bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Selection for mitochondrial quality drives evolution of the germline

Arunas L Radzvilavicius[1,2], Zena Hadjivasiliou[1,2], Nick Lane[1,2]*, Andrew Pomiankowski[1,2]

[1]Centre for Mathematics and Physics in the Life Sciences and Experimental Biology, University College London, Gower Street, London WC1E 6BT, United Kingdom

[2]Research Department of Genetics, Evolution and Environment, University College London, Gower Street, London WC1E 6BT, United Kingdom

*Correspondence to: nick.lane@ucl.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Summary

The origin of the germline-soma distinction is a fundamental unsolved question. Plants and basal metazoans do not have a germline but generate gametes from somatic tissues (somatic gametogenesis), whereas most bilaterians sequester a germline. We develop an evolutionary model which shows that selection for mitochondrial quality drives germline evolution. In organisms with low mitochondrial mutation rates, segregation of mutations over multiple cell divisions generates variation, allowing selection to optimize gamete quality through somatic gametogenesis. Higher mutation rates promote early germline sequestration. Oogamy reduces mitochondrial segregation in early development, improving adult fitness by restricting variation between tissues, but also limiting variation between early-sequestered oocytes, undermining gamete quality. Oocyte variation is restored through proliferation and random culling (atresia) of precursor cells. We predict a novel pathway from basal metazoans lacking a germline to active bilaterians with early sequestration of large oocytes subject to atresia, allowing the emergence of complex developmental processes.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

In distinguishing between the germline and soma, Weismann argued that the division of labour enabled the specialization of cells in somatic tissues, permitting greater organismal complexity[1]. In contrast, germline cells alone retain the capacity to provide genetic information for future generations, and never form somatic cells[2]. Without the specialization enabled by the germline, complex multicellular animals with post-mitotic tissues such as brain might be impossible. But the division of labour cannot account for the origin of the germline, as all plants and many animals (including tunicates, flatworms and Cnidaria[3]) have differentiated tissues but do not sequester a germline, instead generating gametes from pluripotent stem cells in somatic tissues (somatic gametogenesis).

The best known hypothesis for the evolution of the germline relates to selfish competition between the cells of an individual: by resolving evolutionary conflict, the strict distinction between germline and soma stabilises multicellular cooperation[4-6]. According to this theory, plants did not evolve a germline because their rigid cell walls restrict cell movement, limiting the systemic effects of any parasitic cell lines[4]. In contrast, animal cells lack a rigid wall, making them more vulnerable to parasitic cell lines that undermine organismal function[4]. Sequestering a germline theoretically limits this competition. However, because cells in multicellular organisms normally derive from a single cell (unitary development) new selfish mutations must arise within a single generation; if these mutants are inherited then all cells in the offspring will carry the selfish mutation, so there are no longer any non-selfish cells to exploit[7,8]. Conflict can therefore only emerge when there is a high mutation rate and a large number of cell divisions, which is true of most metazoans[7]. But that does not explain why many metazoan groups do not sequester a germline, or why some phyla (such as Ectoprocta and Entoprocta secondarily returned to somatic gametogenesis[9].

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A second line of thinking dating back to Weismann emphasises the protected environment of a sequestered germline[1,10]. By restricting mutations in gametes, the germline enhances germ-cell quality, contrasting with the 'disposability' of the soma[11]. Avoiding the accumulation of mutations in nuclear genes associated with greater metabolic work can theoretically favour differentiation of a germline[12]. But testes have a high metabolic rate and sperm are produced continuously through life, with 30 cell divisions by puberty and 400 divisions at the age of 30 in humans, contributing a large number of new mutations to offspring[13]. From this point of view, the male germline is equivalent to tissues such as the bone marrow, merely specialising in the mass production of cells of a particular type.

Focusing on the origin of germline sequestration therefore shifts the problem specifically to female gametogenesis. Female gametes are sequestered early in development in a transcriptionally repressed state, with meiosis arrested in prophase I and mitochondria in a state of functional quiescence[14]. Strikingly, female gametes (oocytes) are usually large cells packed with mitochondria, with as many as 10^6 copies of mitochondrial DNA in mammalian oocytes[15]. Mitochondrial DNA is usually inherited uniparentally: male mitochondria are either excluded from the zygote or destroyed on entry to the oocyte[16]. These traits point to mitochondrial function as being central to germline evolution. The possibility that selection for mitochondrial quality could have driven the evolution of the germline has been raised before[14,17,18], but never formally addressed. We show that selection for mitochondrial quality can explain many long-puzzling features of germline evolution.

Model of germline evolution

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We consider the evolution of the germline in a model multicellular organism with variable tissue-level differentiation. Gametes are formed either from adult somatic tissues or early in embryogenesis (**Fig. 1a & b**). The quality of gametes and fitness of the organism depends on the number of mitochondrial mutations. These build up due to copying errors in mitochondrial genes at a rate μ_S per cell division (**Fig. 1c**). Early sequestration of a germline restricts the number of cell divisions to gamete production, and therefore the number of copying errors compared with gametes differentiated from adult tissues. Mutations can also result from 'background' damage, μ_B, per unit time (caused by oxidative damage or UV radiation; **Fig. 1d**), as cells can accumulate mitochondrial mutations even when not actively dividing[19-22]. So background mutations affect gametes whether they are derived from an early germline or later from somatic cells.

Unlike nuclear genes, which are clonally transmitted through mitosis, the mitochondrial population doubles and segregates at random to daughter cells at each cell division. Some daughter cells receive more, others fewer, mitochondrial mutations (**Fig. 1e**). With somatic gametogenesis (i.e. late differentiation of germ cells), segregational variation has the capacity to generate gametes that carry few mutants or are even mutation free. The more rounds of cell division, the greater the degree of segregation, the higher the variance in mutation load, and the larger the number of gametes carrying very few or no mitochondrial mutations (**Fig. 2a**). The effect of segregation is dampened with larger numbers of mitochondria per cell, as this diminishes drift[15] (**Fig. 2b**). At higher copying-error rates (high μ_S), segregation still generates increased variance between daughter cells, but the proportion of gametes with few or no mutations falls to nearly zero, making segregation and selection less effective at maintaining mitochondrial quality (**Fig. 2c**). Including selection at the level of mitochondria equates to a change in the mutation rate. For example, removing deleterious mitochondria by

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mitophagy effectively reduces the mutation rate, whereas faster replication of selfish or impaired mitochondria increases the mutation rate. For simplicity we therefore ignore selection at the level of mitochondria. Likewise, although the severity of mitochondrial mutations is variable, we fix each to have a similar effect for ease of analysis.

Overall, the resulting mitochondrial population in somatic cells is subject to a mutational load, which impacts on cell fitness

$$\omega(m)=1 - (m/M)^2,$$

where m is the number of mutant mitochondria and M is the total size of the mitochondrial population in each cell. This fitness function is concave, and assumes that a large number of mutants must accumulate before cell function is significantly undermined (**Fig. 1f**). This is the case for mitochondrial diseases, where mutant load can reach 60–80% before symptoms are detectable[23]. Other possible fitness functions (e.g. linear or convex) do not affect the general outcome of our modelling but produce slightly different thresholds for germline evolution. We therefore concentrate on a concave fitness function as this is the most empirically justifiable.

Another important dimension of fitness in multicellular organisms is the number and quality of tissues, and their mutual interdependence. We take the fitness of a somatic tissue to be the mean fitness of its constituent cells, each of which reflects the accumulation and segregation of mutations as described above. We assume strong negative epistasis between tissues, so that adult fitness is determined by the quality of the worst tissue. If by chance a disproportionate number of mitochondrial mutations segregate into precursors of one tissue, the fitness of the adult is severely impaired; conversely if mutations are distributed equally into the precursor cells of all tissues, organismal fitness is improved, as each tissue would have equivalent

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

quality. This corresponds to bilaterians in which tissue-precursor cells are defined early in embryonic development, and organ function is mutually interdependent[4]. Failure of one organ (e.g. heart, liver) causes death or debilitation of the organism. In contrast, ancestral metazoans such as sponges have different cell types, but do not differentiate distinct tissue precursors early in development[24], so there is unlikely to be significant tissue-specific epistasis effects on adult fitness. Plant tissues fall between sponges and bilaterians. Nascent tissue organization in seeds is equivalent to the embryonic organization of metazoans[25], but the repetitive modularity of adult tissues means that organ failure due to accumulation of mitochondrial mutations (e.g. of a flower, leaf) has a minimal effect on the fitness of the organism. In the model, plants are regarded as having fewer mutually dependent tissues compared with bilaterians, as there is less negative epistasis between the component parts.

Adult fitness thus depends on the number of mitochondrial mutations inherited by the zygote, the number of new mutations arising during development, how these mutations segregate into cells and tissues, and the degree of tissue-level epistasis. As we will show, this combination determines whether it is better to sequester germ cells early in development in a germline, or derive them from adult somatic cells.

