

# Multiple distinct evolutionary mechanisms govern the dynamics of selfish mitochondrial genomes in *Caenorhabditis elegans*

Received: 30 January 2024

Accepted: 13 September 2024

Published online: 19 September 2024

 Check for updatesBryan L. Gitschlag<sup>1,2</sup>✉, Claudia V. Pereira<sup>1</sup>, James P. Held<sup>1</sup>,  
David M. McCandlish<sup>2</sup> & Maulik R. Patel<sup>1,3,4,5</sup>✉

Cells possess multiple mitochondrial DNA (mtDNA) copies, which undergo semi-autonomous replication and stochastic inheritance. This enables mutant mtDNA variants to arise and selfishly compete with cooperative (wildtype) mtDNA. Selfish mitochondrial genomes are subject to selection at different levels: they compete against wildtype mtDNA directly within hosts and indirectly through organism-level selection. However, determining the relative contributions of selection at different levels has proven challenging. We overcome this challenge by combining mathematical modeling with experiments designed to isolate the levels of selection. Applying this approach to many selfish mitochondrial genotypes in *Caenorhabditis elegans* reveals an unexpected diversity of evolutionary mechanisms. Some mutant genomes persist at high frequency for many generations, despite a host fitness cost, by aggressively outcompeting cooperative genomes within hosts. Conversely, some mutant genomes persist by evading inter-organismal selection. Strikingly, the mutant genomes vary dramatically in their susceptibility to genetic drift. Although different mechanisms can cause high frequency of selfish mtDNA, we show how they give rise to characteristically different distributions of mutant frequency among individuals. Given that heteroplasmic frequency represents a key determinant of phenotypic severity, this work outlines an evolutionary theoretic framework for predicting the distribution of phenotypic consequences among individuals carrying a selfish mitochondrial genome.

Major evolutionary transitions follow a similar adaptive logic from solitary to social living as from unicellular to multicellular life, prokaryotic to eukaryotic cells, and molecular replicators to cells and genomes<sup>1–3</sup>. In each case, previously independent entities form a cooperative group that reproduces and undergoes natural selection as a unified collective<sup>1–5</sup>. By contributing to the fitness of other group

members, however, cooperative interactions can select for selfish, cheater entities<sup>6–9</sup>, which occur at each level of biological organization, from the molecular to societal<sup>9–13</sup>. Although cheating confers an advantage within groups, by assuming the benefits without the costs of cooperation<sup>6</sup>, cheaters also face disadvantages for numerous reasons. These include suppression or policing mechanisms that enforce

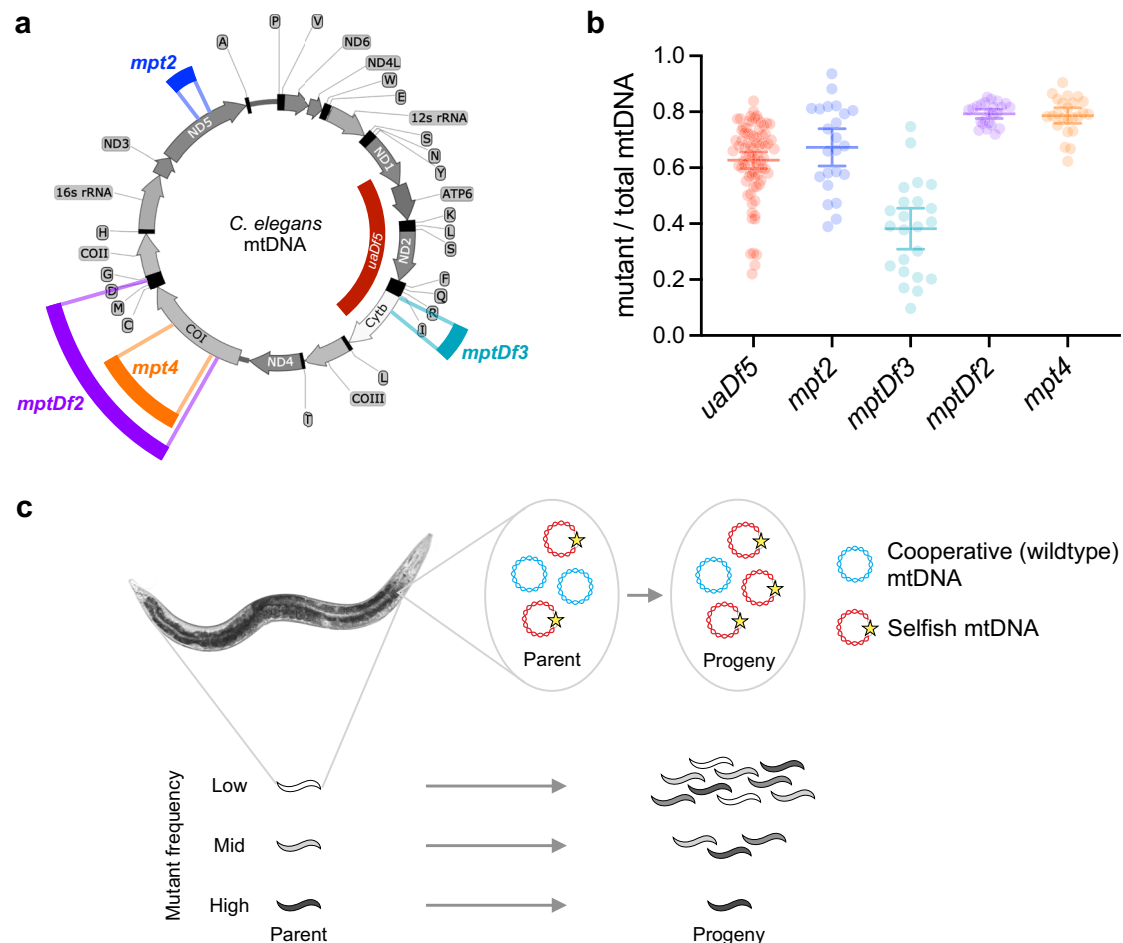
<sup>1</sup>Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA. <sup>2</sup>Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. <sup>3</sup>Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN, USA. <sup>4</sup>Diabetes Research and Training Center, Vanderbilt University School of Medicine, Nashville, TN, USA. <sup>5</sup>Evolutionary Studies, Vanderbilt University, VU Box #34-1634, Nashville, TN, USA. ✉e-mail: [gitschl@cshl.edu](mailto:gitschl@cshl.edu); [maulik.r.patel@vanderbilt.edu](mailto:maulik.r.patel@vanderbilt.edu)

cooperation, as well as negative frequency-dependent selection, that is, declining fitness as cheaters become more frequent<sup>12,14</sup>. Selection at higher levels of biological organization, which rely on cooperation, can also counter the fitness advantage of cheaters, consistent with observations ranging from selfish genetic elements to animal societies<sup>10,14–17</sup>. Thus, different evolutionary mechanisms, such as multilevel and frequency-dependent selection, are thought to help resolve an apparent paradox, namely the coexistence of cooperative and selfish entities<sup>12,14–17</sup>.

Like other major evolutionary transitions, the origin of the eukaryotic cell led to conflicting strategies of cooperation and cheating. This arises from the fact that mitochondria retain their own DNA (mtDNA), a remnant of their bacterial ancestor, whose symbiotic relationship with the nucleus serves as both the foundation of eukaryotic life and as a source of genetic conflicts<sup>18–20</sup>. In contrast to the typically diploid nuclear genome, an animal cell contains hundreds or thousands of mtDNA molecules<sup>21,22</sup>, which replicate asynchronously throughout the cell cycle<sup>23,24</sup>. High mtDNA copy number and relaxed, semi-autonomous replication frequently give rise to a heteroplasmic state, in which a host organism harbors a mixed population of different mtDNA variants. Competition for replication and transmission can select for mutations that confer a selfish advantage over other mtDNA variants within a host, whereas host fitness constitutes a form of group selection favoring cooperative genomes<sup>18,25–28</sup>. Mitochondria thus

serve as an exceptional case study for understanding the dynamics of cooperation and conflict.

Consistent with the conceptual framework of multilevel selection, prior research has uncovered numerous insights into the molecular, physiological, and environmental determinants of the fitness effects of mitochondrial mutations<sup>29–32</sup>. Many such findings can be credited to the mutant genome *uaDf5* in the model species *Caenorhabditis elegans*<sup>26,30–39</sup>. Due to a 3.1-kilobase deletion resulting in the loss of four protein-coding and seven tRNA genes (Fig. 1a), *uaDf5* negatively impacts mitochondrial respiration and host fitness<sup>26,34–36</sup>. Despite these effects, *uaDf5* can persist alongside cooperative (wildtype) mtDNA in a heteroplasmic state for hundreds of generations<sup>38</sup>. How is this accomplished? In agreement with theory<sup>40</sup>, experimental work has shown that *uaDf5* propagates not only in spite of but *because* of its adverse effects, by exploiting physiological stress-resistance mechanisms<sup>32,34,36</sup>, consistent with *uaDf5* being a bona fide cheater<sup>6</sup>. On the one hand, these findings highlight *uaDf5* as a valuable model of a biological cheater. On the other hand, a disproportionate fraction of mechanistic insights have come through studying this genome<sup>18,26,27,29,30,32,34,36,37,39</sup>, which challenges the ability to draw broad conclusions. Moreover, although prior studies have implicated diverse mito-nuclear interactions in the selfish propagation of mitochondrial mutations<sup>26,29,32,34,36,41</sup>, as well as substantial variation in their phenotypic effects<sup>42,43</sup>, this work spans multiple species<sup>18,27,29</sup>. Animal species



**Fig. 1 | A collection of mutant mitochondrial genotypes stably persisting in a heteroplasmic state despite disruption of essential respiratory genes.**

**a** *C. elegans* mitochondrial genome, with the locations and sizes of mutations represented by the color-coded bars: *uaDf5* (red), *mpt2* (blue), *mptDf2* (purple), *mpt4* (orange), *mptDf3* (green). **b** Distribution of heteroplasmic mutant frequencies for age-synchronized adult animals each carrying one of the mutations shown in (a)

(data are provided in Source Data file). Error bars indicate mean and 95% confidence interval of mean:  $0.63 \pm 0.03$  (*uaDf5*);  $0.67 \pm 0.7$  (*mpt2*);  $0.38 \pm 0.07$  (*mptDf3*);  $0.79 \pm 0.02$  (*mptDf2*);  $0.79 \pm 0.03$  (*mpt4*). **c** Schematic of multilevel selection, with a selfish mitochondrial genome undergoing positive intra-organismal (within-host) selection and negative inter-organismal (between-host) selection. Shading indicates frequency of selfish mtDNA.

vary widely in ways that affect mitochondrial genome biology, including nuclear genome composition, development and organization of the germline, and environmental context<sup>29,44</sup>. Although such factors provide important insight on mitochondrial biology across a diversity of taxa and conditions, here we sought to explore the diversity in evolutionary dynamics that can arise solely from variation in mitochondrial genotype, under conditions that control for these other factors.

To address these challenges, we developed an approach that leverages a collection of mutant mitochondrial genotypes within a single species, namely *C. elegans*. In particular, we combine mathematical modeling with laboratory experiments to comprehensively measure the evolutionary forces that govern the population dynamics of these stable heteroplasmies. First, we introduce a stochastic population-genetic model of selfish mitochondrial genomes, which includes different levels of selection (organismal and intra-organismal) plus genetic drift. Theoretical modeling revealed that different evolutionary mechanisms give rise to qualitatively similar outcomes, such as the persistence of a mutant genome at high heteroplasmic frequencies. Despite these qualitative similarities, however, the distribution of mutant frequencies among hosts contains information about the underlying evolutionary forces. To test our theoretical results, we fit our model to empirical data, obtained from experiments designed to disentangle the levels of selection. We applied this approach to a collection of mitochondrial genotypes, encompassing mutations that affect enzyme complexes I, III, and IV of the electron transport chain (ETC), in addition to the well-known mutant mtDNA variant *uaDf5* (Fig. 1a). Importantly, although each mutant genome stably propagates in heteroplasmic lineages, the distribution of mutant frequencies among heteroplasmic individuals varies considerably between these genotypes (Fig. 1b). Indeed, consistent with our theoretical analysis, we find that although cheating is a common feature of mutant mitochondrial genomes, the evolutionary mechanisms comprising their cheating strategy, and which maintain these observed distributions of mutant frequencies, are fundamentally different depending on the affected loci.

