Evolutionary dynamics of cytoplasmic segregation and fusion: Mitochondrial mixing facilitated the evolution of sex at the origin of eukaryotes

Arunas L. Radzvilavicius

Department of Genetics, Evolution and Environment & CoMPLEX, University College London, Gower street, WC1E 6BT London, United Kingdom

Corresponding author: ucbprad@ucl.ac.uk

Abstract

Sexual reproduction is a trait shared by all complex life, but the complete account of its origin is missing. Virtually all theoretical work on the evolution of sex has been centered around the benefits of reciprocal recombination among nuclear genes, paying little attention to the evolutionary dynamics of the multi-copy mitochondrial genome. Here we develop a mathematical model to study the evolution of nuclear alleles inducing cell fusion in an ancestral population of clonal proto-eukaryotes. Segregational drift maintains high mitochondrial variance between clonally reproducing hosts, but the effect of segregation is opposed by cytoplasmic mixing which tends to reduce variation between cells in favor of higher heterogeneity within the cell. Despite the reduced long-term population fitness, alleles responsible for sexual cell fusion can spread to fixation. The evolution of sex requires negative epistatic interactions between mitochondrial mutations and is promoted by strong purifying selection, low mutant load and weak mitochondrial-nuclear associations. We argue that similar conditions were maintained during the late stages of eukaryogenesis, facilitating the evolution of sexual cell fusion and meiotic recombination without compromising the stability of the emerging complex cell.

Keywords: mitochondria, cell fusion, eukaryogenesis, negative epistasis

Highlights

- Sex evolved in a cell that already possessed mitochondria
- Mitochondrial mixing drives the evolution of sexual cell fusion under negative epistasis between deleterious mitochondrial mutations
- Evolution of sex requires strong purifying selection and weak mito-nuclear associations

1. Introduction

Sexual reproduction with gamete fusion and reciprocal recombination is among the traits shared by 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 unfavorable allelic combinations under fluctuating selection, or rescues the genome from the mutational meltdown (Otto, 2009). These views, however, are based on the long-term effects of recombination among nuclear genes, and do not explain when or how these traits first arose. Initial selective forces promoting the origin of sex might have been different from the ones maintaining the meiotic recombination in modern eukaryotes.

It is becoming increasingly clear that sex first appeared as a part of the evolutionary transition towards eukaryotic cell, most likely after the endosymbiotic acquisition of mitochondria (Gross and Bhattacharya, 2010; Lane and Martin, 2010; Speijer et al., 2015). The origin of sex would therefore seem to entail far more than recombination, and a full account of the evolution of sexual reproduction has to consider the complex relationship between mitochondrial symbionts and the host. An aspect of early eukaryotic sex that deserves close attention is the whole-cell fusion with the symmetric transmission of organelle (endosymbiont) genomes. What role did the mitochondrial mixing play at the origin of sex?

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 (Buss, 1987; Michod, 1997; Michod and Nedelcu, 2003). Conflict resolution often involves mechanisms reducing genetic variance within groups of lower-level units, thus eliminating the scope for defection and detrimental competition. In contrast, cell fusion allows for cytoplasmic mixing and horizontal spread of selfish genetic elements, facilitating the evolutionary conflict and reducing host fitness (Hastings, 1992, 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). The evolution of sex might have required prior mechanisms suppressing conflicts related to the fast proliferation of selfish mitochondria. While two mating types and uniparental inheritance (UPI) might eliminate the issue of cytoplasmic conflict in modern eukaryotes (Birky, 1995; Hadjivasiliou et al., 2013; Sato and Sato, 2013; Greiner et al., 2015), the mechanism of asymmetric inheritance would not have been present during the early evolution of sex.

Several authors have recently recognized the importance of mitochondrial symbiosis in 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 cell's viability. As many genes coding for key subunits of the electron transport chain are also located in the nuclear genome, this prompts the evolution of compensatory nuclear modifications that could potentially 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 many ways, the hypothesis does not account for one of the hallmark features of mitochondrial genetics—cytoplasmic segregation (Rand, 2008, 2011)—which together with the purifying selection makes for the efficient elimination of mitochondrial mutations.

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, mitochondrial symbionts could have systematically manipulated the host cell's

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 favorable 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 the cell that already possessed mitochondria (Lane and Martin, 2010; Lane, 2014; Garg and Martin, 2015; Speijer et al., 2015), the conditions favoring 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 selection, mutation rate, epistatic interactions, intracellular competition and the properties of early cell cycles. Here we introduce an infinite-population model to investigate the role of cytoplasmic segregation and mixing at the origin of sexual cell fusion. We study the spread of (proto)nuclear alleles inducing cell fusion, considering only the mitochondrial contribution to the cell's fitness. The results of our modeling suggest a set of conditions under which cytoplasmic mixing promotes the emergence of sex in the form of the eukaryotic cell fusion, and supports the view that mitochondria could have represented one of the driving forces behind the origin of sexual life cycles.

2. Mathematical model for cytoplasmic segregation and the evolution of sexual mixing

2.1 Neutral segregational drift with clonal host reproduction

The eukaryotic cell can be conceptualized as a collective of tightly interacting mitochondria within a cytosol which also contains the haploid host genome. Consider an infinite population of cells, containing M mitochondria each (or endosymbionts, in the early stages of eukaryogenesis) and reproducing clonally. Mitochondria are found in one of two possible states, wild-type or mutant. We model the clonal reproduction 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 the probability

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

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 modeled 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.

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

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

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 mutants have either reached fixation within the cell or were replaced by the wild type mitochondria. It follows then that the probability to reach fixation is equal to the initial mutant frequency f_0. The expression is structurally similar to the outcome of the Wright-Fisher process, used to estimate the effect of mitochondrial segregation in several empirical studies (Solignac et al., 1984; Rand and Harrison, 1986).