Germline evolution depends on mitochondrial mutation rate

The main benefit of early sequestration of germ cells is to reduce the number of cell divisions before gamete production, so reducing the net input of copying errors (μ_S) in gametes. An early germline raises the mean fitness of offspring in the next generation, but comes at the cost of reduced segregational variance. When gametes are derived later in development, there is more chance for segregation to generate gametes with lower numbers of mitochondrial mutations, which facilitates selection in the offspring, improving fitness over generations

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(**Fig. 2**). The tension between these two forces determines the advantage (or disadvantage) of a germline.

When mitochondrial mutation input through copying errors is low, the benefit of increased variance between gametes tends to outweigh the benefit of sequestering an early germline (**Fig. 3a**). In other words, at low μ_S the benefit of curtailing cell division in an early-sequestered germline is less than the benefit of increased variance between gametes (forming some gametes with few or no mutations) generated by somatic gametogenesis (**Fig. 2**). Note that if copying errors were the only form of mutation, then selection would always favour an early germline (**Fig 3a** bottom left corner). However, oocytes sequestered early in development continue to accumulate mutations as a result of background damage (μ_B)[19]. These mutations can only be segregated out through further rounds of cell division, which favours late somatic gametogenesis. Germline sequestration is therefore unlikely to evolve in organisms with low μ_S but relatively high μ_B (**Fig 3a**, bottom right corner), such as those that have long lifespan or live in damaging environments (e.g. high UV exposure). These conditions are true for basal metazoans and plants. Unlike most animal mitochondrial genomes, basal metazoans including sponges, corals and placozoans all have very low copying error rates[26-28] (low μ_S) and long lives (high μ_B) which readily explains why these major phyla lack a germline. Likewise, most plants have low μ_S (10-20 fold lower than in their nuclear genomes[29], and 50-100 times lower than typical animal mitochondrial genomes[30]) while being exposed to high UV (high μ_B) in their phototrophic niche.

Importantly, increasing the number of tissues (**Fig. 3b**) or mitochondria (**Extended Data Fig. 1**), which both correspond to greater complexity, does not enhance the likelihood of germline evolution. The reason again relates to segregation. In organisms with a single tissue, mean

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

adult fitness closely follows the mutation load in the zygote, with some variation induced by mutation accumulation (**Fig. 4a**). In contrast, with multiple tissues, mutations segregate differentially into distinct tissues. As the fitness of an organism depends on the mutational state of the worst tissue, this magnifies the fitness disadvantage of an inherited mutation load. So there is a substantial increase in variance and reduction in mean adult fitness relative to the inherited zygote fitness (**Fig. 4b**). This gives an added advantage to the late production of gametes (**Fig. 3b**), as the extra segregational drift increases the variance in mutation load between gametes. Higher variance between individuals with somatic gametogenesis leads to higher average adult fitness, compared to competitors with early germline sequestration. Increasing mitochondrial number (larger M) also reduces the variance in mutation load between tissue precursor cells, which likewise improves adult fitness (**Fig. 4c**). This again militates against the production of early germline sequestration with less segregational variation produced by fewer cell divisions (**Extended Data Fig. 1**). Overall, then, multiple tissues and high numbers of mitochondria do not drive the evolution of early germline sequestration, but rather reinforce the stability of somatic gametogenesis in plants and many metazoans.

The one factor that reliably drives sequestration of an early germline is an increase in the rate of mitochondrial copying errors, μ_S (**Fig. 3a,** top left). With high μ_S, it is better to avoid mutation accumulation and move gamete production into an early sequestered germline, as segregation is no longer strong enough to generate fitter gametes with few or no mitochondrial mutants (**Fig. 2c**). There are two animal groups in which early germline sequestration is widespread – bilaterians[4,9] and ctenopohores[31]. In line with the prediction of our model, both groups have high mitochondrial mutation rates, 10-50 times faster than their mean nuclear mutation rate[32,33], and these mutations are largely produced by copying errors[34].

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Why did the mitochondrial mutation rate increase in the lineages giving rise to bilaterians and ctenophores? Increasing oxygen levels in the late Neoproterozoic[35] enabled predation for the first time, because energy conservation from aerobic respiration approaches 40% compared with < 10% for fermentation, making predation virtually impossible in anoxic worlds[36,37]. In the early Cambrian, predation drove an evolutionary arms race leading to greater size and physical activity in bilaterians and probably ctenophores[38]. Greater activity increased rates of tissue turnover, protein synthesis and replication of mitochondrial genes, contributing to an elevated μ_S relative to sessile early metazoans[39]. A link between mitochondrial mutation rate and germline evolution is corroborated by the Ceriantharia, which unlike other Anthozoa have fast-evolving mitochondrial DNA[40], suggesting a relatively high μ_S; strikingly, their larvae have gonads and are apparently paedogenetic[41], implying the evolution of germline sequestration in this group. Conversely, the secondary evolution of somatic gametogenesis in Ectoprocta and Entoprocta[9] may have been favoured by a falling μ_S. Little is known about mitochondrial sequence divergence in these groups, but we predict that μ_S will be lower in these groups than in related metazoans that retain a germline.

Oogamy improves adult fitness and gamete quality

Oogamy has typically been interpreted as an outcome of disruptive selection for gamete specialization, with large eggs for provisioning, contrasting with numerous small sperm for success in fertilization[42]. Oogamy is universal in extant metazoan groups and common to multicellular organisms in general, both with and without germline, but is most extreme in complex organisms with a germline[43]. Our analysis specifies more precisely the benefits that flow from the generation of large oocytes with high (or low) mitochondrial numbers and the relationship of this to the evolution of a germline.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Oogamy offers a temporary increase in mitochondrial number (M) that feeds through to the zygote. Imagine that oocytes undergo Q additional rounds of mitochondrial replication without cell division, producing large oocytes with $2^Q M$ mitochondria. After fertilization, the first few rounds of cell division partition the initial mitochondrial population in the zygote between the daughter cells, without the need for further replication, until the baseline M mitochondria is restored (**Fig. 5a**). Note that a large oocyte does not alter the total input of mitochondrial mutations per generation, because replication events lost from early development are simply moved later into gametogenesis. Nor does it impact on the net input of background mutations (μ_B), which is simply a function of lifespan unchanged by oogamy.

The temporary increase in M profoundly dampens the opportunity for segregational variance to build up in the early cell divisions (**Fig. 5b**, left-hand side), which greatly reduces the risk of one tissue inheriting a disproportionate number of mitochondrial mutants. This early suppression of variance is most valuable in organisms with multiple tissues, where oogamy greatly improves adult fitness (**Fig. 5c**). That in turn allows oogamy to readily fix, at least in organisms with somatic gametogenesis (**Fig. 5d,** $N_G = 10$). Once the early divisions of the zygote have allowed cells to regain the standard low level of M mitochondria, the many subsequent cell divisions before the end of development allow plenty of opportunity for segregational variation to generate fitter gametes with few or no mitochondrial mutants (**Fig. 5b**, right-hand side).

Strikingly, plants and basal metazoans have smaller numbers of mitochondria in oocytes (i.e. lower values of Q) than are typical of bilaterians[44-46], suggesting that 'mitochondrial oogamy' is distinct from the large size of oocytes for provisioning. This lower degree of mitochondrial oogamy can be explained by the lower mutual dependence of adult fitness on organ function

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in plants (i.e. less negative epistasis between tissue fitness). In the model this is equivalent to a smaller number of tissues. **Fig. 5e** shows that the optimal Q is indeed lower with fewer tissues (e.g. 4 compared with 8 tissues). Nonetheless, even in plants there is a degree of embryonic tissue determination in early development (e.g. in seeds[25]), so moderate mitochondrial oogamy is predicted. In the case of basal metazoans that lack tissue organization such as sponges, the model suggests that oogamy would not be favoured (**Fig. 5e**, no tissues). However, other factors, in particular the degree of uniparental inheritance of mitochondria, can still favour the spread of oogamy as well as spermatogenesis – the production of small male gametes (**Extended Data Fig. 2**).

