Multilevel selection analysis reveals diversity in selfish mitochondrial genome dynamics

Bryan L. Gitschlag[1], Claudia V. Pereira[1], James P. Held[1], & Maulik R. Patel[1,2,3] *

[1]Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

[2]Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN, USA

[3]Diabetes Research and Training Center, Vanderbilt University School of Medicine, Nashville, TN, USA

**Corresponding author*: maulik.r.patel@vanderbilt.edu

Abstract

Mutant copies of mitochondrial DNA (mtDNA) can behave as selfish entities, propagating within organisms at the expense of host fitness. We previously developed experiments to separately measure the within-host proliferation, and host fitness cost, of selfish mtDNA. By focusing on a single mutant genome, the *uaDf5* variant, our previous work raises the question of generalizability: do selfish genomes uniformly overcome a heavy fitness cost by proliferating to high frequency, or do their population dynamics vary? By applying a standardized multilevel selection analysis across a collection of mitochondrial mutants, we uncover variation in cheating strategies. Mutants that impose a heavy fitness cost nevertheless persist by outcompeting wildtype mtDNA within hosts. Conversely, some mutants fail to outcompete wildtype mtDNA but nevertheless persist, maintaining a frequency range at which host fitness cost is negligible. These findings suggest that the population dynamics of mutant mtDNA depend on the loci affected.

Introduction

Cooperation and cheating represent conflicting adaptive strategies understood to shape the evolution and behavior of complex biological systems (Fisher and Regenberg, 2019, Gulli et al., 2019, Michod et al., 2006, West et al., 2015, Hammerschmidt et al., 2014). In particular, transitions toward larger, more complex forms of life occur when natural selection favors cooperation and loss of conflict between individual life forms, enabling them to become integrated into a collective entity that replicates and undergoes selection as a unified whole (Michod et al., 2006, West et al., 2015, Queller and Strassmann, 2009). However, because cooperators incur the near-term cost of contributing to the fitness of partners, albeit for mutual long-term benefit, cooperation creates conditions that can select for the emergence of selfish, "cheater" entities. Cheating can be defined as occurring when a cooperative partner provides a benefit to the fitness of another entity (the "cheater"), at the expense of its own fitness (Ghoul et al., 2014). Cheaters reap the benefit of belonging to a cooperative relationship without incurring the cost of reciprocating. Examples of cheating can be found across the various levels of scale in the hierarchy of biological organization, including the level of genes (Bravo Nunez et al., 2018, Hammond et al., 2012, Hu et al., 2017, Larracuente and Presgraves, 2012, Schimenti, 2000), cells (Aktipis et al., 2015), and social animals (Riehl and Frederickson, 2016). By exploiting the costly contributions of others without reciprocating, cheaters maintain an advantage within a group. Competition between groups, however, maintains a selection pressure that can select for cooperation and against cheating (Aktipis et al., 2015, de Vargas Roditi et al., 2013, Fiegna and Velicer, 2003, Moreno-Fenoll et al., 2017, Rainey and Rainey, 2003, Wenseleers and Ratnieks, 2004). The combination of conflicting selection pressures across these different levels of scale, termed conflicting multilevel selection, thus resolve an apparent paradox, namely the coexistence between selfish and cooperative entities within a hierarchically structured population.

53

54 The eukaryotic cell was first recognized as the result of a cooperative symbiosis more than fifty
55 years ago (Sagan, 1967). Moreover, mitochondria were hypothesized to experience conflicting
56 evolutionary pressures (Leigh, 1983), even before pathogenic mitochondrial mutations were first
57 identified (Holt et al., 1988). Specifically, natural selection was proposed to occur within
58 individual host organisms in a manner that favors selfish genetic elements, while selection for
59 host fitness was proposed as a form of "group selection," favoring the faithful transmission of a
60 metabolically competent population of mitochondrial genomes (Leigh, 1983). Consistent with
61 this notion, multiple studies in recent years have identified mitochondrial mutations across a
62 wide diversity of life that propagate while inflicting a fitness cost on their hosts (Havird et al.,
63 2019, Klucnika and Ma, 2019). These observations have inspired a widespread interest among
64 mitochondrial geneticists and microbial ecologists in the phenomenon of multilevel selection
65 (Clark et al., 2012, Dubie et al., 2020, Gitschlag et al., 2020, Havird et al., 2019, Klucnika and
66 Ma, 2019, Shou, 2015).

67

68 Consistent with the predictions of conflicting multilevel selection, a deleteriously mutated
69 mitochondrial DNA (mtDNA) variant—the *uaDf5* variant in *Caenorhabditis elegans*—has been
70 previously reported to stably persist in a heteroplasmic state at approximately 60 percent
71 frequency across hundreds of generations (Tsang and Lemire, 2002), despite being deleterious
72 at the level of host fitness (Liau et al., 2007). More recent follow-up studies have used the *uaDf5*
73 heteroplasmy as a model to identify mechanisms underlying the selfish propagation of
74 deleterious mtDNA (Gitschlag et al., 2016, Gitschlag et al., 2020, Lin et al., 2016, Yang et al.,
75 2022). To what extent can the results of such work be generalized? Mutant mitochondrial
76 genomes that propagate selfishly at the expense of host fitness have been observed to arise
77 spontaneously in laboratory populations of yeast (Bernardi, 2005) and *C. elegans* (Dubie et al.,
78 2020, Konrad et al., 2017), suggesting that cheating may represent a relatively common theme

79 in the population dynamics of mitochondrial mutations. On the other hand, while *uaDf5* persists
80 at approximately 60 percent heteroplasmic frequency, another heteroplasmy in *C. elegans*, a
81 spontaneously-arising mutant mtDNA variant labeled the Δ*ctb-1* mitotype, was found to maintain
82 upwards of 90 percent heteroplasmy frequency (Dubie et al., 2020). These findings suggest that
83 mitochondrial mutations differ from one another in their heteroplasmy dynamics, consistent with
84 the observed diversity of phenotypic manifestations among mtDNA variants (Taylor and
85 Turnbull, 2005). Given that selfish mtDNA proliferation has been observed to arise as a result of
86 interactions between the mitochondrial and host genomes (Gitschlag et al., 2016, Gitschlag et
87 al., 2020, Lin et al., 2016, Yang et al., 2022), variation among mtDNA variants in average
88 frequency and in phenotypic effects raise the possibility that the propensity for cheating
89 behavior likely differs between deleterious mtDNA variants.
90
91 To address this possibility, we sought to develop a standardized method for quantifying the
92 cheating behavior of different mutant mtDNA (ΔmtDNA) variants. We recently developed an
93 experimental strategy to isolate and measure the strength of natural selection at each
94 respective level (Gitschlag et al., 2020), namely within hosts (sub-organismal selection) and at
95 the level of overall host fitness (organismal selection). Accordingly, we were able to identify
96 effects of nutrient status on both the selfish proliferation and the host fitness effects of *uaDf5*
97 (Gitschlag et al., 2020). Building on this prior work, here we describe an approach for integrating
98 empirical measurement with theory, in order to model the population dynamics of selfish
99 mitochondrial genomes. Using this approach, we sought to expand the scope of our
100 investigation to include other mtDNA variants, representing mutations in enzyme complexes I,
101 III, and IV of the electron transport chain (ETC). By characterizing their within-host and
102 between-host population dynamics, we find that although cheating is evidently a common
103 feature among deleterious mitochondrial mutations, the specific nature of the cheating "strategy"
104 varies considerably among different mtDNA variants.

105

106

107

108

109 **Results**

110 **An experimental strategy to isolate selection on a selfish mitochondrial genome at**

111 **different levels**

112 We sought to identify broad themes and notable differences between different stably

113 propagating heteroplasmic ∆mtDNA variants. We elected to use the mutant mtDNA variants

114 *mpt1*, *mpt2*, *mptDf3*, *mptDf2*, and *mpt4* (Figure 1A). These encompass mutations in ETC

115 complex I (*mpt1* and *mpt2*), complex III (*mptDf3*), and complex IV (*mpt4* and *mptDf2*). Mutations

116 in *mpt1*, *mpt2*, and *mpt4* are each confined to a single protein-coding gene. The *mptDf2* deletion

117 also deletes transfer RNA (tRNA) genes and the mutation in *mptDf3* deletes the five most 3'

118 nucleotides of the tRNA gene for isoleucine suggesting that these mutations may also disrupt

119 ETC function more broadly by interfering with mitochondrial protein synthesis. Importantly, these

120 mutant variants persist at substantially different heteroplasmic frequencies (Figure 1B),

121 suggesting that they differ with respect to within-host proliferation, host-level selection, or both.

122 We therefore sought to develop a standardized approach for quantifying differences in cheating

123 behavior among ∆mtDNA variants; that is, differences in their within-host proliferation and host

124 fitness consequences.

125

126 We previously measured the impact of sub-organismal selection on the propagation of the

127 ∆mtDNA variant *uaDf5* (Figure 1A, red bar) by quantifying *uaDf5* frequency longitudinally

128 between individual parents and their own respective adult progeny (Figure 2A) (Gitschlag et al.,

129 2020). These data revealed negative frequency-dependent selection, in which *uaDf5*

130 outcompetes wildtype mtDNA for replication across organismal development, but its ability to do

131 so declines as a function of increasing starting (parental) *uaDf5* frequency. Selection against
132 *uaDf5* that occurs strictly at the level of host fitness was measured by competing *uaDf5*-carrying
133 heteroplasmic animals against their homoplasmic (wildtype) counterparts on the same food
134 plate (Figure 2B) (Gitschlag et al., 2020). Non-competing control lines were maintained, which
135 consisted of exclusively heteroplasmic animals, with no wildtype animals to compete against. To
136 control for variation in *uaDf5* frequency due to reasons other than the organismal fitness cost of
137 *uaDf5*, the *uaDf5* frequencies of competing lines were normalized to that of non-competing lines
138 at each generation; any resulting change in the *uaDf5* frequency of the competing lines relative
139 to the non-competing lines can therefore be attributed to an organismal fitness effect of *uaDf5*,
140 since the presence of animals lacking *uaDf5* represents the only distinguishing condition
141 between competing and non-competing lines. Accordingly, we have developed an approach to
142 separately measure the effects of selection on ∆mtDNA at the sub-organismal and organismal
143 levels.
144

145 **Multilevel selection analysis of the *uaDf5* heteroplasmy**

146 Here, we sought to integrate the measurements of sub-organismal and organismal selection
147 into a single theoretical framework describing the overall population dynamics of ∆mtDNA. The
148 Price Equation, which provides such a framework, describes evolution as the covariance
149 between a trait and its reproductive fitness (Frank, 1997, Price, 1972). Importantly, since this
150 equation can accommodate more than one source of covariance, including non-linear
151 interactions (Frank, 2012), it can be used to describe selection at different levels
152 simultaneously:

$$w\Delta z = cov(w,z)_{org} + E[cov(w,z)_{sub}]$$

153 In this form of the equation, *z* refers to a heritable trait that can be quantitatively measured.
154 Since mitochondrial genotype consists of a heteroplasmic mix of mtDNA variants, trait-value *z*
155 can be defined in this case as *uaDf5* frequency. Fitness *w* refers to the net replicative success

156 of *uaDf5* relative to wildtype mtDNA, defined as the ratio of new *uaDf5* copies to new wildtype
157 mtDNA copies per generation. Subscripts $_{org}$ and $_{sub}$ label the covariance between *w* and *z* at the
158 organismal and sub-organismal levels, respectively, with covariance defined by the following
159 formula:

$$cov(w,z) = \frac{\sum(w_i - w_{average})(z_i - z_{average})}{n-1}$$

160 The term *E* in front of the sub-organismal covariance term denotes expected, or average, sub-
161 organismal covariance between *w* and *z*; note that the total sub-organismal change in *uaDf5*
162 frequency is the average number of replicated copies per *uaDf5* template compared to that of
163 wildtype mtDNA. Moreover, since the strength or intensity of selection can be expressed as the
164 selection coefficient *S* (equal to 1 – *w*), the change in *uaDf5* frequency at each respective level
165 of selection can be expressed using the covariance term that relates selection coefficient to
166 *uaDf5* frequency (see Figure 2C for a sample illustration):

$$cov(S, \%\Delta mtDNA) = \frac{\sum(S_i - S_{average})(\%uaDf5_i - \%uaDf5_{average})}{n-1}$$