It can be similarly shown (Appendix A) that neutral segregation increases the homogeneity within the cell. We find that the probability for two lower-level units within the cell to be identical by descent can be expressed as

$$\varphi_n = \frac{\mathrm{Var}(F_n)}{f_0(1-f_0)} = 1 - \left(1 - \frac{1}{2M-1}\right)^n. \quad (3)$$

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.

2.2. Evolution of cytoplasmic mixing

Now consider a full population life cycle with mitochondrial mutation, selection and reproduction (Fig. 1). We assume an ancestral state without sex or cell-cell fusion, where haploid hosts reproduce clonally in a way 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 (3a-3c in Fig. 1) before proceeding to the standard clonal reproduction.

The state of the population at generation t can be represented by a $(M+1)\times 2$ matrix $\mathbf{P}^{(t)}$ with the matrix element $P_{m,j}^{(t)}$ depicting the frequency of cells with m mitochondrial mutants and the nuclear state j ($j = 0, 1$). 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*.

2.2.1 Mutation

Distinct events in the population life cycle can be represented as matrix operations changing the population state $\mathbf{P}^{(t)}$. The state of the population after the mutation step 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 a 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

$$U_{i,j} = \binom{M-j}{i-j}\mu^{i-j}(1-\mu)^{M-i}, i,j \in [0,M]. \quad (4)$$

2.2.2. 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 of individuality. Selection among mitochondria of the same cell is modeled as a random sampling with replacement, with 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 the matrix $\mathbf{W}$ are

$$W_{i,j} = \binom{M}{i}\left[\frac{j(1+\kappa)}{M+j\kappa}\right]^i\left(\frac{M-j}{M+j\kappa}\right)^{M-i}, i,j \in [0,M]. \quad (5)$$

2.2.3. Selection between eukaryotic hosts

Selection on the higher level changes the relative frequencies of genotypes according to the host cell fitness. In matrix notation the population state after selection is

$$\mathbf{P}^{(t,3)} = \frac{(\mathbf{I}\mathbf{w})\mathbf{P}^{(t,2)}}{\mathbf{w}^{\mathrm{T}}\mathbf{P}^{(t,2)}\overline{\mathbf{u}}_2}, \quad (6)$$

where $\mathbf{I}$ is the identity matrix, $\bar{\mathbf{u}}_2$ 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), we assume that the relative fitness of the cell depends only on the number of mitochondrial mutants m, and can be expressed as $\omega(m) = 1 - s(m/M)^{\xi}$. Parameter s here represents the intensity of selection and ξ determines the strength of epistatic interactions between mitochondrial mutations. Empirical studies suggest that in higher eukaryotes $\xi > 1$, leading to the so-called phenotypic threshold effects (Rossignol et al., 2003). The relative effect of each new mutation therefore increases with the overall mutation load.

2.2.4. 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 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

$$\mathbf{P}_{\bullet,0}^{(t,4)} = \mathbf{P}_{\bullet,0}^{(t,3)} \bar{\mathbf{u}}_{M+1}^{\mathrm{T}} \mathbf{P}_{\bullet,0}^{(t,3)} + \mathbf{K}\left(\mathbf{P}_{\bullet,0}^{(t,3)} * \mathbf{P}_{\bullet,1}^{(t,3)}\right) \qquad (7a)$$

$$\mathbf{P}_{\bullet,1}^{(t,4)} = \mathbf{K}\left(\mathbf{P}_{\bullet,1}^{(t,3)} * \mathbf{P}_{\bullet,1}^{(t,3)}\right) + \mathbf{K}\left(\mathbf{P}_{\bullet,0}^{(t,3)} * \mathbf{P}_{\bullet,1}^{(t,3)}\right). \qquad (7b)$$

Asterisk here denotes vector convolution, and $\bar{\mathbf{u}}_{M+1}^{\mathrm{T}}$ is a row vector of $M+1$ ones, so that $\bar{\mathbf{u}}_{M+1}^{T} \mathbf{P}_{\bullet,0}^{(t,3)}$ 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. 1) $K_{i,j} = r(i|M, j/2)$, where $i \in [0, M], j \in [0, 2M]$.

The life cycle ends with a 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. 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. $\bar{\mathbf{u}}_{M+1}^{T} \mathbf{P}_{\bullet,0}^{(0)} = 1$ and $\mathbf{P}_{\bullet,1}^{(0)} = 0$. After the equilibrium is reached at time t_E, the allele *H* is inserted at a small 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 we 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.

3. Mitochondrial variation and the fitness cost of cytoplasmic mixing

3.1. Mitochondrial variation at the segregation–fusion equilibrium

Let us first look into the effect of recurrent cell fusion on the mitochondrial variance between cells generated by cytoplasmic segregation. Starting with a fixed number of mutants per cell m_0 we 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 of our modeling show that the effect of cytoplasmic mixing opposes the constant increase of mitochondrial variance between cells generated by segregational drift (Eq. 2), establishing an intermediate equilibrium (Fig. 2). While drift alone results in diverging cell

lineages with highly clonal intracellular populations of mitochondria, increasing frequency of sexual 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 in the process of selection on the higher level, frequent cell fusion could result in diminished population fitness, which we investigate further.

3.2 Mitochondrial mutation pressure

Here we return to the full population life cycle and analyze 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 an *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. 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. This occurs due to the fact that larger number of segregating units dampens the effect of segregational drift (Eq. 2), reducing the efficacy of selection on the higher level.

3.3. Fast replicating "selfish" mutants

Fitness costs of cytoplasmic mixing can be exacerbated in the presence of so called "selfish" mitochondria—organelles that have gained mutations leading to the faster reproduction rate but at the same time reducing their ability to participate in cooperative interactions. Selection on 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 eukaryotic cell (Schable and Wise, 1988; Taylor et al., 2002; Clark et al., 2012; Bastiaans et al., 2014). This time 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.