## Results

### Population-genetic modeling implicates multiple mechanisms in the evolutionary dynamics of selfish mitochondrial genomes

To understand the complex population genetics of selfish mitochondrial genomes, we developed a population-genetic model that captures natural selection at the intra- and inter-organismal levels, as well as genetic drift within heteroplasmic lineages in order to reflect stochasticity in development and mitochondrial inheritance.

For selfish mtDNA mutants, let us first consider how mutant mtDNA frequency within an individual,  $z$ , changes between parent and progeny (Fig. 1c, top row). To capture this effect, we define the intra-organismal fitness function  $w_{intra}(z)$  to be the net relative replication success of selfish versus cooperative (wildtype) mtDNA, per organismal generation, as a function of  $z$ . Thus, mutant frequency tends to increase when  $w_{intra}(z) > 1$  but tends to decrease when  $w_{intra}(z) < 1$ . We use a three-parameter fitness function, which can accommodate frequency-dependent or frequency-independent fitness effects and either linear or nonlinear frequency-fitness relationships (see Methods for details). While the intra-organismal fitness function and current mutant mtDNA frequency  $z$  determine the expected mutant mtDNA frequency among progeny, to capture stochasticity during development and inheritance, we model these dynamics phenomenologically as consisting of a single generation of a Wright-Fisher process where in each generation mtDNA passes through a developmental bottleneck of size  $N$ . In our modeling framework  $N$  is best viewed as an effective or “virtual” mtDNA bottleneck size, since it reflects the cumulative effect of drift throughout the organismal lifecycle, both during early development and in the adult germline.

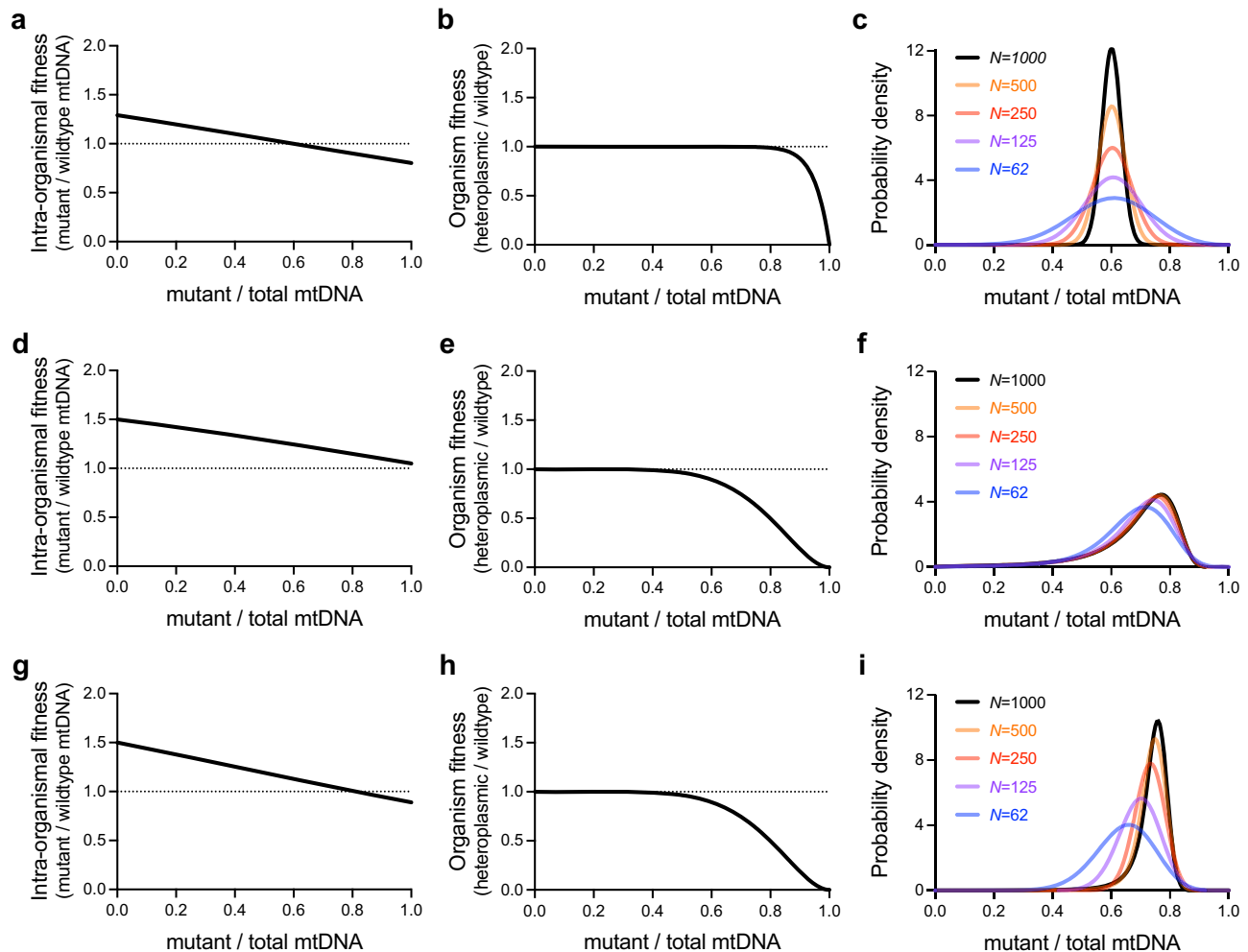
Moving to the organismal level, we treat the inter-organismal (between-host) dynamics deterministically, corresponding to the case of an infinite population of heteroplasmic hosts. We define organism fitness  $w_{org}(z)$  as the expected number of viable offspring per heteroplasmic parent with mutant mtDNA frequency  $z$ , divided by the expected number of viable offspring per homoplasmic-wildtype parent. We assume a model of purifying selection, so that whole-organism fitness is monotonically decreasing with increasing selfish mtDNA frequency (Fig. 1c, bottom row), with organisms fixed for the mutant mtDNA becoming sterile or inviable (that is,  $w_{org}(1) = 0$ ), since the mutations investigated here delete one or more essential respiratory gene(s). Because previous heteroplasmy studies have observed threshold effects, whereby modest shifts in mutant frequency can have large phenotypic consequences<sup>45,46</sup>, we fit a two-parameter fitness function that can accommodate both simple linear and power-law models for how the organism-level fitness cost scales with selfish mtDNA frequency, as well as more complicated threshold-like models (see Methods for details).

Combining our intra- and inter-organismal models, each generation begins with a distribution of mutant mtDNA frequencies among heteroplasmic hosts. These frequencies are shifted by intra-organismal selection and the distribution is broadened by intra-organismal genetic drift. Finally, host organisms reproduce with an expected number of viable offspring determined by their mutant mtDNA frequency. As this process repeats across generations, the distribution of mutant mtDNA frequencies among hosts approaches a characteristic shape (given by the dominant eigenvector of a principal submatrix of the transition matrix describing the above evolution process, see Methods), where the shape of this stable mutant frequency distribution reflects the specific parameter values employed. From this distribution, we can calculate the mean fitness of heteroplasmic carriers relative to their wildtype counterparts, as well as the rate of spontaneous reversion to the homoplasmic-wildtype state due to the de novo loss of the mutant genome via intra-organismal drift. Together, these values determine the timescale over which an established heteroplasmy will be lost, which can either be short or long depending on the specific parameters. Indeed, under this model we observe that selfish mtDNA can persist in a heteroplasmic state for hundreds, or even thousands, of generations under biologically realistic population-genetic conditions.

Examining this model led us to identify a number of distinct evolutionary mechanisms that determine the stable distribution of mutant mtDNA frequencies (Fig. 2). Under the first mechanism, intra-organismal balancing selection, the intra-organismal fitness advantage of the mutant genome vanishes in favor of the wildtype genome at a frequency  $z^*$  that is still too low to encounter significant inter-organismal selection (Fig. 2a, b); that is, there exists a frequency  $z^*$  with  $w_{intra}(z^*) = 1$  and  $w_{org}(z^*) \approx 1$ . The mutant frequency distribution among hosts is therefore centered at approximately  $z^*$  and is subject to negligible inter-organismal selection, with genetic drift determining the width of the distribution (Fig. 2c).

Under the second mechanism, which could be termed strong multilevel selection, a mutant genome maintains an intra-organismal fitness advantage across all heteroplasmic frequencies, so that there exists no frequency  $z^*$  such that  $w_{intra}(z^*) = 1$  (Fig. 2d). Thus, intra-organismal selection constitutively drives mutant frequency up into a range that elicits a strong organism-level fitness cost (Fig. 2e). Under this mechanism, the stable mutant mtDNA frequency distribution is determined predominantly by the forms of the opposing intra- and inter-organismal selection, and is relatively insensitive to changes in the strength of intra-organismal drift, particularly for large mtDNA effective bottlenecks  $N$  (Fig. 2f).

Finally, in the third mechanism, there exists an intra-organismal equilibrium frequency  $z^*$  with  $w_{intra}(z^*) = 1$ , but the organism-level fitness cost intensifies rapidly at frequencies similar to  $z^*$  (Fig. 2g, h).



**Fig. 2 | Theoretical modeling reveals the stable maintenance of a selfish mitochondrial genome through a variety of evolutionary mechanisms.**

**a–c** Under one hypothesized evolutionary mechanism, the intra-organismal fitness advantage of a selfish mitochondrial genome (**a**) may disappear at a frequency too low for the genome to inflict a substantial organism-level fitness cost (**b**). The selfish mtDNA is maintained at a mean frequency set by balancing selection at the intra-organismal level, with genetic drift (a function of the per-generation intra-organismal mtDNA effective bottleneck,  $N$ ) determining the width of the mutant frequency distribution (**c**). **d–f** Under an alternative hypothesized mechanism, a