167 Although the above formula for covariance merely describes how variation in *uaDf5* frequency
168 relates to variation in selection acting on *uaDf5*, this relation can be applied to the empirically
169 measured shifts in *uaDf5* frequency reported previously (Gitschlag et al., 2020) to derive a
170 function that relates selection to *uaDf5* frequency, taking both levels of selection into account
171 (for a more detailed derivation of this equation, see Figure 3—Figure Supplement 1):

$$S_{overall} = S_{sub-organismal} + S_{organismal} = (0.01154 * \%uaDf5 - 0.923) + (10^{0.0181 * \%uaDf5 - 1.81})$$

172 The term shown in the parentheses on the left defines the relationship between *uaDf5*
173 frequency and sub-organismal selection (plotted in Figure 3A). The term shown in the
174 parentheses on the right defines the relationship between *uaDf5* frequency and organismal
175 selection (plotted in Figure 3B, solid red line), relying on two additional assumptions. The first
176 assumption is that selection coefficient would be zero when *uaDf5* frequency is 0 percent, since
177 there is no *uaDf5* to select against. Second, the selection coefficient is assumed to reach 1—

178 fitness reaches 0, reflecting elimination of *uaDf5* from the population within a single
179 generation—as *uaDf5* frequency reaches 100 percent. This assumption rests on the fact that
180 the mutation deletes essential genes and should therefore be lethal if present at 100 percent
181 frequency. Summing the organismal and sub-organismal regressions yields the overall,
182 population-wide relationship between heteroplasmic *uaDf5* frequency and the strength of
183 selection acting on *uaDf5* (Figure 3C).
184
185 The Price Equation framework shows that selection at the sub-organismal and organismal
186 levels balance—the overall selection coefficient approaches 0, reflecting no net difference in
187 fitness between *uaDf5* and wildtype mtDNA—when *uaDf5* is near 60 percent heteroplasmic
188 frequency (Figure 3C). Note that this is the approximate frequency at which *uaDf5* persists in
189 the non-competing populations, in which the conflicting levels of selection are both in effect
190 (Gitschlag et al., 2020). Importantly, this conclusion is compatible with a range of alternate
191 assumptions regarding the impact of rising *uaDf5* frequency on organismal fitness (Figure 3B-
192 C). The mutant genome may become lethal before reaching fixation, at 90 or 95 percent
193 frequency for example. Consistent with this idea, *uaDf5* is rarely observed at or above 90
194 percent frequency (Gitschlag et al., 2016, Gitschlag et al., 2020, Lin et al., 2016, Tsang and
195 Lemire, 2002). Alternatively, the fitness cost of *uaDf5* might not rise smoothly, as depicted in the
196 non-linear regression in Figure 3B. For example, the organismal selection coefficient may rise
197 linearly with increasing *uaDf5* frequency up to and beyond the mean of 60 percent
198 heteroplasmic frequency, perhaps until reaching a critical threshold before imposing a heavy
199 fitness cost, resulting in an abrupt upward shift in the strength of selection. However, taking
200 these alternative assumptions into account had little effect on the overall conclusions (Figure
201 3B-C). In conclusion, we have integrated measurements of selection on *uaDf5* at different levels
202 into a unified theoretical framework that robustly predicts the overall population-level dynamics
203 observed for *uaDf5*.

204

205 **Quantification of tracking sub-organismal frequency and organismal selection amongst**
206 **different heteroplasmies**

207 By measuring sub-organismal and organismal selection using the approach initially developed
208 for the study of *uaDf5*, we were able to isolate changes in mutant mtDNA frequency in a manner
209 specific to each level of selection (Figure 4 and Figure 4—Figure Supplements 1-2). Since each
210 mtDNA variant stably persists in a heteroplasmic population, despite harboring a deleterious
211 mutation in at least one essential gene, they could all be predicted to undergo conflicting
212 multilevel selection similarly to *uaDf5* (Figure 3). Surprisingly, however, the capacity for within-
213 host proliferation, as well as the host-level fitness costs, varies considerably between mutant
214 mitochondrial genomes.

215

216 Neither of the ETC complex I deletions show a significant effect on host fitness (Figure 4B and
217 4D). It should be noted, however, that the two complex I mutations stably persist at different
218 frequencies (Figures 1B, 4A, 4C, and Figure 4—Figure Supplement 2A-B). Consistent with a
219 lack of host fitness effects, these mutations appear to undergo neither within-host proliferation
220 nor negative organismal selection, remaining at similar frequencies across generations in both
221 sub-organismal and organismal experiments (Figure 4A-D). The same is true for the complex III
222 mutation in *mptDf3* (Figure 4E-F). Interestingly, the population-wide frequency of *mptDf3* rose
223 dramatically, albeit in only one of the competition lines (Figure 4F), consistent with random drift
224 toward a higher population-wide fraction of heteroplasmic animals. The complex IV mutations,
225 on the other hand, inflict dramatic fitness effects, being almost entirely outcompeted by
226 homoplasmic (wildtype) animals within six generations. This was true for *mptDf2* (Figure 4H),
227 which affects both complex IV and tRNA genes, as well as *mpt4* (Figure 4J), which only affects
228 the complex IV gene *CO1*. Consistent with these findings, mutations in human *CO1* are
229 associated with several pathologies (Brown et al., 1992, Namslauer and Brzezinski, 2009,

230 Varlamov et al., 2002). These severe host fitness costs predict that for these mutations to
231 persist in a population, they must undergo significant within-host proliferation, sufficient to
232 overcome the strong negative selection at the organismal level. Indeed, this prediction is
233 remarkably consistent with our observation (Figure 4G and 4I). We therefore have identified
234 mutations that behave similarly to *uaDf5*, and mutations that differ, from a collection of five
235 additional heteroplasmic mtDNA variants.
236
237
238
239 **Multilevel selection analysis of a panel of different heteroplasmies**
240 Finally, we sought to integrate these findings using the same theoretical framework developed
241 for *uaDf5* (Figure 3), by summing the regressions between mutant mtDNA frequency and the
242 strength of selection at the sub-organismal and organismal levels (Figure 5). Not surprisingly,
243 when the levels of selection are combined into a single overarching description of their
244 respective population dynamics, there is no overall (population-level) covariance between
245 mutant frequency and selection, at least for *mpt1*, *mpt2*, and *mptDf3* (Figure 5A-I). In other
246 words, the frequency of these mutations does not seem to correlate with their fitness effects, at
247 least within the frequency ranges at which these heteroplasmies were observed to persist. In
248 contrast, the regressions for the mutant genome *mptDf2* (Figure 5J-L) predicts that this genome
249 stably persists at very high frequency—near 80 percent—due to a balance between a strong
250 sub-organismal proliferative advantage and a strong host fitness cost. As with *uaDf5*, this
251 prediction is compatible with alternate assumptions regarding the relationship between
252 heteroplasmy frequency and host fitness cost (Figure 5—Figure Supplement 1); moreover,
253 these findings are highly consistent with the empirical observation that *mptDf2* maintains very
254 high frequency in non-competing lines, where the forces of sub-organismal and organismal
255 selection are both in effect (Figure 1B and Figure 4—Figure Supplement 2D). Although *mpt4*

256 undergoes robust sub-organismal proliferation (Figure 4I and Figure 4—Figure Supplement 4E),
257 the magnitude of proliferation does not appear to co-vary with *mpt4* frequency (Figure 5M),
258 indicating that the combined population-level covariance fails to predict the level at which this
259 mutant genome persists (Figure 5O). However, the substantial host fitness cost of this genome
260 (Figures 4J and 5N) implies a strong selection pressure that acts in opposition to the sub-
261 organismal proliferation, similarly to *uaDf5* and *mptDf2*. In conclusion, we have identified
262 notable similarities and differences amongst the population dynamics of mitochondrial
263 mutations.
264
265
266 **Discussion**
267 Multilevel selection theory provides a powerful explanatory account of cooperator-cheater
268 dynamics. We developed an experimental approach to empirically quantify selection for selfish
269 mitochondrial genomes at the sub-organismal (within-host) level by tracking their frequency
270 longitudinally within isolated parent-progeny lineages. At the organismal (between-host) level,
271 we measure selection against selfish mitochondrial genomes by competing heteroplasmic and
272 wildtype animals on the same food plate and tracking the population-wide frequency of mutant
273 mtDNA. By enabling the empirical measurement of selection at different levels, this approach
274 can be applied to explore cooperator-cheater dynamics broadly. Integrating these empirical
275 measurements into a theoretical model (Figure 3) reveals a more complete, higher-resolution
276 depiction of the overall population dynamics of selfish mtDNA. Apart from confirming that the
277 mutant persists within a given frequency range due to balancing selection among the different
278 levels, measuring selection at each level makes it possible to account for their respective
279 contributions toward experimentally induced shifts in heteroplasmy level, providing a framework
280 for detailed mechanistic follow-up.
281

282 The mutant mtDNA variant *uaDf5* is an excellent model for the study of selfish mitochondrial
283 genome dynamics, for a number of reasons. For example, this mutant rises in frequency over
284 time within hosts (Figure 3A) at the expense of inflicting a fitness cost (Figure 3B). The
285 proliferation of *uaDf5* is also frequency-dependent, a common feature of cheater entities
286 (Dobata and Tsuji, 2013, Dugatkin et al., 2005, Pruitt and Riechert, 2009, Riehl and
287 Frederickson, 2016, Ross-Gillespie et al., 2007). Moreover, *uaDf5* proliferates by exploiting
288 interactions with the host genome that exist to minimize the harmful effects of metabolic and
289 nutrient stress (Gitschlag et al., 2016, Gitschlag et al., 2020, Lin et al., 2016, Yang et al., 2022),
290 or which reason it can be viewed as a bona fide cheater entity. However, mitochondrial
291 mutations are not all alike with respect to their fitness effects. How broadly generalizable, then,
292 are the findings associated with the study of *uaDf5*? We sought to address this question here by
293 identifying noteworthy similarities and differences in the population dynamics of different
294 mitochondrial DNA mutations/deletions.
295
296 **Variation in selfish mitochondrial genome dynamics**
297 By quantitatively measuring the strength of selection acting on mutant mtDNA at the sub-
298 organismal and organismal levels, across a collection of mitochondrial mutations, we identify
299 two noteworthy themes. On one hand, like *uaDf5*, some mutant mtDNA variants propagate in
300 spite of inflicting a heavy fitness burden on their hosts. As expected, their seemingly paradoxical
301 persistence—in other words, their ability to avoid extinction in the face of strong negative
302 organismal selection—appears to coincide with strong positive sub-organismal selection (Figure
303 6, left side). These conflicting multilevel selection dynamics can be observed in three of the six
304 heteroplasmies featured here: *uaDf5*, *mptDf2*, and *mpt4*. Although the molecular details
305 underlying these dynamics are not clear, the only thing that all three mutations have in common
306 is a disruption to ETC complex IV. Consistent with these findings, mutations in the complex IV

307 gene *CO1* are associated with a number of human clinical pathologies (Brown et al., 1992,
308 Namslauer and Brzezinski, 2009, Varlamov et al., 2002). In any case, selfish propagation in
309 which mutant genomes survive negative organismal selection by outcompeting the wildtype
310 mtDNA within individual hosts, appears to be a relatively common, albeit not universal, theme
311 among stably persisting deleterious heteroplasmies.
312
313 In addition to these conflicting multilevel selection dynamics, a second general theme emerges
314 from tracking the levels of some of the mutant genomes featured here. In contrast to *uaDf5*,
315 *mptDf2*, and *mpt4*, not all mutant mitochondrial genomes persist by rising in frequency across
316 development at the expense of a strong host fitness cost. On the contrary, despite carrying
317 deletions in essential metabolic genes, some mutant genomes appear to undergo neutral drift—
318 that is, fluctuations in frequency of existing alleles not attributable to natural selection—even
319 when the heteroplasmic animals compete for several generations against their wildtype
320 counterparts on the same food. This is true for both mutations in ETC complex I—the *mpt1*
321 deletion in *ND1* and the *mpt2* deletion in *ND5*—as well as for *mptDf3*, which affects the complex
322 III gene for cytochrome b (Figure 6, right side). Taken together, these observations suggest that
323 mtDNA variants can persist in a manner consistent with neutral drift, despite carrying mutations
324 that deleteriously affect essential ETC components.
325
326 **Comparison with previous findings**
327 Interestingly, *mptDf3*, which carries a deletion affecting the complex III mtDNA-encoded protein,
328 namely cytochrome b, does not appear to inflict a significant host fitness cost, at least not within
329 the frequency range at which *mptDf3* is observed to persist. This may be difficult to reconcile
330 with previous work showing that an mtDNA variant labeled Δ*ctb-1*—owing to a cytochrome b
331 mutation—behaves much more like *uaDf5*, *mptDf2*, and *mpt4*, in particular by proliferating within
332 hosts while imposing a host fitness cost (Dubie et al., 2020). In fact, in contrast to *uaDf5*, Δ*ctb-1*