Our results confirm that clonal reproduction (low p_H) coupled with purifying selection on the higher level suppresses selfish mitochondrial competition within the cell and maintains high population fitness (Fig. 4). Unable to spread horizontally in the absence of cell fusion, selfish mutants affect only those cell lineages in which rare mutations occur, allowing selection on the higher level to rapidly eliminate the affected individuals. We find, however, a critical frequency p_H, at which the non-cooperative mitochondria start proliferating faster (Fig. 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. 4A-D), weak selection (Fig. 4E-F) and strong epistasis ξ (Fig. 4E-F). With increasing ξ the fitness function becomes flatter at low m, allowing the non-cooperative mitochondria to reach high per-cell frequencies before they get 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. 4A-F).

Interestingly, the equilibrium frequency of selfish mitochondrial mutants is not always a monotonic function of their relative replicative advantage κ (Fig. 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 selection to efficiently suppress their further spread.

4. Invasion of *H* mutants

In this section we 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 1.).

4.1. Epistasis between mitochondrial mutations

We found that the fusion-inducing allele *H* is able to invade, spread to an equilibrium frequency of $p_H < 1$ or reach fixation ($p_H = 1$) (Fig. 5A). The invasion occurs despite the curtailed long-term population fitness due to the lower variance in the number of mitochondrial mutants among invaders (Fig. 5B). The necessary condition for successful invasion is $\xi > 1$, i.e. the negative epistasis between deleterious mitochondrial mutations. The detrimental effect of every new mutation has to increase with the total number of mutations.

Cytoplasmic mixing increases the frequency of intermediate cytoplasmic states, reducing the frequency of cells with extreme mutant numbers (both high and low, Fig. 5C). This reduced variance has a long-term fitness disadvantage due to the weakened response to selection (Fig. 5B). However, with negative epistasis ($\xi > 1$), the intermediate cytotypes have a higher fitness than expected with linear interactions, which gives the invading allele *H* a short-term advantage (Fig. 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. 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.

4.2. Further conditions favoring the spread of H

With recurrent mitochondrial mutation, the allele *H* spreads to high frequencies and fixes more readily under low mutation rates (Fig. 6A-B). Cytoplasmic fusion has a stronger evolutionary advantage with small mitochondrial population sizes, as segregational drift is more efficient at generating mitochondrial variance with small M (Eq. 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 fusion allele *H* (Fig. 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 fixation is also facilitated by strong purifying selection on the higher level s (Fig. 6E-F).

4.3. Alternating life cycles and mito-nuclear linkage

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 the segregational drift, η is likely to affect the evolutionary success of the cell fusion allele *H*. Indeed our results show that increasing 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. 6G-H). As little as 4-8 clonal cell divisions could be enough to prevent the invasion of the cell fusion allele *H* under all reasonable conditions.

The principal reason allowing the allele *H* to invade lies in the way reduced mitochondrial variance affects the 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 weak. This is easiest to achieve 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. $\eta > 1$. 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 mixing, reducing the strength of mitochondrial-nuclear associations.

5. Conclusions and discussion

The origin of sex is among the most perplexing problems in evolutionary biology (Otto, 2009). A new wave of studies have recently recognized that sex must have originated as a part of the evolutionary transition to eukaryotes, and so its evolution is intrinsically tied to mitochondria (Blackstone and Green, 1999; Lane and Martin, 2010; Lane, 2011; Speijer et al., 2015; Havird et al., 2015; Radzvilavicius and Blackstone, 2015). In the present paper we studied the evolution of sex from the perspective of cell fusion, mitochondrial segregation and mixing. Based on previous work we assumed that sex appeared in the cell that already had mitochondria, but not mating types or mechanisms constraining the cytoplasmic inheritance. We showed, that sex evolves under negative epistatic interactions among deleterious mitochondrial mutations and with weak mito-nuclear associations, owing to the short-term advantage of mitochondrial mixing.

Low mutation rates allow cell fusion to be established more readily, but the model does not require any unreasonable assumptions of the mutation rate. In fact, sexual cell fusion spreads to fixation even under mutation rates high enough to drive the evolution of the uniparental inheritance in modern eukaryotes (Radzvilavicius et al., manuscript in preparation). The reason is that strict UPI requires the prior existence of two mating types and strong linkage between UPI and mating-type loci (Hadjivasiliou et al., 2013). This creates strong associations between mitochondrial and nuclear genes in females, in which case the long-term advantage of high mitochondrial variance dominates. Interestingly, owing to the weaker mito-nuclear associations, male-specific nuclear alleles would in fact promote mitochondrial mixing, accounting for the frequent observations of paternal leakage, heteroplasmy and doubly uniparental inheritance (Radzvilavicius et al., manuscript in preparation).

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. 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. Indeed, empirical data reveals substantial purifying selection acting on mitochondrial populations in modern animals (Elson et al., 2004; Stewart et al., 2006; Castellana et al., 2011). Similarly, as the evolutionary transition progressed, selfish mitochondrial competition also had to become suppressed (reduced κ), through several proto-eukaryotic mechanisms of conflict mediation, e.g. honest signaling, membrane uncoupling and reduced mitochondrial genomes (Radzvilavicius and Blackstone, 2015).

Extrapolating from the fitness interactions in modern eukaryotes with mitochondrial threshold effects, it is not unreasonable to assume negative epistatic interactions ($\xi > 1$) between detrimental mutations in multi-copy mitochondrial genomes late in eukaryogenesis. Indeed, with multiple mitochondria per cell and several copies of mtDNA per organelle, a critical amount of deleterious mutations has to accumulate before cellular respiration is significantly impaired (reviewed in Mazat et al., 2001; Rossignol et al., 2003). Mitochondrial endosymbiosis therefore created a unique genetic system with strong synergistic interactions, driving the evolution of sexual cell fusion. This is in stark contrast to the deleterious mutations in the nucleus (or bacterial chromosomes), where negative epistatic interactions are relatively uncommon, disfavoring the so-called mutational-deterministic hypothesis for the maintenance of sex (Kouyos et al., 2007).