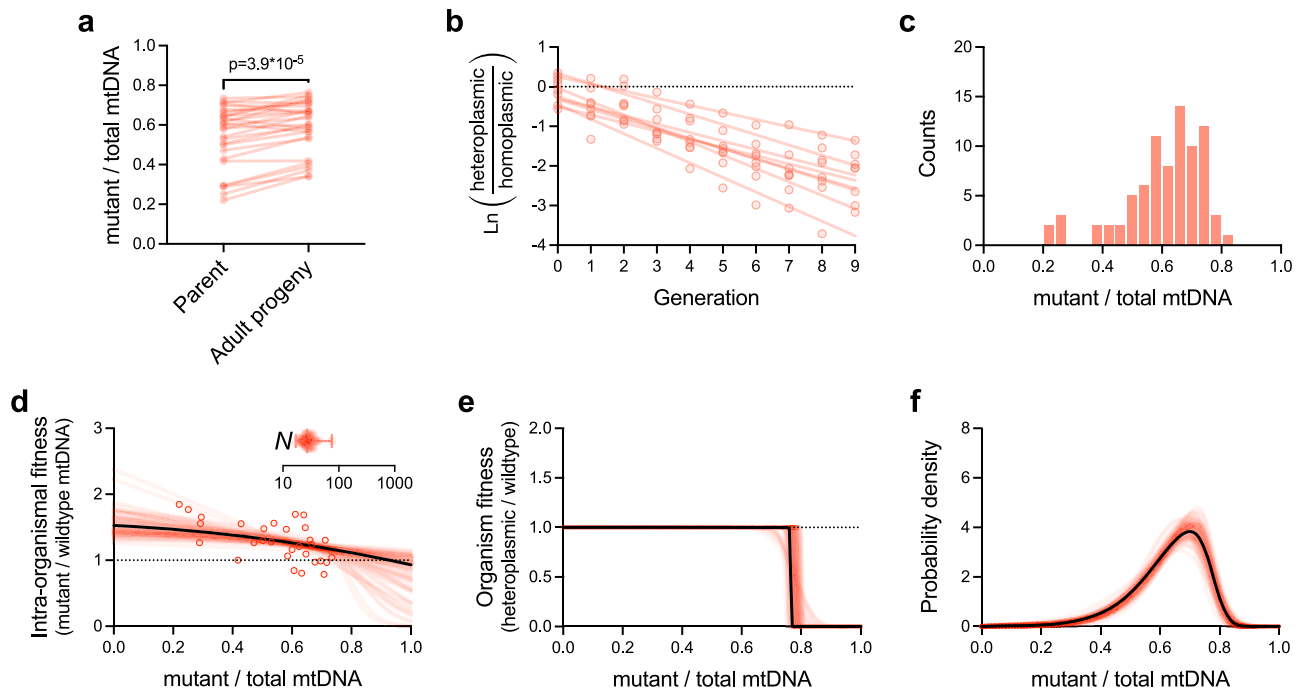
constitutive intra-organismal fitness advantage (**d**) pushes mutant frequency into a range wherein it encounters increasing organism-level fitness cost **e**. Multilevel selection can maintain a broad range of mutant frequencies (**f**), even when genetic drift is minimal. **g–i** Another hypothesized mechanism shows the intra-organismal fitness advantage (**g**) disappear within approximately the same range of mutant frequencies in which organism-level fitness declines most rapidly (**h**). Changes in the strength of intra-organismal genetic drift alter both the width and center of the distribution of selfish mtDNA frequencies (**i**).

Under this mechanism, the shape of the distribution depends on the form of all three forces, namely intra-organismal selection, inter-organismal selection, and drift (Fig. 2i). We therefore call this the mixed mechanism. Together, these findings suggest that mitochondrial mutations can persist at similar heteroplasmic frequencies despite substantial differences in the underlying evolutionary mechanisms. Understanding the mechanisms relevant to any given selfish mitochondrial genome therefore requires that our model be integrated with experiments to measure both selection and drift within hosts, in addition to inter-organismal selection.

### Integrating theory with experiment reveals the evolutionary mechanisms of a selfish mitochondrial genome

As a first application of the approach, we inferred the mechanisms determining the form of the stable mtDNA frequency distribution for the well-characterized selfish mitochondrial genome *uaDf5*. To measure intra-organismal selection in the absence of inter-organismal competition, we tracked mutant frequency longitudinally, between isolated individual parents and their respective age-matched progeny

(Fig. 3a, see Methods)<sup>26</sup>. At the organismal level, fitness is expected to depend on heteroplasmic frequency, which itself is in flux over the course of organismal development due to intra-organismal selection and drift. This greatly limits our ability to assess the inter-organismal component of selection directly by comparing the fitness of different heteroplasmic hosts. To overcome this challenge and estimate the organism-level fitness effects of harboring selfish mtDNA, we instead competed a diverse cohort of heteroplasmic animals against their homoplasmic-wildtype counterparts on the same food plate (Fig. 3b). Since the heteroplasmic hosts are taken from a stock population stably maintaining the selfish mtDNA, the strength of inter-organismal selection against the mitochondrial mutation reflects the overall fitness of the entire heteroplasmic cohort relative to a population lacking the selfish genome, as well as the rate of de novo loss of heteroplasmy by genetic drift. To control for factors other than inter-organismal selection against heteroplasmic hosts, we also propagated non-competed populations as controls, which consist of only heteroplasmic animals. Finally, we also collected information on the shape of the stable mutant mtDNA frequency distribution by measuring *uaDf5*



**Fig. 3 | Integration of benchtop experimentation with theoretical modeling reveals the stable maintenance of a selfish mitochondrial genome via a combination of multilevel selection and genetic drift.** **a–c** Empirical data (provided in Source Data file). **a** Pairwise comparisons of mutant frequency (*uaDf5*) between parent and progeny for the purpose of measuring intra-organismal selection on a selfish mitochondrial genome ( $n = 30$  parent-progeny lineages). Two-tailed Wilcoxon matched-pairs test. **b** Inter-organismal selection against *uaDf5*, measured by directly competing heteroplasmic carriers of the selfish genome against homoplasmic (wildtype) animals on the same food plate, in a mixed population ( $n = 8$  competed populations). Data shown as the log of the heteroplasmic over homoplasmic fraction of each replicate competed population (see Methods); regression lines are fit to the data from each replicate population by least squares. **c** histogram of mutant frequencies of age-synchronized adults ( $n = 81$ ) sampled from a

heteroplasmic stock population (same data as in Fig. 1b). Data in **a–c** are taken from our prior study<sup>26</sup>. **d–f** Maximum-likelihood model of the intra-organismal (**d**) and inter-organismal (**e**) fitness effects, each as a function of mutant frequency, and the most evolutionarily stable mutant frequency distribution (**f**). Each plot in **d–f** shows the maximum-likelihood estimates corresponding to the empirical data in **a–c** (solid black lines in **d–f**) plus 100 simulated data sets for parametric bootstrapping (red lines) to visualize the confidence estimates for each plot. Model parameters specifying intra-organismal selection, inter-organismal selection, and intra-organismal genetic drift (**d**, inset) were collectively estimated using a joint maximum-likelihood approach that combines all data sets in (**a–c**). Error bars in genetic drift plot (**d**, inset) indicate mean, minimum, and maximum bootstrap values. See Supplementary Data 1 for empirical and bootstrap model parameters.

frequency in multiple fertile adults sampled from a heteroplasmic stock population (Fig. 3c).

Using the above data (Fig. 3a–c), we inferred the parameters describing organism-level fitness, intra-organismal fitness and drift by maximum likelihood, employing a parametric bootstrap procedure to assess uncertainty (Fig. 3d–f, maximum likelihood results in black, bootstrap replicates in red; see Methods for details; Supplementary Figs. 2–6 confirm that our computational procedure can recover the true parameters for simulated data). For *uaDf5*, we found strong evidence for negative frequency-dependent intra-organismal selection (Fig. 3d), where the mutant genome has an intra-organismal advantage when present at low heteroplasmic frequency ( $w_{intra}(0) = 1.53$ , 95% bootstrap confidence interval 1.30–1.88), and fitness decreases with frequency so that it likely faces an intra-organismal disadvantage at high frequency ( $w_{intra}(1) = 0.93$  for maximum-likelihood estimate and  $w_{intra}(1) < 1$  for 69% of bootstrap samples;  $w_{intra}(0) > w_{intra}(1)$  for 100% of bootstraps). Although our measure of intra-organismal drift is intended to be phenomenological, our maximum likelihood estimate of  $N = 25$  (Supplementary Fig. 7 and Fig. 3d inset, 95% bootstrap confidence interval 19–42) is consistent with previous reports of the estimated effective mtDNA bottleneck in the developing germline<sup>47</sup>. For inter-organismal selection (Fig. 3e), we infer a strong threshold-like response, with essentially no fitness cost at low heteroplasmic frequency ( $w_{org}(z) > 0.99$  for all  $z < 0.5$  in 100% of bootstrap samples), with the onset of a substantial fitness cost once the mutant reaches sufficiently high frequency (organismal fitness is decreased by half,

$w_{org}(z) = 0.5$ , at  $z = 0.76$ , 95% bootstrap confidence interval 0.75–0.80). Together, these parameters predict a strongly left-skewed stable mutant mtDNA frequency distribution among hosts (Fig. 3f), as observed in the data (Fig. 3c), a moderate growth defect for the heteroplasmic relative to the wildtype sub-population (relative growth rate of 0.77, 95% bootstrap confidence interval 0.74–0.81), and a low rate of de novo loss of heteroplasmy (per-generation probability of mutant loss of  $1.5 \cdot 10^{-4}$ , 95% bootstrap confidence interval  $1.8 \cdot 10^{-6}$ – $7.7 \cdot 10^{-4}$ ). Interestingly, although the data are consistent with either a gradual or sudden loss of organism fitness at high *uaDf5* frequency (Fig. 3e, two clusters of slopes), as noted previously the range of mean organism-level fitness across the bootstraps is quite limited, indicating that this uncertainty in the precise form of the frequency-fitness relationship does not substantially affect our estimate of mean organism-level fitness. Finally, variation in  $N$  is predicted to substantially impact both the spread and location of the stable *uaDf5* frequency distribution (Supplementary Fig. 9a–c), consistent with the mixed-mechanism hypothesis (Fig. 2g–i). These results suggest that intra-organismal selection, inter-organismal selection, and genetic drift are all important in the dynamics of *uaDf5*.

### Diverse evolutionary mechanisms in the dynamics of selfish mitochondrial genotypes

We sought to compare our results for *uaDf5* with a larger panel of selfish mitochondrial genotypes that exhibit a greater diversity of stable mutant frequency distributions (Fig. 1a, b). These include

mutations affecting ETC complex I (*mpt2*), complex III (*mptDf3*), and complex IV (*mptDf2* and *mpt4*), as well as affecting tRNA genes (*mptDf2* and *mptDf3*). We measured selection and collected mutant frequency distribution data (Fig. 4), using the same experimental approach used for *uaDf5*, and applied our theoretical analysis to these data.

Overall, we find a surprising diversity of evolutionary mechanisms at play across this panel (Fig. 5), representing all three of the distinct mechanisms described in Fig. 2. For example, although *mptDf2* and *mpt4* have qualitatively similar stable mutant frequency distributions (Fig. 1b), *mptDf2* is governed by the mixed mechanism, similar to *uaDf5*, whereas the distribution for *mpt4* is determined by strong multilevel selection, with little influence of drift. The difference can be seen in Fig. 5a versus Fig. 5d, where frequency-dependent intra-organismal selection on *mptDf2* exhibits an internal equilibrium above which selection favors the wildtype (maximum likelihood  $z^* = 0.85$ , 95% bootstrap confidence interval  $> 0.83$ ) whereas *mpt4* is constitutively favored by intra-organismal selection under the maximum likelihood fit ( $w_{intra}(z) > 1$  for all  $z$  for 85% of bootstrap samples,  $z^*$  is greater than any empirically observed  $z$  in Fig. 4f for 99% of bootstrap samples, and the few bootstrap replicates exhibiting  $z^* < 1$  have  $w_{org}(z^*) < 0.1$ ). Concordant with this analysis of evolutionary mechanisms, we find that the stable *mptDf2* frequency distribution is sensitive to changes in the mtDNA bottleneck size (Supplementary Fig. 9f, compare Fig. 2i) whereas the stable *mpt4* frequency distribution is relatively insensitive (Supplementary Fig. 9i, compare Fig. 2f). In contrast, we find that *mpt2* persists primarily through frequency-dependent intra-organismal selection (Fig. 5g–i), with a stable mutant frequency distribution centered on  $z^*$  ( $z^* = 0.68$ , 95% bootstrap confidence interval 0.62–0.89), where *mpt2* frequency is too low to inflict a substantial organismal fitness cost ( $w_{org}(z^*) > 0.99$  for maximum-likelihood estimate and  $w_{org}(z^*) > 0.96$  for 95% of bootstrap samples). Although *mpt2* results are consistent with a negligible role for drift (Fig. 5g and Table 1, 95% bootstrap confidence upper limit of  $N = 1000$ ), we also find evidence of only weak selection at most (Fig. 5h and Table 1, relative growth rate  $> 0.99$  for maximum-likelihood estimate and  $> 0.95$  for 95% of bootstrap samples). Furthermore, our analysis of the effects of changing the effective mtDNA bottleneck  $N$  suggests that the width of the stable *mpt2* frequency distribution is highly sensitive to genetic drift (Supplementary Fig. 9l, compare to Fig. 2c). Our modeling results for *mptDf3* are more ambiguous (Fig. 5j–l and Table 1), due to uncertainty in the strength of inter-organismal selection. Like *mpt2* but in contrast to the other genotypes, average *mptDf3* frequency is not significantly different from  $z^*$  (Supplementary Table 2), consistent with frequency-dependent intra-organismal selection. Finally, we find no significant inter-organismal selection against *mptDf3* (Fig. 4k and Table 1, relative growth rate 95% bootstrap confidence upper limit of approximately 0.99), consistent with the maintenance of *mptDf3* via intra-organismal dynamics. We conclude that although different evolutionary mechanisms maintain the heteroplasmic state, they also give rise to idiosyncratic mutant frequency distributions.

