance machinery, and are not adaptive in their own right. In this chapter I present a new mathematical model for the evolution of nuclear alleles controlling the level of asymmetry in cytoplasmic inheritance under the effect of purifying selection. By explicitly considering the role of both sexes in the control of cytoplasmic transmission, I show that with maternal regulation, strict uniparental transmission is the only evolutionarily stable asymmetric state, whereas paternal leakage is stable under paternal control. Cytoplasmic mixing and heteroplasmy are therefore outcomes of the tension between selection on females and males, and occur even if the result is a long-term fitness cost inflicted on both sexes. Competition over the control of cytoplasmic transmission explains the recurrent evolution of strikingly diverse mechanisms involved in controlling asymmetric organelle transmission.

4.2 Introduction

Sexual reproduction in eukaryotes involves the fusion of nuclei from both gametes, but the inheritance of organelles containing their own DNA—mitochondria and chloroplasts—from only one of them (Birky, 1995; Sato and Sato, 2013; Greiner et al., 2015). This pattern of uniparental inheritance (UPI) is nearly universal across eukaryotes, from isogamous protists with equal-sized gametes through to animals and plants with extreme gamete-size asymmetry (i.e. oogamy and diminutive sperm), and is believed to facilitate purifying selection against deleterious mutations (Bergstrom and Pritchard, 1998; Hadjivasiliou et al., 2015; Chapter 2; Chapter 3), restrict inter-genomic conflicts (Eberhard, 1980; Cosmides and Tooby, 1981) and prevent heteroplasmy (Christie et al., 2015). Cytoplasmic mixing, in contrast, reduces inter-individual variation, ultimately impeding the efficacy of selection against defective organelles or selfish genetic elements (Rand, 2008; Bastiaans et al., 2014).

Asymmetric inheritance arises from active mechanisms beyond those due to simple gamete-size difference. Nuclear genes restrict organelle transmission from the incoming mating type gamete (maternal control) or from a mating type's own cytoplasm (paternal control). Multiple attempts at modelling the evolution of asymmetric organelle inheritance have concluded that the lack of long-term linkage between the nuclear genotype and the maternally inherited cytoplasm should prevent the evolution of paternally-controlled organelle destruction (Hastings, 1992; Randerson and Hurst, 1999; Hoekstra, 2011). These studies suggest that maternally-controlled elimination of paternal mitochondria should dominate in nature, which is indeed consistent with some empirical observations. In *Ascidian tunicates*, for instance, male organelles are prevented from entering the oocyte (Ursprung and Schabtach, 1965), while maternal autophagy machinery eliminates paternal mitochondria in *Caenorhadbitis elgans* (Sato and Sato, 2011; Al Rawi et al., 2011; Zhou et al., 2016). More notably, in the fungal plant pathogen *Ustilago maydis*, *lga*2 and *rga*2 genes expressed in mating type *a*2 are

responsible for the selective elimination of the opposite mating type's mitochondrial DNA (mtDNA) after fusion, at the same time protecting the mtDNA of the mating type *a*2 (Fedler et al., 2009).

In other cases, however, paternal mtDNA is eliminated without any involvement of the maternal mating type. For instance, paternal control of cytoplasmic inheritance has been shown to operate in *Drosophila melanogaster*, where mtDNA is actively degraded during spermatogenesis (DeLuca and O'Farrell, 2012). Similar elimination of mtDNA during spermatogenesis has been reported in fish and mice (Nishimura et al., 2006; Luo et al. 2013). In addition, the control of mitochondrial inheritance often involves both parents. Mitochondria in bovine and primate sperm are modified with ubiquitin, which serves as a signal for selective degradation after gamete fusion (Sutovsky et al., 1999). Similarly, in isogamous basidiomycete *Cryptococcus neoformans, SXI1α* in MATα and *SXI2a* in MATa are both required for the uniparental inheritance of MATa mitochondria (Yan et al., 2007). In all these cases, it appears that one mating type is responsible for tagging and the other for recognition and selective degradation of paternal organelles. It is clear that paternal involvement ensuring the asymmetric transmission of mtDNA is more important than current theoretical views predict. The striking diversity of both maternal and paternal mechanisms indicates repeated evolution, and presents an additional puzzle, given the seemingly universal advantage of organelle transmission from one sex.

Another challenge to the current theoretical views came with the advent of next-generation sequencing, demonstrating that paternal leakage of mitochondria and persistent heteroplasmy are not as rare as traditionally thought (Barr et al., 2005; Xu, 2005; McCauley, 2013). Biparental inheritance has been documented in diverse groups of animals, including mammals, arthropods, fish and birds, involving both interspecific (Kondo et al., 1990; Kaneda et al., 1995) and intraspecific matings (Shitara et al., 1998; Schwartz and Vissing, 2002; Zhao et al., 2004; Kvist et al., 2003; Magoulas and Zouros, 1993; Sherengul et al., 2006; Wolff et al., 2013). These

observations show that mechanisms preventing the inheritance of paternal mtDNA might be leaky at best or prone to evasion and failure. While biparental transmission in hybrid crosses could be explained by incompatibilities of molecular organelle tagging and recognition machineries, it is not clear whether paternal leakage within isolated populations is just an episodic breach of a strict rule or is adaptive in its own right. The enigmatic case of doubly-uniparental inheritance in bivalve molluscs, where heteroplasmy in male somatic tissues is common, points towards the latter (Passamonti and Ghiselli, 2009), but the theoretical explanation is lacking.

In this chapter I present a model of the evolution of uniparental inheritance controlled by either of the two isogamous mating types under the effect of purifying selection against mitochondrial mutations. In contrast to previous work, I assume that paternal leakage of mitochondria is an evolvable trait and use an adaptive dynamics approach to specify conditions under which strict UPI or varying degrees of paternal leakage are evolutionarily stable. I show that mitochondrial mixing can be selected under paternal control of cytoplasmic transmission and negative epistasis, and provide the first theoretical explanation for the prevalence of paternal leakage, heteroplasmy and the repeated evolution of diverse mechanisms of uniparental inheritance.

4.3 Mathematical model

Consider an infinite population of haploid unicellular organisms with two mating types, “maternal” *MTmat* and “paternal” *MTpat*, at equal frequencies. Each cell harbours M mitochondria during the haploid stage of the life cycle. A single nuclear locus controls the pattern of mitochondrial inheritance, which can vary from strict uniparental inheritance (no paternal leakage, $\pi = 0$) to complete biparental transmission ($\pi = 1$). The state of a population at any time t can be represented by the $(M + 1)\times 2$ matrix $P^{l(t)}$, where index l denotes a mating type, $l = 0,\ 1$. The matrix element $P_{m,j}^{l(t)}$ then

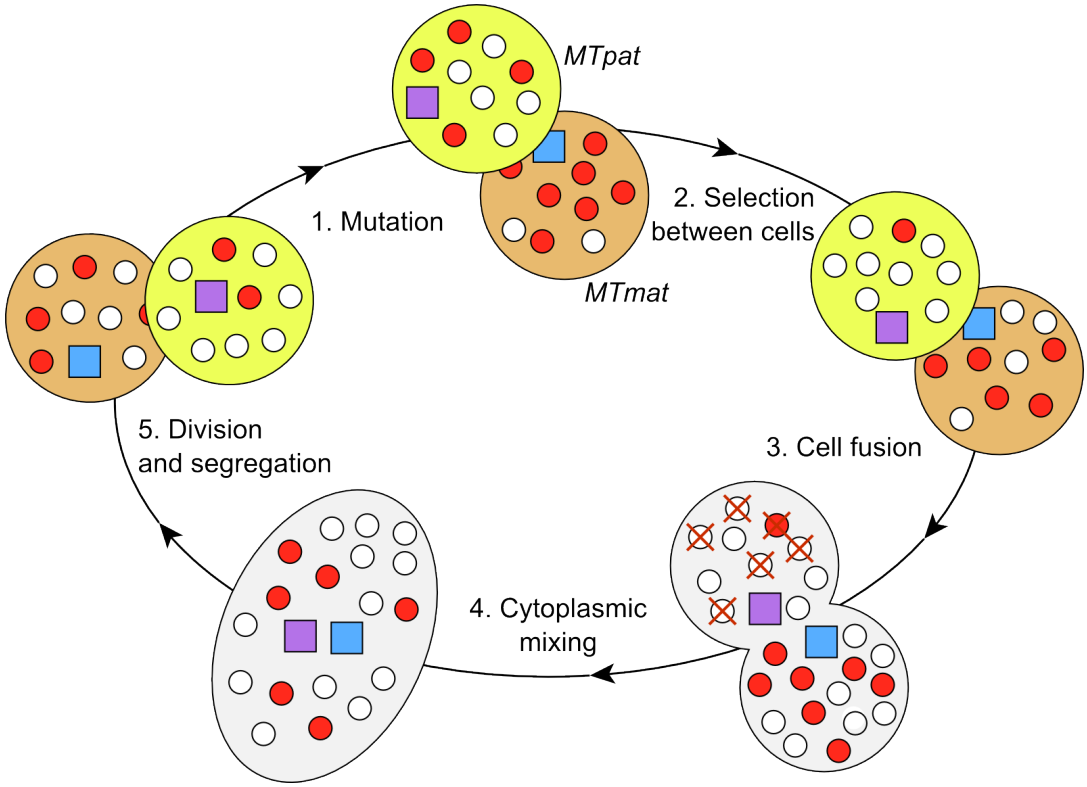

Figure 4.1. Model life cycle. The life cycle consists of discrete steps of mitochondrial mutation (1), selection between eukaryotic cells (2), random mating with cell fusion (3) and mitochondrial mixing (4), and division with random mitochondrial segregation (5). Paternal gamete (*MTpat*) contributes to the zygote only πM out of its M mitochondria, while the rest $(2-\pi)M$ are sampled from the maternal gamete (*MTmat*). The amount of paternal leakage π is controlled by one of the mating types. Wild-type and mutant mitochondria are represented by small circles, while squares depict nuclear mating-type loci.

represents the frequency of cells of mating type l, with m mutant mitochondria and a nuclear state j ($j = 0$ for the wild–type allele and $j = 1$ for a mutant invader).

The life cycle consists of discrete and non-overlapping generations with distinct steps for mutation, selection and mating (Fig. 4.1). I consider two mitochondrial states, wild-type and a mutant state with impaired respiration. Wild–type mitochondria mutate at rate μ, but the reverse transition is ignored as the probability of a back-mutation is much lower. The state of the population at generation t and after the mutation step is therefore $\mathbf{P}^{l(t,1)} = \mathbf{U}\mathbf{P}^{l(t)}$, where $\mathbf{U}$ is a $(M+1)\times(M+1)$ transition matrix. The matrix element $U_{i,j}$ represents the probability that a cell with j mutant mitochondria will have i mutants after the transition,

$$U_{i,j} = \binom{M-j}{i-j} \mu^{i-j} (1-\mu)^{M-i}. \quad (4.1)$$

Mutation is followed by selection, i.e. the change in genotype frequencies according to their fitness values. In our model, a cell's fitness w depends only on the mitochondrial part of the genotype. The updated state of the population after selection is then

$$\mathbf{P}^{l(t,2)} = \frac{(\mathbf{Iw})\mathbf{P}^{l(t,1)}}{\mathbf{w}^{\mathrm{T}}\mathbf{P}^{l(t,1)}\mathbf{u}_2}, \quad (4.2)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{u}_2 = (1,1)^{\mathrm{T}}$ is a column vector of ones and $\mathbf{w}$ is a column vector containing all $M+1$ possible values of mitochondrial fitness,$w_i = 1 - \left(\frac{i}{M}\right)^{\xi}$. Parameter ξ determines the magnitude of epistasis between mitochondrial mutations. $\xi = 1$ corresponds to the simplest case of additive fitness, where the fitness cost of each new mutation does not depend on the total mutational load. With $\xi > 1$ one has negative epistasis, where the fitness cost of several combined mutations is lower than expected under the additive model, but increases with every new mutation. This leads to the mitochondrial threshold effects observed in experimental studies (Shoffner et al., 1990; Miyabayashi et al., 1992; Rossignol et al., 2003).

Individuals of opposite mating types fuse at random, forming a population of diploid zygotes containing $2M$ mitochondria each. Let $\mathbf{z}_{gh}$ be a column vector with the i-th element representing the frequency of zygotes containing i mutant mitochondria and nuclear alleles $g = 0,1$ and $h = 0,1$, inherited from mating types *MTmat* and *MTpat* respectively. Zygote frequencies are then

$$\mathbf{z}_{gh} = (\mathbf{\Phi}^{(\pi)}\mathbf{P}^{0(t,2)}_{\bullet,g}) * (\mathbf{\Psi}^{(\pi)}\mathbf{P}^{1(t,2)}_{\bullet,h}), \quad (4.3)$$

where asterisk denotes vector convolution and the two transition matrices $\mathbf{\Phi}^{(\pi)}$ and $\mathbf{\Psi}^{(\pi)}$ are included to implement the mitochondrial inheritance bias. I assume that the paternal mating type *MTpat* contributes πM mitochondria through sampling without replacement, with $(2-\pi)M$ mitochondria coming from the maternal mating type

MTmat via random sampling with replacement ($\pi = 0, \frac{1}{M}, \ldots, \frac{1-M}{M}, 1$), although alternative sampling methods have been tested and shown to not affect the outcome. The two transition matrices then have elements

$$\Phi_{i,j}^{(\pi)} = \binom{[2-\pi]M}{i} \left(\frac{j}{M}\right)^{i} \left(1 - \frac{j}{M}\right)^{(2-\pi)M-i} \tag{4.4}$$

and

$$\Psi_{i,j}^{(\pi)} = \binom{j}{i} \binom{M-j}{\pi M - i} \binom{M}{\pi M}^{-1}. \tag{4.5}$$

The cell cycle is completed by the two-step meiosis restoring the haploid state. Genotype frequencies at the start of the next generation are then

$$\begin{aligned} \mathbf{P}_{\bullet,v}^{0(t+1)} &= \mathbf{F}_2\mathbf{F}_1(\mathbf{z}_{v0} + \mathbf{z}_{v1}), \\ \mathbf{P}_{\bullet,v}^{1(t+1)} &= \mathbf{F}_2\mathbf{F}_1(\mathbf{z}_{0v} + \mathbf{z}_{1v}), \end{aligned} \tag{4.6}$$

Where $v = 0, 1$, and $\mathbf{F}_1$ and $\mathbf{F}_2$ are transition matrices for the first and second meiotic divisions implemented as mitochondrial sampling without replacement. Their corresponding elements are

$$F_{(1)i,j} = \binom{2j}{i} \binom{4M-2j}{2M-i} \binom{4M}{2M}^{-1}, \tag{4.7}$$

and

$$F_{(2)i,j} = \binom{j}{i} \binom{2M-j}{M-i} \binom{2M}{M}^{-1}. \tag{4.8}$$

To study the dynamics of the system, I consider the invasion of a mutant allele with a value of paternal leakage π different from the resident population. The new allele is inserted into the population at a low frequency $\varphi = 0.005$ at $t = 500$, so that $\mathbf{P}_{\bullet,1}^{l(t)} = \varphi\mathbf{P}_{\bullet,0}^{l(t-1)}$ and $\mathbf{P}_{\bullet,0}^{l(t)} = (1-\varphi)\mathbf{P}_{\bullet,0}^{l(t-1)}$, and its evolution is tracked until an equilibrium is reached. I consider all possible values of the trait π and build the pairwise invasibility plots depicting the sign of the invasion fitness, i.e. the growth rate of the invader subpopulation when rare (Geritz et al., 1998). These plots determine the expected evolutionary outcomes and stable strategies when the trait value changes in small

discrete steps, but also give insight into the dynamics of the system with large effect mutations (Dieckmann, 1997; Geritz et al., 1998).

4.4 With maternal control, strict UPI is the only asymmetric ESS

In the conventional case, the maternal mating type *MTmat* regulates the contribution of paternal mitochondria to the zygote. Under maternal control, the invasion analysis recovers two boundary ESS corresponding to complete uniparental ($\pi = 0$) and biparental ($\pi = 1$) inheritance, and an unstable singular point π^* ($0 < \pi^* < 1$) in between (an evolutionary repeller, Fig. 4.2a). Fully uniparental inheritance is the only evolutionarily stable state with asymmetric inheritance, while intermediate values of paternal leakage cannot be maintained.

If we assume that biparental inheritance of mitochondria is the ancestral state ($\pi = 1$), then the transition to maternally controlled exclusion of paternal mitochondria is not favoured with small effect mutations in π and low mitochondrial mutation rates (Fig. 4.2). With high enough mutational rates and weak epistasis, biparental inheritance loses its local stability and complete uniparental inheritance ($\pi = 0$) becomes the sole evolutionary attractor (Fig. 4.2b). Under these conditions, small changes in π will eventually lead to the evolution of purely uniparental inheritance of mitochondria imposed by the maternal gamete. The value of paternal leakage corresponding to the unstable singular point becomes higher ($\pi^* \to 1$) with increasing mutation rates (μ) and weaker epistatic interactions ($\xi \to 1$) (Fig. 4.2a-d).