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 the meiotic, reciprocal recombination had to evolve in the routine presence of cell fusion events, and the barrier separating sex from prokaryotic recombination might have never been crossed without mitochondria. It is more likely, however, that multiple mechanisms promoting cell fusion acted at the same time, with mitochondrial selection pressure contributing to the ease at which sexual reproduction with cytoplasmic fusion and reciprocal recombination came into the widespread existence.

Finally, we have to consider the possibility that routine cell fusion might have preceded the acquisition of mitochondrial symbionts in the host lineage. Indeed, several species of archaea seem to be capable of cell fusion, facilitating the horizontal gene transfer (Noar et al., 2012, 2013), even though the evidence is indirect. The ancestral ability to fuse would have facilitated the rapid spread of proto-mitochondrial endosymbionts throughout the population. Nevertheless, initially strong selection on the lower level, competition and unmediated conflicts would have hindered the further progress of the evolutionary transition, due to low variation among higher-level units.

Acknowledgements

Andrew Pomiankowski, Nick Lane, Neil W. Blackstone are thanked for their contributions to this work. A. Radzvilavicius is supported by a UCL CoMPLEX PhD award from the Engineering and Physical Sciences Research Council (1206315).

Appendix A

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 equals the initial number of mutants within the cell, $\mathrm{E}(X_n) = x_0$. Variance in the number of mutants can be expressed as

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

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

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

we can further write

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

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

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

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

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

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

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

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 $h_0 = 1$. We then easily find that

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

and

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

446 Comparing this result to the expression for the variance after n generations (Eq. 2), we
447 notice that clonality within the cell is just a normalized mitochondrial variance between host
448 cells,

450 $$f_n = \frac{\mathrm{Var}(p_n)}{p_0\,(1-p_0)}.$$

449

References

1. Bastiaans E., Aanen D.K., Debets A.J.M., Hoekstra R.F., Lestrade B., Maas M.F.P.M., 2014. Regular bottlenecks and restrictions to somatic fusion prevent the accumulation of mitochondrial defects in Neurospora. Phil. Trans. R. Soc. B 369, 20130448.
2. Birky C.W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc. Natl. Acad. Sci. USA 92(25), 11331–11338.
3. Blackstone N.W., Green D.R., 1999. The evolution of a mechanism of cell suicide. Bioessays 21(1), 84–88.
4. Burger G., Plante I., Lonergan K.M., Gray M.W., 1995. The Mitochondrial DNA of the Amoeboid Protozoon, Acanthamoeba castellanii: Complete Sequence, Gene Content and Genome Organization. J. Mol. Biol. 245(5), 522–37.
5. Burger G., Gray M.W., Forget L., Lang B.F., 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol. Evol. 5(2), 418–438.
6. Buss L., 1987. The evolution of individuality. Princeton University Press, Princeton, NJ.
7. 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.
8. Clark K.A., Howe D.K., Gafner K., Kusuma D., Ping S., Estes S., Denver D.R.,2012. Selfish little circles: transmission bias and evolution of large deletion-bearing mitochondrial DNA in *Caenorhabditis briggsae* nematodes. PLoS ONE 7(7), e41433.
9. Dacks J., Roger A.J., 1999. The first sexual lineage and the relevance of facultative sex. J. Mol. Evol. 48(6), 779–783.
10. Elson J.L., Turnbull D.M., 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.
11. Embley T.M., Martin W., 2006. Eukaryotic evolution, changes and challenges. Nature 440, 623–630.
12. Ewens W.J., 2004. Mathematical Population Genetics, I. Theoretical Introduction. Second edition. Springer, New York.
13. Goodenough U., Lin H., Lee J.H., 2007. Sex determination in *Chlamydomonas*. Semin. Cell Dev. Biol. 18(3), 350–361.
14. Goodenough U., Heitman J., 2014. Origins of eukaryotic sexual reproduction. Cold Spring Harb. Perspect. Biol. 6, a016154.
15. Greiner S., Sobanski J., Bock R., 2015. Why are most organelle genomes transmitted maternally? BioEssays 37(1), 80–94.
16. Gross J., Bhattacharya, D., 2010. Uniting sex and eukaryote origins in an emerging oxygenic world. Biol. Direct 5, 53.
17. Hadjivasiliou Z, Lane N, Seymour R.M., Pomiankowski A, 2013. Dynamics of mitochondrial inheritance in the evolution of binary mating types and two sexes. Proc. R. Soc. B 280(1769), 20131920.
18. Hastings I.M., 1992. Population genetic aspects of deleterious cytoplasmic genomes and their effect on the evolution of sexual reproduction. Genet. Res. 59, 215–225.
19. Havird J.C., Hall M.D., and Dowling D.K., 2015. The evolution of sex: a new hypothesis based on mitochondrial mutational erosion. Bioessays 37, 951–958.
20. Hörandl E., Hadacek F., 2013. The oxidative damage initiation hypothesis for meiosis. Plant. Reprod. 26(4), 351–367.
21. Huang D., Meier R., Todd P. A., Chou L. M., 2008. Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. J. Mol. Evol., 66(2), 167–174.