In contrast to somatic gametogenesis, the model analysis indicates that oogamy becomes less advantageous as the germline moves earlier in development (**Fig. 5d**, $N_G = 6$). Oogamy has the down-side that reduced segregation early in development carries over into gametes when they are drawn from an early germline. This reduces variation and the production of gametes with few or no mitochondrial mutations, which keenly weakens the response to selection. This disadvantage comes to outweigh the advantage gained by raising average tissue fitness through the suppression of segregation in early embryos. Yet the observation is that active bilaterians invariably combine early germline sequestration with large oocytes, in the extreme packed with several orders of magnitude more mitochondria than are present in normal cells (10^6 vs. ~10^3)[14,15]. So the model seems at odds with observations. In short, in active bilaterians with differentiated, mutually dependent tissues and high μ_S, there is strong selection for early germline sequestration to limit the accumulation of mutations in gametes (**Fig. 4b**, top left), but equally strong selection for oogamy to minimize variation between tissue precursor cells (**Fig. 5d**), with the two forces opposing each other.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The solution to this paradox is suppression of the accumulation of copying errors (μ_S), while maximising the variance between oocytes via segregation during germline development. In human development > 6 million oogonia are generated by mitotic proliferation in the first half of foetal development, followed by seemingly random apoptotic death of all but 500,000 of them (i.e. > 90%) by the start of puberty (atresia)[47]. A similar pattern is common across most bilaterian metazoans[48]. The proliferation of oogonia helps restore segregational variance among female gametes, which now have the chance of carrying few or no mitochondrial mutations (**Extended Data Fig. 3**). This "over supply" is rectified by eliminating 9 out of every 10 oogonia, the survivors being 'islands' of segregational variance, more different to each other than would be the case if fewer divisions were used to generate an equal number of oocytes. The balance between restricting μ_S through early germline sequestration and increasing variance through oocyte proliferation defines the number of cell divisions before atresia (**Fig. 5f**) – enough to generate variance, but less than the hundreds of cell divisions in the male germline or somatic gametogenesis. Thus increasing Q favours germline sequestration at an intermediate stage, neither early nor late, but after some rounds of cell division to generate variance between gametes (**Fig. 5f**).

Restoring variance necessarily generates new mutations via copying errors (μ_S), but the total build-up of new mutations can be lowered by adaptations that suppress the accumulation of background damage (μ_B) in female germ cells. These adaptations include protecting oocytes in an internal ovary, provisioning them from follicular cells, and repressing the transcription of mitochondrial genes and active respiration, which can produce reactive oxygen species[14,17,18]. Accordingly, early sequestration of large oocytes is most favoured when accompanied by specific lowering of μ_B (**Extended Data Fig. 4)**, giving rise to quiescent oocytes. This is the case in the female germline of virtually all bilaterians[14].

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conclusions

For the last thirty years, the dominant explanation for the evolution of the germline has been that it was essential for the emergence of multicellular organisms through the suppression of selfish conflict between the cells that make up an individual. At its heart, this viewpoint lacks a rationale for the evolutionary stability of numerous organisms that lack a dedicated germline. It also sees the restricted mutational environment of the female germline as of secondary importance.

Here we have turned this paradigm on its head, locating the key driving force in selection for mitochondrial quality. This reflects the interplay of mutation and segregation of mitochondrial mutations in gametes and tissues. With low copying-error rates (μ_S), the combination of segregation and selection improves mitochondrial quality over generations, favouring somatic gametogenesis in plants and simple metazoans that lack multiple tissues or high numbers of mitochondria. Higher μ_S drives early germline sequestration, which ultimately permits greater tissue differentiation and complexity in groups such as bilaterians and ctenophores. The increase in μ_S probably related to the rise in oxygen shortly before the Cambrian explosion[37]. We find that the evolution of complex development, with strong negative epistasis between tissues, requires suppression of mitochondrial variance between tissue-progenitor cells. This is achieved by large oocytes containing high mitochondrial numbers. That in turn risks the loss of segregational variance between oocytes sequestered in the germline early in development. The solution was to reintroduce segregational variance through the mitotic proliferation of oogonia, followed by germline atresia, maximizing the variance between the surviving oocytes.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

These findings set out a completely novel hypothesis for the origin and evolution of germline sequestration, which predicts the traits of groups from plants and sponges to ctenophores and bilaterians. Unlike other hypotheses, selection for mitochondrial quality can account for the stability of both somatic gametogenesis and early germline sequestration. The requirement for segregational variation in the germline also elucidates the potential risks associated with mitochondrial heteroplasmy[16] and the need for strict uniparental inheritance in more active animals[52]. The unusual population genetics of mitochondrial mutation and segregation can uniquely explain the evolution of the germline in complex bilaterians.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Weismann, A. Prof. Weismann's theory of heredity. *Nature* **41,** 317–323 (1890).
2. Extavour, C. G. Evolution of the bilaterian germ line: lineage origin and modulation of specification mechanisms, *Integr. Comp. Biol.* **47**, 770–785 (2007).
3. Seipel, K., Yanze, N. & Schmid, V. The germ line and somatic stem cell gene Cniwi in the jellyfish *Podocoryne carnea*. *Int. J. Dev. Biol*. **48**, 1–7 (2004).
4. Buss, L. W. *The Evolution of Individuality* (Princeton University Press, 1987).
5. Michod, R. E. Cooperation and conflict in the evolution of individuality. II: Conflict mediation. *Proc. R. Soc. Lond. B* **263**, 813–822 (1996).
6. Michod, R. E. & Roze, D. Cooperation and conflict in the evolution of multicellularity. *Heredity* **86**, 1–7 (2001).
7. Queller, D. C. Relatedness and the fraternal major transitions. *Phil. Trans. R. Soc. B.* **355,** 1647–1655 (2000).
8. Grosberg, R. K. & Strathmann, R. R. The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Evol. Syst*. **38**, 621–654 (2007).
9. Blackstone, N. W. & Jasker, B. D. Phylogenetic considerations of clonality, coloniality, and mode of germline development in animals. *J. Exp. Zool. Part B* **297**, 35–47 (2003).
10. Weismann, A. *The Germ-Plasm: A Theory of Heredity* (Walter Scott, Ltd, 1893).
11. Kirkwood, T. B. Evolution of ageing. *Nature* **270**, 301–304 (1977).
12. Goldsby, H. J., Knoester, D. B., Ofria, C. & Kerr B. The evolutionary origin of somatic cells under the dirty work hypothesis. *PLoS Biol.* **12**(5), e1001858 (2014).
13. Crow, J. F. The origins, patterns and implications of human spontaneous mutation. *Nature Rev. Genet*. **1**, 40–47 (2000).
14. de Paula, W. B., Lucas, C. H., Agip, A. N. A., Vizcay-Barrena, G. & Allen, J. F. Energy, ageing, fidelity and sex: oocyte mitochondrial DNA as a protected genetic template. *Phil. Trans. R. Soc. B* **368**, 20120263 (2013).
15. Shoubridge, E. A. & Wai, T. Mitochondrial DNA and the mammalian oocyte. *Curr Top. Dev. Biol.* **77**, 87–111 (2007).

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

17. Allen, J. F. Separate sexes and the mitochondrial theory of ageing. *J. Theoret. Biol.* **180**, 135–140 (1996).

18. Bendich, A. J. Mitochondrial DNA, chloroplast DNA and the origins of development in eukaryotic organism. *Biol. Direct* **5**, 42 (2010).

19. Barritt, J. A., Cohen, J. & Brenner C. A. Mitochondrial DNA point mutation in human oocytes is associated with maternal age. *Reprod. Biomed. Online* **1**, 96–100 (2000).

20. Pang, C. Y., Lee, H. C., Yang, J. H. & Wei, Y. H. Human skin mitochondrial DNA deletions associated with light exposure. *Arch. Biochem. Biophys.* **312**, 534–538 (1994).

21. Krishnan, K. J., Reeve, A. K., Samuels, D. C., Chinnery, P. F., Blackwood J. K., *et al*. What causes mitochondrial DNA deletions in human cells? *Nature Genet.* **40**, 275–279 (2008).

22. Evans, M. D., Dizdaroglu, M. & Cooke, M. S. Oxidative DNA damage and disease: induction, repair and significance. *Mutat. Res*. **567**, 1–61 (2004).

23. Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J. & Letellier, T. Mitochondrial threshold effects. *Biochem. J*. **370,** 751–762 (2003).

24. Leys, S. P. & Hill, A. The physiology and molecular biology of sponge tissues. In *Advances in Marine Biology* (eds Becerro, M. A., Uriz, M. J., Maldonado, M., & Turon, X.) (Academic Press, 2012).

25. Evert, R. F. *Esau's Plant Anatomy. Meristems, Cells and Tissues of the Plant Body: their Structure, Function and Development*, 3rd Edition (Wiley, 2006).

26. Shearer, T. L., van Oppen, M. J. H., Romano, S. L. & Worheide, G. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria) *Mol. Ecol.* **11**, 2475–2487 (2002).

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

28. Huang, D., Meier, R., Todd, P. A. & Chou, L. M. Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. *J. Mol. Evo.* **66**, 167–174 (2008).

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

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

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

31. Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S. Norekian, T. P. Povolotskaya, I. S., *et al*. The ctenophore genome and the evolutionary origins of neural systems. *Nature* **510**, 109–114 (2014).

32. Lavrov, D. V. Key transitions in animal evolution: a mitochondrial DNA perspective. *Integr. Comp. Biol.* **47**, 734–743 (2007).