333 was reported to persist at upwards of 90 percent heteroplasmic frequency for dozens of
334 generations (Dubie et al., 2020). One possible explanation for this apparent discrepancy
335 between the dynamics of Δ*ctb-1* and *mptDf3* is that Δ*ctb-1* contains other mutations in addition
336 to the deletion in cytochrome b, which may contribute to its observed selfish proliferation, or
337 host fitness consequences, or both (Dubie et al., 2020).
338

339 **Hypothesis and future directions**

340 The sub-organismal selection dynamics described here are remarkably consistent with recent
341 mammalian work. In particular, a tRNA point mutation in mouse mtDNA, and three point
342 mutations in human mtDNA, were recently found to undergo an upward shift in frequency from
343 parents to offspring, especially among parents with lower starting mutant levels (Zhang et al.,
344 2021, compare Figure 1C and Figure 8B, D and F to Figure 2B in (Gitschlag et al., 2020). These
345 shifts occur in spite of deleterious phenotypic effects associated with this heteroplasmy (Zhang
346 et al., 2021), suggesting a cheating behavior strikingly similar to mtDNA variants *uaDf5*, *mptDf2*,
347 and *mpt4* in *C. elegans*. Moreover, we previously applied the experiments described in this
348 study to identify conditions and host genotypes that modulate both the sub-organismal
349 proliferation and host fitness cost of *uaDf5* (Gitschlag et al., 2020). Taken together, in addition to
350 an apparent wide prevalence of cheating behavior among animal mitochondrial genomes, these
351 findings further suggest an overlap between the mechanisms underlying cheating and those
352 underlying disease. By applying the multilevel selection analysis described here to a diversity of
353 conditions and host genotypes, we anticipate that it will be possible to identify avenues for
354 mitigating the risk and severity of mitochondrial disease.
355

356 How can a mitochondrial genome lacking an essential gene manage to stably persist with no
357 measurable fitness effect at either level of selection (organismal or sub-organismal)? Although
358 the explanation is not known, these observations raise the interesting possibility that some

359 mutant genomes propagate, not in spite of a deleterious host fitness cost, but rather because of
360 what could be considered a “flying under the radar” strategy. Some mutant genomes seem to
361 successfully persist by maintaining a frequency range that enables them to avoid exposure to
362 negative organismal selection. One possible basis for this phenomenon would invoke a balance
363 between positive and negative selection at the sub-organismal level. For example, these mutant
364 genomes might affect mitochondrial function in a manner that elicits a combination of cellular
365 mechanisms that simultaneously favor and disfavor the mutant genome within individual hosts.
366 If this balance is achieved within a frequency range that is too low to significantly impact host
367 fitness, the mutant mtDNA could undergo neutral drift at the organismal level while maintaining
368 a steady-state heteroplasmy.
369
370 This possibility could be explored in future research by testing host genes for their effect on
371 heteroplasmy dynamics within hosts. For example, previous work has shown that mtDNA
372 replication can be biased in favor of either a mutant or wildtype genome depending on the
373 expression level of the mtDNA polymerase, POLG (Chiang et al., 2019, Lin et al., 2016). This
374 raises the possibility that physiological stress-response mechanisms might bias replication in
375 favor of the mutant genome by influencing the expression or localization of mtDNA replication
376 machinery. For example, the gene *daf-16* encodes a Forkhead box O (FoxO) transcription factor
377 that functions as a master regulator of metabolic homeostasis. Although *daf-16* does not appear
378 to directly regulate POLG expression (Tepper et al., 2013), it nevertheless regulates other
379 aspects of mitochondrial function, including the expression of proteins involved in energy
380 metabolism and stress response (Depuydt et al., 2014, Tepper et al., 2013, Webb et al., 2016).
381 If its attempt to rescue metabolic integrity influences mtDNA replication, even indirectly, then a
382 compensatory effect could result that favors underperforming regions of the mitochondrial
383 network, inadvertently promoting replication of mutant genomes. Consistent with this possibility,
384 *daf-16* mediates rising *uaDf5* frequency across organismal development (Gitschlag et al., 2020).

385 Like *daf-16*, the gene *atfs-1* also regulates a transcriptional response to metabolic stress, in this
386 case by activating the mitochondrial unfolded protein response (UPR^{mt}) (Nargund et al., 2012).
387 Importantly, *atfs-1* gain-of-function mutations are associated with elevations in both *uaDf5*
388 frequency (Gitschlag et al., 2016, Lin et al., 2016) and expression of *polg-1*, the gene encoding
389 POLG (Lin et al., 2016). Together, these considerations highlight *daf-16* and *atfs-1* as attractive
390 candidate genes underlying positive sub-organismal selection for the mutant mitochondrial
391 genomes featured in this study. Conversely, *drp-1* is involved in mitochondrial fission, while
392 *pink-1* and *pdr-1* play central roles in PINK1/Parkin-dependent mitophagy, and both of these
393 processes have been implicated in selection against mutant mtDNA at the sub-organismal level
394 (Gitschlag et al., 2016, Kandul et al., 2016, Lieber et al., 2019, Lin et al., 2016, Suen et al.,
395 2010, Valenci et al., 2015). Does the combined activity, and hence influence, of each of these
396 pathways vary amongst different selfish mitochondrial genomes? This question could be
397 addressed by future research, for example by comparing UPR^{mt} activation and levels of
398 mitophagy, across different heteroplasmies. It would also be interesting to determine how
399 differently, stably-propagating heteroplasmies compare in their sensitivity to the disruption of
400 these pathways.
401
402 **Materials and Methods**
403 **Nematode husbandry**
404 *C. elegans* strains used in this study were maintained on 60-mm standard nematode growth
405 medium (NGM) plates (for measuring sub-organismal selection), or 100-mm NGM plates (for
406 measuring organismal selection), seeded with live OP50-strain *E. coli* bacteria as a food source.
407 Nematode strains were incubated at 20˚C. In addition to the Bristol wildtype strain and the
408 heteroplasmic *uaDf5* strain featured in prior work (Gitschlag et al., 2020), five additional *C.*
409 *elegans* strains were used in this study. These consisted of heteroplasmic mutant genomes

410 *mpt1*, *mpt2*, *mptDf3*, *mptDf2*, and *mpt4*, each crossed into the nuclear background of the Bristol
411 strain.
412
413 **Nuclear genome exchange in heteroplasmic strains.**
414 To ensure that our analysis was not confounded by variation in the nuclear genome, we
415 employed a novel unigametic inheritance method to completely exchange the nuclear genome
416 of each heteroplasmic strain with the nuclear genome of wildtype (Bristol strain) *C. elegans*
417 (method from Artiles *et al.* 2019). Thus, all heteroplasmic strains used in this study have
418 identical nuclear backgrounds. Briefly, each heteroplasmic strain was crossed to the *gpr-1* over-
419 expression strain PD2220 following Mendelian genetics. Next, hermaphrodites of the stable *gpr-*
420 *1* over-expression heteroplasmy strains were crossed to wildtype males. Non-Mendelian
421 hermaphrodite progeny from these crosses in which the paternal nuclear background is
422 unigametically inherited in the germline cell lineage (determined by pharyngeal mosaic
423 patterning) were individually propagated. Stock strains were established from the progeny of
424 these animals as they have a complete wildtype nuclear genomic background and retain the
425 given heteroplasmy.
426
427 **Lysate preparation**
428 To prepare animals for quantification of mutant mtDNA frequency, nematodes were transferred
429 to sterile PCR tubes or 96-well PCR plates containing lysis buffer with 100 µg/mL proteinase K.
430 Lysis buffer consisted of 50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM $MgCl_2$, 0.45% Tween 20,
431 0.45% NP-40 (IGEPAL), and 0.01% gelatin, in deionized water. Volume of lysis buffer varied by
432 worm count: 10 µL for individual adults of the parent generation and 20 µL for pooled adult
433 progeny for measuring sub-organismal selection, and 50 µL for pooled animal lysates from the
434 competition experiments for measuring organismal selection. After transferring worms to lysis
435 buffer, each tube or plate was immediately sealed and incubated at -80˚C for 10 minutes to

436 rupture nematode cuticles, followed by lysis incubation at 60˚C for 60 minutes (90 minutes for
437 pooled nematodes), and then at 95˚C for 15 minutes to inactivate the proteinase K. Nematode
438 lysates were then kept at -20˚C for stable long-term storage until being used for genotyping and
439 quantification.
440
441 **Quantifying mutant mtDNA levels**
442 Levels of each mutant mtDNA variant were quantified as described previously for *uaDf5*
443 (Gitschlag et al., 2020), using droplet digital PCR (ddPCR). Nematodes were lysed as described
444 above, then diluted in nuclease-free water, with a dilution factor varying depending on nematode
445 concentration: 200x for single adults, 1000x for pooled adults (sub-organismal selection
446 experiment), and 20,000x for pooled nematodes of mixed age (organismal selection
447 experiment). Next, 5 µL of each dilute nematode lysate was combined with 0.25 µL of a 10-µM
448 dilution of each primer needed for amplifying each mtDNA variant. The following primers were
449 used for ddPCR amplification of each mutant variant as necessary.
450
451 *mpt1*:
452 Forward primer 1: 5'-TTCTGAAGGTGAAAGGGAGTTAG-3'
453 Reverse primer 1: 5'-ACTGACCTTAATGGTAAAGTATTTG-3'
454 Forward primer 2: 5'-GGGTACTTTCTTCTGCTATAT-3'
455 Reverse primer 2: 5'-CGATAACGAGGGTATGAACTACG-3'
456 *mpt2*:
457 Forward primer 1: 5'-GAAGAAGGTGGTAGCCTTGAGGAC-3'
458 Reverse primer 1: 5'-CGTATAAGAAAAGTCTTGGGATGTTAAG-3'
459 Forward primer 2: 5'-GGATTAATTTTCTCAAGGGGTGCTG-3'
460 Reverse primer 2: 5'-CTTTTTCAAAGACGAAAACTGTAACC-3'
461 *mptDf3*:

462 Forward primer 1: 5'-CCCTGAAGAGGCTAAGAATATTAGG-3'

463 Reverse primer 1: 5'-GGCAATGTCACCAACATCC-3'

464 Reverse primer 2: 5'-CCCAATACAATAACTAGAATAGCTCACG-3'

465 *mptDf2*:

466 Forward primer 1: 5'-GGATTGGCAGTTTGATTAGAGAG-3'

467 Reverse primer 1: 5'-AAGTAACAAACACTAAAACTCCCAAC-3'

468 Forward primer 2: 5'-CGTGCTTATTTTTCGGCTGC-3'

469 Reverse primer 2: 5'-CTTTAACACCTGTTGGCACTG-3'

470 *mpt4*:

471 Forward primer 1: 5'-CGGTGGTTTTGGTAACTG-3'

472 Reverse primer 1: 5'-TCATAGTGTAACACCCGTGAAAATCC-3'

473 Forward primer 2: 5'-TGATCCAAGAACTGGAGGTAATC-3'

474 Reverse primer 2: 5'-CCTGTTGGCACTGCAATAAC-3'

475

476 Mixtures of dilute nematode lysate and primer were combined with 12.5 µL of Bio-Rad QX200™
477 ddPCR EvaGreen Supermix and nuclease-free water to a volume of 25 µL in Eppendorf™ 96-
478 well twin.tec™ PCR plates. Droplet generation and PCR amplification were performed according
479 to manufacturer's protocol. Wildtype and mutant-specific primers were combined in the same
480 reaction, and each droplet was scored as containing either wildtype or mutant mtDNA using the
481 2-dimensional (518 nm and 554 nm dual-wavelength) clustering plot option in the Bio-Rad
482 QuantaSoft™ program.