The local attraction towards either BPI or UPI under maternal control can be explained in terms of costs and benefits of asymmetric mitochondrial transmission under negative epistasis. If the invader has a lower value of paternal leakage π than the resident, there is higher asymmetry of cytoplasmic inheritance and increased mitochondrial variance among the invaders. This boosts the efficacy of purifying

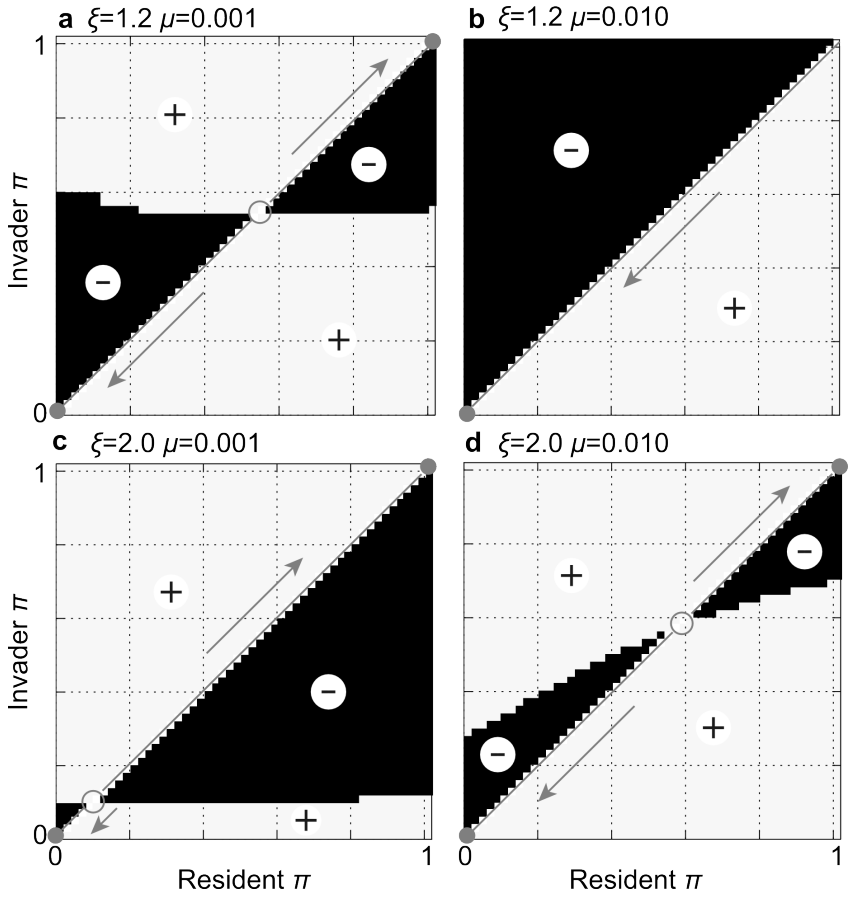

Figure 4.2. Pairwise invasibility plots for the maternally controlled asymmetric inheritance of mitochondria. Values of paternal leakage π for which the invasion is successful are within the regions marked with "+". Arrows show the direction of trait evolution assuming small mutational changes in π. (**a**) For weak epistatic interactions (lower ξ) and low mutation rates, there are two evolutionarily stable states (filled circles), one at which mitochondria are symmetrically inherited from both gametes ($\pi = 1$), and the second at which there is full uniparental inheritance of mitochondria ($\pi = 0$). These are separated by a singular point between $\pi = 0$ and $\pi = 1$ which is an evolutionary repeller (open circle). (**b**) With higher mutation rates the zone of attraction to the asymmetric equilibrium ($\pi = 0$) increases, until the symmetric equilibrium is eliminated. (**c**-**d**) Increasing the degree of epistasis increases the short-term benefit of mixing mitochondria and weakens the attraction of the asymmetric equilibrium. The number of mitochondria was set to $M = 50$.

selection and improves population fitness over several generations, giving the mutant a long-term advantage (Bergstrom and Pritchard, 1998; Hadjivasiliou et al., 2013). More symmetric inheritance of mitochondria with higher values of π in the invader, on the other hand, reduces the efficacy of selection, but increases the frequency of intermediate cytotypes, which under negative epistasis have higher fitness than the

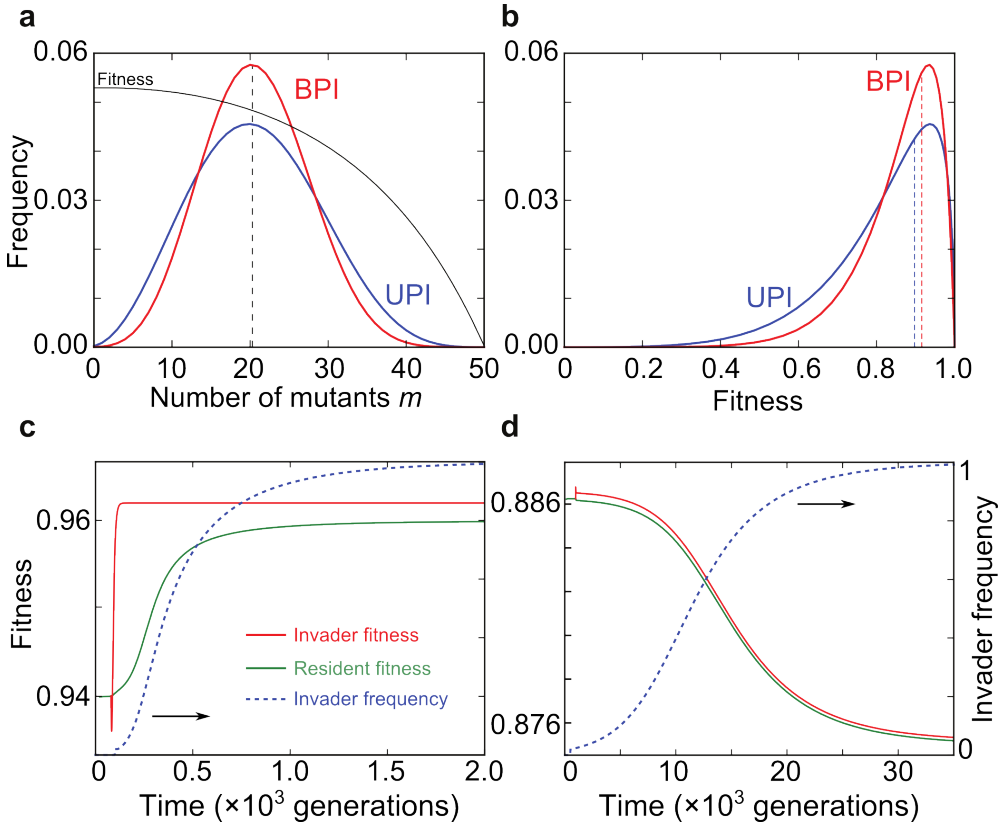

Figure 4.3. Selective effects of more or less paternal leakage shown for the extreme cases of biparental (BPI) and uniparental inheritance (UPI) of mitochondria, $\pi = 1$ and $\pi = 0$ respectively. (**a**) BPI (red) reduces mitochondrial mutation variance among offspring by increasing the frequency of intermediate cytotypes compared to UPI (blue). Note that the mode of inheritance does not alter the mean, only the variance in mutation frequency. The fitness function (grey curve) is concave, and assumes that a large number of mutants must accumulate before cell function is significantly undermined (epistasis $\xi = 3$). (**b**) This gives BPI a short-term mean fitness advantage (dotted lines) because intermediate cytoplasmic states have higher fitness than the mean of the extreme states. Mutation rate $\mu = 0.04$. (**c**) A uniparental mutant $\pi = 0$ invades a population with paternal leakage of $\pi = 0.2$ under maternal control due to the long-term benefit of asymmetric inheritance. (**d**) Biparental invader ($\pi = 1$) fixes within the resident population of $\pi = 0.8$ due to the short-term fitness effects of mitochondrial mixing, even though this reduces the population fitness in the long term. Mutation rate is $\mu = 0.01$, epistasis $\xi = 2$ as in Fig. 4.2d.

linear combination of extreme phenotypes (the fitness of the mix is higher than the mean of the two initial fitness values, Fig. 4.3a). Symmetric inheritance of mitochondria can therefore give the invading allele a short-term fitness advantage (Fig. 4.3b).

The evolutionary success of an invader is determined by the complex interplay between the long- and short-term effects of asymmetric mitochondrial inheritance, and the degree to which the nuclear allele that controls π is associated with the resulting mitochondrial population. Long-term fitness effects are more relevant with strong mito-

nuclear linkage, while short-term effects dominate under weak statistical associations between the two genomes. Since (by definition) the mother passes on more mitochondria, the maternal nuclear alleles are always strongly linked to the mitochondrial population, and this association strengthens as the paternal contribution of mitochondria falls (i.e. with lower values of π). The reverse is true for the paternal nuclear alleles, which leads to radical differences in the selection of maternal and paternal control (see below).

With maternal control, the singular point $0 < \pi^* < 1$ is located where the long- and short-term fitness effects are matched and the fitness landscape as experienced by a nearby mutation is virtually flat. An invader with a lower level of paternal leakage than the resident to the left of the singular point (i.e. $\pi < \pi^*$) increases both the mitochondrial variance and the strength of mito-nuclear linkage, and so invades due to the long-term fitness advantage and more efficient elimination of mitochondrial mutations (Fig. 4.3c). Conversely, to the right of the singular point ($\pi > \pi^*$), an invader allowing more paternal leakage reduces both the variance and strength of the mito-nuclear associations, and thus benefits mostly from short-term effects (Fig. 4.3d). Successful invasion on either side of the singular point therefore makes it an evolutionary repeller under maternal control of cytoplasmic inheritance (Fig. 4.2).

4.5 Paternal leakage evolutionarily stable with paternal control of mitochondrial inheritance

In the reverse case of paternal control, mating type *MTpat* determines what fraction of its own mitochondria is discarded, either before or after gamete fusion. Although the effect on the mitochondrial population of a zygote is the same as under maternal control, the evolutionary dynamics are significantly different. The nuclear gene

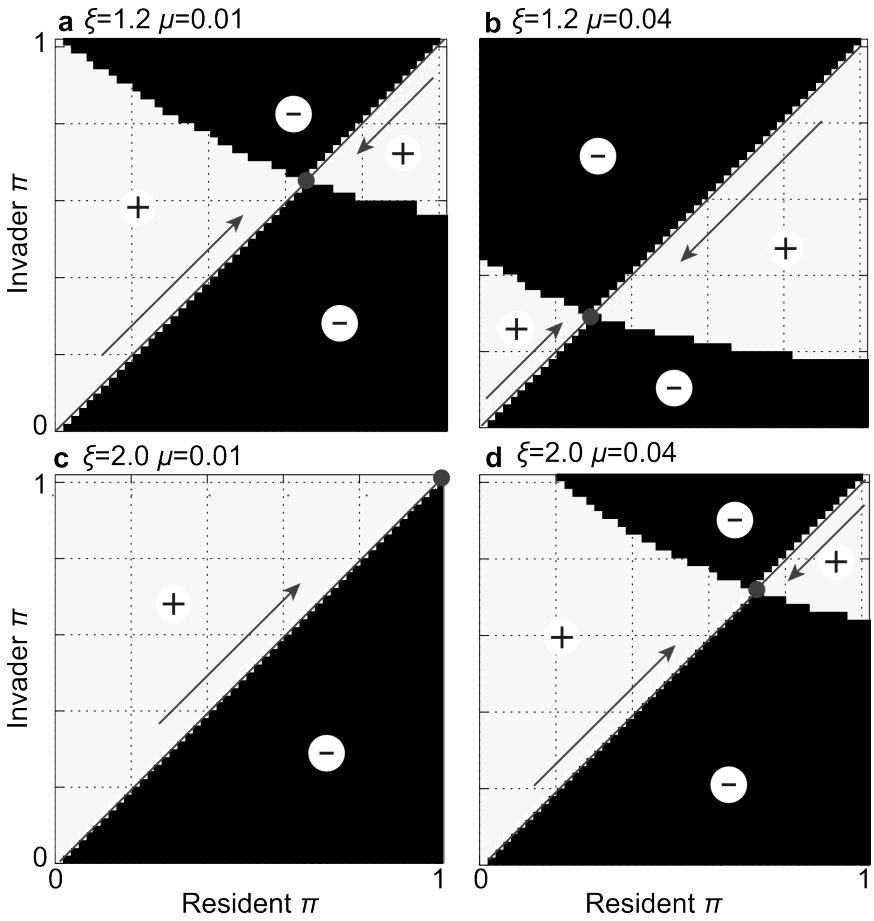

Figure 4.4. Pairwise invasibility plots for the paternally-controlled mitochondrial inheritance. Values of paternal leakage π for which the invasion is successful are within the regions marked with "+". Arrows show the direction of trait evolution, that is, the expected changes in the value of π via recurrent invasion of mutants. (**a**) There is only one evolutionarily stable attractor at a singular point between $\pi = 0$ and $\pi = 1$. (**b**) Increasing mutation rates favour an ESS with less paternal leakage, consistent with the role of uniparental inheritance purging deleterious mitochondrial mutations. (**c**-**d**) Increasing the degree of negative epistasis (higher ξ) increases the short-term benefit of mitochondrial mixing, and favours higher levels of paternal leakage. The number of mitochondria was set to $M = 50$.

restricting paternal cytoplasmic inheritance is now patrilineally inherited and therefore is more weakly associated with its own mitochondrial population than the equivalent matrilineally inherited control allele. And the association weakens as the strength of restriction on paternal inheritance grows (i.e. lower values of π). At the limit of $\pi = 0$, all mitochondria are inherited from the maternal gamete, and therefore there is no association with the paternal nuclear allele.

With paternal control, the analysis again recovers a singular point corresponding to an intermediate level of paternal leakage ($0 < \pi^* < 1$). However, this

point is now an evolutionarily stable attractor (Fig. 4.4), contrary to the case of maternal regulation. This point is a continuously stable equilibrium, which is attractive across the full range of values of π. So an ancestral population with complete biparental inheritance of mitochondria ($\pi = 1$) or one with uniparental inheritance exclusively from the maternal gamete ($\pi = 0$) will evolve intermediate levels of paternal leakage. An ESS with more asymmetric inheritance (lower π) evolves with higher mutation rates (μ) and weaker epistatic interactions ($\xi \to 1$) (Fig. 4.4). Only when the mitochondrial mutation rate is sufficiently low and epistasis sufficiently high do the maternal and paternal control equilibria coincide, with biparental inheritance being the evolutionary outcome (Fig. 4.4c).

These results are explained once again by the balance between long- and short-term benefits of asymmetric inheritance, but this time with paternal control there is an opposite effect on the strength of mito-nuclear linkage. At the singular point π^*, the long-term and short-term effects are in balance. An invader with $\pi < \pi^*$ increases the asymmetry in mitochondrial transmission and the long-term effect of variance, but weakens the genetic associations between the paternal nuclear allele and paternal mitochondria, to the detriment of the invader. Conversely, higher paternal leakage ($\pi > \pi^*$) increases mitochondrial mixing, improving short-term fitness but undermining longer term outcomes. In this case, greater mixing strengthens genetic linkage, which is again harmful to the invader. This means that a singular point must necessarily be stable, since any deviation to either side is deleterious.