33. Pett, W. Ryan, J. F., Pang, K., Mullikin, J. C., Martindale, M. Q., Baxevanis, A. D. & Lavrov, D. V. Extreme mitochondrial evolution in the ctenophore *Mnemiopsis leidyi*: insight from mtDNA and the nuclear genome. *Mitochondrial DNA* **22**, 130–142 (2011).

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

35. Chen, X., Ling, H. F., Vance, D., Shields-Zhou, G. A., Zhu M., *et al*. Rise to modern levels of ocean oxygenation coincided with the Cambrian radiation of animals. *Nature Commun.* **6**, 7142 (2015).

36. Fenchel, T. & Finlay B. J. *Ecology and Evolution in Anoxic Worlds* (Oxford University Press, 1995).

37. Lane, N. *Oxygen: The Molecule that Made the World* (Oxford University Press, 2002).

38. Sperling, E. A., Frieder, C. A. Raman, A. V. Girguis, P. R. Levin, L. A. Knoll A. H. Oxygen, ecology, and the Cambrian radiation of animals. *Proc. Natl. Acad. Sci. USA* **110**, 13446–13451 (2013).

39. Lane, N. On the origin of barcodes. *Nature* **462**, 272–274 (2009).

40. Stampari, S., Maronna, M. M., Kitahara, M. V., Reimer, J. D. & Morandini, A. C. Fast-evolving mitochondrial DNA in *Ceriantharia*: a reflection of hexacorallia paraphyly? *PLoS One* **9**(1), e86612 (2014).

41. Molodtsova, T. N. On the taxonomy and presumable evolutionary pathways of planktonic larvae of Ceriantharia (Anthozoa, Cnidaria). *Hydrobiologia* **530/531**, 261–266 (2004).

42. Bulmer, M. G. & Parker G. A. The evolution of anisogamy: a game-theoretic approach. *Proc. R. Soc. Lond. B Biol. Sci.* **269**, 2381–2388 (2002).

43. Bell, G. The evolution of anisogamy. *J. Theoret. Biol.* **73**, 247–270 (1978).

44. Wang, D.-Y., Zhang, Q., Liu, Y., Lin, Z.-F., Zhang, S.-X., Sun, M.-X. & Sodmergen. The levels of male gametic mitochondrial DNA are highly regulated in angiosperms with regard to mitochondrial inheritance. *Plant Cell* **22**, 2402–2416 (2010).

45. Takanashi, H., Ohnishi, T., Mogi, M., Okamoto, T., Arimura, S. I. & Tsutsumi, N. Studies of mitochondrial morphology and DNA amount in the rice egg cell. *Curr. Genet.* **56**, 33–41 (2010).

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

46. Kuroiwa, H., Ohta, T. & Kuroiwa, T. Studies on the development and three dimensional reconstruction of giant mitochondria and their nuclei in egg cells of *Pelargonium zonale*. *Ait. Protoplasma* **192**, 235–244. 1996.

47. Krakauer, D. C. & Mira, A. Mitochondria and germ-cell death. *Nature* **400**, 125–126 (1999).

48. Saidapur, S. V. Follicular atresia in the ovaries of non-mammalian vertebrates. *Int. Rev. Cytol.* **54**, 225–244 (1978).

49. Carelli, V. Keeping in shape the dogma of mitochondrial DNA maternal inheritance. **11**(5), e1005179 (2015).

50. White, D. J., Wolff, J. N., Pierson, M. & Gemmell, N. J. Revealing the hidden complexities of mtDNA inheritance. *Mol. Ecol.* **17**, 4925–4942 (2008).

51. Birky, C. W., Jr. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc. Natl. Acad. Sci. USA.* **92**, 11331–11338 (1995).

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

53. Hadjivasiliou, Z., Lane, N., Seymour, R. M. & Pomiankowski, A. Dynamics of mitochondrial inheritance in the evolution of binary mating types and two sexes, *Proc. R. Soc. B* **280**, 1920 (2013).

Acknowledgments The authors acknowledge the use of the UCL Legion High Performance Computing Facility (Legion@UCL), and associated support services, in the completion of this work. A.L.R. is supported by a CoMPLEX PhD studentship from the Engineering and Physical Sciences Research Council, Z.H. by an EPSRC Research Fellowship (EP/L504889/1), N.L. by the Leverhulme Trust (RPG-425) and the Provost's Venture Research Prize, and A.P. by grants from the Natural Environment Research Council (NE/G00563X/1) and the Engineering and Physical Sciences Research Council (EP/F500351/1, EP/I017909/1, EP/K038656/1).

Author Contributions All authors designed the study, A.L.R., N.L. and A.P. analyzed the results and wrote the paper. A.L.R. carried out the programming and simulations.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figures

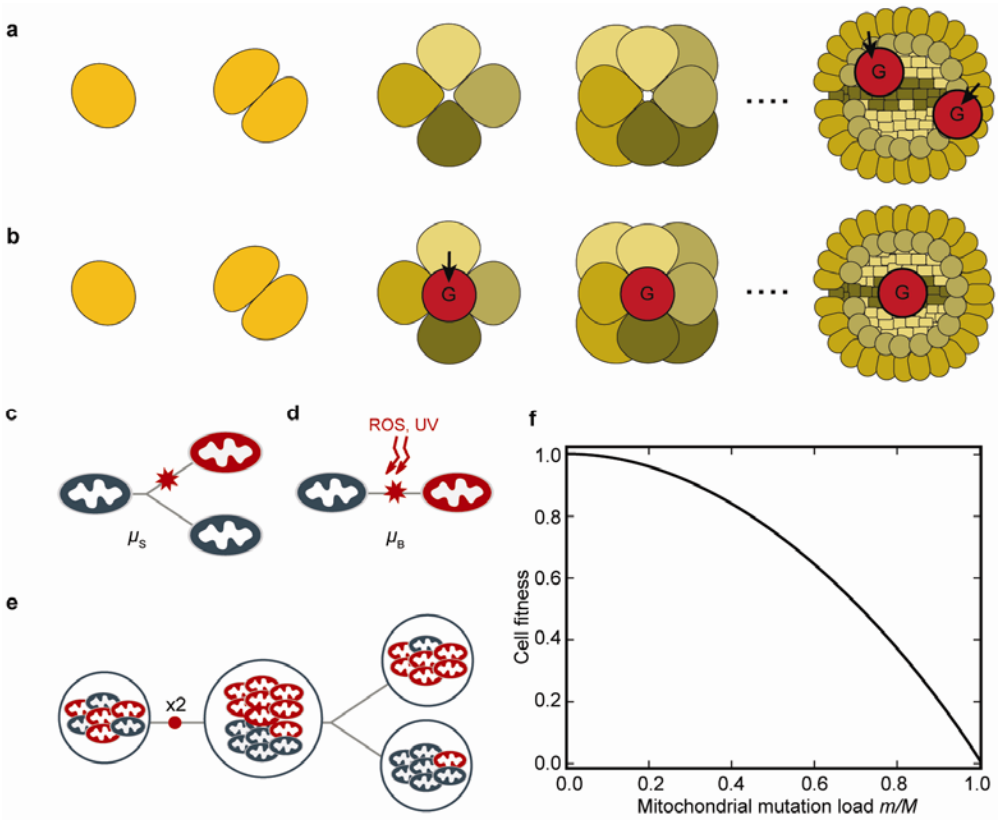


Figure 1. Life cycle of the model multicellular organism

(a) Development from zygote (left-hand side) to adult, showing early tissue differentiation (cells with differing shade) and late formation of gametes (G) from somatic cells in the adult (somatic gametogenesis). **(b)** Equivalent multicellular development depicting sequestration of gametes early in development (early germline). Dotted line indicates further development; adult and gamete cells not drawn to scale. **(c)** Copying errors (μ_S) during replication of mitochondrial DNA. **(d)** Mutations caused by background damage (μ_B) from e.g. ultraviolet radiation (UV) or reactive oxygen species (ROS). **(e)** Doubling followed by random segregation of mitochondrial mutants (red) at cell division increases variance between daughter cells. **(f)** Concave fitness function, in which cell fitness declines non-linearly with the accumulation of mitochondrial mutations ($\mu_S + \mu_B$) as seen in mitochondrial diseases[23].