483

484 **Sub-organismal selection assay**

485 The strength of sub-organismal (within-host) selection on mutant mtDNA was measured
486 longitudinally across isolated parent-progeny lineages, as previously described (Gitschlag et al.,
487 2020). Briefly, multiple L4-stage (late larval) heteroplasmic animals were picked at random

488 under a dissecting microscope from stock populations carrying each heteroplasmy that had
489 been crossed in the Bristol strain (wildtype) nuclear background. These larvae were transferred
490 to fresh NGM plates seeded with OP50 *E. coli* bacteria as a food source and incubated for 2
491 days at 20˚C. The day-2 adults were then segregated by transfer onto individual plates, one
492 adult per plate, and incubated for 4 hours at 20˚C to produce embryos that are age-
493 synchronized to within a four-hour window. Each parent was then individually lysed. After 4 days
494 of continued incubation at 20˚C, the progeny had progressed from embryos to day-2 adults,
495 reaching the same age at which their respective parents were lysed. Adult progeny were lysed
496 at this point to obtain progeny that are age-matched to their parents, to control for age-
497 dependent differences in mutant mtDNA levels. Progeny from each parent were lysed in pools
498 of 3 to minimize the confounding effect of random drift. Each parent-progeny lineage was
499 individually segregated from the rest, to ensure that mutant mtDNA frequency from each
500 progeny lysate was being compared with that of its own respective parent, thereby minimizing
501 the confounding effect of competition between lineages (organismal selection). Mutant mtDNA
502 frequency of parents and progeny was determined for each heteroplasmy using ddPCR as
503 described above, across multiple replicate parent-progeny lineages for statistical power.
504

505 **Experimental evolution (organismal selection)**

506 Selection against mutant mtDNA that occurs strictly at the level of host fitness was measured
507 using an organismal competition experiment similarly to the experiment previously described
508 (Gitschlag et al., 2020). Briefly, for each mutant mtDNA variant, heteroplasmic nematodes
509 carrying mutant mtDNA in the Bristol nuclear background were combined with Bristol-strain
510 nematodes on 10-cm NGM plates seeded with live OP50 *E. coli* bacteria as a food source.
511 Approximately 500 nematodes were transferred to each plate. In addition to 4 replicate
512 competition lines for each heteroplasmy, 4 non-competing control lines were established by
513 transferring only heteroplasmic animals onto their own food plates, with no homoplasmic-

514 wildtype animals to compete against. Every 3 days, the generation for each experimental line
515 was reset; nematodes were washed off the plates using M9 buffer into a sterile 1.7 mL
516 collection tube. Approximately 500 animals of mixed age from each line were transferred to a
517 fresh food plate. Another 500 nematodes were lysed together in a single pooled lysate. This
518 experiment was continued for 6 consecutive generations.
519
520 **Statistical analysis**
521 The effect of initial mutant mtDNA frequency on sub-organismal heteroplasmy dynamics was
522 observed in three previously reported experiments involving *uaDf5* (Gitschlag et al., 2020), and
523 one experiment each for the five additional mutant mtDNA variants featured in this study. Each
524 data point represents a biological replicate, consisting of a parent-progeny pair (sub-organismal
525 selection) or a population of nematodes confined to the same food plate (organismal selection).
526 For each experiment, sample sizes (number of biological replicates) are provided in Figure 3—
527 Source Data (for *uaDf5*) and the Figure 4 figure legend (for the remaining mutant mtDNA
528 variants). Number of animals per replicate is provided in sub-sections **Sub-organismal**
529 **selection assay** and **Experimental evolution (organismal selection)** of the **Materials and**
530 **Methods**. Sample sizes were chosen based on prior observation of sample sizes sufficient for
531 drawing statistical inferences (Gitschlag et al., 2020). Where appropriate, statistical analyses
532 are indicated in the figure legend (Figures 3-5, and Figure 4—Figure Supplement 1); for more
533 information on analysis of data, see Figure 3—Source Data 1-2 and Figure 3—Figure
534 Supplement 1. The Wilcoxon matched-pairs sign-ranks was used for the longitudinal measure of
535 sub-organismal mutant mtDNA dynamics (Figure 4A, 4C, 4E, 4G, and 4I); this was due to the
536 pairing of parent and progeny inherent in the experimental design. Moreover, the Wilcoxon test
537 was chosen because it is non-parametric; that is, it does not require assuming the distribution of
538 mutant mtDNA frequency. Linear regression analysis was used for comparing sub-organismal
539 selection versus starting (parental) mutant mtDNA frequency and for tracking the multi-

540 generational decline in mutant mtDNA frequency due to organismal selection (Figures 3, 5,
541 Figure 4—Figure Supplement 1, Figure 4—Figure Supplement 2, and Figure 5—Figure
542 Supplement 1). No statistical testing was performed to identify or remove outliers. For the sub-
543 organismal selection assay (Figure 4A, 4C, 4E, for which N<24), biological replicates were
544 omitted from final analysis if the ddPCR amplification failed to generate a clear, unambiguous
545 measurement of mutant mtDNA levels in the corresponding plate-well.
546
547
548
549
550
551
552

Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Genetic reagent (*Caenorhabditis elegans*)	*him-8*(e1489); ∆mtDNA(*uaDf5*)/+	Caenorhabditis Genetics Center	RRID:WB-STRAIN:WBStrain00024106	LB138
Genetic reagent (*Caenorhabditis elegans*)	ccTi1594 [mex-5p::GFP::gpr-1::smu-1 3'UTR + Cbr-unc-119(+), III: 680195] III. umnIs27 [myo-2::GFP + NeoR, III: 8856215 (intergenic)] III	Artiles et al. 2019		PD2220

Genetic reagent (*Caenorhabditis elegans*)	∆mtDNA(*uaDf5*)/+ in Bristol nuclear background	Gitschlag et al., 2020		See Materials and Methods: Genetic crosses and genotyping
Genetic reagent (*Caenorhabditis elegans*)	∆mtDNA(*mpt1*)/+ in Bristol nuclear background	This study		See Materials and Methods: Nuclear genome exchange in heteroplasmic strains
Genetic reagent (*Caenorhabditis elegans*)	∆mtDNA(*mpt2*)/+ in Bristol nuclear background	This study		See Materials and Methods: Nuclear genome exchange in heteroplasmic strains
Genetic reagent (*Caenorhabditis elegans*)	∆mtDNA(*mptDf3*)/+ in Bristol nuclear background	This study		See Materials and Methods: Nuclear genome exchange in heteroplasmic strains
Genetic reagent (*Caenorhabditis elegans*)	∆mtDNA(*mptDf2*)/+ in Bristol nuclear background	This study		See Materials and Methods: Nuclear genome exchange in heteroplasmic strains
Genetic reagent (*Caenorhabditis elegans*)	∆mtDNA(*mpt4*)/+ in Bristol nuclear background	This study		See Materials and Methods: Nuclear genome exchange in heteroplasmic strains
Genetic reagent (*Escherichia coli*)	OP50 strain	Caenorhabditis Genetics Center	RRID: WB-STRAIN:WBStrain00041971	
Sequence-based reagent	PCR primers used in this study	This study		See Materials and Methods: Quantifying mutant mtDNA levels
Other	Eppendorf™ 96-well twin.tec™ PCR plates	Fisher Scientific	Cat#951020303	
Other	QX200 ddPCR EvaGreen Supermix	Bio-Rad	Cat#1864034	

Other	Automated Droplet Generation Oil for EvaGreen	Bio-Rad	Cat#1864112	
Other	DG32 Automated Droplet Generator Cartridges	Bio-Rad	Cat#1864108	
Other	Droplet Reader Oil for ddPCR	Bio-Rad	Cat#1863004	
Software, algorithm	QuantaSoft™	Bio-Rad	Cat#1864011	
Software, algorithm	Prism 8 for macOS	GraphPad Software, Inc.	RRID:SCR_002798	Version 8.1.2

553 **Acknowledgments**

554 The *uaDf5* heteroplasmic strain was kindly provided by the Caenorhabditis Genetics Center
555 (CGC). We thank O. Thompson and R. Waterston at University of Washington for identifying the
556 *mpt1*, *mpt2*, *mptDf3*, *mptDf2*, and *mpt4* deletions from the Million Mutation Project worm
557 collection. We thank Patrick Abbot, Ann T. Tate and Allison Hilbun of Vanderbilt University for
558 valuable guidance and assistance with data analysis. This work was supported by the NIH
559 research project grants (R01 GM123260 and GM145378), the Discovery Award (PR170792)
560 from the Department of Defense's Congressionally Directed Medical Research Program, the
561 Vanderbilt University Medical Center Diabetes Research and Training Center Pilot & Feasibility
562 Grant, the NIH-sponsored Cellular, Biochemical and Molecular Sciences Training Program
563 (5T32GM008554-18), and the Ruth L. Kirschstein National Research Service Award Individual
564 Predoctoral Fellowship grant (1F31GM125344).

565

566
567
568
569
570
571
572
573
574
575
576
577
578
579

References

580 AKTIPIS, C. A., BODDY, A. M., JANSEN, G., HIBNER, U., HOCHBERG, M. E., MALEY, C. C. &
581 WILKINSON, G. S. 2015. Cancer across the tree of life: cooperation and cheating in
582 multicellularity. *Philos Trans R Soc Lond B Biol Sci,* 370.
583 BERNARDI, G. 2005. Lessons from a small, dispensable genome: the mitochondrial genome of
584 yeast. *Gene,* 354**,** 189-200.
585 BRAVO NUNEZ, M. A., NUCKOLLS, N. L. & ZANDERS, S. E. 2018. Genetic Villains: Killer Meiotic
586 Drivers. *Trends Genet,* 34**,** 424-433.
587 BROWN, M. D., YANG, C. C., TROUNCE, I., TORRONI, A., LOTT, M. T. & WALLACE, D. C. 1992. A
588 mitochondrial DNA variant, identified in Leber hereditary optic neuropathy patients,
589 which extends the amino acid sequence of cytochrome c oxidase subunit I. *Am J Hum*
590 *Genet,* 51**,** 378-85.
591 CHIANG, A. C., MCCARTNEY, E., O'FARRELL, P. H. & MA, H. 2019. A Genome-wide Screen
592 Reveals that Reducing Mitochondrial DNA Polymerase Can Promote Elimination of
593 Deleterious Mitochondrial Mutations. *Curr Biol,* 29**,** 4330-4336 e3.
594 CLARK, K. A., HOWE, D. K., GAFNER, K., KUSUMA, D., PING, S., ESTES, S. & DENVER, D. R. 2012.
595 Selfish little circles: transmission bias and evolution of large deletion-bearing
596 mitochondrial DNA in Caenorhabditis briggsae nematodes. *PLoS One,* 7**,** e41433.
597 DE VARGAS RODITI, L., BOYLE, K. E. & XAVIER, J. B. 2013. Multilevel selection analysis of a
598 microbial social trait. *Mol Syst Biol,* 9**,** 684.