Finally, we sought to explore the potential for long-term maintenance of the heteroplasmic state. We reasoned that high heteroplasmic frequency and a strong intra-organismal advantage (in the case of *uaDf5*, *mptDf2*, and *mpt4*) limits the probability of de novo mutant loss in finite populations, while a minimal organism-level fitness cost (in the case of *mpt2* and *mptDf3*) limits the probability of successful invasion by a de novo homoplasmic lineage. Using the inferred per-generation probability of mutant loss due to drift (Table 1), together with the organism-level fitness benefit of the homoplasmic-wildtype state and previously published theory<sup>48</sup>, we calculated likely persistence times of the heteroplasmic state, defined as the number of generations until the probability of successful invasion is  $> 0.5$ , for various values of effective organism population size  $N_e$ . Prior estimates of  $N_e$  typically range from 200 to  $10^4$  for wild

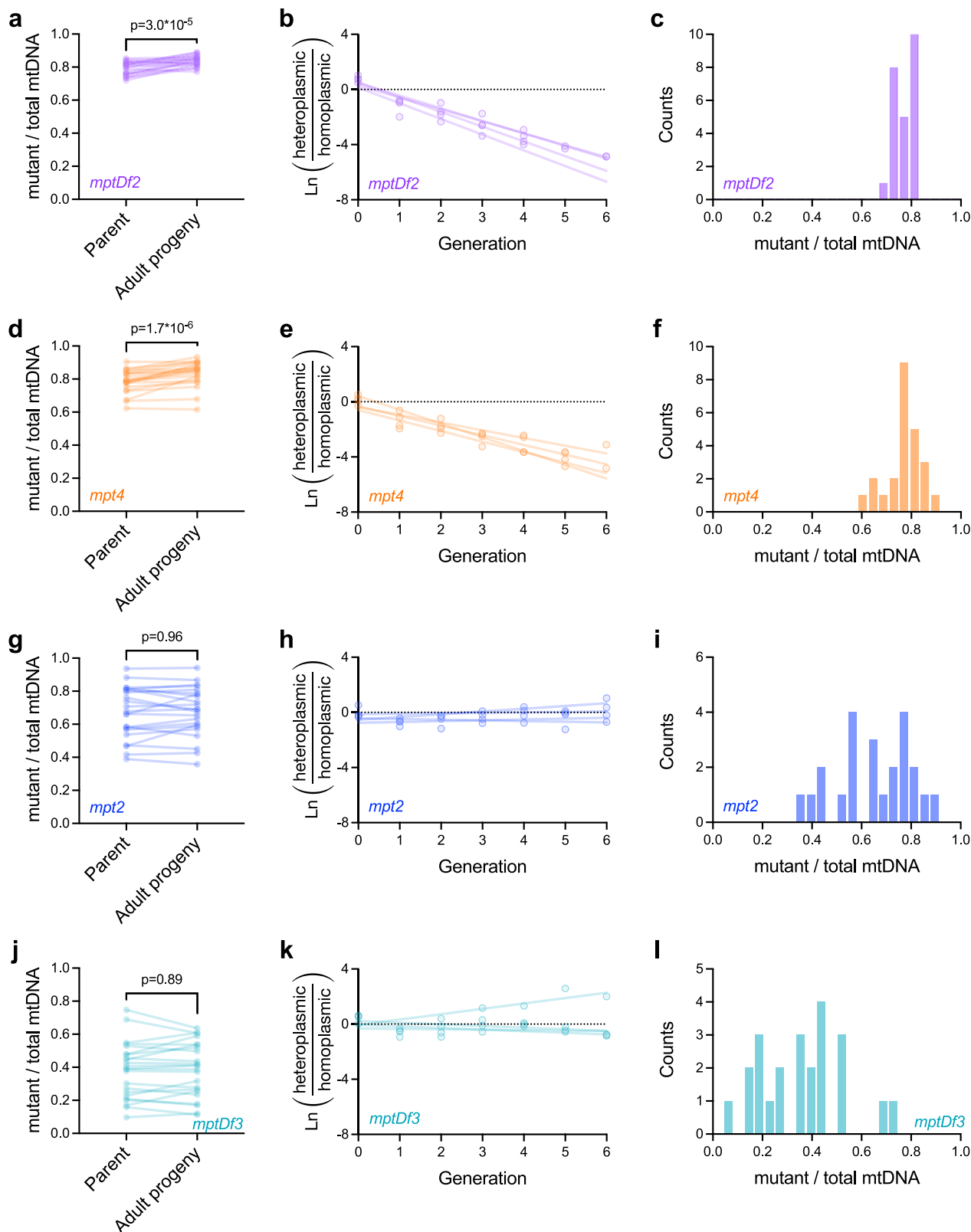
populations of *C. elegans*<sup>49</sup>, whereas maintenance of laboratory strains involves the frequent transfer of small founding populations to fresh food plates, resulting in an approximate per-generation bottleneck on the order of tens or hundreds of *C. elegans*<sup>50</sup>. Modeling similar variation in  $N_e$ , we estimate a wide range of persistence times, reflecting joint uncertainties in organism-level fitness and probability of loss (Table 1). Unsurprisingly, our estimates are most consistent with long-term persistence of the heteroplasmic state in small to mid-sized populations (Supplementary Table 1). With an  $N_e$  of 1000, for example, our *uaDf5* data are consistent with a median time till invasion on the order of hundreds of generations (95% bootstrap confidence upper limit of 649 generations). Likewise, our results for *mptDf2*, *mpt4*, and *mpt2* are consistent with a median time till invasion of  $1.6 \times 10^{19}$ ,  $7 \times 10^{19}$ , and  $3.3 \times 10^8$  generations, respectively (95% bootstrap confidence upper limits). For even smaller populations with an  $N_e$  of 100, the data for *mptDf3* are likewise consistent with a median time till invasion on the order of hundreds of generations (95% bootstrap confidence upper limit of 104 generations). We conclude that conditions influencing organism population size on one hand, and conflicting selection pressures on the other, jointly determine which mtDNA mutations are able to persist on long-term evolutionary timescales.

## Discussion

Major evolutionary transitions require cooperative interactions that incentivize group members to assume the benefits of cooperation without the cost of reciprocating<sup>6–9</sup>. How might cheaters stably persist despite deleteriously affecting the cooperators on which they depend? This question carries broad implications ranging from evolutionary theory to the social and biomedical sciences, as prior exploration of cooperation and cheating spans multiple systems and different levels of scale<sup>12,14,51</sup>. We sought to systematically compare separate instances of cooperator-cheater dynamics, by leveraging a collection of selfish mitochondrial genotypes existing in a uniform genetic and environmental background.

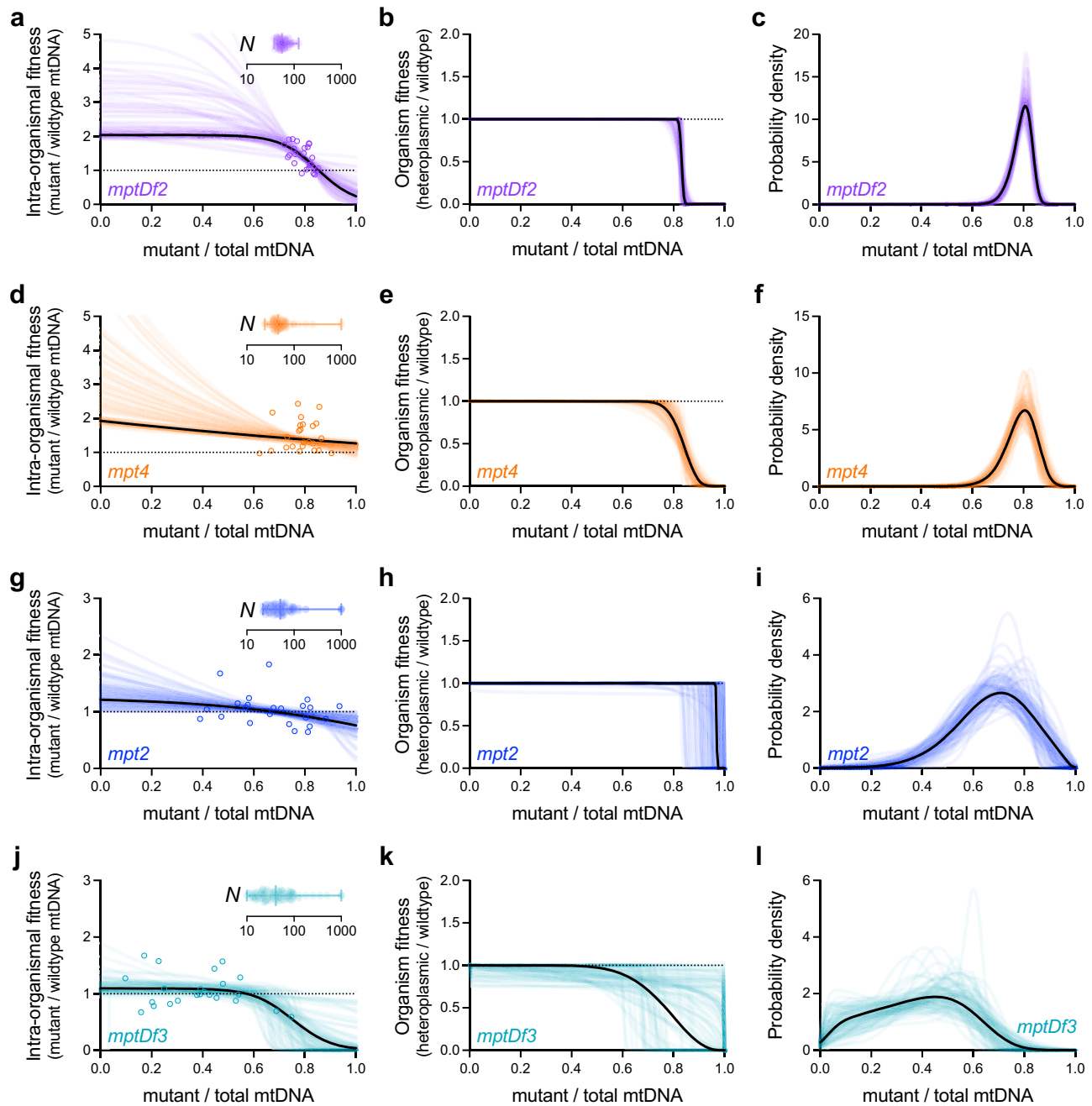
The selfish proliferation of mitochondrial mutations is responsible for a number of human diseases, affecting an estimated 1 in 4300 individuals, while an additional 1 in 200 healthy humans carry potentially disease-causing mitochondrial mutations<sup>46</sup>. The same mutation may thus cause disease in one individual but not another, due to variation in mutant frequency<sup>45,46</sup>. Consistent with this, disease-causing mutations vary substantially in their ability to propagate from mother to offspring<sup>43,52</sup>. Predicting the inheritance and development of mtDNA-associated disorders therefore requires a deep understanding of the underlying evolutionary forces, and how they explain observed mutant frequencies. To this end, prior work generally consists of either theoretical modeling or empirical study of mitochondrial mutations<sup>29,43,53–55</sup>, making it difficult to integrate theory with experiment. Although some modeling studies incorporate empirical data<sup>29,55–57</sup>, these focus on heteroplasmy dynamics within organisms, or within parent-progeny lineages, making it difficult to combine the levels of selection into a complete evolutionary picture. To address these challenges, we developed a hybrid approach that integrates theoretical modeling with experiments designed to individually probe the levels of selection and the mutant frequency distribution.