22. 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, 12944–8.
23. Kouyos R. D., Silander O. K., Bonhoeffer S., 2007. Epistasis between deleterious mutations and the evolution of recombination. Trends Ecol. Evol. 22(6), 308–315.
24. Kuijper B., Lane N., Pomiankowski A., 2015. Can paternal leakage maintain sexually antagonistic polymorphism in the cytoplasm? J. Evol. Biol. 28, 468–480.
25. Lane N., Martin W., 2010. The energetics of genome complexity. Nature 467, 929–934.
26. Lane N., 2011. Energetics and genetics across the prokaryote–eukaryote divide. Biology Direct 6, 35.
27. Lane N., 2014. Bioenergetic constraints on the evolution of complex life. Cold Spring Harb. Perspect. Biol. 6(5), a015982.
28. Lynch M., Koskella B., Schaack S., 2006. Mutation pressure and the evolution of organelle genomic architecture. Science 311, 1727–1730.
29. Lynch M., Sung W., Morris K., Coffey N., Landry C.R., Dopman E.B., Dickinson W.J., Okamoto K., Kulkarni S., Hartl D.L., Thomas W.K., 2008. A genome-wide view of the spectrum of spontaneous mutations in yeast. Proc. Natl. Acad. Sci. USA 105, 9272–9277.
30. Marais G.A., Calteau A., Tenaillon O., 2008. Mutation rate and genome reduction in endosymbiotic and free-living bacteria. Genetica 134, 205–210.
31. Martin W., Muller M., 1998. The hydrogen hypothesis for the first eukaryote. Nature 392(6671), 37–41.
32. Martin W.F., Garg S., Zimorski V., 2015. Endosymbiotic theories for eukaryote origin. Phil. Trans. R. Soc. B 370, 20140330.
33. Maynard Smith J., Szathmáry E., 1995. The Major Transitions in Evolution. Oxford University Press, Oxford, UK.
34. Mazat J.P., 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. Biochim. Biophys. Acta. 1504(1), 20–30.
35. Michod R.E., 1997. Cooperation and conflict in the evolution of individuality. I. Multilevel selection of the organism. Am. Nat. 149, 607–645.
36. Michod R.E., Nedelcu A.M., 2003. On the reorganization of fitness during evolutionary transitions in individuality. Integr. Comp. Biol. 43, 64–73.
37. Naor A., Lapierre P., Mevarech M., Papke R.T., Gophna U., 2012. Low species barriers in halophilic archaea and the formation of recombinant hybrids. Curr. Biol. 22(15), 1444–1448.
38. Naor A., Gophna U., 2013. Cell fusion and hybrids in Archaea: prospects for genome shuffling and accelerated strain development for biotechnology. Bioengineered 4(3), 126–129.
39. Nelson-Sathi S., Sousa F.L., Roettger M., Lozada-Chávez N., Thiergart T., Janssen A., Bryant D., Landan G., Schönheit P., Siebers B., McInerney J.O., Martin W.F., 2015. Origins of major archaeal clades correspond to gene acquisitions from bacteria. Nature 517(7532), 77–80.
40. Otto S., 2009.. Am. Nat. 174, S1–S14.
41. Palmer J.D., Herbon L.A., 1988. Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. J. Mol. Evol. 28, 87–97.
42. Radzvilavicius A.L., Blackstone N.W., 2015. Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. J. R. Soc. Interface 12: 20150584.
43. Rand D.M., 2008. Mitigating Mutational Meltdown in Mammalian Mitochondria. PLoS Biol. 6(2), e35.
44. Rand D.M., 2011. The units of selection of mitochondrial DNA. Annu. Rev. Ecol. Syst. 32, 415–448.

45. Rand D.M., Harrison R.G., 1986. Mitochondrial DNA Transmission Genetics in Crickets. Genetics 114(3), 955–970.
46. Randerson J. P., Hurst L. D., 1999. Small sperm, uniparental inheritance and selfish cytoplasmic elements: a comparison of two models. Journal of Evolutionary Biology 12, 1110–1124.
47. Ramesh M.A., Malik S.B., Logsdon J.M. Jr., 2005. A phylogenomic inventory of meiotic genes. Curr. Biol. 15(2), 185–191.
48. Rossignol R., Faustin B., Rocher C., Malgat M., Mazat J.P., Letellier T., 2003. Mitochondrial threshold effects. Biochem. J. 370, 751–762.
49. 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.
50. Schable P.S., Wise R.P., 1998. The molecular basis of cytoplasmic male sterility and fertility restoration. Trends Plant Sci. 3, 175–180.
51. Shearer, T.L., Van Oppen, M.J.H., Romano, S.L., Wörheide, G., 2002. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). Molecular Ecology, *11*(12), 2475–2487.
52. Sloan D.B., Oxelman B., Rautenberg A., Taylor D.R., 2009. Phylogenetic analysis of mitochondrial substitution rate variation in the angiosperm tribe Sileneae. BMC Evol. Biol. 9, 260.
53. Smith D.R., Keeling P.J., 2015. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proc. Natl. Acad. Sci. USA 112(33), 10177–10184.
54. Solignac, M., Génermont, J., Monnerot, M., Mounolou, J. C., 1984. Genetics of mitochondria in Drosophila: mtDNA inheritance in heteroplasmic strains of D. mauritiana. Mol. Gen. Genet. 197(2), 183–188.
55. 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.
56. Stewart J.B., Freyer C., Elson J.L., Wredenberg A., Cansu Z., Trifunovic A., Larsson, N.G., 2008. Strong purifying selection in transmission of mammalian mitochondrial DNA. PLoS Biol. 6(1), e10.
57. Taylor D.R., 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.