599 DEPUYDT, G., XIE, F., PETYUK, V. A., SMOLDERS, A., BREWER, H. M., CAMP, D. G., 2ND, SMITH,
600 R. D. & BRAECKMAN, B. P. 2014. LC-MS proteomics analysis of the insulin/IGF-1-
601 deficient Caenorhabditis elegans daf-2(e1370) mutant reveals extensive restructuring of
602 intermediary metabolism. *J Proteome Res,* 13**,** 1938-56.
603 DOBATA, S. & TSUJI, K. 2013. Public goods dilemma in asexual ant societies. *Proc Natl Acad Sci*
604 *U S A,* 110**,** 16056-60.
605 DUBIE, J. J., CARAWAY, A. R., STOUT, M. M., KATJU, V. & BERGTHORSSON, U. 2020. The conflict
606 within: origin, proliferation and persistence of a spontaneously arising selfish
607 mitochondrial genome. *Philos Trans R Soc Lond B Biol Sci,* 375**,** 20190174.
608 DUGATKIN, L. A., PERLIN, M., LUCAS, J. S. & ATLAS, R. 2005. Group-beneficial traits, frequency-
609 dependent selection and genotypic diversity: an antibiotic resistance paradigm. *Proc Biol*
610 *Sci,* 272**,** 79-83.
611 FIEGNA, F. & VELICER, G. J. 2003. Competitive fates of bacterial social parasites: persistence and
612 self-induced extinction of Myxococcus xanthus cheaters. *Proc Biol Sci,* 270**,** 1527-34.
613 FISHER, R. M. & REGENBERG, B. 2019. Multicellular group formation in Saccharomyces
614 cerevisiae. *Proc Biol Sci,* 286**,** 20191098.
615 FRANK, S. A. 1997. The Price Equation, Fisher's fundamental theorem, kin selection, and causal
616 analysis. *Evolution,* 51**,** 1712-1729.
617 FRANK, S. A. 2012. Natural selection. IV. The Price equation. *J Evol Biol,* 25**,** 1002-19.
618 GHOUL, M., GRIFFIN, A. S. & WEST, S. A. 2014. Toward an evolutionary definition of cheating.
619 *Evolution,* 68**,** 318-31.
620 GITSCHLAG, B. L., KIRBY, C. S., SAMUELS, D. C., GANGULA, R. D., MALLAL, S. A. & PATEL, M. R.
621 2016. Homeostatic Responses Regulate Selfish Mitochondrial Genome Dynamics in C.
622 elegans. *Cell Metab,* 24**,** 91-103.
623 GITSCHLAG, B. L., TATE, A. T. & PATEL, M. R. 2020. Nutrient status shapes selfish mitochondrial
624 genome dynamics across different levels of selection. *Elife,* 9.
625 GULLI, J. G., HERRON, M. D. & RATCLIFF, W. C. 2019. Evolution of altruistic cooperation among
626 nascent multicellular organisms. *Evolution,* 73**,** 1012-1024.
627 HAMMERSCHMIDT, K., ROSE, C. J., KERR, B. & RAINEY, P. B. 2014. Life cycles, fitness decoupling
628 and the evolution of multicellularity. *Nature,* 515**,** 75-9.
629 HAMMOND, T. M., REHARD, D. G., XIAO, H. & SHIU, P. K. 2012. Molecular dissection of
630 Neurospora Spore killer meiotic drive elements. *Proc Natl Acad Sci U S A,* 109**,** 12093-8.
631 HAVIRD, J. C., FORSYTHE, E. S., WILLIAMS, A. M., WERREN, J. H., DOWLING, D. K. & SLOAN, D. B.
632 2019. Selfish Mitonuclear Conflict. *Curr Biol,* 29**,** R496-R511.
633 HOLT, I. J., HARDING, A. E. & MORGAN-HUGHES, J. A. 1988. Deletions of muscle mitochondrial
634 DNA in patients with mitochondrial myopathies. *Nature,* 331**,** 717-9.
635 HU, W., JIANG, Z. D., SUO, F., ZHENG, J. X., HE, W. Z. & DU, L. L. 2017. A large gene family in
636 fission yeast encodes spore killers that subvert Mendel's law. *Elife,* 6.
637 KANDUL, N. P., ZHANG, T., HAY, B. A. & GUO, M. 2016. Selective removal of deletion-bearing
638 mitochondrial DNA in heteroplasmic Drosophila. *Nat Commun,* 7**,** 13100.
639 KLUCNIKA, A. & MA, H. 2019. A battle for transmission: the cooperative and selfish animal
640 mitochondrial genomes. *Open Biol,* 9**,** 180267.
641 KONRAD, A., THOMPSON, O., WATERSTON, R. H., MOERMAN, D. G., KEIGHTLEY, P. D.,
642 BERGTHORSSON, U. & KATJU, V. 2017. Mitochondrial Mutation Rate, Spectrum and

643 Heteroplasmy in Caenorhabditis elegans Spontaneous Mutation Accumulation Lines of
644 Differing Population Size. *Mol Biol Evol,* 34**,** 1319-1334.
645 LARRACUENTE, A. M. & PRESGRAVES, D. C. 2012. The selfish Segregation Distorter gene
646 complex of Drosophila melanogaster. *Genetics,* 192**,** 33-53.
647 LEIGH, E. G. 1983. When does the good of the group override the advantage of the individual?
648 *Proc Natl Acad Sci U S A,* 80**,** 2985-9.
649 LIAU, W. S., GONZALEZ-SERRICCHIO, A. S., DESHOMMES, C., CHIN, K. & LAMUNYON, C. W. 2007.
650 A persistent mitochondrial deletion reduces fitness and sperm performance in
651 heteroplasmic populations of C. elegans. *BMC Genet,* 8**,** 8.
652 LIEBER, T., JEEDIGUNTA, S. P., PALOZZI, J. M., LEHMANN, R. & HURD, T. R. 2019. Mitochondrial
653 fragmentation drives selective removal of deleterious mtDNA in the germline. *Nature,*
654 570**,** 380-384.
655 LIN, Y. F., SCHULZ, A. M., PELLEGRINO, M. W., LU, Y., SHAHAM, S. & HAYNES, C. M. 2016.
656 Maintenance and propagation of a deleterious mitochondrial genome by the
657 mitochondrial unfolded protein response. *Nature,* 533**,** 416-9.
658 MICHOD, R. E., VIOSSAT, Y., SOLARI, C. A., HURAND, M. & NEDELCU, A. M. 2006. Life-history
659 evolution and the origin of multicellularity. *J Theor Biol,* 239**,** 257-72.
660 MORENO-FENOLL, C., CAVALIERE, M., MARTINEZ-GARCIA, E. & POYATOS, J. F. 2017. Eco-
661 evolutionary feedbacks can rescue cooperation in microbial populations. *Sci Rep,* 7**,**
662 42561.
663 NAMSLAUER, I. & BRZEZINSKI, P. 2009. A mitochondrial DNA mutation linked to colon cancer
664 results in proton leaks in cytochrome c oxidase. *Proc Natl Acad Sci U S A,* 106**,** 3402-7.
665 NARGUND, A. M., PELLEGRINO, M. W., FIORESE, C. J., BAKER, B. M. & HAYNES, C. M. 2012.
666 Mitochondrial import efficiency of ATFS-1 regulates mitochondrial UPR activation.
667 *Science,* 337**,** 587-90.
668 PRICE, G. R. 1972. Extension of Covariance Selection Mathematics. *Annals of Human Genetics,*
669 35**,** 485-&.
670 PRUITT, J. N. & RIECHERT, S. E. 2009. Frequency-dependent success of cheaters during foraging
671 bouts might limit their spread within colonies of a socially polymorphic spider. *Evolution,*
672 63**,** 2966-73.
673 QUELLER, D. C. & STRASSMANN, J. E. 2009. Beyond society: the evolution of organismality.
674 *Philos Trans R Soc Lond B Biol Sci,* 364**,** 3143-55.
675 RAINEY, P. B. & RAINEY, K. 2003. Evolution of cooperation and conflict in experimental bacterial
676 populations. *Nature,* 425**,** 72-4.
677 RIEHL, C. & FREDERICKSON, M. E. 2016. Cheating and punishment in cooperative animal
678 societies. *Philos Trans R Soc Lond B Biol Sci,* 371**,** 20150090.
679 ROSS-GILLESPIE, A., GARDNER, A., WEST, S. A. & GRIFFIN, A. S. 2007. Frequency dependence
680 and cooperation: theory and a test with bacteria. *Am Nat,* 170**,** 331-42.
681 SAGAN, L. 1967. On the origin of mitosing cells. *J Theor Biol,* 14**,** 255-74.
682 SCHIMENTI, J. 2000. Segregation distortion of mouse t haplotypes the molecular basis emerges.
683 *Trends Genet,* 16**,** 240-3.
684 SHOU, W. 2015. Acknowledging selection at sub-organismal levels resolves controversy on pro-
685 cooperation mechanisms. *Elife,* 4.

686 SUEN, D. F., NARENDRA, D. P., TANAKA, A., MANFREDI, G. & YOULE, R. J. 2010. Parkin
687 overexpression selects against a deleterious mtDNA mutation in heteroplasmic cybrid
688 cells. *Proc Natl Acad Sci U S A,* 107**,** 11835-40.
689 TAYLOR, R. W. & TURNBULL, D. M. 2005. Mitochondrial DNA mutations in human disease. *Nat*
690 *Rev Genet,* 6**,** 389-402.
691 TEPPER, R. G., ASHRAF, J., KALETSKY, R., KLEEMANN, G., MURPHY, C. T. & BUSSEMAKER, H. J.
692 2013. PQM-1 complements DAF-16 as a key transcriptional regulator of DAF-2-mediated
693 development and longevity. *Cell,* 154**,** 676-690.
694 TSANG, W. Y. & LEMIRE, B. D. 2002. Stable heteroplasmy but differential inheritance of a large
695 mitochondrial DNA deletion in nematodes. *Biochem Cell Biol,* 80**,** 645-54.
696 VALENCI, I., YONAI, L., BAR-YAACOV, D., MISHMAR, D. & BEN-ZVI, A. 2015. Parkin modulates
697 heteroplasmy of truncated mtDNA in Caenorhabditis elegans. *Mitochondrion,* 20**,** 64-70.
698 VARLAMOV, D. A., KUDIN, A. P., VIELHABER, S., SCHRODER, R., SASSEN, R., BECKER, A., KUNZ,
699 D., HAUG, K., REBSTOCK, J., HEILS, A., ELGER, C. E. & KUNZ, W. S. 2002. Metabolic
700 consequences of a novel missense mutation of the mtDNA CO I gene. *Hum Mol Genet,*
701 11**,** 1797-805.
702 WEBB, A. E., KUNDAJE, A. & BRUNET, A. 2016. Characterization of the direct targets of FOXO
703 transcription factors throughout evolution. *Aging Cell,* 15**,** 673-85.
704 WENSELEERS, T. & RATNIEKS, F. L. 2004. Tragedy of the commons in Melipona bees. *Proc Biol*
705 *Sci,* 271 Suppl 5**,** S310-2.
706 WEST, S. A., FISHER, R. M., GARDNER, A. & KIERS, E. T. 2015. Major evolutionary transitions in
707 individuality. *Proc Natl Acad Sci U S A,* 112**,** 10112-9.
708 YANG, Q., LIU, P., ANDERSON, N. S., SHPILKA, T., DU, Y., NARESH, N. U., LI, R., ZHU, L. J., LUK, K.,
709 LAVELLE, J., ZEINERT, R. D., CHIEN, P., WOLFE, S. A. & HAYNES, C. M. 2022. LONP-1 and
710 ATFS-1 sustain deleterious heteroplasmy by promoting mtDNA replication in
711 dysfunctional mitochondria. *Nat Cell Biol,* 24**,** 181-193.
712 ZHANG, H., ESPOSITO, M., PEZET, M. G., ARYAMAN, J., WEI, W., KLIMM, F., CALABRESE, C.,
713 BURR, S. P., MACABELLI, C. H., VISCOMI, C., SAITOU, M., CHIARATTI, M. R., STEWART, J.
714 B., JONES, N. & CHINNERY, P. F. 2021. Mitochondrial DNA heteroplasmy is modulated
715 during oocyte development propagating mutation transmission. *Sci Adv,* 7**,** eabi5657.
716
717
718
719
720
721
722
723

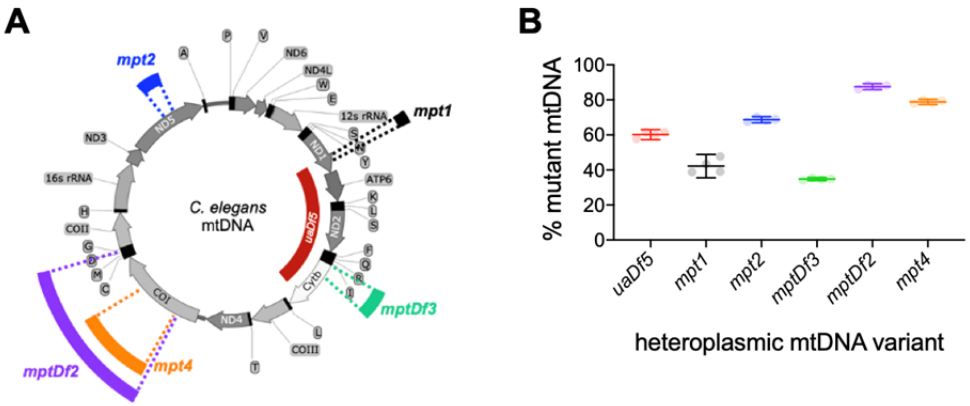


Figure 1. Mitochondrial mutations featured in this study. (**A**) Map of the mitochondrial genome in *C. elegans* showing location and size of each heteroplasmic mutation featured in this study. The relative sizes and locations of protein and ribosomal RNA genes are represented by the gray arrows; thick black bars represent tRNA genes; thin black lines represent noncoding regions. The *uaDf5* deletion featured in previous work (Gitschlag et al., 2020) is shown in dark red. Colored bars around the perimeter of the genome represent approximate locations and sizes of the remaining mtDNA deletions used in this study. Mutations are color-coded corresponding to the visualization of data in panel (**B**) and in subsequent figures in this study. (**B**) Heteroplasmic mutant mtDNA frequency sampled from N=4 lines stably maintaining each mutation featured in (**A**), each crossed into the wildtype (Bristol strain) nuclear background, and quantified using droplet digital PCR.