4.6 Large mutational effects

The analysis above considers small mutational steps. Previously published analyses considered large-effect mutations, but limited to two states, i.e. π = 0 or 1 (Hastings, 1992; Hadjivasiliou et al., 2012; 2013; Christie et al., 2015). Allowing arbitrary large-

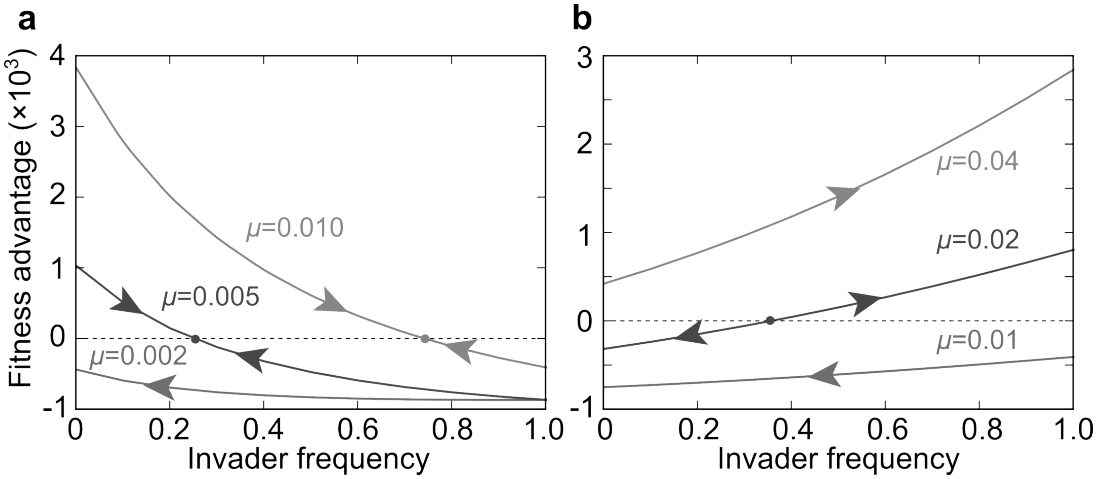

Figure 4.5. The fitness advantage of the invader is frequency-dependent. (**a**) With maternal control, the fitness advantage of the invader declines with its frequency, leading to a stable equilibrium between 0 and 1. This leads to protected dimorphic states where the mutant invades but does not completely replace the resident, but these dimorphisms are not necessarily evolutionarily stable (Fig. 4.6). (**b**) Under paternal control of cytoplasmic transmission, the fitness advantage increases with allele frequency. The outcome of the invasion of paternal alleles is therefore always either fixation or extinction, i.e. dimorphic states cannot be established. In both cases the long-term advantage of asymmetric inheritance increases with the mitochondrial mutation rate μ, since asymmetric inheritance increases the efficacy of purifying selection against deleterious mitochondrial mutations. $M = 50$, $\xi = 1.5$, mutant with $\pi = 0.5$ invading biparental population with $\pi = 1$.

effect mutations, different dynamics with dimorphic states are uncovered. Under maternal control these arise because the fitness advantage of an invading allele is subject to negative frequency-dependence (Fig. 4.3c, d and Fig. 4.5a). An invader with a lower level of paternal leakage than the resident benefits from long-term fitness effects, but due to random mating these fitness benefits spread into the resident population. As the frequency of the invader rises, so does the fitness of the resident, reducing the invader's advantage (Fig. 4.3c). Likewise, the fitness advantage of a mutant with higher π than the resident is greatest at low frequencies (Fig. 4.3d).

With small mutational changes in π, the negative frequency-dependence is sufficiently weak that the invader always goes to fixation and leads to unitary stable states at π = 0 or 1 (Fig. 4.2). But with large effect mutations, a successful invader does not always replace the resident allele, spreading to an intermediate frequency instead (Fig. 4.6). Dimorphic states always lie away from the main diagonal of the pairwise-invasibility plots, and therefore cannot be reached via small mutational

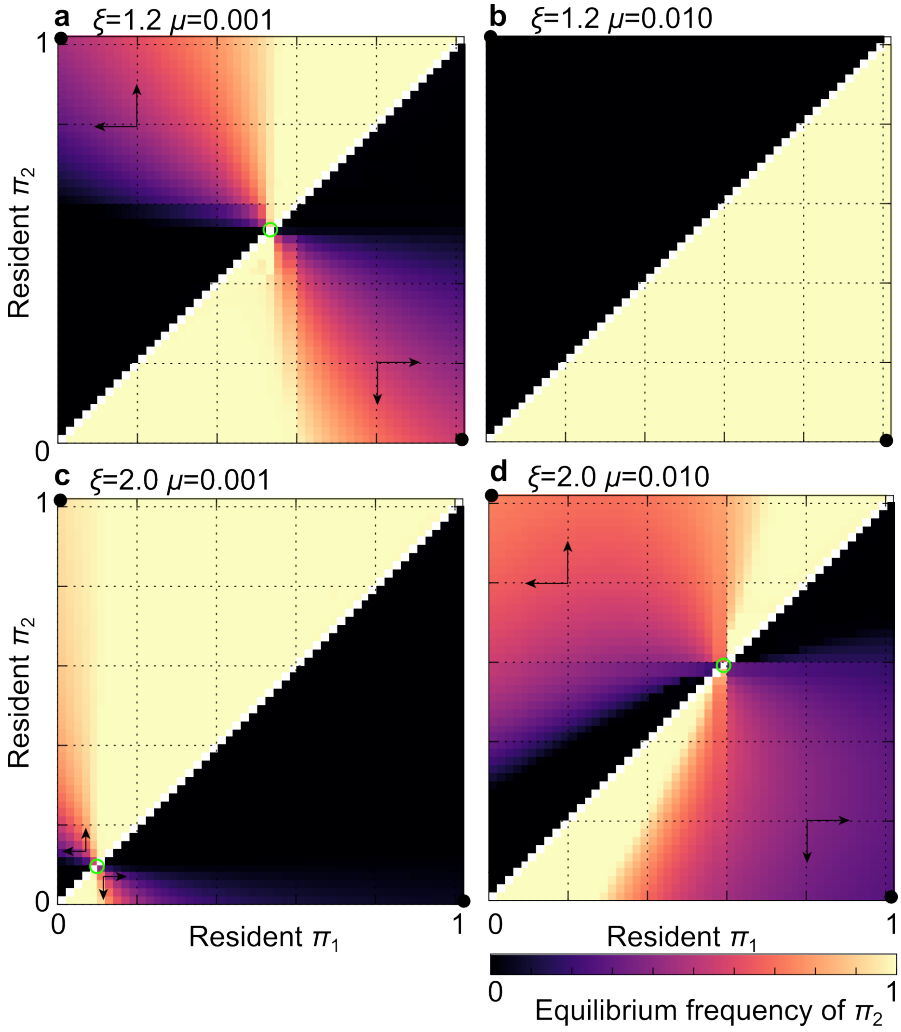

Figure 4.6. Dimorphic states with maternal control of mitochondrial inheritance, where the alleles for two distinct values of paternal leakage π_1 and π_2 coexist. Color represents the equilibrium frequency of the maternal allele for paternal leakage π_2 in a protected dimorphism with π_1. Arrows indicate the direction of evolutionary change in resident trait values, via recurrent invasions of a third mutant π_3. Given that $\pi_1 < \pi_2$, invader π_3 replaces the resident π_1 if $\pi_3 < \pi_1$; the invader replaces π_2 if $\pi_3 > \pi_2$, until the population reaches the evolutionarily stable coalition of coexisting $\pi_1 = 0$ and $\pi_2 = 1$ (filled circles). (**a**) The dimorphic states lie away from the main diagonal (except for the close vicinity of the repulsive singular point). (**b**) Increasing mutation rates favour an evolutionarily stable coalition with higher frequencies of $\pi_1 = 0$, which under weak epistasis can completely displace the biparental allele. (**c-d**) Increasing the degree of negative epistasis (higher ξ) increases the short-term benefit of mixing mitochondria and reduces the frequency of the strictly uniparental $\pi_1 = 0$ at the dimorphic ESS.

changes in π (Fig. 4.6). When a further invader is introduced, if it spreads, it drives out one of the existing dimorphic alleles. The final stable endpoint is a dimorphic population consisting of individuals with complete uniparental ($\pi = 0$) and biparental ($\pi = 1$) inheritance, with the frequency of uniparental-inheritance alleles increasing with higher mutation rates and weaker epistasis (Fig. 4.6). In contrast, under paternal control the

fitness advantage of a mutant increases with its frequency (Fig. 4.5b). This means that all successful invaders eventually reach fixation, and stable dimorphic states do not exist. This is true for both small and large mutational change.

Of greater general interest, large mutational effects were assumed in previous studies which concluded that paternal control of mitochondrial inheritance cannot evolve due to the lack of association between the nuclear modifier and the inherited mitochondria (Hastings, 1992; Randerson and Hurst, 1999; Hadjivasiliou et al., 2013). These studies considered whether a discrete mutational state with complete restriction of paternal inheritance (π = 0) invaded the ancestral condition of biparental inheritance (π = 1). As there is no association of the paternal control nuclear modifier with its mitochondria under strict uniparental inheritance, it cannot be favoured by selection. As I have shown here, with a continuous distribution of mutational states, paternal control can be associated with the inherited mitochondria over many generations, and this can favour asymmetric transmission with paternal leakage and persistent heteroplasmy, depending on the mutation rate and epistasis.

4.7 Discussion

Current theoretical views do not account for the active role of males in destroying their own organelles and cannot explain paternal leakage as anything more than a sporadic deviation from the rule of strict uniparental inheritance (UPI), in spite of its common occurrence. In this chapter I analysed the evolutionary stability of asymmetric inheritance of mitochondria, controlled by either of the two mating types, assuming small mutational steps and negative epistatic interactions between mitochondrial mutations. The analysis shows not only that paternally-regulated asymmetric transmission of mitochondria can evolve, but also that it is inherently associated with paternal leakage and heteroplasmy.

Consider that both maternal and paternal control of mitochondrial inheritance is possible. With low mutation rates, fully symmetrical biparental transmission is evolutionarily stable for both modes of inheritance regulation (Fig. 4.7). At higher mutation rates or with weak epistatic interactions, however, the two sex-specific ESS's diverge (Fig. 4.7): maternal control favours strict uniparental inheritance ($\pi_{\mathrm{mat}} = 0$), while paternal regulation favors an equilibrium with paternal leakage $0 < \pi_{\mathrm{pat}} < 1$. As the mutation rate increases, these different equilibria become close but do not necessarily coincide, especially with stronger epistasis (Fig. 4.7a vs. 4.7b). This divergence implies competition over the control of cytoplasmic inheritance and cycles of repeatedly evolving maternal and paternal control mechanisms.

With sufficiently high mutation rates, maternal control favours no transmission of paternal mitochondria (i.e. $\pi_{\mathrm{mat}} = 0$) and establishment of strict UPI ($\mu > 0.005$ in Fig. 4.7a or $\mu > 0.02$ in Fig. 4.7b). This could be achieved by the selective destruction of paternal mitochondria in the zygote, while protecting the maternal organelles as is known to occur in *Ustilago maydis* (Fedler et al., 2009). Now consider a newly evolving mechanism of paternal control that affects the number of sperm mitochondria surviving within the zygote and subverting part of the maternal organelle-destruction machinery. This could be achieved, for instance, by placing a ubiquitin tag on only a subpopulation of mitochondria during mammalian spermatogenesis. According to our results, the paternal mutation protecting some of its mitochondria would again spread to fixation, leading to an evolutionarily stable state with paternal leakage and persistent heteroplasmy. The short-term fitness advantage of mitochondrial mixing allows paternal-control mutations to fix, even though paternal leakage reduces the variance in the mutation load, hinders purifying selection against defective mitochondria and poses a long-term fitness cost to both mating types. The paternally regulated state could persist, or the cycle could start again with a new maternal mutation recognizing *MTpat*'s mitochondria and restoring strict maternal UPI. The competition over the

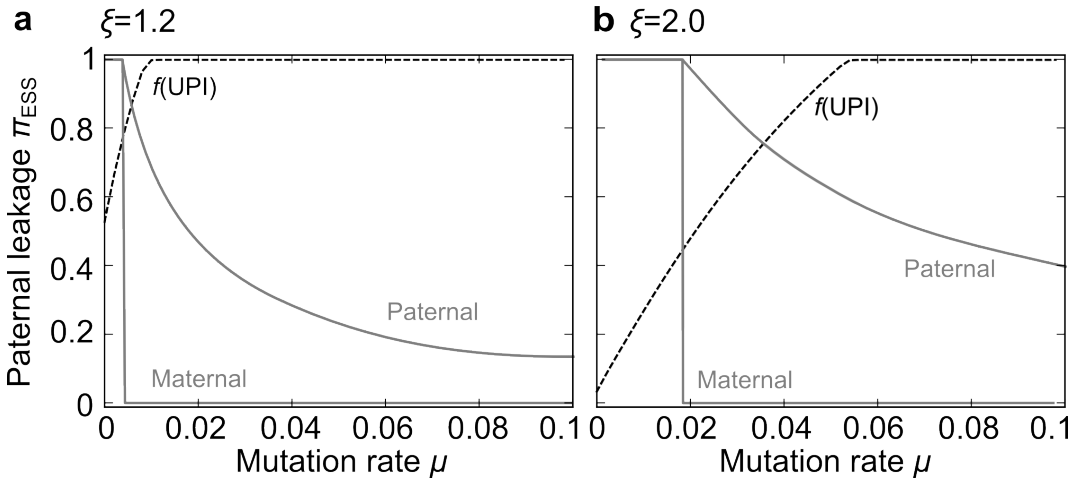

Figure 4.7. Evolutionarily stable value of paternal leakage π_{ESS} depends on which mating type controls the destruction of paternal organelles. (**a**) The two evolutionarily stable values $\pi_{\mathrm{ESS}}(\mathrm{materal})$ and $\pi_{\mathrm{ESS}}(\mathrm{paternal})$ coincide at the fully symmetric state of biparental inheritance with low mutation rate μ, but diverge as asymmetric inheritance becomes favoured with increasing mutation rates. (**b**) Strong epistasis ($\xi = 2.0$) favours more frequent mitochondrial mixing and higher levels of paternal leakage. Dotted lines indicate the frequency of the UPI allele $f(\mathrm{UPI})$ in a stable dimorphic state if large changes in π are allowed. The number of mitochondria per cell was set to $M = 50$.

control of cytoplasmic inheritance with males benefiting from mitochondrial leakage and females favouring strict UPI, would therefore lead to the repeated evolution of molecular mechanisms recognizing or protecting paternal organelles, accounting for the high diversity of maternal and paternal UPI mechanisms observed in nature.

These results have implications for the evolution of oogamy as a means to control the extent of cytoplasmic mixing. Among its various benefits (Parker et al., 1972; Radzvilavicius et al., 2015), the development of a large female gamete packed with mitochondria can be viewed as a way to enforce the highly asymmetric transmission of mtDNA. Assuming that the size of the sperm mitochondrial population cannot readily be increased, oogamy provides a reliable mechanism of UPI, independent of the males' ability to resist the active removal of their organelles after the gamete union, and limiting the extent of possible heteroplasmy.

A central conclusion of this chapter is that paternal leakage can evolve as an indirect consequence of purifying selection against deleterious mitochondrial mutations, and the intrinsic differences in statistical associations between mitochondria

and the nuclei of the two mating types. We should not exclude the possibility, however, that paternal leakage could provide a more direct sex-specific benefit (Wade and McCauley, 2005; Wade and Brandvain, 2009; Kuijper et al., 2015). These sex-specific effects could be utilized through the preferential segregation of mitochondrial haplotypes into distinct somatic tissues, cases of which are known (Magnacca and Brown, 2010; Burgstaller et al., 2014). The most striking example of such segregation occurs in bivalve molluscs with doubly-uniparental inheritance (DUI), where M-type mitochondria are transmitted exclusively through males' germline (Passamonti and Ghiselli 2009). Remarkably, male (but not female) somatic tissues remain heteroplasmic in bivalves with DUI, which according to our model should indeed correspond to the ESS under paternal control of mitochondrial transmission.

In contrast with previous analyses (Hadjivasiliou et al., 2012, 2013; Kuijper et al., 2015), paternal leakage in the present analysis is treated as an evolvable trait. It is subject to indirect selection that acts against deleterious mutations in the mitochondrial population. We can conclude that the evolutionarily optimal pattern of mitochondrial transmission critically depends on which mating type (or sex) controls the number of paternal mitochondria transmitted to the zygote. Males benefit from paternal leakage, since mitochondrial mixing increases the mean fitness of their progeny. Strict UPI is favoured by the maternal mating type, due to its long-term effect of increasing the efficacy of purifying selection. Tension between selection on males and females explains the seemingly unstable evolutionary pattern of UPI with multiple origins and reversals, the numerous mechanisms involved in the asymmetric transmission of mitochondrial genes, persistent heteroplasmy and paternal leakage. Our analysis therefore offers a simple way of understanding the extraordinary variation in the patterns of mitochondrial transmission around the central tendency towards uniparental inheritance.

CHAPTER 5. SELECTION AGAINST MITOCHONDRIAL MUTATIONS SHAPED THE EVOLUTION OF METAZOAN GERMLINE ARCHITECTURE

5.1 Summary

Basal metazoans such as corals and sponges generate gametes from pluripotent stem cell populations that also produce somatic cells, whereas most bilaterians sequester a unipotent germ cell lineage early in development. In spite of several long-standing explanations, the selective forces responsible for the evolution of early predetermined germline with a limited number of mitotic divisions remain unclear. In this chapter I propose that the mode of germline development in metazoans is determined by selection against mitochondrial mutations. A simple mathematical model supports the hypothesis, showing that the evolutionarily stable number of germline cell divisions depends on mitochondrial mutation rates. In organisms with low mitochondrial DNA copying-error rates, segregation of mutations over multiple cell divisions generates variation and allows selection to optimize gamete quality through high numbers of germline cell divisions. The new hypothesis successfully explains the absence of germline sequestration in basal metazoans and plants and the germline structure differences between males and females.