The variant *uaDf5* is an exemplary model selfish mitochondrial genome, undergoing positive intra-organismal selection at the expense of host fitness<sup>26,30,32,34–36,38</sup>. We expand the understanding of *uaDf5* in numerous key ways. Prior studies have found that *uaDf5* proliferates not merely in spite of, but at least partly because of, its negative effect on mitochondrial function<sup>32,34,36</sup>. Conversely, hosts are equipped with mechanisms that limit *uaDf5* proliferation<sup>32,34,36,37,39</sup>. Together, these findings suggest opposing selection forces at the intra-organismal level, in addition to inter-organismal selection. Consistent with this, and with previous reporting<sup>26,38</sup>, our modeling results confirm negative frequency-dependent selection for *uaDf5*, with rising



**Fig. 4 | Measurements of intra-organismal selection, organismal selection, and mutant frequency distributions across a panel of putatively selfish mitochondrial genome variants. a** *mptDf2* frequencies in isolated parent-progeny lineages ( $n=24$ ) to measure intra-organismal selection, similar to Fig. 3d. **b** Relative proportions of *mptDf2* heteroplasmic versus homoplasmic-wildtype hosts per generation in populations ( $n=4$ ) of animals competing on the same food plates, to measure inter-organismal selection; regression lines are fit to the data from each

replicate population by least squares. **c** Histogram of *mptDf2* frequencies in age-synchronized adults sampled from heteroplasmic stock populations (the same individuals used as parent data points in **a**). **d–f** Similar to **a–c** but for *mpt4* (**d**,  $n=24$ ; **e**,  $n=4$ ). **g–i** Similar to **a–c** but for *mpt2* (**g**,  $n=23$ ; **h**,  $n=4$ ). **j–l** Similar to **a–c** but for *mptDf3* (**j**,  $n=23$ ; **k**,  $n=4$ ). Two-tailed Wilcoxon matched-pairs test (left column). All data shown here are provided in the Source Data file.



**Fig. 5 | Integrating theoretical modeling with experiment shows the maintenance of selfish mitochondrial genomes in a heteroplasmic state through multiple distinct evolutionary mechanisms.** Maximum-likelihood estimates of intra- and inter-organismal fitness effects (left and center columns, respectively), each as a function of mutant frequency, together with the resulting most evolutionarily stable mutant frequency distribution (right column). Similar to bottom row of Fig. 3, but for the 4 mitochondrial genotypes shown in Fig. 4: *mptDf2* (a–c, Supplementary Data 2), *mpt4* (d–f, Supplementary Data 3), *mpt2* (g–i, Supplementary Data 4), and *mptDf3* (j–l, Supplementary Data 5). Each plot shows

maximum-likelihood estimates corresponding to the empirical data shown in Fig. 4 (solid black lines) and 100 parametric bootstrap data sets (colored lines) to visualize confidence estimates. Model parameters specifying intra-organismal selection, organismal selection, and intra-organismal drift (left column, insets), were collectively estimated using a joint maximum-likelihood approach that combines all 3 corresponding empirical data sets in Fig. 4. Error bars in genetic drift plots (left column, insets) indicate mean, minimum and maximum bootstrap values. See Supplementary Data 2–5 for empirical and bootstrap model parameters.

frequency leading to loss of its intra-organismal advantage, in addition to an organism-level fitness cost. More strikingly, we establish theoretically that the shape of the mutant frequency distribution contains information about the underlying evolutionary forces. Thus, using *uadF5* data, we show how the combination of multilevel selection and genetic drift explains a previously observed peculiarity, namely the skewed frequency distribution that concentrates most individuals toward the high end<sup>34</sup>.

How generalizable are the results concerning *uadF5*? To answer this, we expanded our analysis to other stably propagating heteroplasmies, each of which features a mutation known to deleteriously affect at least one essential gene. Like *uadF5*, some variants persist over time despite a heavy organism-level fitness cost, due to a strong intra-organismal advantage. These dynamics can be observed in three of the five heteroplasmies—*uadF5*, *mptDf2*, and *mpt4*—and are remarkably consistent with recent mammalian work. For example, some mtDNA



**Table 1 | Population genetic characteristics of five selfish mitochondrial genomes**

Genotype	$w_{intra}(0)$	$w_{intra}(1)$	$z(w_{org} = 0.5)$	Growth rate	$N$	Average $z$	SD $z$	Probability of loss
<i>uaDf5</i>	1.53	0.93	0.76	0.77	25	0.63	0.12	$1.5 \times 10^{-4}$
	1.30–1.88	0.07–1.13	0.75–0.80	0.74–0.81	19–42	0.60–0.65	0.10–0.14	$1.8 \times 10^{-6}$ – $7.7 \times 10^{-4}$
<i>mptDf2</i>	2.04	0.24	0.83	0.41	52	0.79	0.038	$2.1 \times 10^{-11}$
	1.94–5.92	0.02–1.11	0.82–0.84	0.35–0.46	40–94	0.78–0.80	0.030–0.071	$1.7 \times 10^{-23}$ – $8.3 \times 10^{-6}$
<i>mpt4</i>	1.93	1.26	0.84	0.47	47	0.79	0.063	$4.4 \times 10^{-9}$
	1.81–6.17	0.88–1.33	0.81–0.86	0.43–0.53	30–177	0.76–0.82	0.045–0.094	$4.6 \times 10^{-24}$ – $5.9 \times 10^{-6}$
<i>mpt2</i>	1.21	0.76	0.97	>0.99	46	0.67	0.15	$6.4 \times 10^{-5}$
	1.01–2.02	0.21–0.99	0.83–1	0.93–1	23–1000	0.61–0.72	0.10–0.19	$1.1 \times 10^{-9}$ – $1.8 \times 10^{-3}$
<i>mptDf3</i>	1.10	0.04	0.77	0.93	23	0.38	0.18	0.02
	0.97–1.47	0–0.87	0.60–1	0.75–0.99	12–500	0.32–0.46	0.12–0.21	$3.1 \times 10^{-4}$ –0.06

Each entry: maximum likelihood estimate (first line), 95% bootstrap confidence interval (second line).

$w_{intra}$ : relative fitness of mutant genome (intra-organismal level).

$w_{org}$ : relative fitness of mutant genome (organismal level).

$N$ : effective mtDNA bottleneck size per organism, per generation.

$z$ : average mutant mtDNA frequency among heteroplasmic hosts.

mutations in mice and humans rise in frequency from parents to offspring, in a manner consistent with negative frequency-dependent selection<sup>43</sup>, suggesting a cheating behavior in human mitochondrial disease similar to our findings. For some variants, however, we empirically observed no inter-organismal selection, despite deletions in respiratory genes in the mutant genomes. This is true for *mpt2* and possibly *mptDf3*, containing deletions in the genes *ND5* (ETC complex I) and cytochrome b (complex III), respectively, although *mptDf3* results are consistent with a mild organism-level fitness cost. How might a lineage lacking an essential gene stably persist with no empirically observed fitness effects? We reason that stably persisting heteroplasmies require at least weak fitness effects, since truly neutral genotypes can be expected to readily drift to either extinction or fixation. Indeed, our modeling reveals evidence of a weak but nonzero intra-organismal fitness advantage for *mpt2* and *mptDf3*, particularly when present at low frequency (Table 1). Moreover, our theoretical results show that a mutant genome can selfishly persist via frequency-dependent intra-organismal selection rather than multilevel selection. Specifically, the mutation confers a weak intra-organismal advantage that disappears at a frequency that is still low enough that the organism-level fitness cost is negligible. One possible molecular basis for this is that some mutant genomes are more vulnerable to cheater-suppression mechanisms, namely mitochondrial autophagy, consistent with prior empirical work<sup>34,36,37,39,58–60</sup>. Alternatively, the dynamics of mitochondrial fission and fusion may permit the diffusion of gene products throughout the mitochondrial network, resulting in the genetic complementation of a mutation by nearby wildtype genomes. In either case, such scenarios may enable a mutation to selfishly persist in a heteroplasmic state—that is, at the cost of wildtype genomes—by remaining neutral or nearly neutral at the organismal level. This is consistent with prior reporting that the persistence of cheater entities in other systems is often maintained by frequency-dependent selection<sup>12</sup>.

We note two key limitations of this study, which easily serve as a basis for future work. First, in addition to opposing selection forces, the introduction of selfish genomes by de novo mutation, a factor not considered in this study, suggests that selfish mitochondrial genome dynamics represent a persistent phenomenon. Consistent with this, previous mutation-accumulation experiments have identified selfish mtDNA variants in *C. elegans* and the closely-related species *C. briggsae*<sup>25,61</sup>. Together with our study, these findings suggest that a heteroplasmic state may evolutionarily persist by a balance of selection forces, recurring reintroduction of selfish mtDNA, or a combination of these factors.

Conditions beyond the mitochondria represent a second limitation. While our study employs a single host genotype, previous

research shows the nuclear genome to influence heteroplasmy dynamics. For example, genetic regulators of mitochondrial biogenesis in response to stress, such as ATFS-1 and FoxO (DAF-16 in *C. elegans*), as well as expression level of mtDNA replication machinery, are important determinants of selfish mtDNA propagation<sup>26,32,34,36,41</sup>. Moreover, proteins with a more direct role in mitochondrial biogenesis and mtDNA replication, such as POLG and the mtDNA-associated protein TFAM, or that promote binding of replication machinery to mtDNA, reportedly mediate selfish mtDNA proliferation<sup>30,32,36</sup>. Conversely, genes involved in mitochondrial fission and autophagy reportedly contribute to intra-organismal selection *against* mutant mtDNA<sup>34,36,37,39,58–60</sup>. Given these findings, we propose that host genotypes and environmental factors, especially those linking physiological stress with mtDNA replication and turnover, will be important determinants of the cooperator-cheater dynamics described here and should be considered in future research.

## Methods

### Nematode husbandry

*C. elegans* strains used in this study were maintained on 60-mm standard nematode growth medium (NGM) plates (for measuring intra-organismal selection), or 100-mm NGM plates (for measuring inter-organismal selection), seeded with live OP50-strain *E. coli* bacteria as a food source. Nematode strains were incubated at 20 °C. In addition to the Bristol wildtype strain and the heteroplasmic *uaDf5* strain featured in prior work<sup>26</sup>, five additional *C. elegans* strains were used in this study. These consisted of heteroplasmic mutant genomes *mptDf2*, *mpt4*, *mpt2*, and *mptDf3*, each crossed into the nuclear background of the Bristol strain.