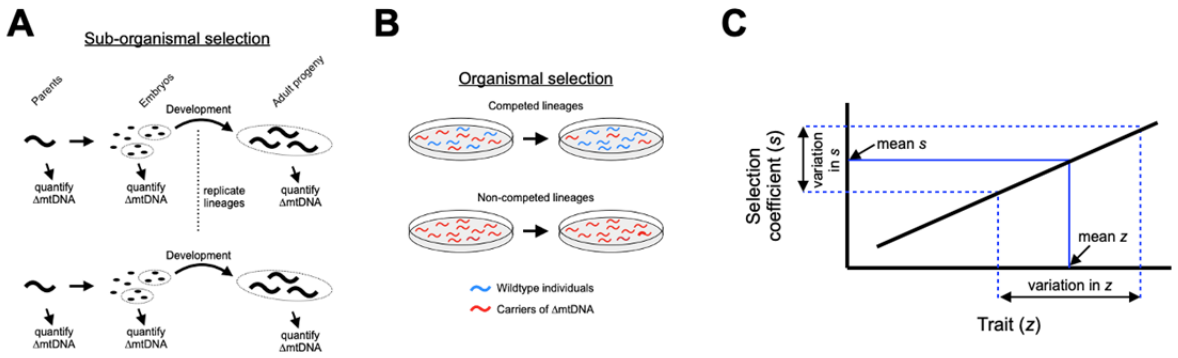


Figure 2. Experimental approach to characterize multilevel selection dynamics of mutant mtDNA. (**A**) Longitudinal experiment, featuring measurement of mutant mtDNA between individual parents and their own respective adult progeny, to quantify sub-organismal selection. (**B**) Competition experiment designed to quantify organismal selection against mutant mtDNA. (**C**) Graphical depiction of covariance between a quantifiable trait value and selection coefficient. Covariance implies that variation in a measured trait value *z*, such as mutant mtDNA frequency, corresponds to variation in the selection coefficient *S*, a measure of the strength of the selection acting on *z*. A slope of zero indicates that variation in the trait does not correspond to variation in the fitness effects of the trait.

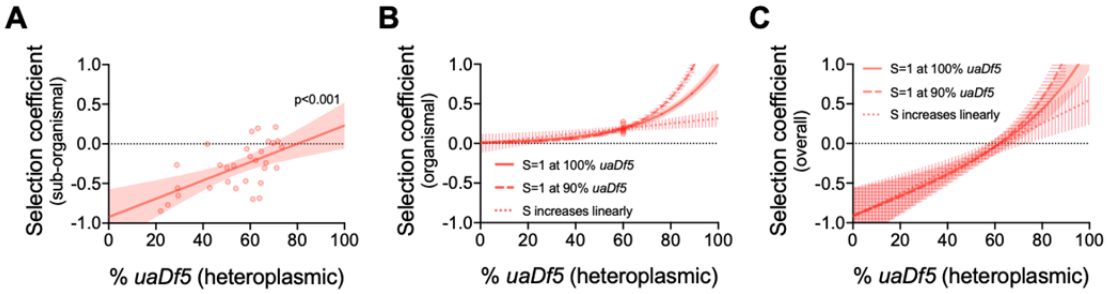
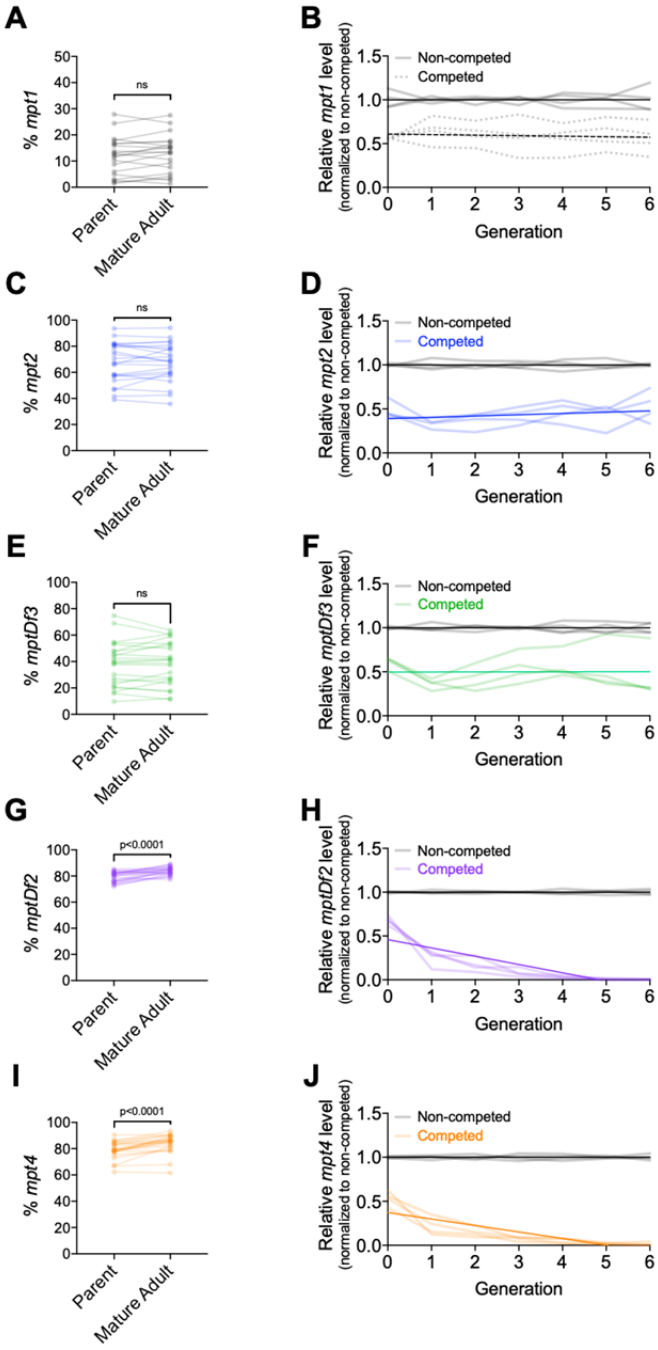


Figure 3. Integration of the multilevel selection dynamics of *uaDf5* using the Price Equation. (**A**) Linear regression between *uaDf5* frequency and the strength of sub-organismal selection, obtained from converting the previously measured intergenerational shifts in *uaDf5* frequency (Gitschlag et al., 2020) to selection coefficients (red circles) and plotting them as a function of parental *uaDf5* frequency. Red shaded regions represents 95% confidence interval. F-test for non-zero slope (sample size and raw data shown in Figure 3—Source Data 1). (**B**) Non-linear regression between heteroplasmic (not population-wide) *uaDf5* frequency and the strength of organismal selection, obtained from converting the previously measured declines in *uaDf5* frequency from the organismal competition experiment (Gitschlag et al., 2020) to selection coefficients (red circles). Solid regression and shaded region represents the best fit and 95% confidence interval, respectively, assuming the organismal selection coefficient reaches 1 (lethality) when *uaDf5* frequency is 100%. For comparison, dashed line and horizontally-striped region represents the best fit and 95% confidence interval, respectively, assuming the organismal selection coefficient reaches 1 when *uaDf5* frequency is 90%. Dotted line and vertically-striped region represents the best fit and 95% confidence interval, respectively, assuming the selection coefficient increases linearly with *uaDf5* frequency (raw data shown in Figure 3—Source Data 2). (**C**) Overall population-level

regression (summation of sub-organismal and organismal regressions in panels **A**-**B**), showing the population-level consequences of each organismal covariance model assumed in panel (**B**). For each graph, the black horizontal dotted line represents no net selection. Above the dotted horizontal line: selection favors wildtype mtDNA. Below the dotted horizontal line: selection favors *uaDf5*.



796 **Figure 4.** Data from experiments designed to measure influence of sub-organismal and organismal levels
797 of selection on mtDNA heteroplasmies. (**A**-**B**) Intergenerational shifts in *mpt1* frequency at the levels of
798 sub-organismal (**A**) and organismal (**B**) selection. Wilcoxon matched-pairs signed-rank test, N=21 (**A**).
799 Linear regression analysis (**B**). (**C**-**D**) Same as above but for the *mpt2* heteroplasmy. N=23 (**C**). (**E**-**F**)
800 Same as above but for the *mptDf3* heteroplasmy. N=23 (**E**). (**G**-**H**) Same as above but for the *mptDf2*
801 heteroplasmy. N=24 (**G**). (**I**-**J**) Same as above but for the *mpt4* heteroplasmy. N=24 (**I**). Raw data shown
802 in Figure 4—Source Data 1 and Figure 4—Source Data 2.

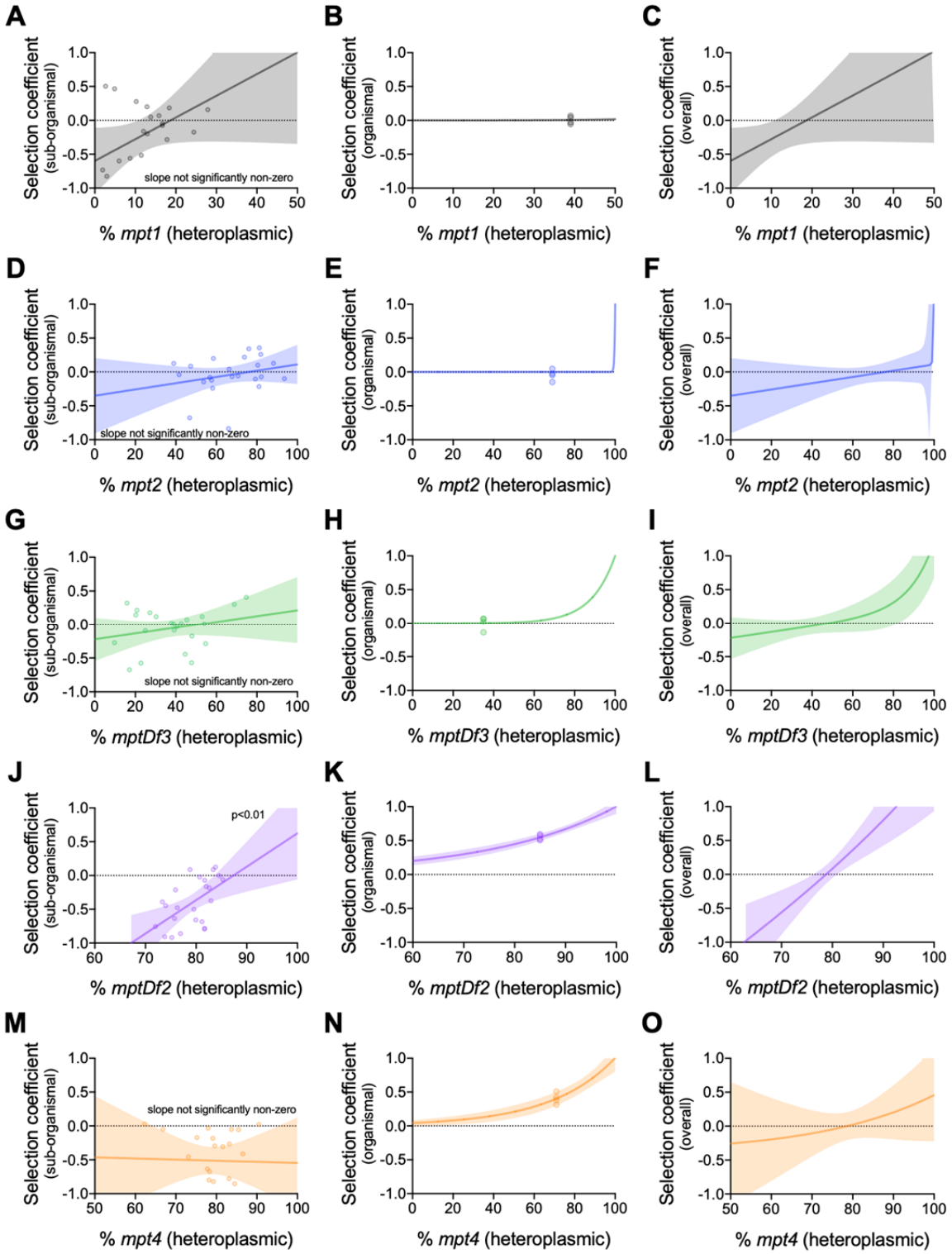


Figure 5. Integration of the multilevel selection dynamics of all five heteroplasmies using the Price Equation framework. For each graph shown above, the horizontal dotted line represents no net selection. Above the dotted horizontal line: selection favors wildtype mtDNA. Below the dotted horizontal line:

807 selection favors ∆mtDNA. (**A**-**C**) The plotted regressions between heteroplasmic frequency and selection
808 coefficient at sub-organismal, organismal, and overall (combined) levels, using data from Figure 4A-B.
809 Same analysis as in Figure 3 but applied to the *mpt1* heteroplasmy. (**D**-**F**) Same as above but for *mpt2*
810 (data from Figure 4C-D). (**G**-**I**) Same as above but for *mptDf3* (data from Figure 4E-F). (**J**-**L**) Same as
811 above but for *mptDf2* (data from Figure 4G-H). (**M**-**O**) Same as above but for *mpt4* (data from Figure 4I-J).
812 F-test for non-zero slope (panels **A**, **D**, **G**, **J**, and **M**; sample sizes and mutant frequency data shown in
813 Figure 4 and Figure 4—Figure Supplement 1). As before (see Figure 3), the organismal and combined
814 ("overall") plots respectively represent the 95% confidence ranges that result from assuming the selection
815 coefficient reaches 1 (lethality) when the mutant is at 100 percent frequency. For visual simplicity, only
816 this assumption is shown here; see Figure 5—Figure Supplement 1 for consequences of alternately
817 assuming the organismal selection coefficient reaches 1 when mutant mtDNA frequency is 90% or
818 increases linearly with mutant mtDNA frequency. Note that confidence ranges capable of accommodating
819 a horizontal regression indicate a slope that is not significantly non-zero, implying no significant
820 relationship between mutant mtDNA frequency and selection coefficient, at least within the mutant
821 frequency range in which the data was collected.
822
823
824
825
826
827
828
829
830
831
832
833
834

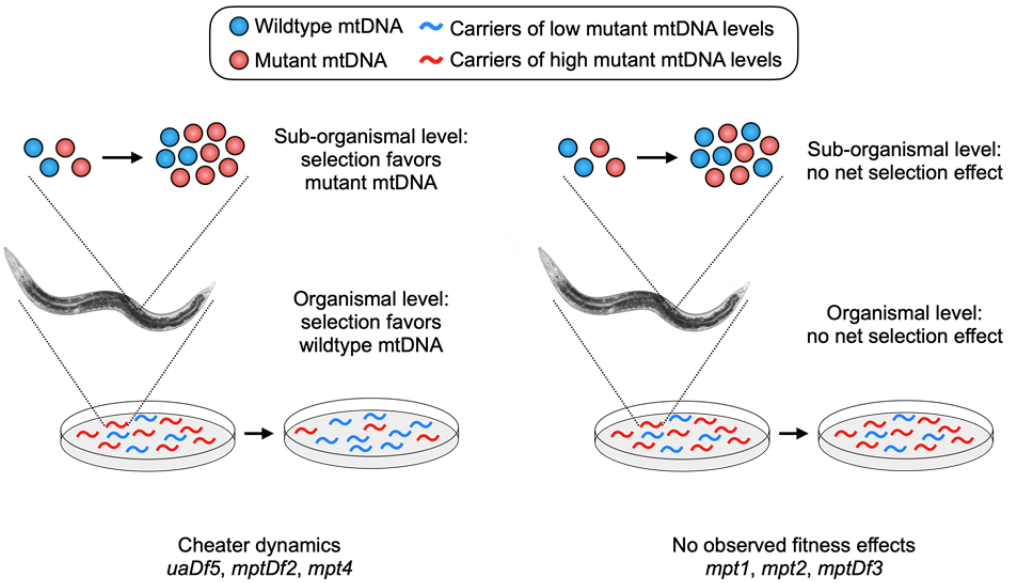


Figure 6. Alternate population dynamics depending on mitochondrial genotype. **Left**: sub-organismal or within-host dynamics favor the preferential proliferation of mutant over wildtype mtDNA across individual parent-progeny lineages, when the effect of organismal selection is minimized. Simultaneously, selection for host fitness strongly favors wildtype over mutant mtDNA. These dynamics are observed for heteroplasmic mtDNA variants *uaDf5*, *mptDf2*, and *mpt4*. In contrast to *uaDf5* and *mptDf2*, however, the sub-organismal proliferation of *mpt4* does not appear to be frequency-dependent (Figure 5M). **Right**: for some heteroplasmic variants (*mpt1*, *mpt2*, and *mptDf3*), neither sub-organismal dynamics nor selection at the organismal level favor a significant shift in mutant mtDNA frequency, at least at the frequencies at which these mutant genomes were observed to stably persist (Figure 1B).

Figure 3—Figure Supplement 1. Mathematical framework for multilevel selection analysis of *uaDf5*.

Here, the regressions shown in Figure 3 are derived from the data obtained from the measurements of organismal and sub-organismal selection on *uaDf5* (Gitschlag et al., 2020), using the general framework of the Price Equation. Hereafter, *uaDf5* is referred to as ∆mtDNA to denote a deletion-bearing genome, to

emphasize that this framework can be applied to other mutant mtDNA (Figure 5). We begin with the Price Equation:

$$w\Delta z = cov(w,z)_{org} + E[cov(w,z)_{sub}]$$

Where:
w = reproductive fitness, or number of replication events per generation, for one trait relative to another
z = measurable value of a trait (in this case, mutant mtDNA frequency)
$cov(w,z)_{org}$ = covariance between w and z contributed by selection at the organismal level
$E[cov(w,z)_{sub}]$ = expected or average covariance between w and z contributed by sub-organismal selection

Note that the version of the Price Equation shown above is equivalent to the form shown as equation A17 in Price, 1972. At each level of selection, organismal and sub-organismal, the term for the covariance between w and z can be defined as follows:

$$cov(w,z) = \frac{\sum(w_i - w_{average})(z_i - z_{average})}{n-1}$$

Where:
w_i = fitness w of trait z in the ith replicate parent-progeny lineage from the sample or population
$w_{average}$ = average w across the sample or population
z_i = value of trait z in the parent of the ith replicate parent-progeny lineage from the sample or population
$z_{average}$ = average value of z across the sample or population
n = number of parent-progeny lineages sampled

Since the selection coefficient, which measures the strength of selection acting on a trait, is directly to fitness, the Price Equation can be expressed using covariance terms between selection coefficient, S, and the value of a trait, z. Defining trait-value as the heteroplasmic (sub-organismal or within-host) frequency of mutant mtDNA (*%Δ mtDNA*) yields the following form of the Price Equation:

$$Change\ in\ \%\Delta mtDNA = cov(S, \%\Delta mtDNA)_{org} + E[cov(S, \%\Delta mtDNA)_{sub}]$$

Replacing w and z in the Price Equation with S and *%Δ mtDNA* yields the following formula for covariance:

$$cov(S, \%\Delta mtDNA) = \frac{\sum(S_i - S_{average})(\%\Delta mtDNA_i - \%\Delta mtDNA_{average})}{n-1}$$

Where:
S_i = selection coefficient of ΔmtDNA in the ith replicate of each experiment (sub-organismal and organismal)
$S_{average}$ = average S across the sample or population of each experiment
$\%\Delta mtDNA_i$ = heteroplasmic ΔmtDNA frequency from the ith parent (sub-organismal) or line (organismal)
$\%\Delta mtDNA_{average}$ = average heteroplasmic ΔmtDNA frequency across all replicates

In the sub-sections that follow, the individual regressions are derived for the strength of sub-organismal and organismal selection as a function of heteroplasmic ΔmtDNA frequency, shown using the data from (Gitschlag et al., 2020). (see also Figure 3—Source data files). The multilevel selection analyses in Figure 5 are performed using the same approach and data shown in Figure 4.

Sub-organismal regression (Figure 3A)

For selection acting on ∆mtDNA at the sub-organismal level, selection coefficient, *S*, is calculated for a given change in ∆mtDNA frequency between parent and progeny using the following equation:

$$S = 1 - w = 1 - \frac{\%\Delta mtDNA_{g+1}/\%\Delta mtDNA_g}{\%wildtype\ mtDNA_{g+1}/\%wildtype\ mtDNA_g}$$

<u>Where:</u>
w = average replicative fitness of ∆mtDNA relative to wildtype mtDNA
g = parental generation

The following regression, relating selection coefficient to heteroplasmic ∆mtDNA frequency, can be derived using the covariance relationship defined on the previous page and the data shown in Figure 2—Source data 1:

$$S = slope(S, \%\Delta mtDNA) * \%\Delta mtDNA + intercept(S, \%\Delta mtDNA)$$

Note that this equation takes the form of the slope-intercept equation, *y=mx+b*, in which *y* corresponds to selection coefficient, *S*, and *x* corresponds to heteroplasmic ∆mtDNA frequency, denoted *%∆mtDNA*. The slope, *m*, can be calculated from its relationship to covariance and variance. In particular, the covariance between *S* and *%∆mtDNA* is equal to the variance of *%∆mtDNA* multiplied by the slope of the regression (Price, 1970):

$$cov(S, \%\Delta mtDNA) = var(\%\Delta mtDNA) * slope(S, \%\Delta mtDNA)$$

Here, the term *var*(*S,%∆mtDNA*) represents the variance of *%∆mtDNA*, calculated as follows:

$$var(\%\Delta mtDNA) = \frac{\sum(\%\Delta mtDNA_i - \%\Delta mtDNA_{average})^2}{n-1}$$

Note that the formula for variance is similar to the formula for covariance, with one exception: the errors or residuals in *S* (S_i - $S_{average}$) are replaced by the errors in *%∆mtDNA* ($\%\Delta mtDNA_i$ - $\%\Delta mtDNA_{average}$). In other words, in contrast to the covariance of *S* and *%∆mtDNA*, the variance of *%∆mtDNA* is equivalent to the covariance of *%∆mtDNA* with itself. The term *slope*(*S,%∆mtDNA*), shown above, represents the slope of the regression obtained by plotting *S* as a function of *%∆mtDNA*. The slope can be calculated by rearranging the relation between covariance, variance, and slope, as follows:

$$slope(S, \%\Delta mtDNA) = \frac{cov(S, \%\Delta mtDNA)}{var(\%\Delta mtDNA)} = \frac{\sum(S_i - S_{average})(\%\Delta mtDNA_i - \%\Delta mtDNA_{average})}{\sum(\%\Delta mtDNA_i - \%\Delta mtDNA_{average})^2}$$

Entering the empirically derived values for *S* and *%∆mtDNA* (Figure 3—Source data 1) yields a slope of 0.01154. Next, the slope-intercept equation can be rearranged to solve for the intercept, as follows:

$$intercept(S, \%\Delta mtDNA) = S - slope(S, \%\Delta mtDNA) * \%\Delta mtDNA$$

By entering the slope, together with any pair of corresponding values of *S* and *%∆mtDNA* that fall on the regression line (such as $S_{average}$ and $\%\Delta mtDNA_{average}$, taken from Figure 3—Source data 1), the intercept is calculated to be -0.923 as follows:

$$intercept(S, \%\Delta mtDNA) = S - slope(S, \%\Delta mtDNA) * \%\Delta mtDNA = -0.289 - 0.01154 * 54.96 = -0.923$$

Accordingly, the covariance between selection coefficient and ∆mtDNA frequency at the sub-organismal level yields the following relation:

$$S = 0.01154 * \%\Delta mtDNA - 0.923$$

This equation defines the regression line shown in Figure 3A. The confidence interval surrounding the regression (defining the shaded region) is calculated for each value of ∆mtDNA frequency, from 0 to 100, using the following formula:

$$95\%\ confidence\ interval = t_{n-2} * \sigma_R \sqrt{\frac{1}{n} + \frac{(\%\Delta mtDNA^* - \%\Delta mtDNA_{ave})^2}{{\sigma_{RS}}^2 (n-1)}}$$