5.2 Introduction

Higher animals generate reproductive cells in dedicated germ-cell lineages, separating primordial germ cells from the rest of the stem cell populations early in embryogenesis (Extavour, 2007). Basal metazoans, such as sponges and corals, on the other hand, generate gametes from the long-lived pluripotent stem cell lineages that also give rise to terminally differentiated somatic cells (Blackstone and Jasker, 2003; Extavour and Akam, 2003; Extavour, 2007). Based on the phylogenetic distribution of character states, it has been proposed that the last metazoan common ancestor produced gametes in a similar fashion, and did not segregate a dedicated germline from somatic stem cells (Blackstone and Jasker, 2003). While several hypotheses exist to explain the evolutionary advantages of reproductive division of labour into an immortal germline and “disposable” soma (Weismann, 1890; Kufopanou, 1994; Queller, 2000; Bendich, 2010; Simpson, 2012; Goldsby et al., 2014), very little is known about the evolutionary forces responsible for the origin of bilaterian-like germline architecture. Particularly intriguing are the sex-specific features of the germline, where a limited number of oocytes is produced in a relatively low number of germline cell divisions, while sperm are produced continuously through the male’s adulthood.

In his seminal work on biological individuality as a derived animal trait, Buss (1983, 1987) proposed that germline sequestration in the early stages of embryonic development protects the multicellular organism from proliferation of non-cooperative defector cell lineages and their transmission across generations. Stable multicellularity requires cooperation between somatic and germ cell lineages, but selection on the lower level could favour cells that evolve selfish traits and increase their own replication rate to the detriment of the multicellular organism. Once a single lineage is set aside as a progenitor of all reproductive cells, selfish mutations that arise in a much larger somatic cell population are not included in the gametes and do not survive past a single generation (Michod and Roze, 2001). The only way for somatic cells to increase the

reproductive fitness of their kin is therefore by increasing the fitness of the multicellular organism they belong to through cooperation. In Buss' view, germline sequestration was one of the key innovations in multicellular development, allowing the evolution of complex multicellular organisms with highly specialized somatic tissues. Frank (1996) argued that selection for host control over endosymbiont reproduction would similarly favour early germline sequestration, enforcing cooperation between symbionts in somatic cells. While germline segregation reduces the chances for a selfish mutation to access the germline or found a new organism clonally, selfish conflict does not explain why early branching metazoans with no pre-determined germ-cell lineage manage to maintain relatively complex multicellularity, even though opportunities for selfish mutation to arise in multiple stem cell divisions abound. This implies that evolutionary conflict might not be the key force driving the evolution of early germline sequestration.

Multipotent stem cell lineages giving rise to both reproductive and somatic tissues must necessarily exist in development of any multicellular organism, but how long multipotency is maintained differs significantly across metazoan groups (Juliano et al., 2010). Basal metazoans maintain these lineages throughout their lengthy lifetimes (Seipel et al., 2003; Funayama, 2010; Muller et al., 2004; Juliano et al., 2010), while bilaterians often lose pluripotency early in development (Extavour 2007), implying, by definition, an early origin of a unipotent germ cell lineage. Arguably, the number of germline cell divisions can be effectively regulated only in the latter case, and the real evolutionary advantage of the strict germline-soma distinction might in fact lie in the reduced number of mitotic germline stem cell divisions (Michod and Roze, 2001). Particularly illuminating in this regard are the sex-specific features of the germline architecture, where the number of mitotic germline divisions can differ significantly between males and females (Drost and Lee, 1995; Crow, 2000).

Focusing on the evolution of germline by reducing the number of cell replication cycles shifts the problem specifically to bilaterian females, that, in contrast to male

spermatogenesis, very rarely maintain germline stem cells throughout adulthood (Spradling et al., 2011). Female gametes of higher animals are sequestered early in embryogenesis in a transcriptionally repressed state, with meiosis arrested in prophase I and mitochondria in a state of functional quiescence (dePaula et al., 2013). Strikingly, female gametes are usually large cells packed with mitochondria, with as many as 10^6 copies of mitochondrial DNA in mammalian oocytes (Shoubridge and Wai, 2007). Mitochondrial DNA is usually inherited uniparentally: male mitochondria are either excluded from the zygote or destroyed on entry to the oocyte (Sato and Sato, 2013; Chapter 4). Hence, from a mitochondrial point of view, male gametes are relieved of the constraints that operate on female germ cells, and do not need to function as a sequestered and protected environment. These traits point to mitochondrial function as being central to germline development and evolution. The possibility that selection for mitochondrial quality could have contributed to the evolution of the female germline has been raised before (Allen, 1996; Bendich, 2010; dePaula et al., 2013), but never formally addressed.

5.3 Hypothesis: Purifying selection against mitochondrial mutations drives evolution of metazoan germline architecture

Cell division in multicellular organisms is preceded by replication of both nuclear and mitochondrial genomes. DNA replication is a major source of deleterious mutations through copying errors and therefore repeated cell division inevitably leads to mutation accumulation. Within a single generation and in the absence of recombination, the number of mutations in the nuclear genome can only increase with time, due to both copying errors and background damage. While the mean number of mutations in mitochondrial genome also increases with time, random segregation of mitochondrial

mutations at every cell division generates variance in the number of mutations per cell, with some daughter cells containing fewer mutations than the parent cell.

I propose that the mode of germline development in metazoans is determined by the competing effects of (1) the increasing *mean* number of mitochondrial mutations and (2) increasing *variance* in the number of independently segregating mtDNA mutations between gametes. Under low mtDNA copying error rates, mitochondrial variance generated via repeated stem cell division together with selection at the level of the organism is predicted to favour gamete production from the long-lived pluripotent stem cell lineages, characteristic of plants and basal metazoans. Increasing mtDNA mutation rates would then select for reduced numbers of germline cell divisions, leading to the derived state of early germline sequestration. As mitochondria in metazoans are inherited mostly maternally, these constraints do not apply to the male germline. While a similar argument can be made for mutation accumulation in the nuclear genome, the large numbers of germline stem cell replication cycles in males and basal metazoans suggests that nuclear mutations do not play a primary role in the evolution of germline development.

5.4 Model of the germline evolution in multicellular organisms

Consider an infinite population of multicellular organisms, each starting its development from a single cell containing *M* independently segregating mitochondria and undergoing N symmetric cell divisions (Fig 5.1). Similar to the models developed in previous chapters, the population state can be represented by two $(M+1)\times 2$ matrices $\mathbf{P}^l$, $l = 0, 1$ denoting the mating type. The matrix element $P^l_{m,j}$ denotes the

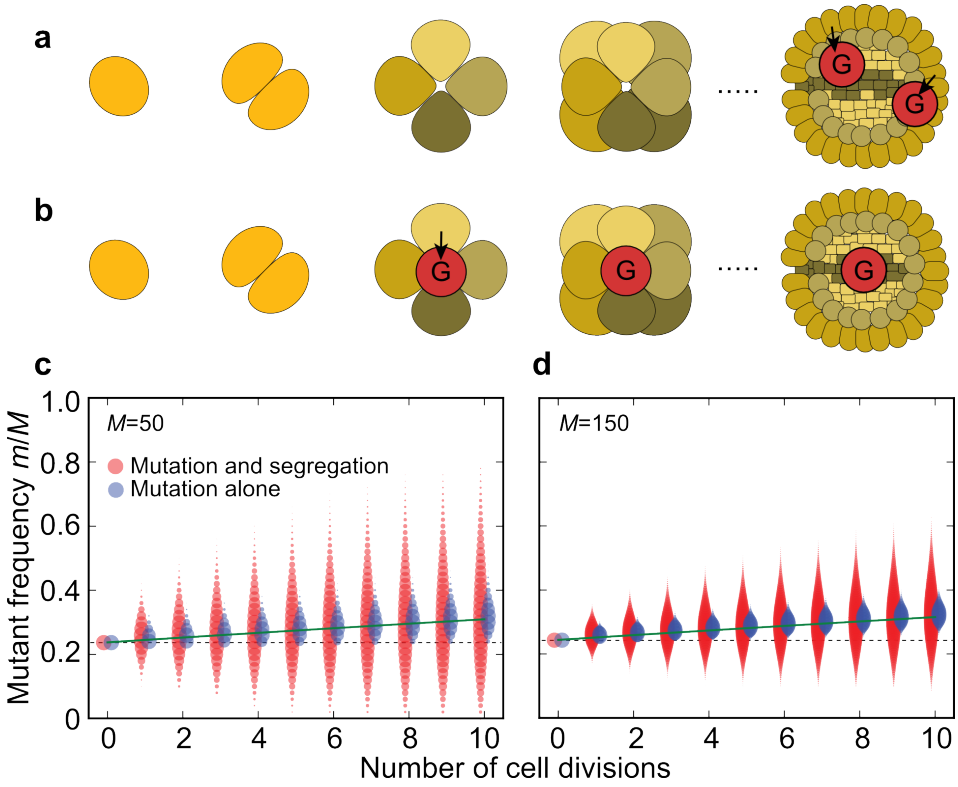

Figure 5.1. Modelling mitochondrial mutation dynamics in development of a multicellular organism. The number of cell divisions needed to produce gametes G is an evolvable trait. Large G indicates gamete production from the multipotent stem cell population as in basal metazoans (**a**), while low values of G are characteristic of bilaterians that set aside their germ cell populations early in embryonic development (**b**). The mean number of mitochondrial mutations increases with every cell division due to copying errors, but segregational drift also increases variance in the mutation frequency (**c**). Increasing number of segregating units (M) reduces the variance, but not the mean (**d**). Copying error rate is set to $\mu = 0.01$ while the initial mutation frequency is set to 0.24.

frequency of zygotes in a nuclear state $j = 0, 1$ (wild-type or mutant) and with m mutant mitochondria at the start of the generation.

Mitochondrial replication introduces new mitochondrial mutations due to mtDNA copying errors at a rate μ per cell division, a process represented by the transition matrix $\mathbf{U}$ with elements

$$U_{i,j} = \binom{M-j}{i-j} \mu^{i-j} (1-\mu)^{M-i}. \quad (5.1)$$

I also consider 'background' mitochondrial mutations due to oxidative or UV damage that arise even in the absence of replication, represented by the transition matrix $\mathbf{V}$ of the same binomial form as $\mathbf{U}$, but with a different mutation rate β (per generation).

Every cell within an organism grows and splits into two, segregating its mitochondria randomly into the two daughter cells. Random drift due to the partitioning of mitochondria via sampling without replacement can be represented by the stochastic matrix $\mathbf{S}$, so that

$$S_{i,j} = \binom{2j}{i}\binom{2M-2j}{M-i}\binom{2M}{M}^{-1}. \quad (5.2)$$

Within a single generation, the probability distribution for the number of mutant mitochondria between the daughter cells after n divsions can be expressed as $\mathbf{f}_m(n) = (\mathbf{SU})^n\mathbf{e}^{(m)}$, where $\mathbf{e}^{(m)}$ is the initial population state for a cell containing m mutant mitochondria, that is $\mathrm{e}_i^{(m)} = \delta(i,m)$ (Kronecker delta). As the number of cell divisions increases, both the mean frequency of mutations within the cell and its variance increases (Fig. 5.1c), so that some cells contain fewer mutants than the zygote. Note that increasing mitochondrial population size within the cell M dampens segregational drift and reduces variance in the mutant frequency (Fig. 5.1d).

The fitness of an adult individual after N cell divisions is calculated as a mean value of its constituent cell fitness values, each of which is itself a function of the frequency of mutant mitochondria within the cell. For the sake of simplicity, I assume that mutations due to background damage accumulate mostly after the embryonic development is complete, which is true if the lifespan of the individual significantly exceeds the duration of embryogenesis. Starting with a zygote containing m mutant mitochondria, adult fitness after N cell divisions can be approximated as

$$\omega(N,m) = \mathbf{w}^{\mathrm{T}}\mathbf{V}(\mathbf{SU})^N\mathbf{e}^{(m)}. \quad (5.3)$$

The row vector $\mathbf{w}^{\mathrm{T}}$ represents the cellular fitness function and has elements $w_m = 1 - s(m/M)^{\xi}$, where s is selection strength and ξ is the strength of epistatic interactions between deleterious mutations. As in preceding chapters, I will only consider $\xi \geq 1$.

Selection changes genotype frequencies according to adult fitness, which is here taken as the mean fitness of all somatic cells. The updated population state after selection is

$$\mathbf{P}^{l'} = \frac{\{\mathbf{I}[\mathbf{w}^{\mathrm{T}}\mathbf{V}(\mathbf{SU})^{N}]^{\mathrm{T}}\}\mathbf{P}^{l}}{\mathbf{w}^{\mathrm{T}}\mathbf{V}(\mathbf{SU})^{N}\mathbf{P}^{l}\overline{\mathbf{u}}_{2}} = \widehat{\Omega}[\mathbf{P}^{l}] \tag{5.4}$$

Here $\overline{\mathbf{u}}_2$ is a column vector of ones, so that $\mathbf{P}^{l}\overline{\mathbf{u}}_2$ gives gamete frequencies independent of their nuclear alleles and the denominator is then simply the average fitness of a mating type l, the part of the nominator within curly brackets is the diagonal square matrix of adult fitness values, and $\widehat{\Omega}$ is selection operator. Gamete frequencies after selection are then

$$\begin{aligned} \mathbf{\Gamma}^{l}_{\bullet,0} &= \mathbf{V}(\mathbf{SU})^{G}\widehat{\Omega}[\mathbf{P}^{l}]_{\bullet,0}, \\ \mathbf{\Gamma}^{l}_{\bullet,1} &= \mathbf{V}(\mathbf{SU})^{G'}\widehat{\Omega}[\mathbf{P}^{l}]_{\bullet,1.} \end{aligned} \tag{5.5}$$

Here G is the number of germline cell divisions in wild-type organisms and G' is the number of germline cell divisions in the mutant invader.

Haploid gametes of the opposite mating types fuse at random. I only consider the case of uniparental inheritance of mitochondria from mating type 0. The zygote frequencies therefore can be expressed as

$$\mathbf{z}_{qr} = \mathbf{R}\mathbf{\Gamma}^{0}_{\bullet,q}\left(\overline{\mathbf{u}}^{\mathrm{T}}_{M+1}\mathbf{\Gamma}^{1}_{\bullet,r}\right). \tag{5.6}$$

A pair of indices (qr) here denotes the diploid nuclear state of the zygote, so that a column vector $\mathbf{z}_{qr}$ contains frequencies of zygotes with nuclear allele $q = 0, 1$ inherited from mating type 0 and $r = 0, 1$ inherited from mating type 1. The stochastic matrix $\mathbf{R}$ corresponds to resampling from M to $2M$ mitochondria in cell fusions with uniparental inheritance, so that

$$R_{i,j} = \binom{2M}{i}\left(\frac{j}{M}\right)^{i}\left(1 - \frac{j}{M}\right)^{2M-i}. \tag{5.7}$$

The haploid life cycle concludes with two meiotic subdivisions,

$$\mathbf{P}^{0}_{\bullet,v} = \mathbf{F}_2\mathbf{F}_1(\mathbf{z}_{v0} + \mathbf{z}_{v1}), \tag{5.8}$$

$$\mathbf{P}^{1}_{\bullet,v} = \mathbf{F}_2\mathbf{F}_1(\mathbf{z}_{0v} + \mathbf{z}_{1v}),$$

with $v = 0,\ 1$ and the two transitions matrices $\mathbf{F}_1$ and $\mathbf{F}_2$being the same as in Chapter 4 (Eq. 4.7 and Eq. 4.8).

To determine the evolutionarily stable number of germline cell divisions G_{ESS}, I consider the invasion of a mutant allele for the number of germline cell divisions $0 \leq G \leq N$ different from the one fixed in the resident population. The new allele is inserted into the subpopulation of mating type 0 at a small frequency $\varphi = 0.005$ and its evolution is tracked until either fixation or extinction. Only the case of strict uniparental inheritance of mitochondria is analysed here, and so nuclear modifiers reducing or increasing G are neutral in mating type 1. I consider all possible values of G and build the pairwise invasibility plots depicting the sign of the invasion fitness, i.e. the growth rate of the invader subpopulation when rare (Geritz et al., 1998).

5.5 Evolutionary stable strategy of germline development depends on μ and β

The main benefit of reducing the number of cell divisions in the germline is to reduce the net input of copying errors (μ) in gametes (Fig. 5.1c, d). Setting gametes aside early in development rises the mean offspring fitness in the next generation, but comes at a cost of reduced segregational variance. When gametes are derived later in development, there is a higher chance for segregational drift to generate germ cells with lower or higher numbers of mitochondrial mutations than the organism mean, which facilitates selection among offspring and improves population fitness over generations. The tension between these two forces determines the mode of germline sequestration favoured by purifying selection against mitochondrial mutations.