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

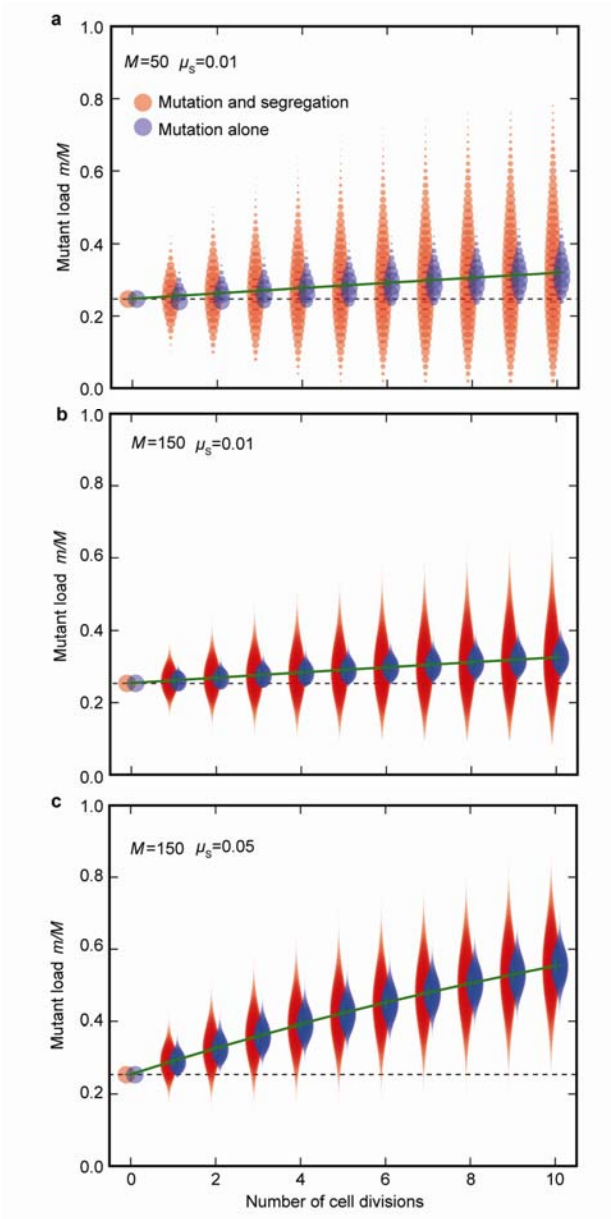


Figure 2. Segregation of mitochondrial mutants at cell division generates variance

At each cell division, the variance in mitochondrial mutant load (*m*/*M*) between daughter cells (red plots) increases due to mutational input as well as random segregation, generating cells with both more (above dotted line) and fewer (below dotted line) mutations than the zygotic cell. The blue plots show the variance in mutant load at each division due to mutation alone, without any random segregation. **(a)** The mutant load increases slowly with each cell division (green line) when the rate of copying errors is low (μ_S = 0.01). **(b)** Increasing

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mitochondrial number (*M*) decreases the variance in mutant load per cell (red plots) as the effects of random segregational drift are dampened. Fewer high-quality daughter cells lacking mitochondrial mutations are now generated. **(c)** Increasing copying errors ($\mu_S = 0.05$) increases the mutant load in daughter cells; segregation no longer generates many high-quality daughter cells with fewer mutations than the zygote. Data were derived from a starting number of mutants in zygotes $m = 0.24M$, and then run iteratively through successive cell divisions as described in the first section of Methods.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

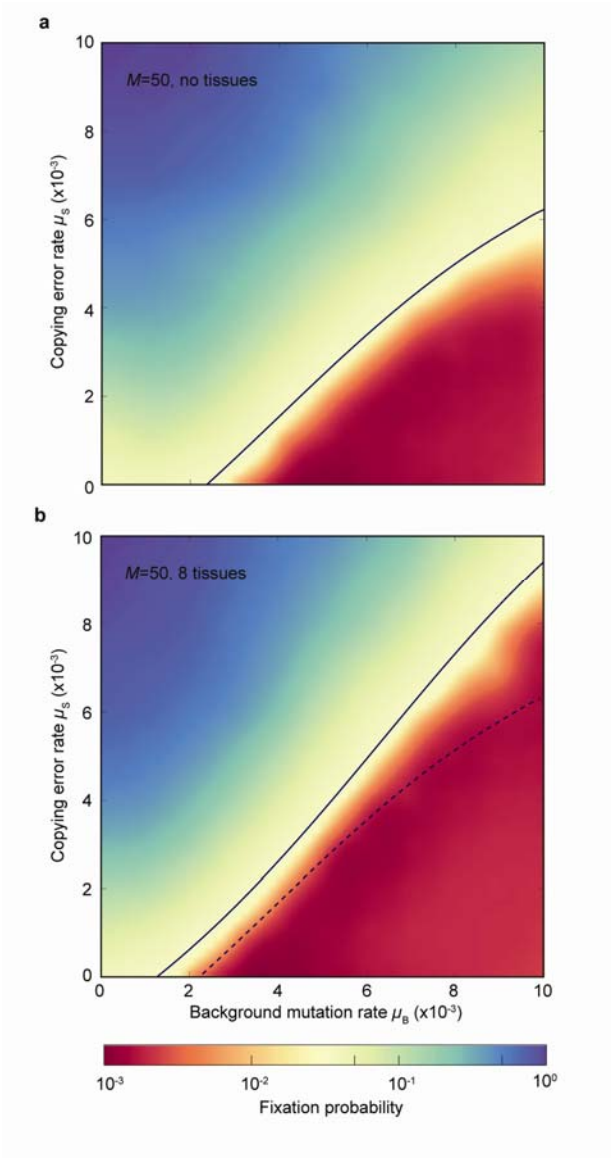


Figure 3. Germline evolution depends on mitochondrial mutation rate

(a) Heat map showing fixation probability of an allele encoding early germline sequestration (introduced at a frequency of 0.05, see Methods) in a simple organism that lacks tissue differentiation, in relation to the rate of copying-errors (μ_S) and background damage (μ_B). Early germline sequestration is favoured by higher μ_S and lower μ_B (blue, top left). The early germline allele is selected against in organisms with low μ_S and high μ_B (red, bottom right), conditions that instead favour somatic gametogenesis. The solid line represents neutrality. **(b)**

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Increasing the number of tissues to 8 makes it harder to fix an early germline – the region shaded in red expands (solid line versus dotted line) so germline fixation now requires higher μ_S and lower μ_B compared with (a).

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

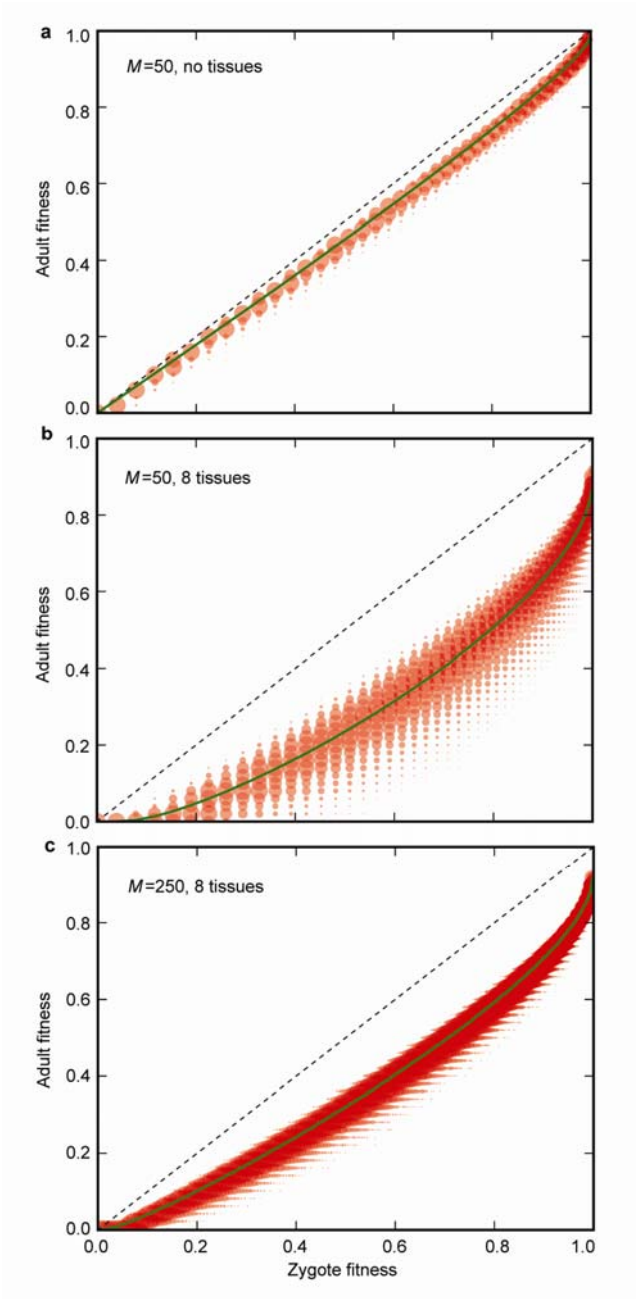


Figure 4. Mitochondrial segregation undermines adult fitness

Adult fitness is a function of zygote fitness (the mutation load inherited), mutational input and random segregation during development. **(a)** In organisms with no tissue differentiation adult fitness is similar to zygote fitness, as the number of new mutations accumulating within a single generation is limited, and variance in mutant load between cells within a tissue has no effect on adult fitness. **(b)** In organisms with early differentiation of multiple tissues, adult