### Nuclear genome exchange in heteroplasmic strains

To ensure that our analysis was not confounded by variation in the nuclear genome, we completely exchanged the nuclear genome of each heteroplasmic strain with the nuclear genome of wildtype (Bristol strain) *C. elegans*, using a previously published unigametic inheritance method<sup>62</sup>. This method enables the complete replacement of the nuclear genome within two generations, by leveraging the activity of *gpr-1*, which encodes a G-protein regulator that regulates the forces exerted on the microtubules during mitosis. Over-expression of *gpr-1* increases the pulling forces on the pronuclei during prometaphase, resulting in the segregation of the paternal and maternal genomes into separate embryonic cell lineages<sup>62</sup>. The germline of the hermaphrodite consequently inherits the nuclear genome of only one parent, allowing us to bypass the need for multiple generations of backcrossing. Thus, all heteroplasmic strains used in this study have identical nuclear backgrounds. Briefly, each heteroplasmic strain was crossed to the *gpr-*

I over-expression strain PD2220 following Mendelian genetics. Next, hermaphrodites of the stable *gpr-1* over-expression heteroplasmy strains were crossed to wildtype males. Non-Mendelian hermaphrodite progeny from these crosses in which the paternal nuclear background is unigametically inherited in the germline cell lineage (determined by fluorescent pharyngeal mosaicism patterning, as described in ref. 62) were individually propagated. Stock strains were established from the progeny of these animals as they have a complete wildtype nuclear genomic background and retain the given heteroplasmy.

### DNA preparation

To prepare animals for quantification of mutant mtDNA frequency, nematodes were transferred to sterile PCR tubes or 96-well PCR plates containing lysis buffer with 100 µg/mL proteinase K. Lysis buffer consisted of 50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM MgCl<sub>2</sub>, 0.45% Tween 20, 0.45% NP-40 (IGEPAL), and 0.01% gelatin, in deionized water. Volume of lysis buffer varied by worm count: 10 µL for individual adults of the parent generation and 20 µL for pooled adult progeny for measuring intra-organismal selection, and 50 µL for pooled animal lysates from the competition experiments for measuring inter-organismal selection. After transferring worms to lysis buffer, each tube or plate was immediately sealed and incubated at -80 °C for 10 minutes to rupture nematode cuticles, followed by lysis incubation at 60 °C for 60 minutes (90 minutes for pooled nematodes), and then at 95 °C for 15 minutes to inactivate the proteinase K. Nematode lysates were then kept at -20 °C for stable long-term storage until use for genotyping and quantification.

### Quantifying mtDNA genotype frequencies

Mutant mtDNA frequencies were quantified as described previously for *uaDf5*<sup>26</sup>, using droplet digital PCR (ddPCR). Nematodes were lysed as described above and diluted in nuclease-free water with a dilution factor varying depending on nematode concentration: 200x for single adults, 1000x for pooled adults (intra-organismal selection experiment), and 20,000x for pooled nematodes of mixed age (inter-organismal selection experiment). For PCR amplification, 5 µL of diluted lysate was combined with 0.25 µL of 10-µM of each oligonucleotide primer as needed depending on genotype:

**mptDf2.** Forward 1: 5'-GGATTGGCAGTTTGATTAGAGAG-3'  
Reverse 1: 5'-AAGTAACAAACTAAAACCTCCCAAC-3'  
Forward 2: 5'-CGTGCTTATTTTTGGGCTGC-3'  
Reverse 2: 5'-CTTTAACACCTGTTGGCACTG-3'

**mpt4.** Forward 1: 5'-CGGTGGTTTTGGTAACTG-3'  
Reverse 1: 5'-TCATAGTGTAACACCCGTGAAAATCC-3'  
Forward 2: 5'-TGATCCAAGAACTGGAGGTAATC-3'  
Reverse 2: 5'-CCTGTTGGCACTGCAATAAC-3'

**mpt2.** Forward 1: 5'-GAAGAAGGTGGTAGCCTTGAGGAC-3'  
Reverse 1: 5'-CGTATAAGAAAAGTCTTGGGATGTTAAG-3'  
Forward 2: 5'-GGATTAATTTTCTCAAGGGGTGCTG-3'  
Reverse 2: 5'-CTTTTTCAAGACGAAAACCTGTAACC-3'

**mptDf3.** Forward 1: 5'-CCCTGAAGAGGCTAAGAATATTAGG-3'  
Reverse 1: 5'-GGCAATGTCACCAACATCC-3'  
Reverse 2: 5'-CCCAATACAATAACTAGAACTAGCTCAGC-3'

Mixtures of dilute lysate and primer were combined with 12.5 µL of Bio-Rad QX200™ ddPCR EvaGreen Supermix and nuclease-free water to a volume of 25 µL in Eppendorf™ 96-well twin.tec™ PCR plates. Droplet generation and PCR amplification were performed according to manufacturer protocol. Wildtype and mutant-specific primers were combined in the same reaction, and each droplet was scored as containing either wildtype or mutant mtDNA using the 2-dimensional

(518 nm and 554 nm dual-wavelength) clustering plot option in the Bio-Rad QuantaSoft™ program.

### Intra-organismal selection assay

The strength of intra-organismal (within-host) selection on mutant mtDNA was measured longitudinally across isolated parent-progeny lineages, as previously described<sup>26</sup>. Briefly, multiple L4-stage (late larval) heteroplasmic animals were picked at random under a dissecting microscope from stock populations carrying each heteroplasmy that had been crossed into the Bristol strain (wildtype) nuclear background. These larvae were transferred to fresh NGM plates seeded with live OP50 *E. coli* bacteria as a food source and incubated for 2 days at 20 °C to allow adult maturation. The day-2 adults were individually segregated by transferring each onto a fresh food plate and incubated for 4 hours at 20 °C to produce embryos that are age-synchronized to within a four-hour window. Each parent was then individually lysed. After 4 days of continued incubation at 20 °C, the progeny had progressed from embryos to day-2 adults, reaching the same age at which their respective parents were lysed. Adult progeny were lysed at this point to obtain progeny that are age-matched to their parents, to control for age-dependent differences in mutant mtDNA levels. Progeny from each parent were lysed in pools of 3 to minimize the confounding effect of random drift. Each parent-progeny lineage was individually segregated from the rest, to ensure that mutant mtDNA frequency from each progeny lysate was being compared with that of its own respective parent, thereby minimizing the confounding effect of competition between lineages (organismal selection). Mutant mtDNA frequency of parents and progeny was determined for each heteroplasmy using ddPCR as described above, across multiple replicate parent-progeny lineages for statistical power.

### Experimental evolution (inter-organismal selection)

Selection against mutant mtDNA that occurs strictly at the level of host fitness was measured using an inter-organismal competition experiment similarly to the experiment previously described<sup>26</sup>. Briefly, for each mutant mtDNA variant, heteroplasmic nematodes carrying mutant mtDNA in the Bristol nuclear background were combined with Bristol-strain nematodes on 10-cm NGM plates seeded with live OP50 *E. coli* bacteria as a food source. Approximately 500 nematodes were transferred to each plate. In addition to 4 replicate competition lines for each heteroplasmy, 4 non-competed control lines were established by transferring only heteroplasmic animals onto their own food plates, with no homoplasmic-wildtype animals to compete against. Every 3 days, the generation for each experimental line was reset; nematodes were washed off the plates using M9 buffer into a sterile 1.7 mL collection tube. Approximately 500 animals of mixed age from each line were transferred to a fresh food plate. Another 500 nematodes were lysed together in a single pooled lysate. To find the volume of suspension corresponding to approximately 500 nematodes, we estimated the number of nematodes per µL of buffer. Nematodes from these collection tubes were pipetted in fixed volumes onto microscope slides and visually counted under a dissecting microscope. This was done in triplicate. This experiment was continued for 6 consecutive generations.

The strength of selection was derived from the decline of the population-wide frequency of the mutant genome in the competed relative to the non-competed lines. We reasoned that the overall evolutionary dynamics of selfish mtDNA is the same between the non-competed control populations and the heteroplasmic fractions of the competed populations, since the presence or absence of homoplasmic-wildtype animals is the only difference. Specifically, at each generation, the population-wide mutant mtDNA frequency in each competed line was divided by the mutant frequency in the non-competed control lines, to estimate the remaining heteroplasmic

fraction of each competed population. To validate this approach, we previously estimated the heteroplasmic fraction by also sampling individuals from the competed populations, providing a similar estimate of selection (compare Fig. 2D and E of ref. 26), and in practice mtDNA frequency changes in non-competed controls are minor (these lines are taken from stock populations stably maintaining the heteroplasmic genomes), consistent with our assumption that the heteroplasmic fraction of the competed populations maintains a stable frequency distribution.

### Population-genetic model

To treat the evolutionary dynamics of multilevel selection and drift in a theoretical population-genetic framework, we constructed a model that couples a stochastic, frequency-dependent Wright-Fisher model for the evolution of mutant mtDNA frequencies within individuals together with a frequency-independent, deterministic model of inter-organismal selection. The intra-organismal fitness function measures the fitness of mutant mtDNA relative to wildtype mtDNA within an individual. We model the intra-organismal fitness function as a sigmoid function of mutant frequency:

$$w_{intra}(z) = \frac{\delta}{1 + e^{\gamma z + \epsilon}} \quad (1)$$

where parameters  $\delta$ ,  $\gamma$ , and  $\epsilon$  control the overall scale of fitness variation, the degree of frequency-dependence, and the position of the inflection point, respectively, as a function of mutant mtDNA frequency. The organism-level fitness function is modeled in terms of the fitness cost of carrying a selfish mitochondrial genome at frequency  $z$ :

$$w_{org}(z) = 1 - \frac{\int_0^z x^{\alpha-1}(1-x)^{\beta-1} dx}{\int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx} \quad (2)$$

This is identical to the complementary cumulative distribution function of a beta distribution with parameters  $\alpha$  and  $\beta$ , so that we necessarily have  $w_{org}(0) = 1$  (wildtype individuals have fitness 1) and  $w_{org}(1) = 0$  (individuals fixed for the selfish mutant are either sterile or nonviable). Importantly, when  $\beta = 1$  this expression reduces to  $w_{org}(z) = 1 - z^\alpha$ , corresponding to a fitness defect that scales as  $z$  to the power  $\alpha$ , or a linear fitness decline when both  $\alpha = 1$  and  $\beta = 1$ . This expression can thus encompass a variety of simple models for how mutant mtDNA frequency impacts organismal fitness, in addition to more complex threshold-like models. Note also that while the organism-level fitness function depends on the mutant mtDNA frequency of the individual, in our model inter-organismal selection is frequency-independent in that the fitness of an individual does not depend on the composition of the remainder of the population. See Supplementary Fig. 1 for a visualization of the effects of varying each of these parameters.

To combine this intra- and inter-organismal selection with stochastic inheritance arising from intra-organismal genetic drift, we construct an  $N + 1$  by  $N + 1$  matrix,

$$M(i, j) = p\left(j; N, \frac{i w_{intra}(i/N)}{i w_{intra}(i/N) + (N - i)}\right) \times w_{org}(j/N), \quad (3)$$

where the indices  $i$  and  $j$  run from 0 to  $N$ ,  $p(j; N, q)$  is the binomial probability of a total of  $j$  successes out of  $N$  attempts with expected mutant mtDNA frequency  $q$  among the offspring of a parent with mutant mtDNA frequency  $i/N$  (which itself is given by  $i w_{intra}(i/N) / (i w_{intra}(i/N) + (N - i))$ ). Thus  $M(i, j)$  gives the expected proportion of progeny with a mutant mtDNA frequency of  $j/N$  for a parent with a mutant mtDNA frequency  $i/N$ . Writing the probability distribution of heteroplasmic mutant frequencies at time  $t$  as a length  $N$  vector  $f_t$ , the

time evolution of this probability distribution is given by:

$$f_{t+1}^T = \frac{f_t^T M}{f_t^T M \mathbf{1}_N}, \quad (4)$$

where  $M$  is the  $N$  by  $N$  principal submatrix of  $M$  obtained by omitting the zeroth row and column, and  $\mathbf{1}_N$  is the length  $N$  vector of all 1s. By the Perron-Frobenius theorem<sup>63</sup>, as  $t$  goes to infinity,  $f_t$  converges to a stable distribution  $f^*$  given by the dominant left eigenvector of  $M$ . The corresponding eigenvalue gives the expected number of heteroplasmic offspring per heteroplasmic parent under the stable distribution  $f^*$ . The expected number of wildtype offspring per heteroplasmic parent (spontaneous reversion from heteroplasmy to homoplasmic-wildtype) under the stable distribution  $f^*$  is given by:

$$r = (f_{N+1}^*)^T M e_0, \quad (5)$$

where  $f_{N+1}^*$  is the vector of length  $N + 1$  whose zeroth entry is 0 and whose remaining entries are given by  $f^*$ , and  $e_0$  is the length  $N + 1$  vector whose zeroth entry is 1 and whose remaining entries are 0.