Numerical analysis of the model recovers a single global attractor $0 \leq G_{\mathrm{ESS}} \leq N$, indicating the uninvadable endpoint of germline evolution (Fig. 5.2). Starting in any

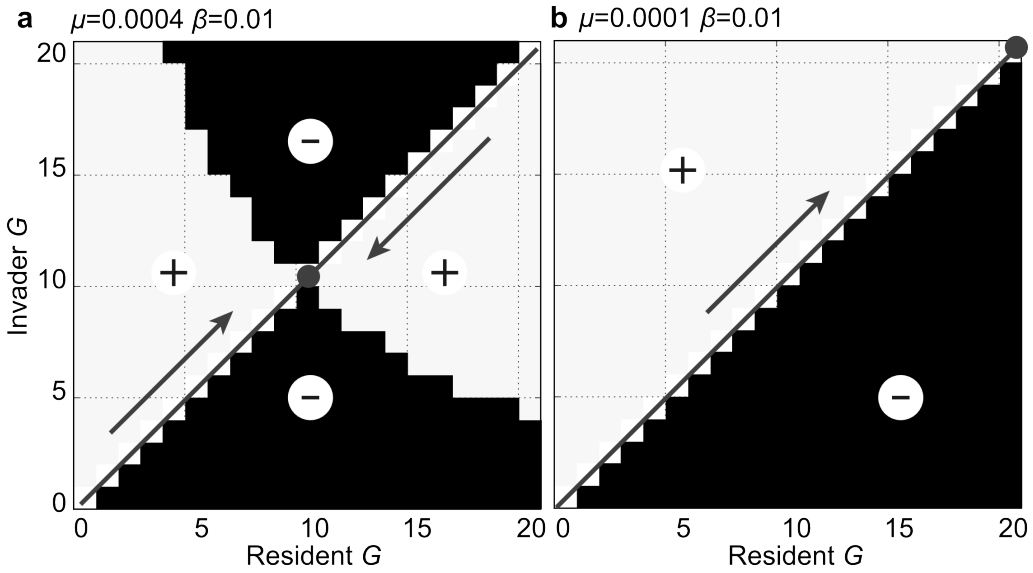

Figure 5.2. Pairwise invasibility plots for the number of germline cell divisions G. Values of G for which the invasion is successful are within the regions marked with "+". Arrows show the direction of trait evolution assuming small changes in G, filled circles indicate positions of evolutionarily stable attractors corresponding to the number of germline cell divisions expected to evolve in the infinite population (**a**). Large values of G are evolutionarily stable under low copying error rates μ (**b**). The number of mitochondria per cell is set to $M = 50$, background mutation rate per generation is $\beta = 0.01$, $\xi = 2$, $s = 1$.

monomorphic state, the population will eventually evolve towards the ESS indicated by the singular point in pairwise invasibility plots, at which point no further change is possible. The invasion attempt in all cases results in either extinction or fixation of the invader allele, that is, dimorphic states do not exist, and the pairwise invasibility plots reflect both the invasion fitness and the equilibrium frequency of the invader allele (0 or 1).

When mitochondrial mutation input through copying errors (μ) is low, the benefit of increased variance between gametes tends to outweigh the benefit of curtailing germline cell division early in development, favouring high values of G_{ESS} (Fig. 5.3a). Note that if copying errors were the only kind of mutation, then selection would always favour setting germ cells aside early in development. However, gametes continue to accumulate new mutations as a result of background damage ($\beta > 0$), irrespective of the number of cell divisions and mitochondrial replication cycles, and even after the germline development is complete (Barritt et al., 2000). These mutations can only be

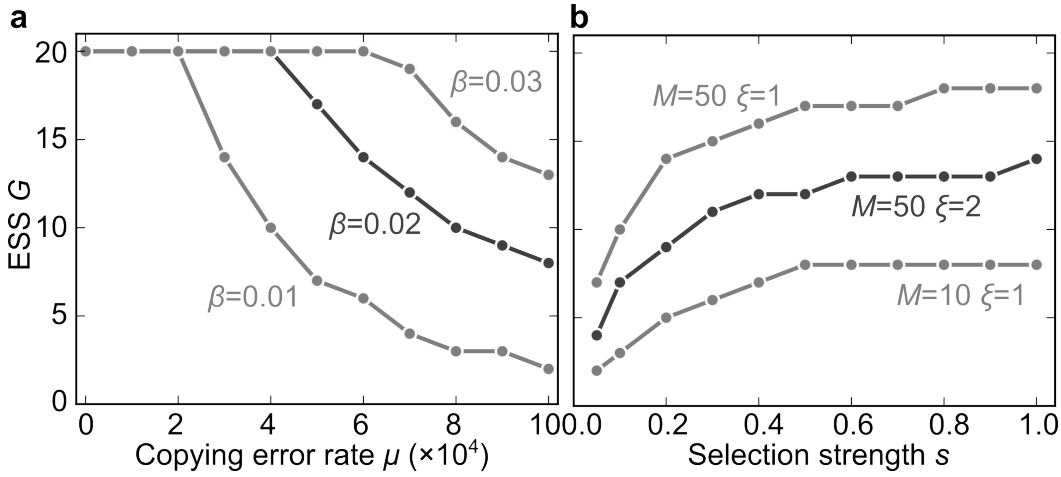

Figure 5.3. Evolutionarily stable number of germline cell divisions G depends on mitochondrial mutation rates and selection strength. Increasing mtDNA copying error rate favours early germline sequestration, while high background mutation rates β favour high numbers of germline cell divisions, i.e. production of gametes from constantly dividing multipotent stem cells throughout organism's adult life (**a**). $M = 50$, $\xi = 2$ and $s = 1$ in the left panel. Increasing M dampens the segregational drift, reducing variance in m/M and giving an additional advantage to late germline sequestration (**b**). The benefit of late gamete production (high G at the ESS) is most apparent under strong purifying selection ($s \to 1$), where the additional segregational variance could increase the efficacy of purifying selection. As in previous chapters, increasing epistasis ξ (the curvature of the concave fitness function) favours alleles associated with less variable mutation load among offspring due to the short-term fitness benefit of reduced genetic variance. Mutation rates are set to $\mu = 0.0003$ and $\beta = 0.01$, the total number of cell divisions in soma is $N = 20$.

segregated out though further rounds of cell division, favouring the increase in G_{ESS} with rising β (Fig. 5.3a).

Not all mitochondrial mutations are necessarily lethal if fixed, and the efficacy of purifying selection for certain kinds of mutations can be low. This effect in the model is captured by selection strength s, which can vary from 0 to 1. Interestingly, the results show that weak selection favours low numbers of germline cell divisions at the ESS, while strong selection is needed to maintain the advantage of late germ cell determination (Fig. 5.3b). The benefit of segregational drift in multiple cell divisions under high values of G is the greater efficacy of selection acting on individuals of highly variable mitochondrial fitness. This benefit cannot be effectively realized under weak purifying selection, in which case early germ cell sequestration is expected to evolve.

The strength of segregational drift depends on the mitochondrial population size M. With large values of M, cell division generates less variance in the frequency of mitochondrial mutations per cell m/M (Fig. 5.1c, d), hindering purifying selection on the organism level. The loss of variance due to high M can be compensated by more germline cell divisions, and the results indeed show that G_{ESS} increases with the number of mitochondria M. The benefit of the additional variance in m/M generated throughout germline development also depends on the strength of epistatic interactions between mitochondrial mutations, as the shape of the fitness function determines how changing variance affects mean adult fitness (see Chapter 2). Akin to the findings discussed in previous chapters where strong negative epistasis ξ gave a short-term advantage to nuclear modifiers reducing variance in m/M, increasing ξ in the present model favours germline sequestration earlier in development (lower G_{ESS}, Fig. 5.3b).

5.6 Discussion

Current theoretical explanations for the evolutionary advantage of the germline-soma differentiation have it that setting aside a dedicated germ cell lineage reduces the scope for evolutionary conflict within a multicellular organism, and reduces the mutational load in the nuclear genome (Buss, 1987; Michod and Roze, 2001; Goldsby et al., 2014). These views, however, cannot explain the lack of germline sequestration in basal metazoans nor the profound differences in germline architecture between males and females in higher animals. Here I developed a new hypothesis that locates the key driving force for the evolution of a dedicated germline in purifying selection against faulty mitochondria. Basal metazoans with low mtDNA copying error rates benefit from segregational variance generated via a large number of germline cell divisions, while high mitochondrial mutation rate per cell division favours early germ cell sequestration. Since mitochondria in most metazoans are inherited maternally, the

constraints on mitochondrial quality in the male germline are less stringent, explaining the sex-specific aspects of the germline architecture. In humans, for example sperm are produced continuously through life with around 400 germline cell divisions by the age of 30, while oocyte production requires only around 24 cell divisions (Drost and Lee, 1995; Crow, 2000).

The mathematical model developed here suggests that with low copying-error rates μ and high background mutation rates β, the combination of segregational drift and purifying selection should improve mitochondrial quality over generations and favour high numbers of germline cell division cycles, eliminating the need for the early germline sequestration. Strikingly, these conditions appear to be true for basal metazoans and plants. Unlike vertebrate mitochondrial genomes, early branching metazoans including sponges, corals and placozoans all have very low mtDNA evolution rates (Shearer et al., 2002; Hellberg, 2006; Huang et al., 2008) and long lives (implying high per-generation background mutation rate β) which readily explains why these major phyla lack a dedicated germline. Likewise, most plants likely have low μ, with mitochondrial evolution rates 50-100 times lower than in animals (Knoop, 2004; Galtier, 2011), while being exposed to high levels of UV radiation in their phototrophic niche, plausibly contributing to high background mutation rate β.

Early germline sequestration is widespread in bilaterians and ctenophores (Blackstone and Jasker, 2003; Moroz et al., 2013). In line with the prediction of the model, both groups have high mitochondrial mutation rates, 10-50 times faster than their mean nuclear evolution rate (Lavrov, 2007; Pett et al., 2011). The metabolic quiescence of oocyte mitochondria reported by de Paula et al. (2013) can be interpreted as a mechanism reducing background damage β, which in our model indeed favours reduced number of germline cell divisions (Fig. 5.3a). While admittedly the overall trend of mitochondrial evolution rates across animals and plants does not allow us to differentiate between μ and β, there is some evidence that mitochondrial

evolution in bilaterians is dominated by copying errors, while background, or oxidative damage is more pronounced in plants and basal metazoans (Hellberg, 2006; Ameur et al., 2011). A more systematic evaluation of μ and β across the metazoan and plant groups would serve as an ultimate test for the present hypothesis.

It has been previously suggested that metazoans derive from an ancestor specifying the fate of its germ cells in late post-embryonic development (Extavour, 2007; Funayama, 2010). Sponges and most of cnidarians use similar strategies of gamete production from populations of endodermally derived pluripotent stem cells capable of giving origin to both somatic and reproductive cells (Agata et al., 2006). It is also very likely that gamete production from somatic or stem cell pools takes place in *Placozoa* (Blackstone, 2009), and could have been present in early lineages predating the Cambrian explosion (Mitchell et al., 2015). If the hypothesis presented here is correct, the evolutionary transition to the germline-soma distinction characteristic of bilaterians was driven by the increase in mitochondrial copying error rate μ.

Why did mitochondrial copying error rate increase in the lineage leading to bilaterians, and, very likely, ctenophores? An interesting possibility is that rising oxygen levels in late Neoproterozoic (Chen et al., 2015) allowed the evolution of predation, rising physical activity and larger body sizes of certain pre-bilaterian lineages (Sperling et al., 2013). Greater activity could have increased rates of tissue turnover, protein synthesis and the frequency of genome replication, inevitably rising the effective copying error rate per cell division μ. The fundamental need to reduce the number of germline cell divisions could have culminated in the complete germline-soma distinction, allowing further somatic differentiation and the evolution of complex developmental processes characteristic of modern bilaterians.

CHAPTER 6. CLOSING REMARKS

6.1 Utility of abstract models in evolutionary biology

The modern evolutionary synthesis of the first half of the 20th century can be attributed in large part to the theoretical work of Sewall Wright, Ronald Fisher and JBS Haldane, uniting Darwinian selection with Mendelian heredity and genetics (Huxley, 1942; Crow, 1987). The trio laid the foundations of the classical population genetics by studying mathematical models that in spite of multiple simplifications and approximations to maintain analytical traceability, elegantly captured the general trends of gene-frequency change under selection, mutation, segregation and drift. The explanatory power and the impact of the earliest theoretical work in population genetics can hardly be overstated, with the mathematical models of Fisher and others now considered a fundamental component of modern evolutionary theory.

But a handful of empirically inclined evolutionary biologists of the time doubted the utility of mathematical gene-pool models. Ernst Mayr, Haldane's close friend and frequent correspondence partner (Rao and Nanjundiah, 2010), sharply criticized the highly reductionist approach of population genetics, and questioned whether mathematical theory can provide any novel contributions to the general understanding of evolutionary processes (Mayr, 1959). Mayr compared the mathematical models of Haldane and colleagues to random sampling from a bag full of coloured beans, referring to them as "beanbag geneticists" (Mayr, 1963):

> "*The Mendelian was apt to compare the genetic contents of a population to a bag full of colored beans. Mutation was the exchange of one kind of bean for another. This conceptualization has been referred to as "beanbag genetics". Work in population and*

developmental genetics has shown, however, that the thinking of beanbag genetics is in many ways quite misleading. To consider genes as independent units is meaningless from the physiological as well as the evolutionary viewpoint..."

Mayr's criticism was largely unfounded, however, and could have stemmed from the plain misunderstanding regarding the true purpose of mathematical models, and unjustified expectations (Borges, 2008). While empirical insights guide modelling, the ultimate goal of mathematical modelling, as that of science in general, is to reduce the complexity of the world making interpretation and discovery easier, and to uncover regularities or laws capable of approximating the system's future behaviour. Models are almost never intended to simulate the complex world as it is. Simplification of reality through modelling is intentional and the reductionism is meaningful, as it removes much of the complexity that is not strictly necessary to understand the most general rules governing the behaviour of a system, nor to make predictions of its future evolution. At the same time, models permit a broader exploration of what is theoretically possible, giving better understanding of why some outcomes and states are universal in nature, while others never occur.

Nearly six decades later, the philosophical underpinnings that fuelled the Mayr-Haldane dispute remain relevant (Servedio et al., 2014). The mathematical models developed in this thesis represent little more than a branch of beanbag-genetics, and as such are prone to the same kind of criticism that Wright, Fisher and many others were forced to endure. But not unlike Haldane (1964), I remain convinced of the explanatory and predictive potential of abstract mathematical models, provided that the readers' expectations are reasonable and that the intention behind developing these models is not misinterpreted.

It therefore needs to be stressed that theoretical models developed in this thesis were not designed to simulate any particular real-world system. Parameter values, such as mutation rates, selection coefficients, epistasis strength or the number

of germline cell divisions, do not (and need not) correspond to empirically measured values. Instead, these mathematical models should be treated as artificial complex systems abstracting the biological world, and designed to test specific hypotheses, to uncover hidden links between selective forces and parameters given a set of simplifying assumptions. Without mathematical models, abstract as they are, the one would be left with vague verbal assertions and conjectures, that, given the fundamental nature of problems investigated, would be nearly impossible to test empirically.

6.2 Mitochondrial mutation dynamics provide a unified account of the evolution of eukaryotic sex

I started this thesis with several fundamental assumptions—hypotheses on the origin and evolution of first eukaryotic lineages, that despite being backed by some empirical and theoretical work, remain criticized (Lynch and Marinov, 2015; Booth and Doolittle, 2015). Perhaps most central to this work is the assumption of mitochondrial endosymbiosis arising early in prokaryote-eukaryote transition and being largely responsible for the further genetic and energetic transformations that shaped the nascent eukaryotic lineage, most notably, sex with whole-cell fusion. I also assumed that later in eukaryotic evolution, there has been a constant selective pressure to maintain the quality of mitochondrial genomes, strong enough to select for nuclear modifiers altering the organism life cycle and developmental programmes. Distancing myself from the solely nuclear perspective on the evolution of eukaryotic sex, I focused on the dynamics of mitochondrial mutations instead.

Within the bioenergetic framework for understanding eukaryotic genome evolution (Lane and Martin, 2010; Lane et al., 2013; Allen, 1993; 2015) these assumptions are easily justifiable, and in fact form a part of the so-called evolutionary synthesis of bioenergetics and genetics, in which chemiosmotic energy transduction is seen as one of the central and most conserved aspects of life (Lane et al., 2013). The

assumption of a fundamental requirement for mitochondrial genome quality allowed me to develop a unified, self-consistent theoretical framework for the evolution of eukaryotic sexual traits, from the origin of sex itself to sexual dimorphism in higher eukaryotes and divergent sex-specific strategies of germline development. Within this framework, sex in the form of proto-eukaryotic host cell fusion evolves because cytoplasmic mixing masks detrimental fitness effects of mutant endosymbionts or faulty organelles, i.e. with cytoplasmic mixing, the fitness of the offspring in the next generation is on the average higher than their parents'. Recombination among proto-nuclear chromosomes then follows, but only as a consequence, and not as a primary driver of cell fusion as the current theory predicts. I therefore suggest that evolutionary forces responsible for the origin of sex were different from the advantages of recombination that maintain sexual life cycles in modern eukaryotes (Otto, 2009).