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

fitness is undermined by segregational variance, as some tissue-precursor cells receive a higher mutant load than others, and adult fitness depends on the function of the worst tissue. **(c)** Increasing the number of mitochondria decreases the variance in mutant load between tissue-precursor cells, and so reduces the loss in adult fitness caused by random segregation. Parameter values $\mu_S = 0.01$, $\mu_B = 0.005$, 10 cell divisions to adulthood, and a lifespan equivalent to 40 cell division cycles.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

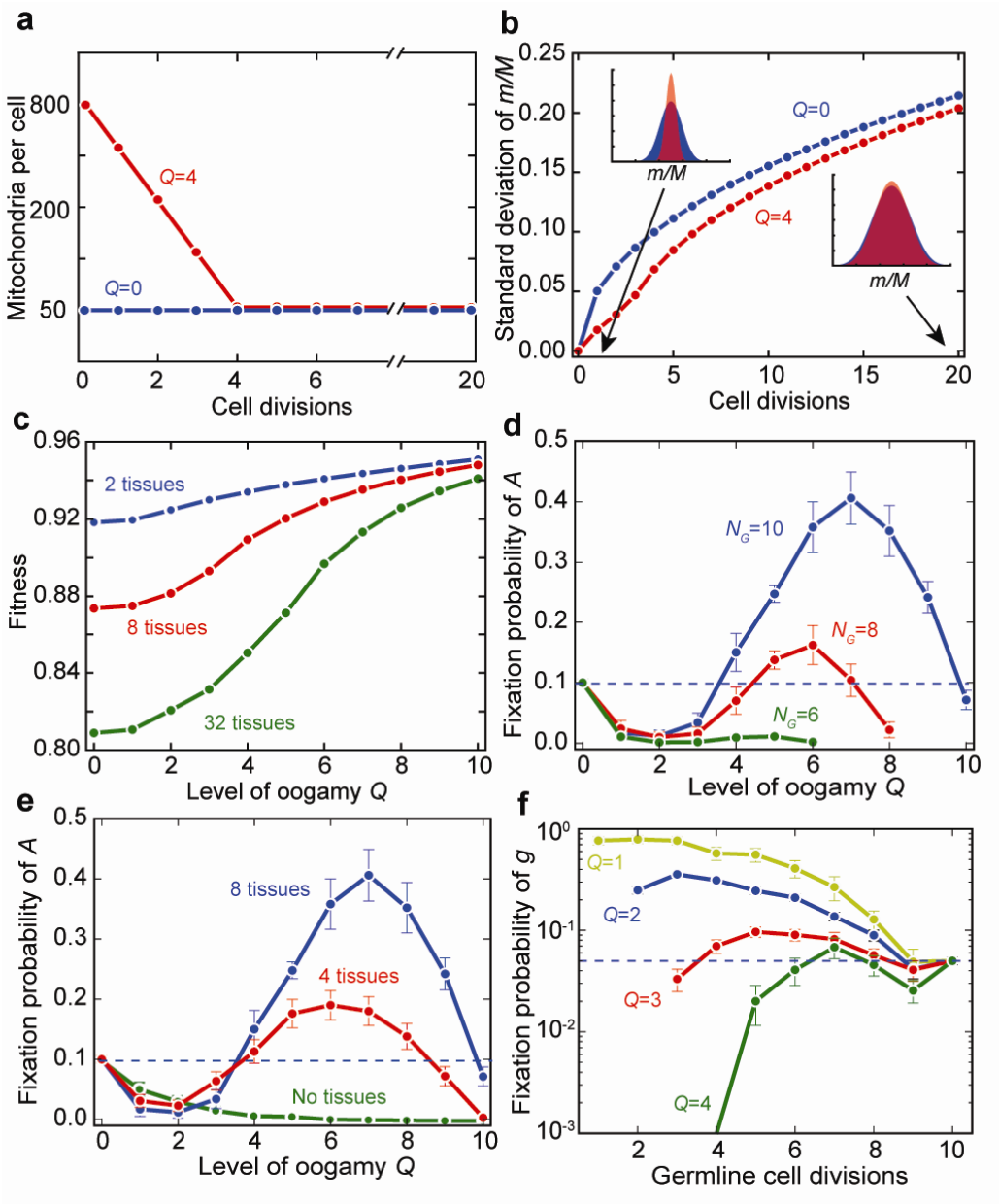


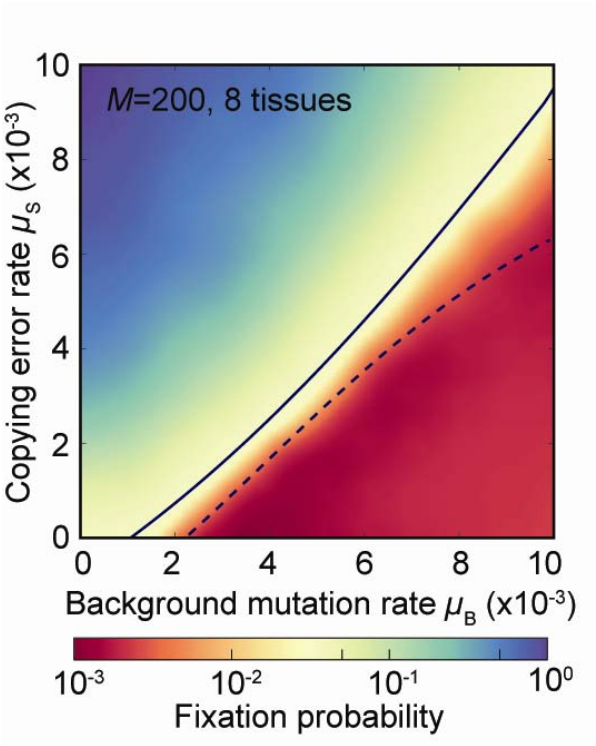
Figure 5. Oogamy improves adult fitness by temporarily suppressing variance

(a) The mitochondria in the zygote are partitioned into daughter cells at each cell division. Under isogamy ($Q = 0$), the zygote contains the same number M mitochondria as normal somatic cells. In contrast, with oogamy the zygote contains a much larger number of mitochondria which are partitioned without further replication until the standard mitochondrial number (M) is restored. For example, when $M = 50$ and $Q = 4$, the zygote has $2^Q M = 800$ mitochondria, which are partitioned to daughter cells over 4 rounds of cell

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

division to restore M. **(b)** A large oocyte (Q = 4, red) suppresses the variance in mutation load (m/M) in the first few rounds of cell division (left inset) compared with a small oocyte (Q = 0, blue). This early difference in variance is virtually lost after 20 rounds of cell division (right inset). The initial number of mitochondrial mutants in the zygote is $2^Q M/2$, and the segregation is modelled as described in the Methods without the further accumulation of mutations. **(c)** The early reduction in variance produced by oogamy improves adult fitness in organisms with multiple tissues. The more tissues, the greater Q needs to be to offset the loss of adult fitness. The initial mutant load in the zygote is $2^Q M/5$, $\mu_S = 0.01$, $\mu_B = 0.005$. **(d)** The fixation probability of an allele A specifying oogamy Q depends on the timing of gamete differentiation (N_G). The allele A fixes with somatic gametogenesis ($N_G = 10$) as more rounds of cell division generates greater variance between gametes, improving fitness over generations. But A does not fix with early germline sequestration ($N_G = 6$ or less), as fewer rounds of cell division do not generate enough variance between gametes for selection to efficiently eliminate mutants. Mutation rates are set to $\mu_S = 0.01$, $\mu_B = 0.005$. **(e)** With somatic gametogenesis ($N_G = 10$) the allele A encoding oogamy Q fixes more readily with a larger number of tissues and should not fix in organisms lacking tissues (but see Extended Data Fig. 3). Parameter values $\mu_S = 0.01$, $\mu_B = 0.005$. **(f)** The forces of mutation and segregation conflict in organisms with multiple tissues, where high Q improves adult fitness but high μ_S forces early germline sequestration. An allele g that specifies the number of cell divisions before gamete formation favours an intermediate number of cell divisions, in which mutation accumulation and segregational variance are balanced depending on Q. Parameter values $\mu_S = 0.02$, $\mu_B = 0.001$. In all panels, $M = 50$.