### Maximum likelihood inference

We estimate the model parameters for each mutant mtDNA genotype using a joint maximum likelihood approach (that is, a combined fit to the data from the intra-organismal selection experiment, inter-organismal selection experiment, and the sampled stable mutant frequency distribution) with parameter uncertainty assessed via a parametric bootstrap approach. Model fitting was conducted using a custom Python script. Code and source data available at <https://github.com/bgitschlag/MiSelf> and in Supplementary Code 1. All source data are also provided in the Source Data file.

For intra-organismal selection experiment, we model the mutant frequency within the progeny,  $z_{obs,t+1}$ , as a function of the mutant frequency  $z_t$  of the parent as:

$$z_{obs,t+1} = \frac{z_t w_{intra}(z_t)}{z_t w_{intra}(z_t) + (1 - z_t)} + e_{intra}, \quad (6)$$

where  $e_{intra}$  is a normally distributed random variable with mean 0 and variance given by the free parameter  $\sigma_{intra}^2$ . For the inter-organismal selection experiment, the population-wide fraction of heteroplasmic individuals,  $\phi$ , changes across generations due to both the organism-level fitness cost and spontaneous loss of the mutant genome,  $r$ . At each generation,  $\phi$  experiences a growth rate, relative to the wildtype, of  $\kappa = (f^*)^T M \mathbf{1}_N$ , and a rate of mutant loss  $r$ , which collectively determine the fraction of heteroplasmic individuals in each subsequent generation:

$$\phi_{t+1} = \frac{\phi_t \kappa}{\phi_t \kappa + \phi_t r + (1 - \phi_t)}. \quad (7)$$

Because in the special case where mutant frequency declines solely due to inter-organismal selection (that is,  $r = 0$ ),  $\ln(\phi/(1 - \phi))$  decreases linearly in time, we model  $\phi$  on a log-odds scale:

$$\ln \frac{\phi_{obs,t}}{1 - \phi_{obs,t}} = \ln \frac{\phi_t}{1 - \phi_t} + e_{org}, \quad (8)$$

where  $\phi_{obs,t}$  is the empirically observed  $\phi$  at time  $t$ ,  $e_{org}$  is a normally distributed random variable with mean 0 and variance given by the free parameter  $\sigma_{org}^2$ . In addition, the initial frequency  $\phi_0$  is also inferred as a free parameter. Finally, we determine the likelihood of the observed draws from  $f^*$ . Because our theoretical distribution  $f^*$  is a discrete distribution (all possible frequencies are multiples of  $1/N$ ) but

our observations are continuous quantities, we construct a continuous analog of the discrete distribution  $f^*$  by applying a Gaussian smoother. We then calculate the likelihood under the resulting probability density function. Specifically, we convolve the discrete distribution described by  $f^*$  with a Gaussian random variable with mean 0 and standard deviation  $1/N$  (the standard deviation equal to the gaps between the discrete frequencies allowable under the Wright-Fisher model). This produces a distribution where the gaps between the multiples of  $1/N$  are smoothly filled without producing a substantial change to the overall shape of the distribution; see also our results in Supplementary Figs. 5 and 6, which indicate that our overall inference procedure, which includes the smoothing step, can accurately recover the ground-truth fitness functions and  $N$ .

In summary, our statistical model has free parameters,  $N$ ,  $\delta$ ,  $\gamma$ , and  $\epsilon$  (controlling intra-organismal selection and drift),  $\sigma_{intra}^2$  (controlling the noise variance of the intra-organismal selection experiment),  $\alpha$  and  $\beta$  (controlling the form of inter-organismal selection), and  $\sigma_{org}^2$  (controlling the noise variance in the inter-organismal selection experiment). Because all of these parameters are continuous except for  $N$ , we define a search space of discrete  $N$  values, and for each  $N$ , we numerically maximized the likelihood with respect to all other parameters. The initial search space for  $N$  consists of every integer between 10 and 96 but expands if the likelihood values for each  $N$  do not display a clear maximum on the interior of this range, using the same procedure described below for the bootstraps. We then selected the value of  $N$  that maximized the likelihood.

The maximum-likelihood parameter values were found using an optimization algorithm that conducts local rather than global optimization. To protect against the algorithm getting stuck in local optima, we repeated the maximum-likelihood inference multiple times for each empirical data set (representing mitochondrial genotypes *uaDf5*, *mptDf2*, *mpt4*, *mpt2*, and *mptDf3*), with the algorithm initialized at different regions of parameter space that correspond to different biological scenarios (the same sets of parameters in Supplementary Fig. 5a), as well as a neutral initialization represented by the line at  $w_{intra}=1$  for all  $z$  ( $\gamma=0$ ,  $\delta=2$ ,  $\epsilon=0$ ). For added robustness, we also performed the optimization procedure via a recursive approach when iterating over different values of  $N$ . According to this approach, for each value of  $N$ , the optimization algorithm is run twice, initialized on both the aforementioned neutral point (together with  $\alpha=1$ ,  $\beta=1$ ) and on the maximum-likelihood parameters from the previous  $N$ . This approach provides the optimizer with an initial guess in a biologically plausible region of parameter space, but also allows the optimizer to switch to a different region of parameter space if a better solution is achieved starting from the neutral initialization. The overall maximum-likelihood result are the parameters found to have the highest likelihood from among all approaches and initializations described above (in practice, these were consistently obtained via the recursive method).

To assess the robustness of the above procedure, we also implemented a model featuring a more complex error model for the intra-organismal selection experiment, which specifically models the effects of genetic drift on the deviations between observed and expected offspring mutant frequencies rather than treating these errors phenomenologically as Gaussian with unknown variance  $\sigma_{intra}^2$ . In particular, for each parental frequency  $z_t$ , we defined the likelihood of  $Z_{obs,t+1}$  as the corresponding density of a beta distribution whose mean is the expected  $z_{t+1}$  given by the intra-organismal fitness function and whose variance is given by the variance under the Wright-Fisher model:

$$\text{Var}_{z_{t+1}} = \frac{z_{\text{exp},t+1}(1 - z_{\text{exp},t+1})}{pN} \quad (9)$$

Here,  $p$  is the number of progeny sampled per parent (see section **Intra-organismal selection assay**), which accounts for the influence

of pooling the offspring prior to determining the mutant frequency. We compare these two alternative methods of evaluating the likelihood for the intra-organismal data, with respect to the estimation of  $N$  (Supplementary Fig. 7) and with respect to the estimation of the fitness functions and mutant frequency distributions (Supplementary Fig. 8). We find that these results between the two methods are primarily very similar, except for *uaDf5* where the beta distribution method favors a lower value of  $z^*$  and a shallower slope for  $w_{org}(z)$ , as also seen in a subset of the bootstrap replicates in Fig. 3d and e.

To estimate our confidence in these parameter estimates, we performed parametric bootstrapping by generating 100 simulated data sets per genotype, using the sample sizes, maximum-likelihood population-genetic parameters, and error estimates for the corresponding empirical data set. Specifically, to simulate the intra-organismal selection experiment, parent mutant frequencies  $z_t$  were randomly drawn from the Gaussian smoothed version of  $f^*$  and then the corresponding progeny frequencies  $z_{t+1}$  were randomly drawn according to our inferred Wright-Fisher process. To simulate the inter-organismal selection experiment, the expected heteroplasmic fraction  $\phi_t$  at each generational time  $t$  was calculated using the maximum-likelihood estimate of  $\phi_0$  combined with the  $r$  and  $\kappa$  values determined from our maximum-likelihood estimates. We then converted these expected fractions  $\phi_t$  onto a log odds scale, added Gaussian noise with variance  $\sigma_{org}^2$  and finally converted them back to a frequency scale. The samples from the mutant frequency distribution were simulated by drawing from the Gaussian smoothed version of  $f^*$  with sample sizes equaling that of the corresponding empirical data set. For each simulated data set, model parameters were re-estimated using the maximum-likelihood approach described above, with the exception that the optimization algorithm was initialized on the maximum-likelihood parameters for the empirical data. However, for some bootstrap simulations the maximum likelihood estimate of  $N$  was realized at or near the upper boundary of our initial search space. In such cases we thus examined larger values of  $N$ , up to 100 and then increasing by powers of  $\sqrt{2}$  (125, 177, 250, 354, 500, etc.). We continued in this manner until the maximum-likelihood estimate of  $N$  is no longer among the 2 consecutive largest values of  $N$  examined so far, or until reaching a maximum-likelihood  $N=1000$ , at which point we concluded that the magnitude of drift approximates that of the high- $N$  limit.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All source data needed to reproduce our results are publicly available and provided in this study. Data from experiments measuring *uaDf5* frequency and selection on *uaDf5* (Fig. 3a–c) were previously published<sup>26</sup>. These and the experimental data encompassing the remaining four genotypes (Fig. 4) are provided in the Source Data. These data can also be downloaded in a code-readable format at <https://github.com/bgitschlag/MiSelf>. The modeling parameters generated from the maximum-likelihood inference on the empirical and bootstrap data are provided in Supplementary Data Files 1–5. Source data are provided with this paper.

## Code availability

This study was conducted using the custom MiSelf Python software, included with this study as Supplementary Code 1. The code is also publicly available at <https://github.com/bgitschlag/MiSelf>.