Figure legends

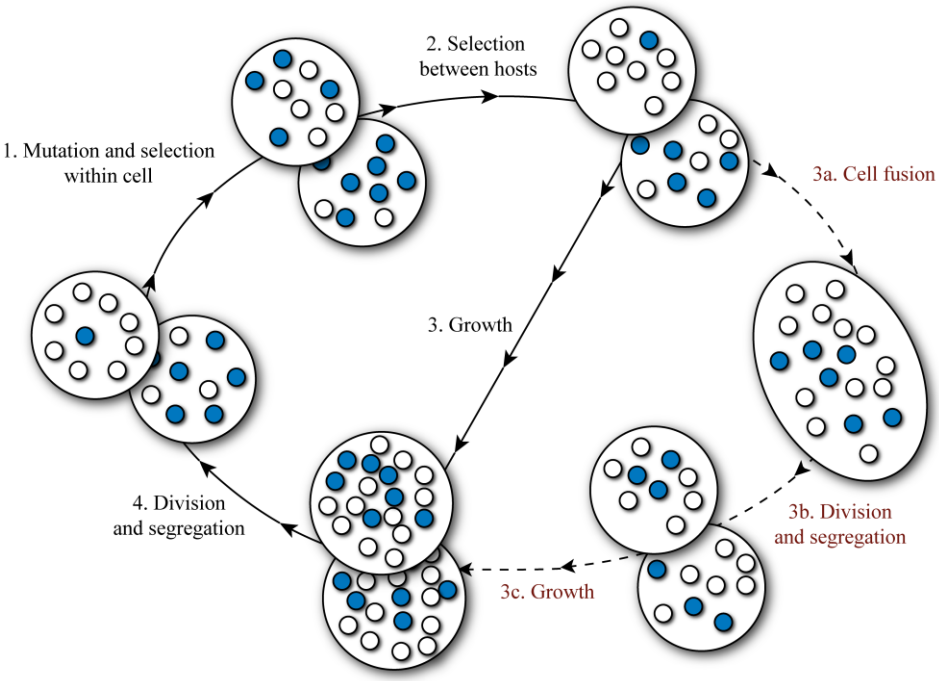


Figure 1. Schematic of the population life cycle. Shaded circles represent deleterious mitochondrial mutants within the host cell, cooperative organelles are left blank. Steps 1–4 (solid arrows) represent the life cycle of clonally reproducing individuals. Steps 3a–3c (dashed arrows) occur only if one of the host cells meeting at random is a carrier of the cell fusion allele *H*.

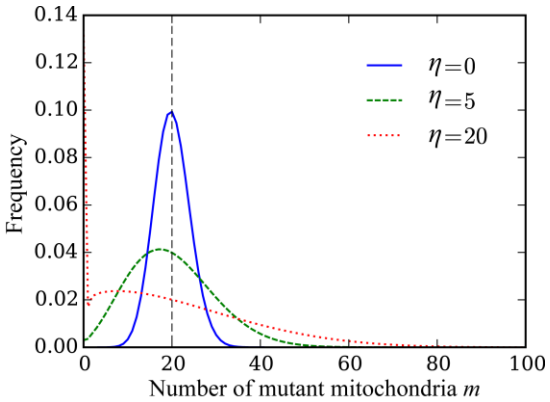
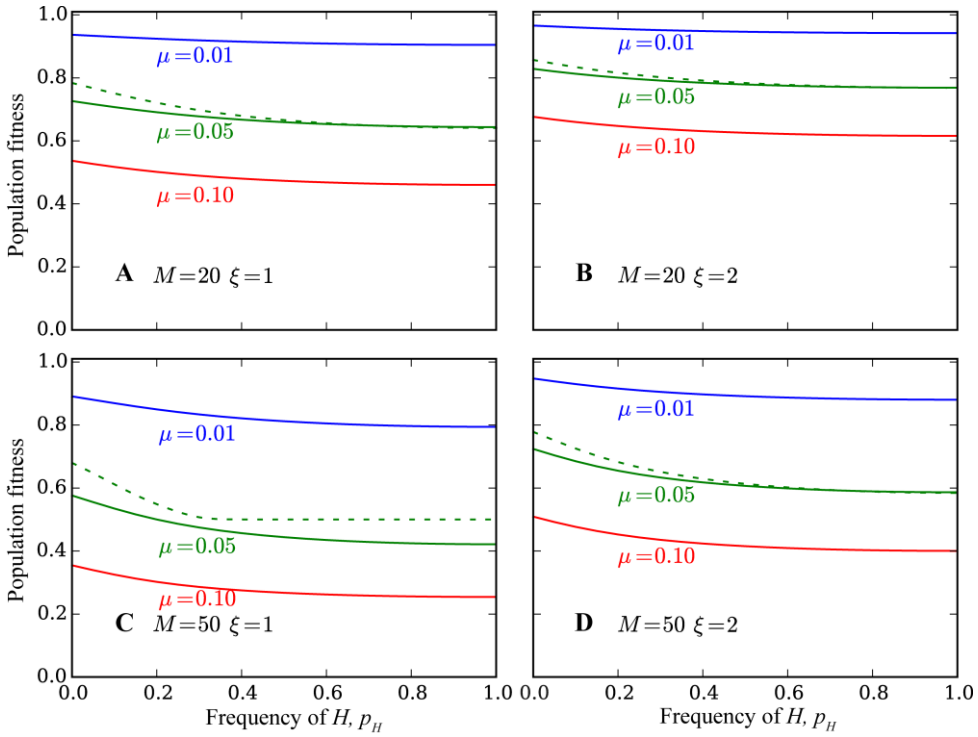


Figure 2. Cytoplasmic fusion opposes the effect of mitochondrial segregation with neutral mutations. Variance in the number of mutant mitochondria per cell increases during the multiple rounds of clonal reproduction, but is reduced by cell fusion, resulting in an equilibrium when the two modes of reproduction alternate in time. η is the number of consecutive clonal generations without cytoplasmic mixing. The initial number of mutants per cell is $m_0 = 20$ (dashed vertical line).



606

607 **Figure 3.** Frequent cytoplasmic mixing reduces the mean population fitness under
608 mitochondrial mutation pressure. *H* is the cell fusion allele. M is the total number of
609 mitochondria per cell, μ is the mitochondrial mutation rate and ξ is the strength of epistatic
610 interactions. Selection strength is set to $s = 1$, except for $\mu = 0.05$, where dotted lines show
611 the effect of weaker selection with $s = 0.5$.