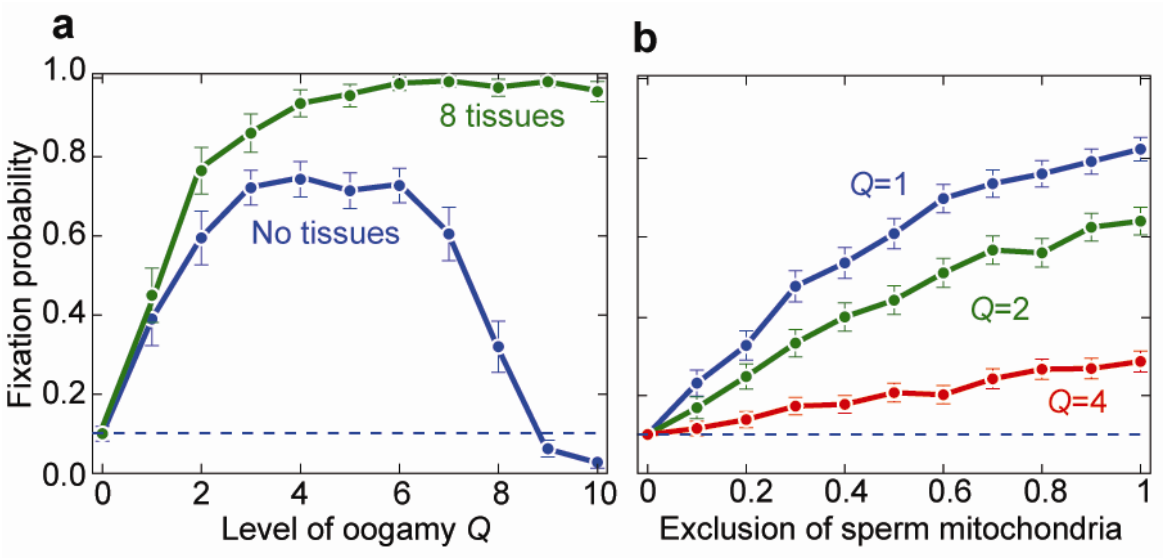
bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Extended Data Figure 1. Greater complexity does not favour germline sequestration

Heat map showing the fixation probability of an allele encoding early germline sequestration (introduced at a frequency of 0.05, see Methods) in an organism with 8 tissues and 200 mitochondria. Early germline sequestration is favoured by higher μ_S and lower μ_B (blue, top left). The allele is selected against in organisms with low μ_S and high μ_B (red, bottom right), instead favouring somatic gametogenesis. The solid line represents neutrality, and the dotted line gives neutrality for the case shown in Fig. 3a, with no tissues and M=50. Increasing the level of complexity (more tissues and mitochondria) therefore does not favour early germline sequestration.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



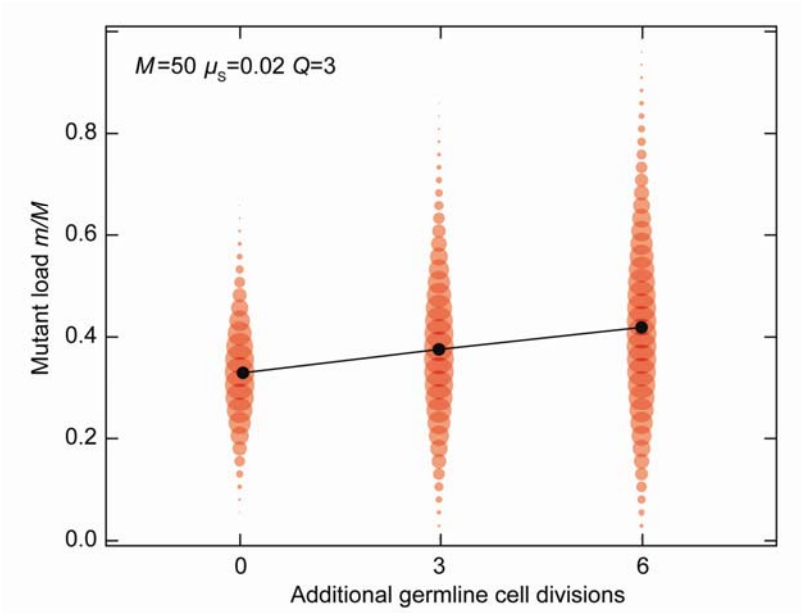
Extended Data Fig. 2. The evolution of anisogamy and strict uniparental inheritance is favoured in basal metazoans without tissues

(a) Oogamy (*Q*) is favoured even in organisms with no tissue differentiation if the starting point is biparental inheritance (BPI) of mitochondria. This is because oogamy produces a degree of partial uniparental inheritance (UPI) of mitochondria, which increases the variance between gametes and homogenizes the mitochondrial population within individual gametes. BPI is not an unreasonable starting point in basal metazoans as next generation sequencing shows that heteroplasmy and paternal leakage are not uncommon[50]; the disparate mechanisms that enforce UPI imply that strict UPI arose on multiple occasions and are subject to frequent turn-over[51], hence some degree of BPI must be common over evolutionary time; and previous modelling work shows that UPI is less strictly enforced in organisms with low mitochondrial numbers and low mutation rates[52,53], conditions that apply to basal metazoans such as sponges, which have very low, plant-like mitochondrial mutation rates[26-28,30]. Very high levels of oogamy (*Q*) are not favoured in the absence of tissues as this restricts segregational variance between gametes even with somatic gametogenesis, and there is no benefit to restricting early variance between tissues (which is why higher levels of *Q* are still favoured

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

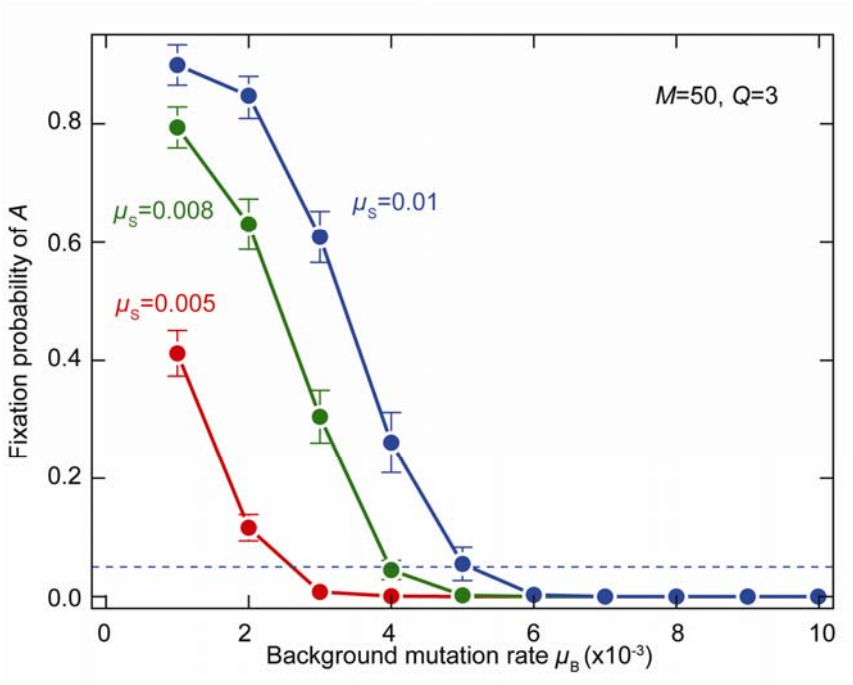
with multiple tissues). **(b)** The exclusion of male mitochondria from the zygote is favoured because that also increases the degree of UPI. The effect weakens with higher Q as the fractional contribution of sperm falls simply because there are fewer of them. The evolution of oogamy combined with exclusion of male mitochondria from the zygote is an alternative pathway to the evolution of strict UPI. We propose that this pathway could account for the evolution of anisogamy with stable UPI in basal metazoans. Parameter values $\mu_S = 0.001$, $\mu_B = 0.005$ and $M = 50$.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Extended Data Fig. 3. Restoration of germline segregational variance after oogamy. Additional cell divisions in the female germline restore the loss of segregational variance as a result of oogamy Q. In this example, $Q = 3$ and there is a high error rate ($\mu_S = 0.02$). Additional cell divisions generate a greater mutational load, but also considerable segregational variance. Under these conditions, the balance between mutation accumulation and variance favours an intermediate germline ($N_G = 6$, i.e. 3 additional germline cell divisions, once 3 rounds of early cell division returns cell content to M mitochondria) rather than an early germline ($N_G = 3$, no additional germline cell divisions) or a late germline ($N_G =$ 9 or more). Random atresia at $N_G = 6$ retains the same level of variance in the surviving cells between a smaller number of gametes. Other parameter values $\mu_B = 0.005$, $M = 50$.