## References

1. Szathmari, E. & Smith, J. M. The major evolutionary transitions. *Nature* **374**, 227–232 (1995).

2. Vanchurin, V., Wolf, Y. I., Katsnelson, M. I. & Koonin E. V. Toward a theory of evolution as multilevel learning. *Proc. Natl Acad. Sci. USA* **119**, e2120037119 (2022).
3. West, S. A., Fisher, R. M., Gardner, A. & Kiers, E. T. Major evolutionary transitions in individuality. *Proc. Natl Acad. Sci. USA* **112**, 10112–10119 (2015).
4. Carmel, Y. & Shavit, A. Operationalizing evolutionary transitions in individuality. *Proc. Biol. Sci.* **287**, 20192805 (2020).
5. Michod, R. E. The group covariance effect and fitness trade-offs during evolutionary transitions in individuality. *Proc. Natl Acad. Sci. USA* **103**, 9113–9117 (2006).
6. Ghoul, M., Griffin, A. S. & West, S. A. Toward an evolutionary definition of cheating. *Evolution* **68**, 318–331 (2014).
7. Gore, J., Youk, H. & van Oudenaarden, A. Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459**, 253–256 (2009).
8. Sexton, D. J. & Schuster, M. Nutrient limitation determines the fitness of cheaters in bacterial siderophore cooperation. *Nat. Commun.* **8**, 230 (2017).
9. Strassmann, J. E., Zhu, Y. & Queller, D. C. Altruism and social cheating in the social amoeba *Dictyostelium discoideum*. *Nature* **408**, 965–967 (2000).
10. Aktipis, C. A., et al. Cancer across the tree of life: cooperation and cheating in multicellularity. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, <https://doi.org/10.1098/rstb.2014.0219> (2015).
11. Bravo Nunez, M. A., Nuckolls, N. L. & Zanders, S. E. Genetic villains: killer meiotic drivers. *Trends Genet* **34**, 424–433 (2018).
12. Riehl, C. & Frederickson, M. E. Cheating and punishment in cooperative animal societies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150090 (2016).
13. Zanders, S. E. & Unckless, R. L. Fertility costs of meiotic drivers. *Curr. Biol.* **29**, R512–R520 (2019).
14. Patten, M. M., Schenkel, M. A. & Agren J. A. Adaptation in the face of internal conflict: the paradox of the organism revisited. *Biol. Rev. Camb. Philos. Soc.* **98**, 1796–1811(2023).
15. de Vargas Roditi, L., Boyle, K. E. & Xavier, J. B. Multilevel selection analysis of a microbial social trait. *Mol. Syst. Biol.* **9**, 684 (2013).
16. Dobata, S. & Tsuji, K. Public goods dilemma in asexual ant societies. *Proc. Natl Acad. Sci. USA* **110**, 16056–16060 (2013).
17. Fiegna, F. & Velicer, G. J. Competitive fates of bacterial social parasites: persistence and self-induced extinction of *Myxococcus xanthus* cheaters. *Proc. Biol. Sci.* **270**, 1527–1534 (2003).
18. Havird, J. C. et al. Selfish mitonuclear conflict. *Curr. Biol.* **29**, R496–R511 (2019).
19. Munoz-Gomez, S. A. et al. Site-and-branch-heterogeneous analyses of an expanded dataset favour mitochondria as sister to known Alphaproteobacteria. *Nat. Ecol. Evol.* **6**, 253–262 (2022).
20. Sagan, L. On the origin of mitosing cells. *J. Theor. Biol.* **14**, 255–274 (1967).
21. Castellani, C. A., Longchamps, R. J., Sun, J., Guallar, E. & Arking, D. E. Thinking outside the nucleus: mitochondrial DNA copy number in health and disease. *Mitochondrion* **53**, 214–223 (2020).
22. Chen, Z., et al. Mitochondrial DNA segregation and replication restrict the transmission of detrimental mutation. *J. Cell Biol.* **219**, e201905160 (2020).
23. Birky, C. W. Relaxed and stringent genomes - why cytoplasmic genes dont obey Mendels laws. *J. Hered.* **85**, 355–365 (1994).
24. Chatre, L. & Ricchetti, M. Prevalent coordination of mitochondrial DNA transcription and initiation of replication with the cell cycle. *Nucleic Acids Res* **41**, 3068–3078 (2013).
25. Dubie, J. J., Caraway, A. R., Stout, M. M., Katju, V. & Bergthorsson, U. The conflict within: origin, proliferation and persistence of a spontaneously arising selfish mitochondrial genome. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190174 (2020).
26. Gitschlag, B. L., Tate, A. T. & Patel, M. R. Nutrient status shapes selfish mitochondrial genome dynamics across different levels of selection. *Elife* **9**, e56686 (2020).
27. Klucznika, A. & Ma, H. A battle for transmission: the cooperative and selfish animal mitochondrial genomes. *Open Biol.* **9**, 180267 (2019).
28. Leigh, E. G. When does the good of the group override the advantage of the individual? *Proc. Natl Acad. Sci. USA* **80**, 2985–2989 (1983).
29. Pereira, C. V., Gitschlag B. L. & Patel M. R. Cellular mechanisms of mtDNA heteroplasmy dynamics. *Critical Rev. Biochem. Mol. Biol.* 1–16, <https://doi.org/10.1080/10409238.2021.1934812> (2021).
30. Schwartz, A. Z. A., Nance, J. Germline TFAM levels regulate mitochondrial DNA copy number and mutant heteroplasmy in *C. elegans*. *MicroPubl Biol* **2023**, <https://doi.org/10.17912/micropub.biology.000727> (2023).
31. Schwartz, A. Z. A., Tsyba, N., Abdu, Y., Patel, M. R. & Nance, J. Independent regulation of mitochondrial DNA quantity and quality in *Caenorhabditis elegans* primordial germ cells. *Elife* **11**, e80396 (2022).
32. Yang, Q. et al. LONP-1 and ATFS-1 sustain deleterious heteroplasmy by promoting mtDNA replication in dysfunctional mitochondria. *Nat. Cell Biol.* **24**, 181–193 (2022).
33. Ahier, A. et al. Affinity purification of cell-specific mitochondria from whole animals resolves patterns of genetic mosaicism. *Nat. Cell Biol.* **20**, 352–360 (2018).
34. Gitschlag, B. L. et al. Homeostatic responses regulate selfish mitochondrial genome dynamics in *C. elegans*. *Cell Metab.* **24**, 91–103 (2016).
35. Liao, W. S., Gonzalez-Serricchio, A. S., Deshommes, C., Chin, K. & LaMunyon, C. W. A persistent mitochondrial deletion reduces fitness and sperm performance in heteroplasmic populations of *C. elegans*. *BMC Genet* **8**, 8 (2007).
36. Lin, Y. F. et al. Maintenance and propagation of a deleterious mitochondrial genome by the mitochondrial unfolded protein response. *Nature* **533**, 416–419 (2016).
37. Meshnik, L. et al. Mutant *C. elegans* mitofusin leads to selective removal of mtDNA heteroplasmic deletions across generations to maintain fitness. *BMC Biol.* **20**, 40 (2022).
38. Tsang, W. Y. & Lemire, B. D. Stable heteroplasmy but differential inheritance of a large mitochondrial DNA deletion in nematodes. *Biochem Cell Biol.* **80**, 645–654 (2002).
39. Valenci, I., Yonai, L., Bar-Yaacov, D., Mishmar, D. & Ben-Zvi, A. Parkin modulates heteroplasmy of truncated mtDNA in *Caenorhabditis elegans*. *Mitochondrion* **20**, 64–70 (2015).
40. Capps, G. J., Samuels, D. C. & Chinnery, P. F. A model of the nuclear control of mitochondrial DNA replication. *J. Theor. Biol.* **221**, 565–583 (2003).
41. Chiang, A. C., McCartney, E., O’Farrell, P. H. & Ma, H. A genome-wide screen reveals that reducing mitochondrial DNA polymerase can promote elimination of deleterious mitochondrial mutations. *Curr. Biol.* **29**, 4330–4336 e4333 (2019).
42. Taylor, R. W. & Turnbull, D. M. Mitochondrial DNA mutations in human disease. *Nat. Rev. Genet* **6**, 389–402 (2005).
43. Zhang, H. et al. Mitochondrial DNA heteroplasmy is modulated during oocyte development propagating mutation transmission. *Sci. Adv.* **7**, eabi5657 (2021).
44. Chiaratti, M. R. & Chinnery, P. F. Modulating mitochondrial DNA mutations: factors shaping heteroplasmy in the germ line and somatic cells. *Pharm. Res* **185**, 106466 (2022).
45. Schon, E. A., DiMauro, S. & Hirano, M. Human mitochondrial DNA: roles of inherited and somatic mutations. *Nat. Rev. Genet* **13**, 878–890 (2012).
46. Stewart, J. B. & Chinnery, P. F. The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nat. Rev. Genet* **16**, 530–542 (2015).

47. Konrad, A. et al. Mitochondrial mutation rate, spectrum and heteroplasmy in *Caenorhabditis elegans* spontaneous mutation accumulation lines of differing population size. *Mol. Biol. Evol.* **34**, 1319–1334 (2017).
48. Patwa, Z. & Wahl, L. M. The fixation probability of beneficial mutations. *J. R. Soc. Interface* **5**, 1279–1289 (2008).
49. Teterina, A. A. et al. Genomic diversity landscapes in outcrossing and selfing *Caenorhabditis* nematodes. *PLoS Genet* **19**, e1010879 (2023).
50. Stiernagle, T. Maintenance of *C. elegans*. *WormBook*, 1-11 (2006).
51. Shou, W. Acknowledging selection at sub-organismal levels resolves controversy on pro-cooperation mechanisms. *Elife* **4**, e10106 (2015).
52. Wei, W. et al. Germline selection shapes human mitochondrial DNA diversity. *Science* **364**, eaau6520 (2019).
53. Bergstrom, C. T. & Pritchard, J. Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. *Genetics* **149**, 2135–2146 (1998).
54. Colnaghi, M., Pomiankowski, A. & Lane, N. The need for high-quality oocyte mitochondria at extreme ploidy dictates mammalian germline development. *Elife* **10**, e69344 (2021).
55. Johnston, I. G. & Burgstaller, J. P. Evolving mtDNA populations within cells. *Biochem Soc. Trans.* **47**, 1367–1382 (2019).
56. Giannakis, K., Broz, A. K., Sloan, D. B. & Johnston, I. G. Avoiding misleading estimates using mtDNA heteroplasmy statistics to study bottleneck size and selection. *G3 (Bethesda)* **13**, (2023).
57. Royrvik, E. C. & Johnston, I. G. MtDNA sequence features associated with ‘selfish genomes’ predict tissue-specific segregation and reversion. *Nucleic Acids Res* **48**, 8290–8301 (2020).
58. Kandul, N. P., Zhang, T., Hay, B. A. & Guo, M. Selective removal of deletion-bearing mitochondrial DNA in heteroplasmic *Drosophila*. *Nat. Commun.* **7**, 13100 (2016).
59. Lieber, T., Jeedigunta, S. P., Palozzi, J. M., Lehmann, R. & Hurd, T. R. Mitochondrial fragmentation drives selective removal of deleterious mtDNA in the germline. *Nature* **570**, 380–384 (2019).
60. Suen, D. F., Narendra, D. P., Tanaka, A., Manfredi, G. & Youle, R. J. Parkin overexpression selects against a deleterious mtDNA mutation in heteroplasmic cybrid cells. *Proc. Natl Acad. Sci. USA* **107**, 11835–11840 (2010).
61. Phillips, W. S. et al. Selfish mitochondrial DNA proliferates and diversifies in small, but not large, experimental populations of *Caenorhabditis briggsae*. *Genome Biol. Evol.* **7**, 2023–2037 (2015).
62. Artilles, K. L., Fire, A. Z. & Frokjaer-Jensen, C. Assessment and maintenance of unigametic germline inheritance for *C. elegans*. *Dev. Cell* **48**, 827–839 e829 (2019).
63. Seneta E. *Non-negative matrices and Markov chains*, Rev. print. edn. Springer (2006).

## Acknowledgements

All strains were kindly provided by the *Caenorhabditis* Genetics Center (CGC). We thank O. Thompson and R. Waterston at the University of Washington for identifying the *mptDf2*, *mpt4*, *mpt2*, and *mptDf3* genotypes from the Million Mutation Project worm collection. We thank Patrick Abbot and Ann T. Tate of Vanderbilt University for valuable conceptual insights and guidance with experimental design. This work was supported by the NIH research project grants R01 GM123260 and R35 GM145378 (M.R.P.), Discovery Award (PR170792) from the

Department of Defense’s Congressionally Directed Medical Research Program (M.R.P.), the Vanderbilt University Medical Center Diabetes Research and Training Center Pilot & Feasibility Grant (M.R.P.), the NIH-sponsored Cellular, Molecular and Biochemical, and Molecular Sciences Training Program (5T32GM008554-18, B.L.G.), the Ruth L. Kirschstein National Research Service Award Individual Predoctoral Fellowship (1F31GM125344, B.L.G.), R35 GM133613 (D.M.M.), an Alfred P. Sloan Research Fellowship and additional funding from the Simons Center for Quantitative Biology and Cold Spring Harbor Laboratory (D.M.M.).

## Author contributions

B.L.G., D.M.M., and M.R.P. designed the study. B.L.G., C.V.P., and J.P.H. performed experiments to collect the data. B.L.G. and D.M.M. performed mathematical modeling. B.L.G., D.M.M., and M.R.P. analyzed the data. B.L.G. prepared the first manuscript draft. All authors revised and edited the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52596-9>.

**Correspondence** and requests for materials should be addressed to Bryan L. Gitschlag or Maulik R. Patel.

**Peer review information** *Nature Communications* thanks Iain Johnston and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024