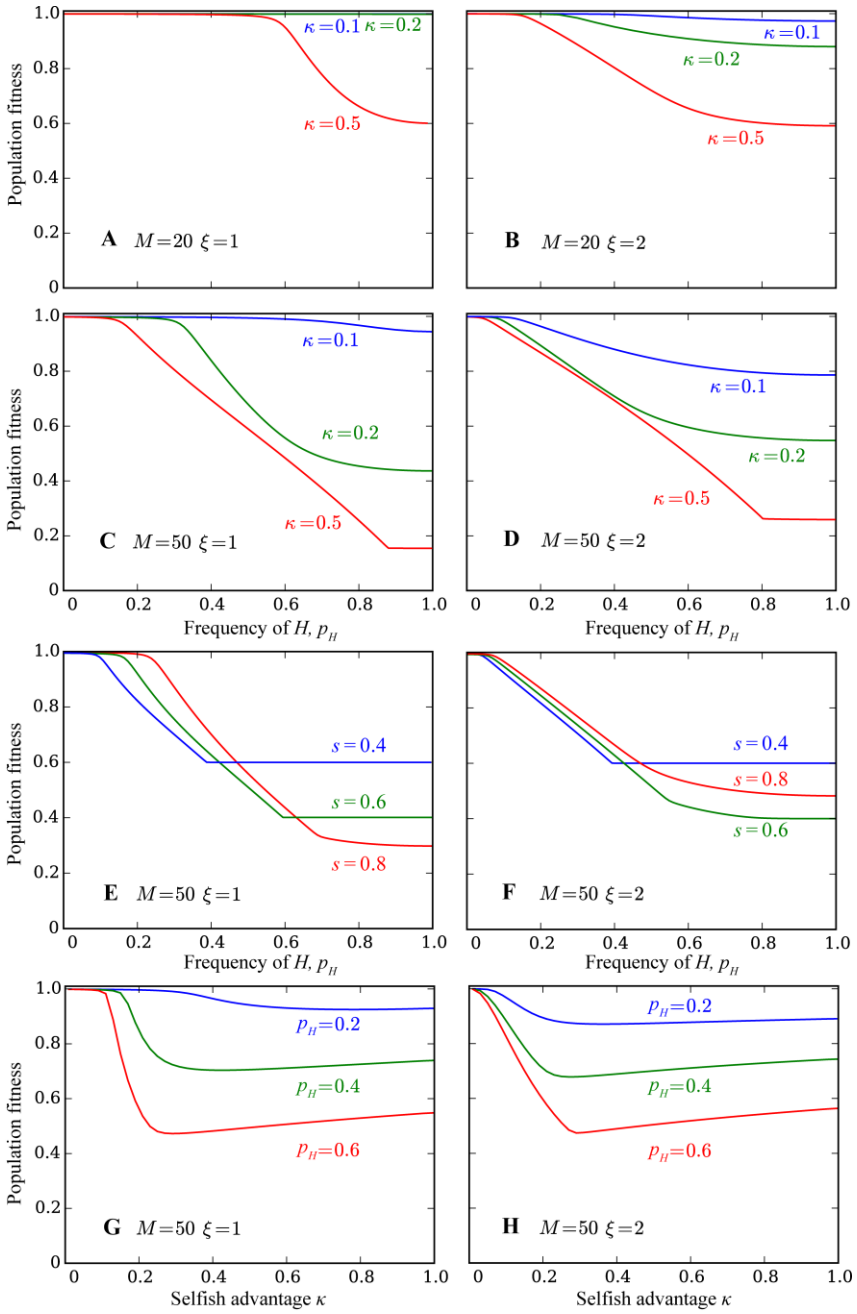


Figure 4. Fitness costs of cytoplasmic mixing can be exacerbated in the presence of selfish mitochondrial mutants. Cooperative interactions within groups of mitochondria break down easier in larger groups (higher M) with strong epistasis (A-D) and weak selection on the higher level s (E-F). Excessive 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. *H* is the cell fusion allele.