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Extended Data Fig. 4. Restricting background mutations favours germline fixation. The evolution of early germline sequestration ($N_G = 3$) and oogamy ($Q = 3$) is favoured by low background mutation rates (μ_B). The allele *A* specifying early germline sequestration is more likely to be fixed at higher values of μ_S, as high μ_S favours restricting the number of rounds of cell division, but at the cost of restricting variance. Germline fixation also depends on the background mutation rate (μ_B). Germline sequestration does not fix if $\mu_B > 6 \times 10^{-3}$, but fixes readily at lower μ_B, even when μ_S is low. Adaptations that favour low μ_B are characteristic of bilaterians germlines, including sequestration of oocytes in an internal ovary, repression of transcription and translation, and suppression of mitochondrial respiration[14].

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Segregational drift generates variance in mitochondrial mutation load. Here we review the mitochondrial dynamics at the cell level, showing that random segregation at every cell division generates variance in the mutational load (Figures 2a-d and 4a-b). We start with a cell containing M mitochondria in total, out of which m_0 are mutant. The initial state of an infinite population is then represented by a state vector $\mathbf{p}^{(0)}$, with all of its $M+1$ entries set to zero, except for the m_0-th, which is set to one, i.e. $p_i^{(0)} = \delta(i, m_0)$. A single cell cycle consists of mitochondrial mutation, replication and cell division. Mitochondrial mutation is a Bernoulli trial with success probability μ while segregation is modelled as a simple random sampling without replacement. After n cell divisions the population state vector becomes

$$\mathbf{p}^{(n)} = (\mathbf{KJ})^n \mathbf{p}^{(0)}.$$

Here $\mathbf{J}$ is the $(M+1) \times (M+1)$ matrix with elements representing the transition probabilities due to mutation,

$$J_{q,m} = p_{mut}(q; m, M) = \binom{M-m}{q-m} \mu^{q-m} (1-\mu)^{M-q} .$$

Similarly, $\mathbf{K}$ is the matrix of same dimensions representing the transformation due to random sampling,

$$K_{q,m} = p_{seg}(q; m, M) = \frac{\binom{2m}{k}\binom{2M-2m}{M-k}}{\binom{2M}{M}}.$$

In Figure 2 we show distributions $\mathbf{p}^{(n)}$ for $n = 0 \ldots 10$.

Variance in the mutant load after n cell divisions can be expressed analytically in the illustrative case of $\mu = 0$, where only segregational drift is accounted for. Let X_n be a random variable denoting the number of mutant mitochondria within a cell in the developing tissue or embryo after n cell divisions, and x_n its actual realization. For sampling from the

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

hypergeometric probability distribution, the population mean equals the initial number of mutants, $\mathrm{E}(X_n) = x_0$, while the variance can be decomposed as

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

For sampling without replacement from the hypergeometric probability distribution, the conditional variance is

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

and so

$$\mathrm{Var}(X_n) = \mathrm{E}\left(\frac{X_{n-1}(M - X_{n-1})}{2M - 1}\right) + \mathrm{Var}(X_{n-1})$$

$$= \mathrm{E}\left(\frac{MX_{n-1}}{2M - 1}\right) - \mathrm{E}\left(\frac{X_{n-1}^2}{2M - 1}\right) + \mathrm{Var}(X_{n-1})$$

$$= \frac{M}{2M - 1}\mathrm{E}(X_{n-1}) - \frac{1}{2M - 1}\mathrm{E}(X_{n-1}^2) + \mathrm{Var}(X_{n-1})$$

$$= \frac{x_0 M}{2M - 1} - \frac{1}{2M - 1}[\mathrm{Var}(X_{n-1}) + x_0^2] + \mathrm{Var}(X_{n-1})$$

$$= \frac{x_0(M - x_0)}{2M - 1} + \left(1 - \frac{1}{2M - 1}\right)\mathrm{Var}(X_{n-1}).$$

Here x_0 is the initial number of mutants within a cell. This is a recurrence relation of the form $h_n = Hh_{n-1}$ with $h_n = Var(X_n) - x_0(M - x_0)$ and $= 1 - \frac{1}{2M-1}$. With the boundary condition $\mathrm{Var}(X_0) = 0, h_0 = -x_0(M - x_0)$, 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

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

which corresponds to Figure 4b.

In oogamous matings, the zygote contains $2^Q M$ mitochondria and undergoes Q divisions without mitochondrial replication. Changes in variance follow a similar trend, but this time the recurrence relation is

$$\mathrm{Var}(P_n) = \frac{p_0(1-p_0)}{2^{1+Q-n}M-1} + \mathrm{Var}(P_{n-1})\left(1-\frac{1}{2^{1+Q-n}M-1}\right), \qquad 1 \le n \le Q.$$

The recurrence relation can be used directly to produce Figure 5b For a large initial number of mitochondria $\frac{1}{2^Q M} \to 0$ and the approximate solution is

$$\mathrm{Var}(P_n) = \frac{p_0(1-p_0)}{2^Q M}(2^n-1), \qquad 1 \le n \le Q.$$

Finite-population simulation model for the evolution of germline. To study the evolutionary dynamics of alleles responsible for germline sequestration we developed a multi-level agent-based model, implemented as a set of simulation routines in ANSI/ISO C++ (*https://github.com/ArunasRadzvilavicius/GermlineEvolution*). The model population consists of $N = 500$ multicellular individuals, with equal numbers of males and females. A single individual is represented as an object containing a variable number of cells, either undifferentiated or grouped into tissues while each cell is an object containing the number of mutant and wild type mitochondria. Generations are discrete and non-overlapping. Organism development proceeds by iterating through the population and modifying these objects according to a set of predefined rules. The population is initialized in a random state, and the simulation proceeds as follows:

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1. Start the generation with a population of N zygotes;
2. For each organism in the population, repeat until development is complete:
    a. For every somatic cell within an organism:
        i. Apply background mitochondrial mutation as a Bernoulli event with the success probability μ_B;
        ii. If the number of mitochondria is M, apply mitochondrial mutation as a singular Bernoulli trial with success probability μ_S, then duplicate the mitochondrial population;
        iii. Partition mitochondria into two daughter cells by sampling without replacement, which in practice is implement by drawing a random number form the hypergeometric probability distribution;
    b. If the number of cell divisions equals G, set aside a single primordial germ cell by copying a randomly chosen stem cell, which will not undergo further mitotic divisions;
    c. If the number of cell divisions equals T, assign each of 2^T stem cells to its own somatic tissue;
    d. If the number of cell divisions equals L, stop the development. For each cell apply the background mitochondrial mutation $S = 40$ times independently, each time with probability μ_B;
3. Apply selection by sampling $N/2$ males and $N/2$ females from the population with replacement, linearly weighted according to their mitochondrial fitness;
4. In female germ cells, repeat Q times:
    a. Apply mitochondrial mutation as a binomial event with probability μ_S, duplicate the mitochondrial population

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5. Complete gametogenesis by executing two meiotic cell divisions, with mitochondrial sampling without replacement; male gametes contain $M/2$ and eggs contain $2^Q M/2$ mitochondria;
6. Fuse random gamete pairs of opposite sexes, making sure that the zygote contains $2^Q M$ mitochondria.

Gametes of opposite sexes meet at random and fuse. Parameter v controls the amount of mitochondria inherited from the father (paternal leakage). At fertilization (step 6), the zygote's mitochondrial population of $2^Q M$ is formed by sampling $(1 - v)M/2$ mitochondria from a sperm cell, and the rest from the oocyte.

Evolution of the nuclear alleles and fixation statistics. Mating type, the number of germline cell divisions G, the log-size of the zygote Q and the degree of uniparental inheritance v are all traits of an organism controlled by a set of loci in the nuclear genome, which is assumed to be diploid. The mating type locus is heterogametic *ZW* in females, while the other mating types is homogametic *ZZ*.

We determine the selective advantage of an invading allele in a Monte Carlo simulation, by numerically calculating its fixation probability within a finite population. With the population at equilibrium, the invading allele is introduced at a low frequency $f_0 = 0.05$ (or 0.1), and its fate is tracked until either fixation or extinction. This requires $10^3 - 10^4$ repetitions of the calculation depending on the mutation rates and the associated levels of noise. This pattern is assessed relative to the fixation probability of a neutral allele, which simply equals its initial frequency f_0. The error bars in figures represent the 95% confidence interval of the binomial

bioRxiv preprint doi: https://doi.org/10.1101/026252; this version posted August 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

proportion via the Gaussian approximation, that is $\Delta p = \pm 1.96\sqrt{\frac{p(1-p)}{n}}$, where n is the number of trials and p is the measured rate of fixation. An allele is deemed to be evolutionarily advantageous if its fixation probability exceeds the chance of fixation of the neutral allele.

The same general procedure is applied to determine the fate of modifiers coding for the number of germline cell divisions, the level of mitochondrial oogamy or mitochondrial exclusion. The germ cell differentiation locus is expressed in both sexes, and is assumed to be autosomal, with the invading allele assumed to be dominant. We also examined other dominance states, but the main conclusions of this work remained unaffected. The other two loci are *W* linked and unaffected by dominance considerations.