Biparental inheritance of endosymbiont or organelle genes inevitably becomes deleterious in a long term, because cytoplasmic mixing lowers between-group variance in the mutational load and reduces the efficacy of purifying selection at the level of the host cell. Two mating types arise, with the coupling of mitochondrial inheritance locus to one of the mating types increasing variance and improving the efficacy of selection in the long term—the effect that is most significant under high mitochondrial mutation rates. Evolution of the two genetically determined mating types breaks the symmetry of sex, and lays the foundation for the further evolution of sexual dimorphism and sexual conflict.

The extent of mitochondrial mixing at fertilization can be controlled by alleles linked to the maternal mating type contributing most of the zygote's mitochondria, or to the opposite (paternal) mating type. I showed that strict uniparental inheritance evolves under maternal control, favouring the complete destruction of paternal cytoplasmic genes, whereas paternal control could favour mitochondrial mixing and support stable heteroplasmy. Since the evolutionary interests of two mating types or two sexes (later in evolution) diverge, competition over the control of cytoplasmic

inheritance can arise. These new findings suggest that inter-sexual competition could have been the main driving force behind the evolution of the extraordinary diversity of cellular mechanisms responsible for discarding paternal mitochondria.

With the advent of multicellularity, the fundamental requirement to maintain healthy mitochondrial populations led to the evolution of developmental programmes that facilitate purifying selection at the level of an organism, at the same time avoiding excessive mutation accumulation in female gametes. A restricted number of germline cell replication cycles, for example, could evolve in response to increasing mtDNA replication error rates in the metazoan lineage leading to bilaterians. If the "background" damage to mitochondrial DNA dominates, however, unrestricted stem cell division within the germline will remain evolutionarily stable, as it maintains high variance in the mutant load between gametes and improves the response to purifying selection between offspring. Interestingly, as males in higher metazoans do not pass on their mitochondria to future generations, these developmental constraints do not apply to spermatogenesis. As a consequence, even under high mitochondrial mutation rates, males can maintain actively dividing germline stem cell populations throughout adulthood, specializing in the mass-production of gametes rather than sequestration of a limited number of gamete precursor cells early in development.

6.3. Implications for astrobiology

One of the central questions in the field of astrobiology is what are the most fundamental shared principles that drive the evolution of life on Earth, and what are the most likely properties of cellular life evolving elsewhere in the universe (Billings et al., 2006). Although we are still far from detecting any signs of life beyond Earth or even estimating the probability of its existence, with the abundance of recent discoveries of planets and moons that might possess properties supporting biological life, the issue has never been more relevant. Arguably, abstract mathematical

modelling might become the major contributor towards understanding the most basic shared characteristics of cellular life, independent of the location of its origin.

Assuming that cellular life is in general likely to rely on pH gradients and chemiosmosis for energy transduction (Russell et al., 2014), by the conjectures of Lane and Martin (2010) it follows that complex multicellular life forms with large genomes could evolve only by the means of internalization of chemiosmotic membranes, either through endosymbiosis or some alternative mechanism. Based on these arguments, Lane (2015) further proposed that simple microbial life might be relatively common in the universe, given the plausible abundance of extra-terrestrial hydrothermal vent-like structures. But because of the strong bioenergetic constraints on what is achievable through natural selection, that does not necessarily imply the successful progression from bacterial to complex eukaryote-like life, which requires a rare evolutionary transition producing internal energy-generating membranes supported by their own local “genomic outposts”—the kind of transition that cannot result from small genetic changes gradually selected over time (Maynard Smith and Szathmáry, 1995).

While somewhat speculative, these assumptions suggest that the evolutionary pathway from asexual microbial life to highly complex multicellular organisms with two sexes, as detailed in the preceding chapters, could be a shared feature of all complex life and apply equally well to cellular life forms elsewhere in the universe—a consequence of strong bioenergetic constraints exerted on the seemingly unlimited potential of nuclear genetics. The evolution of sexual life cycles would be a must *en route* to complex life due to the fundamental need to maintain the integrity of large nuclear genomes, and could evolve through cell fusion initially masking the detrimental effects of mutations within cytoplasmic genes. This would create a selective pressure for the emergence of mechanisms restricting cytoplasmic inheritance and establishing the population with two mating types. The evolutionary transition to multicellularity would create new selective forces acting on developmental programmes, resulting in the fundamental differences in germline development between sexes: female-like

organisms protecting the cytoplasmic genes in early sequestered gametes and males specializing in mass production of reproductive cells that do not contribute their cytoplasmic genes to the next generation. The evolution of secondary sexual characteristics, sexual conflict and sexual selection would follow.

As a final closing remark, it needs to be reiterated that the power of mathematical modelling in biology lies largely in its ability to reveal what is possible, without necessarily discovering the ultimate truth. The theoretical considerations presented in this thesis suggest one evolutionary pathway from primordial life to the biological complexity observed today, but many alternative pathways are possible. Given the enormous complexity of biological systems, it is very likely that multiple selective pressures and multiple intertwined trajectories have contributed to the evolution of phonotypes discussed in this work, obscuring the general picture. It is up to future research, theoretical and empirical, to identify the dominant pathways, and to bring us even closer to a complete understanding of the universal principles governing the evolution of complex life.

Acknowledgements

I am sincerely grateful to my supervisors, Andrew Pomiankowski and Nick Lane, for their guidance and inspiration, for sharing their experience and welcoming new ideas, and for their hard work and patience that ultimately helped me to become a better researcher. It has been a genuine privilege to work with POM and Nick, and it is my sincere hope that the completion of my doctorate will not be the end of our collaborative endeavours. I would also like to thank professor Neil W. Blackstone for hosting me at the Northern Illinois University during my short but very memorable and illuminating visit to the Midwest, and for helping me shape some of the ideas that in one way or another are expressed in this thesis.

Finally, I am grateful to the small community of CoMPLEX, and the broader one of UCL, for welcoming me to the rainy London and making my research here possible. It has been an honour and an inspiration to roam the celebrated halls of University College, where the legacy of JBS Haldane, George Price, Ronald Fisher and John Maynard Smith—some of the greatest minds in evolutionary biology—can still be felt.

References

A

Agata K, Nakajima E, Funayama N, Shibata N, Saito Y, Umesono Y. 2006. Two different evolutionary origins of stem cell systems and their molecular basis. *Semin. Cell Dev. Biol.* 17, 503-509.

Al Rawi S, Louvet-Vallée S, Djeddi A, Sachse M, Culetto E, Hajjar C, Boyd L, Legouis R, Galy V. 2011. Postfertilization autophagy of sperm organelles prevents paternal mitochondrial DNA transmission. *Science* 334(6059), 1144-1147.

Alcock F, Clements A, Webb C, Lithgow T. 2010. Tinkering inside the organelle. *Science* 327(5966), 649-650.

Allen JF. 1993. Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. *J. Theor. Biol.* 165(4), 609-631.

Allen JF. 1996. Separate sexes and the mitochondrial theory of ageing. *J. Theor. Biol.* 180(2), 135-140.

Allen JF. 2015. Why chloroplasts and mitochondria retain their own genomes and genetic systems: colocation for redox regulation of gene expression. *Proc. Natl. Acad. Sci. USA* 112(33), 10231-10238.

Ameur A, Stewart JB, Freyer C, Hagstrom E, Ingman M, Larsson NG, Gyllensten U. 2011. Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *PLoS Genet.* 7(3), e1002028.

Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Pontén T, Alsmark UCM, Podowski RM, Näslund AK, Eriksson AS, Winkler HH, Kurland CG. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* 396(6707), 133-140.

B

Barr CM, Neiman M, Taylor DR. 2005. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *New Phytol.* 168(1), 39-50.

Barritt JA, Cohen J, Brenner CA. 2000. Mitochondrial DNA point mutation in human oocytes is associated with maternal age. *Reprod. Biomed. Online* 1, 96-100.

Bastiaans E, Aanen DK, Debets AJM, Hoekstra RF, Lestrade B, Maas MFPM. 2014. Regular bottlenecks and restrictions to somatic fusion prevent the accumulation of mitochondrial defects in *Neurospora*. *Phil. Trans. R. Soc. B* 369, 20130448.

Bell G. 1978. The evolution of anisogamy. *J. Theor. Biol.* 73(2), 247-270.

Belous A, Knox C, Nicoud IB, Pierce J, Anderson C, Pinson CW, Chari RS. 2003. Reversed activity of mitochondrial adenine nucleotide translocator in ischemia-reperfusion. *Transplantation* 75(10), 1717-1723.

Bergstrom CT, Pritchard J. 1998. Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. *Genetics* 149(4), 2135-2146.

Billiard S, López-Villavicencio M, Devier B, Hood ME, Fairhead C, Giraud T. 2011. Having sex, yes, but with whom? Inferences from fungi on the evolution of anisogamy and mating types. *Biol. Rev. Camb. Philos. Soc.* 86(2), 421-442.

Billings L, Cameron V, Claire M, Dick GJ, Domagal-Goldman SD, Javaux EJ, Johnson OJ, Laws C, Race MS, Rask J, Rummel JD, et al. 2006. The astrobiology primer: an outline of general knowledge - version 1, 2006. Astrobiology 6, 735-813.

Birky CW. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc. Natl. Acad. Sci. USA* 92(25), 11331-11338.

Birky CW. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu. Rev. Genet.* 35(1), 125-148.

Blackstone NW. 1995. Perspective: a units-of-evolution perspective on the endosymbiont theory of the origin of the mitochondrion. *Evolution* 49(5), 785-796.

Blackstone NW. 2009. A new look at some old animals. *PLoS Biol.* 7(1), e1000007.

Blackstone NW. 2013. Why did eukaryotes evolve only once? Genetic and energetic aspects of conflict and conflict mediation. *Philos. Trans. R. Soc. Lond. B* 368(1622), 20120266.

Blackstone NW, Green DR. 1999. The evolution of a mechanism of cell suicide. *Bioessays* 21(1), 84-88.

Blackstone NW, Jasker BD. 2003. Phylogenetic considerations of clonality, coloniality, and mode of germline development in animals. *J. Exp. Zool. B Mol. Dev. Evol.* 297(1), 35-47.

Blackstone NW. 2016. An evolutionary framework for understanding the origin of eukaryotes. *Biology* 5(2), 18.

Booth A, Doolittle WF. 2015. Eukaryogenesis, how special really? *Proc. Natl. Acad. Sci. USA* 112(33), 10278-10285.

Borges RM. 2008. Commentary: the objection is sustained: a defence of the defense of beanbag genetics. *Int. Journal Epidemiol.* 37(3), 451-454.

Breton S, Stewart DT. 2015. Atypical mitochondrial inheritance patterns in eukaryotes. *Genome* 58(10), 423-431.

Buchet K, Godinot C. 1998. Functional F1-ATPase essential in maintaining growth and membrane potential of human mitochondrial DNA-depleted ρ cells. *J. Biol. Chem.* 273(36), 22983-22989.

Bulmer MG, Parker GA. 2002. The evolution of anisogamy: a game-theoretic approach. *Proc. R. Soc. Lond. B Biol. Sci.* 269(1507), 2381-2388.

Burger G, Plante I, Lonergan KM, Gray MW. 1995. The mitochondrial DNA of the amoeboid protozoon, *Acanthamoeba castellanii*: complete sequence, gene content and genome organization. *J. Mol. Biol.* 245(5), 522-537.

Burger G, Gray MW, Forget L, Lang BF. 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol. Evol.* 5(2), 418-438.

Burgstaller JP, Johnston IG, Jones NS, Albrechtová J, Kolbe T, Vogl C, Futschik A, Mayrhofer C, Klein D, Sabitzer S, Blattner M. 2014. mtDNA segregation in heteroplasmic tissues is common in vivo and modulated by haplotype differences and developmental stage. *Cell Rep.* 7(6), 2031-2041.

Buss LW. 1983. Evolution, development, and the units of selection. *Proc. Natl. Acad. Sci. USA* 80, 1387-1391.

Buss LW. 1987. The Evolution of Individuality (Princeton University Press, Princeton, NJ).

C

Cao L, Shitara H, Horii T, Nagao Y, Imai H, Abe K, Hara T, Hayashi JI, Yonekawa H. 2007. The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. *Nature Genet.* 39(3), 386-390.

Cao L, Shitara H, Sugimoto M, Hayashi JI, Abe K, Yonekawa H. 2009. New evidence confirms that the mitochondrial bottleneck is generated without reduction of mitochondrial DNA content in early primordial germ cells of mice. *PLoS Genet.* 5(12), e1000756.

Castellana S, Vicario S, Saccone C. 2011. Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein coding genes. *Genome Biol. Evol.* 3, 1067-1079.

Cavalier-Smith T. 1983. A 6 kingdom classification and a unified phylogeny. In Schwemmler W, Schenk HEA (eds), *Endocytobiology II* (De Gruyter, Berlin, Germany), 1027-1034.

Cavalier-Smith T. 2006. Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium. *Proc. R. Soc. Lond. B* 273(1596), 1943-1952.

Charlesworth D, Charlesworth B. 1979. The evolution and breakdown of S-allele systems. *Heredity* 43, 41-55.

Chen X, Ling HF, Vance D, Shields-Zhou GA, Zhu M, Poulton SW, Och LM, Jiang SY, Li D, Cremonese L, Archer C. 2015. Rise to modern levels of ocean oxygenation coincided with the Cambrian radiation of animals. *Nature Commun.* 6, 7142.

Christie JR, Schaerf TM, Beekman M. 2015. Selection against heteroplasmy explains the evolution of uniparental inheritance of mitochondria. *PLoS Genet*. 11(4), e1005112.

Clark KA, Howe DK, Gafner K, Kusuma D, Ping S, Estes S, Denver DR. 2012. Selfish little circles: transmission bias and evolution of large deletion-bearing mitochondrial DNA in *Caenorhabditis briggsae* nematodes. *PLoS ONE* 7(7), e41433.

Cooper BS, Burrus CR, Ji C, Hahn MW, Montooth KL. 2015. Similar efficacies of selection shape mitochondrial and nuclear genes in both *Drosophila melanogaster* and *Homo sapiens*. *G3,* 5(10), 2165-2176.

Cosmides LM, Tooby J. 1981. Cytoplasmic inheritance and intragenomic conflict. *J. Theor. Biol.* 89(1), 83-129.

Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM. 2008. The archaebacterial origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* 105(51), 20356-20361.

Cree LM, Samuels DC, de Sousa Lopes SC, Rajasimha HK, Wonnapinij P, Mann JR, Dahl HHM, Chinnery PF. 2008. A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nature Genet.* 40(2), 249-254.

Crow JF. 1987. Population genetics history: a personal view. *Annu. Rev. Genet.* 21(1), 1-22.

Crow JF. 2000. The origins, patterns and implications of human spontaneous mutation. *Nature Rev. Genet.* 1(1), 40-47.

Czárán TL, Hoekstra RF. 2004. Evolution of sexual asymmetry. *BMC Evol. Biol.* 4(1), 1.

D

Dacks J, Roger AJ. 1999. The first sexual lineage and the relevance of facultative sex. *J. Mol. Evol.* 48(6), 779-783.

DeLuca SZ, O'Farrell PH. 2012. Barriers to male transmission of mitochondrial DNA in sperm development. *Dev. Cell* 22(3), 660-668.

Dey G, Thattai M, Baum B. 2016. On the archaeal origins of eukaryotes and the challenges of inferring phenotype from genotype. *Trends Biol.* 26(7), 476-485.

Dieckmann U. 1997. Can adaptive dynamics invade? *Trends Ecol. Evol.* 12(4), 128-131.

Dolezal P, Likic V, Tachezy J, Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. *Science* 313(5785), 314-318.

Driessen AJ, Nouwen N. 2008. Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem.* 77, 643-667.

Drost JB, Lee WR. 1995. Biological basis of germline mutation: comparisons of spontaneous germline mutation rates among *Drosophila*, mouse, and human. *Environ. Mol. Mutagen.* 25(S2), 48-64.

E

Eberhard WG. 1980. Evolutionary consequences of intracellular organelle competition. *Q. Rev. Biol.* 55(3), 231-249.

Elson JL, Turnbull DM, Howell N. 2004. Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am. J. Hum. Genet.* 74(2), 229–238.

Embley M, Van Der Giezen M, Horner DS, Dyal PL, Foster P. 2003. Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. *Philos. Trans. R. Soc. Lond. B* 358(1429), 191-203.

Embley M, Martin W. 2006. Eukaryotic evolution, changes and challenges. *Nature* 440, 623-630.

Eshel I. 1983. Evolutionary and continuous stability. *J. Theor. Biol.* 103(1), 99-111.

Extavour CG. 2007. Evolution of the bilaterian germ line: lineage origin and modulation of specification mechanisms. *Integr. Comp. Biol.* 47, 770-785.

Extavour CG, Akam M. 2003. Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* 130, 5869-5884.

Ewens WJ. 2004. *Mathematical Population Genetics, I. Theoretical Introduction.* Second edition (Springer, New York).