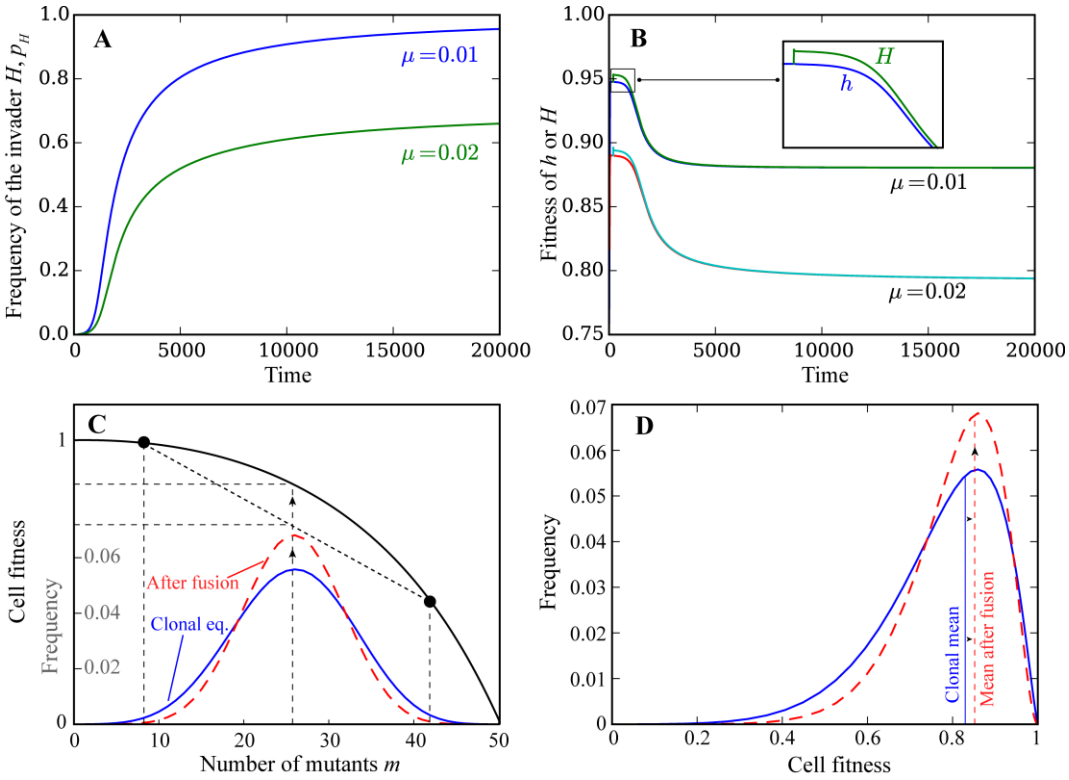


Figure 5. Invasion of the cell fusion allele *H* into an ancestrally clonal population. A nuclear allele inducing cytoplasmic fusion invades and can reach fixation (A), but its spread results in a decline of the long-term population fitness (B). A singular cell fusion-division event increases the frequency of intermediate cytotypes, and decreases 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 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 $M = 50$, $\kappa = 0$.

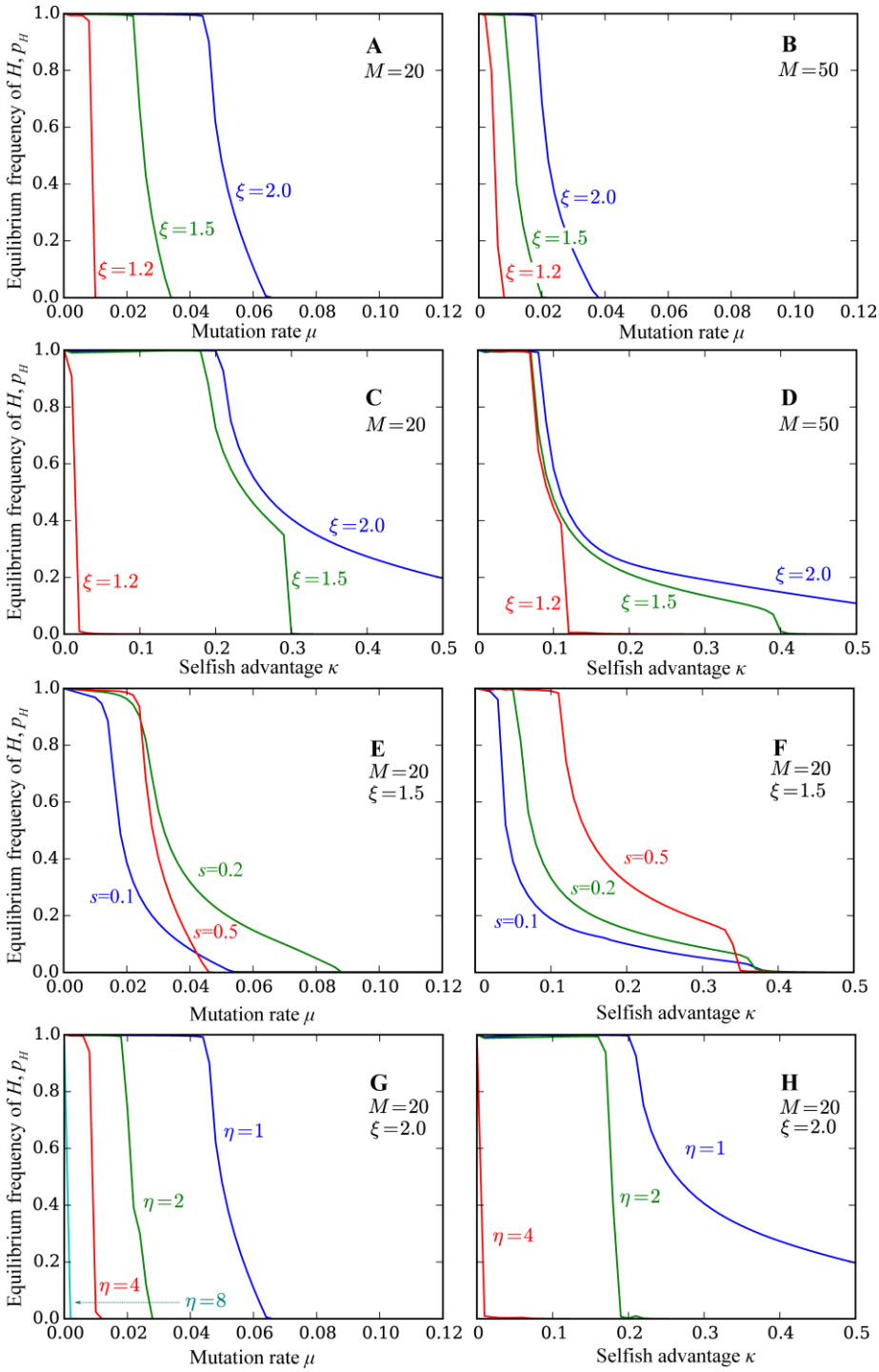


Figure 6. Conditions favoring the evolution of sexual cell fusion. Reduced mutation rates, small mitochondrial populations and negative epistatic interactions all promote the evolution of sexual cell fusion under mitochondrial mutation pressure (A-B). In the presence of selfish mitochondrial mutants, competition on the lower level has to be suppressed before cell fusion can be established (C-D). Cellular fusion evolves easier under strong purifying selection on the higher level (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.