F

Fan W, Waymire KG, Narula N, Li P, Rocher C, Coskun PE, Vannan MA, Narula J, MacGregor GR, Wallace DC. 2008. A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations. *Science*, 319(5865), 958-962.

Fedler M, Luh KS, Stelter K, Nieto-Jacobo F, Basse CW. 2009. The *a*2 mating-type locus genes *lga*2 and *rga*2 direct uniparental mitochondrial DNA (mtDNA) inheritance and constrain mtDNA recombination during sexual development of *Ustilago maydis*. *Genetics* 181(3), 847-860.

Fisher RA. 1930. *The genetical theory of natural selection* (Oxford University Press, Oxford).

Frank SA. 1996. Host control of symbiont transmission: the separation of symbionts into germ and soma. *Am. Nat*. 148(6), 1113-1124.

Fröls S. 2013. Archaeal biofilms: widespread and complex. *Biochem. Soc. Trans.* 41(1), 393-398.

Funayama N. 2010. The stem cell system in demosponges: insights into the origin of somatic stem cells. *Dev. Growth Differ.* 52(1), 1-14.

G

Galtier N. 2011. The intriguing evolutionary dynamics of plant mitochondrial DNA. *BMC Biol.* 9(1), 61.

Garg SG, Martin WF. 2016. Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome Biol. Evol.* 8(6), 1950-1970.

Geritz SA, Mesze G, Metz JA. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* 12(1), 35-57.

Giraud MF, Velours J. 1997. The absence of the mitochondrial ATP synthase δ subunit promotes a slow growth phenotype of rho- yeast cells by a lack of assembly of the catalytic sector F1. *Eur. J. Biochem.* 245(3), 813-818.

Goldsby HJ, Knoester DB, Ofria C, Kerr B. 2014 The evolutionary origin of somatic cells under the dirty work hypothesis. *PLoS Biol*. 12(5), e1001858.

Goodenough U, Lin H, Lee JH. 2007. Sex determination in *Chlamydomonas*. *Semin. Cell Dev. Biol.* 18(3), 350-361.

Goodenough U, Heitman J. 2014. Origins of eukaryotic sexual reproduction. *Cold Spring Harb. Perspect. Biol.* 6(3), a016154.

Gould SB, Garg SG, Martin WF. 2016. Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. *Trends Microbiol.* 24(7), 525-534.

Gray MW, Burger G, Lang BF. 1999. Mitochondrial evolution. *Science* 283(5407), 1476-1481.

Greiner S, Sobanski J, Bock R. 2015. Why are most organelle genomes transmitted maternally? *BioEssays* 37(1), 80-94.

Gross J, Bhattacharya D. 2009. Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective. *Nature Rev. Genet.* 10(7), 495-505.

Gross J, Bhattacharya D. 2010. Uniting sex and eukaryote origins in an emerging oxygenic world. *Biol. Direct* 5(1), 53.

Guerrero R, Pedrós-Alió C, Esteve I, Mas J, Chase D, Margulis L. 1986. Predatory prokaryotes: predation and primary consumption evolved in bacteria. *Proc. Natl. Acad. Sci. USA*, 83(7) 2138-2142.

Guy L, Saw JH, Ettema TJ. 2014. The archaeal legacy of eukaryotes: a phylogenomic perspective. *Cold Spring Harb. Perspect. Biol.* 6(10), a016022.

H

Hadjivasiliou Z, Pomiankowski A, Seymour RM, Lane N. 2012. Selection for mitonuclear co-adaptation could favour the evolution of two sexes. *Proc. R. Soc. Lond. B* 279(1734), 1865-1872.

Hadjivasiliou Z, Lane N, Seymour RM, Pomiankowski A. 2013. Dynamics of mitochondrial inheritance in the evolution of binary mating types and two sexes. *Proc. R. Soc. Lond. B,* 280(1769), 20131920.

Hadjivasiliou Z, Iwasa Y, Pomiankowski A. 2015. Cell-cell signalling in sexual chemotaxis: a basis for gametic differentiation, mating types and sexes. *J. R. Soc. Interface* 12(109), 20150342.

Hadjivasiliou Z, Pomiankowksi A. 2016. Gamete signalling underlies the evolution of mating types and their number. *Phil. Trans. R. Soc. B* 371(1706), 20150531.

Haig D. 2016. Intracellular evolution of mitochondrial DNA (mtDNA) and the tragedy of the cytoplasmic commons. *BioEssays* 38(6), 549-555.

Halliwell B. 2006. Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life. *Plant Physiol.* 141(2), 312-322.

Hamilton WD. 1980. Sex versus non-sex versus parasite. *Oikos* 35, 282-290.

Hamilton WD, Axelrod R, Tanese R. 1990. Sexual reproduction as an adaptation to resist parasites (a review). *Proc. Natl. Acad. Sci. USA* 87(9), 3566-3573.

Hartfield M, Keightley PD. 2012. Current hypotheses for the evolution of sex and recombination. *Integr. Zool.* 7, 192-209.

Hastings IM. 1992. Population genetic aspects of deleterious cytoplasmic genomes and their effect on the evolution of sexual reproduction. *Genet. Res.* 59(03), 215-225.

Havird JC, Hall MD, Dowling DK. 2015. The evolution of sex: a new hypothesis based on mitochondrial mutational erosion. *Bioessays* 37, 951-958.

Heitman J. 2015. Evolution of sexual reproduction: a view from the fungal kingdom supports an evolutionary epoch with sex before sexes. *Fungal Biol. Rev.* 29(3), 108-117.

Hellberg ME. 2006. No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. *BMC Evol. Biol.* 6(1), 24.

Henze K, Martin W. 2003. Evolutionary biology: essence of mitochondria. *Nature* 426(6963), 127-128.

Hillis WD. 1993. Why physicists like models and why biologists should. *Curr. Biol.* 3(2), 79-81.

Hoekstra RF. 1982. On the asymmetry of sex: evolution of mating types in isogamous populations. *J. Theor. Biol.* 98, 427-451.

Hoekstra RF. 2011. Nucleo-cytoplasmic conflict and the evolution of gamete dimorphism. In Togashi T, Cox PA (eds), *The evolution of anisogamy* (Cambridge Univeristy Press, Cambridge), 111-130.

Hörandl E, Hadacek F. 2013. The oxidative damage initiation hypothesis for meiosis. *Plant Reprod.* 26(4), 351-367.

Horner DS, Hirt RP, Kilvington S, Lloyd D, Embley TM. 1996. Molecular data suggest an early acquisition of the mitochondrion endosymbiont. *Proc. R. Soc. Lond. B* 263(1373), 1053-1059.

Huang D, Meier R, Todd PA, Chou LM. 2008. Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. *J. Mol. Evo.* 66, 167-174.

Hurst LD. 1990. Parasite diversity and the evolution of diploidy, multicellularity and anisogamy. *J. Theor. Biol.* 144(4), 429-443.

Hurst LD. 1995. Selfish genetic elements and their role in evolution: the evolution of sex and some of what that entails. *Philos. Trans. R. Soc. Lond. B* 349(1329), 321-332.

Hurst LD, Hamilton WD. 1992. Cytoplasmic fusion and the nature of sexes. *Proc. R. Soc. Lond. B* 247(1320), 189-194.

Hutson V, Law R. 1993. Four steps to two sexes. *Proc. R. Soc. Lond. B* 253(1336), 43-51.

Huxley J. 1944. *Evolution: the modern synthesis* (George Allen and Unwin, London).

I

Itoh T, Martin W, Nei M. 2002. Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. *Proc. Natl. Acad. Sci. USA* 99, 12944-12948.

Iwasa Y, Sasaki A. 1987. Evolution of the number of sexes. *Evolution* 41(1), 49-65.

J

Jaenike J. 1978. An hypothesis to account for the maintenance of sex within populations. *Evol. Theory* 3, 191-194.

Juliano CE, Swartz SZ, Wessel GM. 2010. A conserved germline multipotency program. *Development* 137(24), 4113-4126.

K

Kaneda H, Hayashi J, Takahama S, Taya C, Lindahl KF, Yonekawa H. 1995. Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. *Proc. Natl. Acad. Sci. USA* 92(10), 4542–4546.

Kant I. 1998. Guyer P, Wood AW (eds), *Critique of pure reason* (Cambridge University Press, Cambridge).

Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková R, Eme L, Novák L, Žárský V, Barlow LD, Herman EK, Soukal P. 2016. A eukaryote without a mitochondrial organelle. *Curr. Biol.* 26(10), 1274-1284.

Keightley PD, Otto SP. 2006. Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* 443(7107), 89-92.

Kirk DL. 2006. Oogamy: inventing the sexes. *Curr. Biol. 16*(24), R1028-R1030.

Klinger CM, Spang A, Dacks JB, Ettema TJ. 2016. Tracing the archaeal origins of eukaryotic membrane-trafficking system building blocks. *Mol. Biol. Evol.* 33(6), 1528-1541.

Klingenberg M. 2008. The ADP and ATP transport in mitochondria and its carrier. *BBA-Biomembranes* 1778(10), 1978-2021.

Knoop V. 2004. The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.* 46(3), 123-139.

Knowlton N. 1974. A note on the evolution of gamete dimorphism. *J. Theor. Biol.* 46(1), 283-285.

Kondo R, Satta Y, Matsuura ET, Ishiwa H, Takahata N, Chigusa SI. 1990. Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. *Genetics* 126(3), 657-663.

Kondrashov AS. 1993. Classification of hypotheses on the advantage of amphimixis. *J. Hered.* 84(5), 372-387.

Koonin EV. 2006. The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate? *Biol. Direct* 1(1), 22.

Koonin EV. 2009. Intron-dominated genomes of early ancestors of eukaryotes. *J. Heredity* 100(5), 618-662.

Koonin EV. 2010. The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biol*. 11(5), 209.

Koonin EV. 2015. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Phil. Trans. R. Soc. B* 370(1678), 20140333.

Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM, Makarova KS, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Rogozin IB. 2004. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5(2), R7.

Koufopanou V. 1994. The evolution of soma in the Volvocales. *Am. Nat.* 143(5), 907-931.

Koumandou VL, Wickstead B, Ginger ML, van der Giezen M, Dacks JB, Field MC. 2013. Molecular paleontology and complexity in the last eukaryotic common ancestor. *Crit. Rev. Biochem. Mol. Biol.* 48(4), 373-396.

Kouyos RD, Silander OK, Bonhoeffer S. 2007. Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol.* 22(6), 308-315.

Krakauer DC, Mira A. 1999. Mitochondria and germ-cell death. *Nature* 400(6740), 125-126.

Ku C, Nelson-Sathi S, Roettger M, Garg S, Hazkani-Covo E, Martin WF. 2015. Endosymbiotic gene transfer from prokaryotic pangenomes: inherited chimerism in eukaryotes. *Proc. Natl. Acad. Sci. USA* 112(33), 10139-10146.

Kuijper B, Lane N, Pomiankowski A, 2015. Can paternal leakage maintain sexually antagonistic polymorphism in the cytoplasm? *J. Evol. Biol.* 28, 468-480.

Kulawiak B, Höpker J, Gebert M, Guiard B, Wiedemann N, Gebert N. 2013. The mitochondrial protein import machinery has multiple connections to the respiratory chain. *BBA-Bioenergetics* 1827(5), 612-626.

Kunji ER. 2004. The role and structure of mitochondrial carriers. *FEBS Letters* 564(3), 239-244.

Kurland CG, Andersson SGE. 2000. Origin and evolution of the mitochondrial proteome. *Microbiol. Mol. Biol. Rev.* 64(4), 786-820.

Kutik S, Guiard B, Meyer HE, Wiedemann N, Pfanner N. 2007. Cooperation of translocase complexes in mitochondrial protein import. *J. Cell Biol.* 179(4), 585-591.

Kvist L, Martens J, Nazarenko AA, Orell M. 2003 Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Mol. Biol. Evol.* 20(2), 243-247.

L

Lane N. 2011a. Energetics and genetics across the prokaryote-eukaryote divide. *Biol. Direct* 6, 35.

Lane N. 2011b. Mitonuclear match: optimizing fitness and fertility over generations drives ageing within generations. *Bioessays* 33(11), 860-869.

Lane N. 2014. Bioenergetic constraints on the evolution of complex life. *Cold Spring Harb. Perspect. Biol.* 6(5), a015982.

Lane N. 2015. The vital question: why is life the way it is? Profile books (London).

Lane N, Martin W. 2010. The energetics of genome complexity. *Nature* 467(7318), 929-934.

Lane N, Martin WF, Raven JA, Allen JF. 2013. Energy, genes and evolution: introduction to an evolutionary synthesis. *Phil. Trans. R. Soc. B* 368(1622), 20120253.

Lavrov DV. 2007. Key transitions in animal evolution: a mitochondrial DNA perspective. *Integr. Comp. Biol.* 47, 734-743.

Lehtonen J, Jennions MD, Kokko H. 2012. The many costs of sex. *Trends Ecol. Evol.* 27(3), 172-178.

Lehtonen J, Parker GA, Schärer L. 2016. Why anisogamy drives ancestral sex roles. *Evolution* 70(5), 1129-1135.

Luo SM, Ge ZJ, Wang ZW, Jiang ZZ, Wang ZB, Ouyang YC, Hou Y, Schatten H, Sun QY. 2013. Unique insights into maternal mitochondrial inheritance in mice. *Proc. Natl. Acad. Sci. USA* 110(32), 13038-13043.

Lynch M, Koskella B, Schaack S. 2006. Mutation pressure and the evolution of organelle genomic architecture. *Science* 311, 1727-1730.

Lynch M, Sung W, Morris K, Coffey N, Landry CR, Dopman EB, Dickinson WJ, Okamoto K, Kulkarni S, Hartl DL, Thomas WK. 2008. A genome-wide view of the

spectrum of spontaneous mutations in yeast. *Proc. Natl. Acad. Sci. USA* 105, 9272-9277.

Lynch M, Marinov GK. 2015. The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. USA* 112(51), 15690-15695.

M

Magnacca KN, Brown MJ. 2010. Tissue segregation of mitochondrial haplotypes in heteroplasmic Hawaiian bees: implications for DNA barcoding. *Mol. Ecol. Resour.* 10(1), 60-68.

Magoulas A, Zouros E. 1993. Restriction–site heteroplasmy in anchovy (*Engraulis encrasicolus*) indicates incidental biparental inheritance of mitochondrial DNA. *Mol. Biol. Evol.* 10(2), 319-325.

Makarova KS, Wolf YI, Mekhedov SL, Mirkin BG, Koonin EV. 2005. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. *Nucleic Acids Res.* 33(14), 4626-4638.

Makarova KS, Yutin N, Bell SD, Koonin EV. 2010. Evolution of diverse cell division and vesicle formation systems in Archaea. *Nature Rev. Microbiol.* 8(10), 731-741.

Marais GA, Calteau A, Tenaillon O. 2008. Mutation rate and genome reduction in endosymbiotic and free-living bacteria. *Genetica* 134, 205-210.

Marcon E, Moens PB. 2005. The evolution of meiosis: recruitment and modification of somatic DNA-repair proteins. *Bioessays* 27(8), 795-808.

Martijn J, Ettema TJ. 2013. From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. *Biochem. Soc. Trans.* 41(1), 451-457.

Martin WF, Müller M. 1998. The hydrogen hypothesis for the first eukaryote. *Nature* 392(6671), 37-41.

Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. *Phil. Trans. R. Soc. B* 370(1678), 20140330.

Maynard Smith J, Szathmáry E. 1995. The Major Transitions in Evolution (Oxford University Press, Oxford).

Mayr, E. 1963. *Animal species and evolution* (Harvard University Press, Cambridge Massachusetts).

Mazat JP, Rossignol R, Malgat M, Rocher C, Faustin B, Letellier T. 2001. What do mitochondrial diseases teach us about normal mitochondrial functions...that we already knew: threshold expression of mitochondrial defects. *BBA-Bioenergetics* 1504(1), 20-30.

McCauley DE. 2013. Paternal leakage, heteroplasmy, and the evolution of plant mitochondrial genomes. *New Phytol.* 200(4), 966-977.

Mereschkowsky C. 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Centralbl.* 25, 593-604. English translation in Martin W, Kowallik KV. 1999. *Eur. J. Phycol.* 34, 287-295.

Mereschkowsky C. 1910. Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol. Centralbl.* 30, 278-288.

Michod RE. 1997. Cooperation and conflict in the evolution of individuality. I. Multilevel selection of the organism. *Am. Nat.* 149, 607-645.

Michod RE. 2007. Evolution of individuality during the transition from unicellular to multicellular life. *Proc. Natl. Acad. Sci. USA* 104, 8613-8618.

Michod RE, Roze D. 2001. Cooperation and conflict in the evolution of multicellularity. *Heredity* 86(1), 1-7.

Michod RE, Nedelcu AM. 2003. On the reorganization of fitness during evolutionary transitions in individuality. *Integr. Comp. Biol.* 43, 64-73.

Milani L, Ghiselli F. 2015. Mitochondrial activity in gametes and transmission of viable mtDNA. *Biol. Direct* 10, 22.

Mitchell EG, Kenchington CG, Liu AG, Matthews JJ, Butterfield NJ. 2015. Reconstructing the reproductive mode of an Ediacaran macro-organism. *Nature* 524(7565), 343-346.

Miyabayashi S, Hanamizu H, Nakamura R, Endo H, Tada K. 1992. Defects of mitochondrial respiratory enzymes in cloned cells from MELAS fibroblasts. *J. Inherit. Metab. Dis.* 15(5), 797-802.

Moriyama Y, Kawano S. 2003. Rapid, selective digestion of mitochondrial DNA in accordance with the matA hierarchy of multiallelic mating types in the mitochondrial inheritance of *Physarum polycephalum*. *Genetics* 164(3), 963-975.

Moroz LL, Kocot KM, Citarella MR, Dosung S, Norekian TP, Povolotskaya IS, Grigorenko AP, Dailey C, Berezikov E, Buckley KM, et al. 2014. The ctenophore genome and the evolutionary origins of neural systems. *Nature* 510, 109-114.

Müller HJ. 1964. The relation of recombination to mutational advance. *Mutat. Res.-Fund. Mol. Mech. Mut.* 1(1), 2-9.

Müller M. 1988. Energy metabolism of protozoa without mitochondria. *Annu. Rev. Microbiol.* 42(1), 465-488.

Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu RY, van der Giezen M, Tielens AG, Martin WF. 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Molec. Biol. Rev.* 76(2), 444-495.

Müller WA, Teo R, Frank U. 2004. Totipotent migratory stem cells in a hydroid. *Dev. Biol.* 275(1), 215-224.

Murphy MP. 2009. How mitochondria produce reactive oxygen species. *Biochem. J.* 417(1), 1-13.

N

Narra HP and Ochman H. 2006. Of what use is sex to bacteria? *Curr. Biol.* 16(17), R705-R710.

Nelson-Sathi S, Dagan T, Landan G, Janssen A, Steel M, McInerney JO, Deppenmeier U, Martin WF. 2012. Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of *Haloarchaea*. *Proc. Natl. Acad. Sci. USA* 109(50), 20537-20542.

Nelson-Sathi S, Sousa FL, Roettger M, Lozada-Chávez N, Thiergart T, Janssen A, Bryant D, Landan G, Schönheit P, Siebers B, McInerney JO. 2015. Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 517(7532), 77-80.

Neupert W, Herrmann JM. 2007. Translocation of proteins into mitochondria. *Annu. Rev. Biochem.* 76, 723-749.

Ng WL, Bassler BL. 2009. Bacterial quorum-sensing network architectures. *Annu. Rev.Genet.* 43, 197.

Nishimura Y, Yoshinari T, Naruse K, Yamada T, Sumi K, Mitani H, Higashiyama T, Kuroiwa T. 2006. Active digestion of sperm mitochondrial DNA in single living sperm revealed by optical tweezers. *Proc. Natl. Acad. Sci. USA* 103(5), 1382-1387.

O

Otto SP. 2009. The evolutionary enigma of sex. *Am. Nat.* 174(S1), S1-S14.

P

Page SL, Hawley RS. 2003. Chromosome choreography: the meiotic ballet. *Science* 301(5634), 785-789.

Palmer JD, Herbon LA. 1988. Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. *J. Mol. Evol.* 28, 87-97.

Palmieri F, Monné M. 2016. Discoveries, metabolic roles and diseases of mitochondrial carriers: a review. *BBA-Mol. Cell Res*. 1863(10), 2362-2378

Parker GA, Baker RR, Smith VGF. 1972. The origin and evolution of gamete dimorphism and the male-female phenomenon. *J. Theor. Biol.* 36(3), 529-553.

Parker GA. 1978. Selection on non-random fusion of gametes during the evolution of anisogamy. *J. Theor. Biol.* 73(1), 1-28.

Partridge L, Hurst LD. 1998. Sex and conflict. *Science* 281(5385), 2003-2008.

Passamonti M, Ghiselli F. 2009. Doubly uniparental inheritance: two mitochondrial genomes, one precious model for organelle DNA inheritance and evolution. *DNA Cell Biol.* 28(2), 79-89.

de Paula WB, Agip ANA, Missirlis F, Ashworth R, Vizcay-Barrena G, Lucas CH, Allen JF. 2013a. Female and male gamete mitochondria are distinct and complementary in transcription, structure, and genome function. *Genome Biol. Evol.* 5(10), 1969-1977.

de Paula WB, Lucas CH, Agip ANA, Vizcay-Barrena G, Allen JF. 2013b. Energy, ageing, fidelity and sex: oocyte mitochondrial DNA as a protected genetic template. *Proc. R. Soc. Lond. B Biol. Sci.* 368(1622), 20120263.

Pett W, Ryan JF, Pang K, Mullikin JC, Martindale MQ, Baxevanis AD, Lavrov DV. 2011. Extreme mitochondrial evolution in the ctenophore *Mnemiopsis leidyi*: insight from mtDNA and the nuclear genome. *Mitochondrial DNA* 22, 130-142.

Pittis AA, Gabaldón T. 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* 531(7592), 101-104.

Poole AM, Gribaldo S. 2014. Eukaryotic origins: how and when was the mitochondrion acquired? *Cold Spring Harb. Perspect. Biol.* 6(12), a015990.

Poole AM, Neumann N. 2011. Reconciling an archaeal origin of eukaryotes with engulfment: a biologically plausible update of the Eocyte hypothesis. *Res. Microbiol.* 162(1), 71-76.

Q

Queller DC. 2000. Relatedness and the fraternal major transitions. *Phil. Trans. R. Soc. B*. 355, 1647-1655.

R

Radzvilavicius AL. 2016a. Evolutionary dynamics of cytoplasmic segregation and fusion: mitochondrial mixing facilitated the evolution of sex at the origin of eukaryotes. *J. Theor. Biol.* 404, 160-168.

Radzvilavicius AL. 2016b. Mitochondrial genome erosion and the evolution of sex. *Bioessays* 38(10), 941-942.

Radzvilavicius AL, Hadjivasiliou Z, Lane N, Pomiankowski A. 2015. Selection for mitochondrial quality drives the evolution of sexes with a dedicated germline. *bioRxiv* doi: http://dx.doi.org/10.1101/026252.

Radzvilavicius AL, Blackstone NW. 2015. Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. *J. R. Soc. Interface* 12(111), 20150584.

Ramesh MA, Malik SB, Logsdon JM. 2005. A phylogenomic inventory of meiotic genes: evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. *Curr. Biol.* 15(2), 185-191.

Rand DM. 2008. Mitigating mutational meltdown in mammalian mitochondria. *PLoS Biol.* 6(2), e35.

Rand DM. 2011. The units of selection of mitochondrial DNA. *Annu. Rev. Ecol. Syst.* 32, 415-448.

Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Mol. Biol. Evol.* 13(6), 735-748.

Randerson JP, Hurst LD. 1999. Small sperm, uniparental inheritance and selfish cytoplasmic elements: a comparison of two models. *J. Evol. Biol.* 12, 1110-1124.

Randerson JP, Hurst LD. 2001. A comparative test of a theory for the evolution of anisogamy. *Proc. R. Soc. Lond. B* 268(1469), 879-884.

Rao V, Nanjundiah V. 2010. J.B.S. Haldane, Ernst Mayr and the beanbag genetics dispute. *J. Hist. Biol.* 44(2), 233-281.

Rivera MC, Jain R, Moore JE, Lake JA. 1998. Genomic evidence for two functionally distinct gene classes. *Proc. Natl. Acad. Sci. USA* 95(11), 6239-6244.

Rossignol R, Faustin B, Rocher C, Malgat M, Mazat JP, Letellier T. 2003. Mitochondrial threshold effects. *Biochem. J.* 370, 751-762.

Roze D, Rousset F, Michalakis Y. 2005. Germline bottlenecks, biparental inheritance and selection on mitochondrial variants: a two-level selection model. *Genetics* 170(3), 1385-1399.

Russell MJ, Barge LM, Bhartia R, Bocanegra D, Bracher PJ, Branscomb E, Kidd R, McGlynn S, Meier DH, Nitschke W, Shibuya T. 2014. The drive to life on wet and icy worlds. *Astrobiology* 14(4) 308-343.

S

Sagan L. 1967. On the origin of mitosing cells. *J. Theoret. Biol.* 14, 225-274.

Samson RY, Bell SD. 2009. Ancient ESCRTs and the evolution of binary fission. *Trends Microbiol.* 17(11), 507-513.

Sato M, Sato K. 2011. Degradation of paternal mitochondria by fertilization-triggered autophagy in *C. elegans* embryos. *Science* 334(6059), 1141-1144.

Sato M, Sato K. 2013. Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. *BBA-Mol. Cell Res.* 1833(8), 1979-1984.

Schable PS, Wise RP. 1998. The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends Plant Sci.* 3, 175-180.

Schurko AM, Logsdon JM. 2008. Using a meiosis detection toolkit to investigate ancient asexual “scandals” and the evolution of sex. *Bioessays* 30(6), 579-589.

Seipel K, Yanze N, Schmid V. 2003. The germ line and somatic stem cell gene Cniwi in the jellyfish *Podocoryne carnea*. *Int. J. Dev. Biol.* 48(1), 1-7.

Servedio MR, Brandvain Y, Dhole S, Fitzpatrick CL, Goldberg EE, Stern CA, Yeh DJ. 2014. Not just a theory - the utility of mathematical models in evolutionary biology. *PLoS Biol.* 12(12), e1002017.

Shearer TL, van Oppen MJH, Romano SL, Worheide G. 2002 Slow mitochondrial DNA sequence evolution in the *Anthozoa* (*Cnidaria*). *Mol. Ecol.* 11, 2475-2487.

Sherengul W, Kondo R, Matsuura ET. 2006. Analysis of paternal transmission of mitochondrial DNA in *Drosophila*. *Genes Genetic Syst.* 81(6), 399-404.

Shitara H, Hayashi JI, Takahama S, Kaneda H, Yonekawa H. 1998. Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. *Genetics* 148(2), 851-857.

Shoffner JM, Lott MT, Lezza AMS, Seibel P, Ballinger SW, Wallace DC.1990. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA Lys mutation. *Cell* 61(6), 931-937.

Shoubridge EA, Wai T. 2007. Mitochondrial DNA and the mammalian oocyte. *Curr. Top. Dev. Biol.* 77, 87-111.

Schwartz M, Vissing J. 2002. Paternal inheritance of mitochondrial DNA. *New Engl. J. Med.* 347(8), 576-580.

Simon JC, Delmotte F, Rispe C, Crease T. 2003. Phylogenetic relationships between parthenogens and their sexual relatives: the possible routes to parthenogenesis in animals. *Biol. J. Linnean Soc.* 79(1), 151-163.

Simpson C. 2011. The evolutionary history of division of labour. *Proc. R. Soc. Lond. B* 279(1726), 116-21.

Sloan DB, Oxelman B, Rautenberg A, Taylor DR. 2009. Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe *Sileneae*. *BMC Evol. Biol.* 9, 260.

Smith DR, Keeling PJ. 2015. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proc. Natl. Acad. Sci. USA* 112(33), 10177-10184.

Sousa FL, Neukirchen S, Allen JF, Lane N, Martin WF. 2016. Lokiarchaeon is hydrogen dependent. *Nature Microbiol.* 1, 16034

Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, van Eijk R, Schleper C, Guy L, Ettema TJ. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521(7551), 173-179.

Speijer D, Lukeš J, Eliáš M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci. USA* 112(29), 8827-8834.

Sperling EA, Frieder CA, Raman AV, Girguis PR, Levin LA, Knoll AH. 2013. Oxygen, ecology, and the Cambrian radiation of animals. *Proc. Natl. Acad. Sci. USA* 110, 13446-13451.

Spradling A, Fuller MT, Braun RE, Yoshida S. 2011. Germline stem cells. *Cold Spring Harb. Perspect. Biol.* 3(11), a002642.

Stams AJ, Plugge CM. 2009. Electron transfer in syntrophic communities of anaerobic bacteria and archaea. *Nature Rev. Microbiol.* 7(8), 568-577.

Stechmann A, Hamblin K, Pérez-Brocal V, Gaston D, Richmond GS, Van der Giezen M, Clark CG, Roger AJ. 2008. Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. *Curr. Biol.* 18(8), 580-585.

Stewart JB, Freyer C, Elson JL, Wredenberg A, Cansu Z, Trifunovic A, Larsson NG. 2008. Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biol.* 6(1), e10.

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G. 1999. Development: ubiquitin tag for sperm mitochondria. *Nature* 402(6760), 371-372.

T

Taylor DR, Zeyl C, Cooke E. 2002. Conflicting levels of selection in the accumulation of mitochondrial defects in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 99, 3690-3694.

Tielens AG, Rotte C, van Hellemond JJ, Martin W. 2002. Mitochondria as we don't know them. *Trends Biochem. Sci.* 27(11), 564-572.

Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Rev. Genet.* 5(2), 123-135.

Tovar J, León-Avila G, Sánchez LB, Sutak R, Tachezy J, Van der Giezen M, Hernández M, Müller M, Lucocq JM. 2003. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature* 426(6963), 172-176.

Trivers RL. 1972. In Campbell B (ed), *Sexual selection and the descent of man,1871-1971* (Aldine-Atherton, Chicago), 136-179.

U

Ursprung H, Schabtach E. 1965. Fertilization in tunicates: loss of the paternal mitochondrion prior to sperm entry. *J. Exp. Zool.* 159(3), 379-383.

Uyenoyama MK. 1988. On the evolution of genetic incompatibility systems. II. Initial increase of strong gametophytic self-incompatibility under partial selfing and half-sib mating. *Am. Nat.* 131(5), 700-722.

V

Van der Giezen M, Tovar J. 2005. Degenerate mitochondria. *EMBO Reports* 6(6), 525-530.

Von Dohlen CD, Kohler S, Alsop ST, McManus WR. 2001. Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature* 412(6845), 433-436.

Vrijenhoek RC. 1998. Animal clones and diversity. *Bioscience* 48(8), 617-628.

W

Wade MJ, McCauley DE. 2005. Paternal leakage sustains the cytoplasmic polymorphism underlying gynodioecy but remains invasible by nuclear restorers. *Am. Nat.* 166(5), 592–602.

Wade MJ, Brandvain Y. 2009. Reversing mother's curse: selection on male mitochondrial fitness effects. *Evolution* 63(4), 1084-1089.

Wai T, Teoli D, Shoubridge EA. 2008. The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nature Genet.* 40(12), 1484-1488.

Wallin IE. 1927. Symbionticism and the origin of species (Williams and Wilkins, Baltimore, MD).

West SA, Fisher RM, Gardner A, Kiers ET. 2015. Major evolutionary transitions in individuality. *Proc. Natl. Acad. Sci. USA* 112(33), 10112-10119.

Weismann A. 1890. Prof. Weismann's theory of heredity. *Nature* 41, 317-323.

Weinreich DM, Rand DM. 2000. Contrasting patterns of nonneutral evolution in proteins encoded in nuclear and mitochondrial genomes. *Genetics* 156(1), 385-399.

Williams TA, Foster PG, Nye TM, Cox CJ, Embley TM. 2012. A congruent phylogenomic signal places eukaryotes within the *Archaea*. *Proc. R. Soc. Lond. B* 279(1749), 4870-4879.

Wolff JN, Nafisinia M, Sutovsky P, Ballard JW. 2013. Paternal transmission of mitochondrial DNA as an integral part of mitochondrial inheritance in metapopulations of *Drosophila simulans*. *Heredity* 110(1), 57-62.

X

Xu J. 2005. The inheritance of organelle genes and genomes: patterns and mechanisms. *Genome* 48(6), 951-958.

Y

Yan Z, Hull CM, Sun S, Heitman J, Xu J. 2007. The mating type-specific homeodomain genes SXI1α and SXI2a coordinately control uniparental mitochondrial inheritance in *Cryptococcus neoformans*. *Curr. Genet.* 51(3), 187-195.

Yutin N, Makarova KS, Mekhedov SL, Wolf YI, Koonin EV. 2008. The deep archaeal roots of eukaryotes. *Mol. Biol. Evol.* 25(8), 1619-1630.

Z

Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström D, Juzokaite L, Vancaester E, Seitz KW, Anantharaman K, Starnawski P, Kjeldsen KU, Stott MB. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541, 353-358.

Zhao X, Li N, Guo W, Hu X, Liu Z, Gong G, Wang A, Feng J, Wu C. 2004. Further evidence for paternal inheritance of mitochondrial DNA in the sheep (*Ovis aries*). *Heredity* 93(4), 399-403.

Zhou Q, Li H, Li H, Nakagawa A, Lin JL, Lee ES, Harry BL, Skeen-Gaar RR, et al. 2016. Mitochondrial endonuclease G mediates breakdown of paternal mitochondria upon fertilization. *Science* 353(6297), 394-399.