Where:
t_{n-2} = critical value from Student's t-table for *n-2*, two-tailed, α=0.05
σ_R = standard deviation of the residuals for selection coefficient *S*
n = sample size
$\%\Delta mtDNA^*$ = each value of ∆mtDNA frequency at which confidence interval is being calculated
$\%\Delta mtDNA_{ave}$ = average ∆mtDNA frequency across all samples
σ_{RS} = standard deviation of the residuals of *S*, with each residual defined as $S_{actual} - S_{predicted}$

Organismal regression (Figure 3B)

For selection acting on ∆mtDNA at the organismal level, in other words the selection against ∆mtDNA that arises strictly from the cost that ∆mtDNA imposes on host fitness, the selection coefficient, *S*, is calculated from the change in ∆mtDNA frequency across a population consisting of heteroplasmic hosts (carrying ∆mtDNA) competed against their wildtype counterparts on the same food plate. In contrast to the sub-organismal selection experiment, where *S* is calculated from the change in ∆mtDNA frequency between parents and their immediate progeny, the organismal competition experiment involved measuring the change in ∆mtDNA frequency across multiple successive generations. Accordingly, the selection coefficient can be calculated as follows:

$$S = 1 - w = 1 - e^{slope[g, LN(\%\Delta mtDNA/\%WT\ mtDNA)]}$$

Where:
LN = natural logarithm
%∆mtDNA = population-wide (not heteroplasmic) ∆mtDNA frequency
%WT mtDNA = population-wide wildtype mtDNA frequency (equal to 100 – *%∆mtDNA*)
g = time, measured in number of generations
LN(*%∆mtDNA* / *%WT mtDNA*)] = natural logarithm of the ratio of *%∆mtDNA* to *%WT mtDNA*
slope[*g*,*LN*(*%∆mtDNA* / *%WT mtDNA*)] = slope of *LN*(*%∆mtDNA* / *%WT mtDNA*)] versus time, as follows:

$$slope[g, LN(\%\Delta mtDNA/\%WT\ mtDNA)] = \frac{\sum\left[LN\left(\frac{\Delta}{W}\right)_i - LN\left(\frac{\Delta}{W}\right)_{average}\right](g_i - g_{average})}{\sum(g_i - g_{average})^2}$$

Where:
$LN(\Delta/W)_i$ = *LN*(*%∆mtDNA* / *%WT mtDNA*) for the *i*th generation
$LN(\Delta/W)_{average}$ = average value of $LN(\%\Delta mtDNA_I / \%WT\ mtDNA)$ across all generations
g_i = the *i*th generation
$g_{average}$ = average g_i value across all generations (equal to 5.5 for generations 1 to 10)

The selection coefficient of ∆mtDNA is calculated for both the competed and non-competed populations in the organismal competition experiment. Although ∆mtDNA persists within a stable frequency range in non-competed populations, ∆mtDNA undergoes slight variation in frequency, perhaps due to drift, variations in sub-organismal selection, and population structure (during population expansion at the onset of each new generation, adults become rapidly outnumbered by embryos and larvae, which tend to carry

lower ΔmtDNA than adults, see "Non-competing" section in Figure 3—Source data 2). In any case, to control for sources other than organismal selection (that is, variation in ΔmtDNA frequency occurring for reasons other than the fitness differences between heteroplasmic animals and their wildtype counterparts), the average selection coefficient across all replicate non-competed populations (containing exclusively heteroplasmic animals) is subtracted from the selection coefficient of each replicate competed population, yielding the selection coefficient for organismal selection against ΔmtDNA:

$$S_{organismal} = S_{competed} - \bar{S}_{non-competed}$$

Where:
$S_{organismal}$ = selection coefficient, S, strictly at the level of organismal selection
$S_{competed}$ = S across a competed line
$\bar{S}_{non\text{-}competed}$ = average S across all non-competed lines

Deriving a regression between selection coefficient and heteroplasmic ΔmtDNA frequency is trickier at the organismal level, since the relationship is non-linear. In particular, a large deletion such as in ΔmtDNA results in the loss of essential genes and is therefore expected to be lethal when present at very high frequency (fitness w = 0, or S = 1). However, the organismal selection coefficient is approximately only 0.1892 at a heteroplasmic frequency of 60% ΔmtDNA. Moreover, S is expected to be 0 when ΔmtDNA frequency is zero, since no selection is occurring against ΔmtDNA among hosts that lack ΔmtDNA. The appropriate regression will therefore constitute a non-linear trend-line that rises sharply at progressively higher ΔmtDNA frequency, passing through S=0 when *%ΔmtDNA*=0, S=0.189 when *%ΔmtDNA*=60, and S=1 when *%ΔmtDNA*=100. An ideal candidate regression would therefore be logarithmic, in which rather than being directly proportional to ΔmtDNA frequency, S is equal to some quantity raised to the power of an exponent (which itself is proportional to ΔmtDNA frequency). The least-squares regression method can be used to fit a non-linear regression. As the name suggests, this method seeks to minimize the square of the errors, or residuals, along the dependent variable (in this case S), which can be calculated as shown:

$$SSQ = \sum [S_i - f(\%\Delta mtDNA)]^2$$

Where:
SSQ = sum of the squares of residuals
S_i = organismal selection coefficient for replicate line i in the organismal competition experiment
f(*%ΔmtDNA*) = predicted value of S, defined as a function of *%ΔmtDNA*
S_i – f(*%ΔmtDNA*) = error, or residual, between actual and predicted values of S

The least-squares method, computed here using the GraphPad Prism software package, fits a regression by finding the curve with a shape that minimizes the value of SSQ. When assuming S=0 at 0% ΔmtDNA and S=1 at 100% ΔmtDNA, together with the empirically derived value of S=0.1892 at 60% ΔmtDNA, the least-squares method identifies the following regression to minimize SSQ:

$$S = 10^{0.0181 * \%\Delta mtDNA - 1.81}$$

However, S may approach 1 at a ΔmtDNA frequency lower than 100%. Consistent with this possibility, hosts are almost never observed to carry ΔmtDNA at or above 90% frequency. To account for the uncertainty in how the fitness cost of ΔmtDNA scales with ΔmtDNA frequency, the following regression was obtained using the least-squares method when assuming that S=1 at 90% instead of 100% ΔmtDNA:

$$S = 10^{0.0241 * \%\Delta mtDNA - 2.17}$$

Finally, S might rise modestly until reaching a particular (albeit unknown) ΔmtDNA frequency threshold, followed by an abrupt shift to higher S in a sort of stepwise fashion. To account for this uncertainty, the assumption that S reaches 1 at a given ΔmtDNA frequency was eliminated, enabling a standard linear regression between S=0 at 0% ΔmtDNA and S=0.1892 at 60% ΔmtDNA:

$$S = 0.003153 * \%\Delta mtDNA$$

The confidence interval for the non-linear regressions can be calculated using the following formula:

$$95\%\ confidence\ interval = t_{n-1}\sqrt{CM * dS^2}$$

Where:
t_{n-1} = critical value from student's t-table for *n-1*
CM = covariance matrix value
dS = derivative of the non-linear regression *S* = *f*(*%Δ*mtDNA*) (Figure 3—Source data 2)
n = sample size

The confidence interval for the linear regressions can be calculated in a similar manner as in the case of the linear regressions between *S* and *%Δ*mtDNA* at the sub-organismal level (see the previous sub-section):

$$95\%\ confidence\ interval = t_{n-2} * \sigma_R \sqrt{\frac{1}{n} + \frac{(\%\Delta mtDNA^* - \%\Delta mtDNA_{ave})^2}{\sigma_{\%\Delta mtDNA}{}^2 (n-1)}}$$

Where:
t_{n-2} = critical value from student's t-table for *n-2*
σ_R = standard deviation of the residuals for selection coefficient *S* (Figure 3—Source data 2)
$\%\Delta mtDNA^*$ = each value of ΔmtDNA frequency at which the confidence interval is being calculated
$\%\Delta mtDNA_{ave}$ = average ΔmtDNA frequency across input values (*S*=0.1892 at 60% and *S*=0 at 0% ΔmtDNA)
$\sigma_{\%\Delta mtDNA}$ = standard deviation of ΔmtDNA frequency

Overall (combined) regression (Figure 3C)

According to the general framework of the Price Equation, the overall relationship between ΔmtDNA frequency and the strength of selection—when both the sub-organismal and organismal levels are taken into account—results from summing the regressions as well as their confidence intervals. The overall relationship is therefore defined as the following (assuming $S_{organismal}$=1 at 100% ΔmtDNA):

$$S_{overall} = S_{sub-organismal} + S_{organismal} = (0.01154 * \%\Delta mtDNA - 0.923) + (10^{0.0181 * \%\Delta mtDNA - 1.81})$$

Assuming $S_{organismal}$=1 at 90% ΔmtDNA:

$$S_{overall} = S_{sub-organismal} + S_{organismal} = (0.01154 * \%\Delta mtDNA - 0.923) + (10^{0.0241 * \%\Delta mtDNA - 2.17})$$

Assuming $S_{organismal}$ rises linearly from 0% ΔmtDNA to 60% ΔmtDNA and beyond:

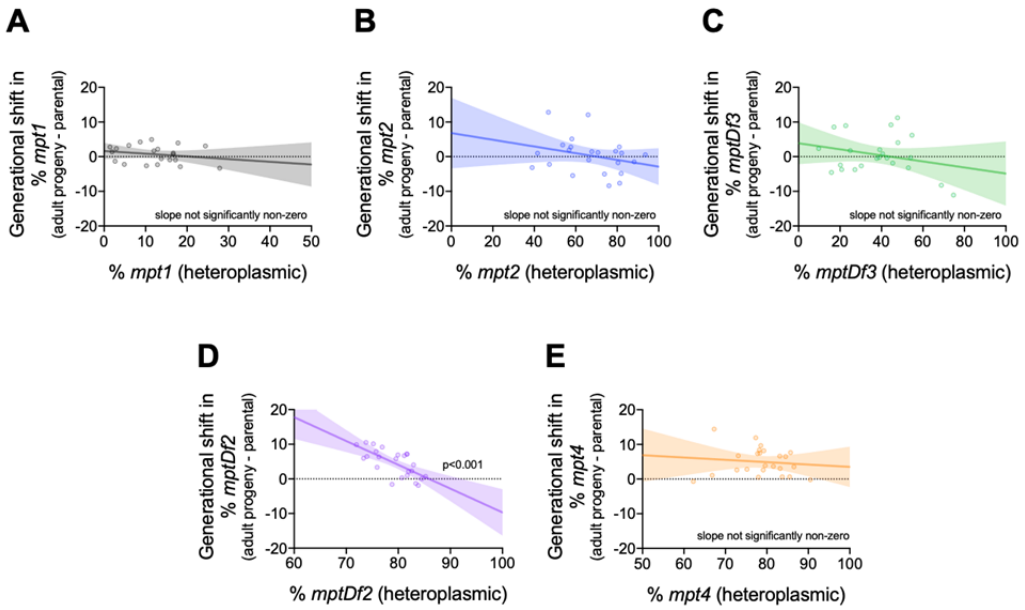
$$S_{overall} = S_{sub-organismal} + S_{organismal} = (0.01154 * \%\Delta mtDNA - 0.923) + (0.003153 * \%\Delta mtDNA)$$

Finally, for each regression, the combined confidence interval can be calculated as follows:

$$95\%\ confidence\ interval_{overall} = \frac{1}{2} * \sqrt{(2 * 95\%CI_{sub-org})^2 + (2 * 95\%CI_{org})^2}$$

Where:
$95\%CI_{sub-org}$ = confidence interval for the regression of *S* and *%Δ*mtDNA* at the sub-organismal level
$95\%CI_{org}$ = confidence interval for the regression of *S* and *%Δ*mtDNA* at the organismal level

1093



1094

1095 **Figure 4—Figure Supplement 1.** Intergenerational shifts in mutant mtDNA frequency at the sub-
1096 organismal level, similar to Figure 2B in (Gitschlag et al., 2020) and Figures 1C and 8B,D,F in (Zhang et
1097 al., 2021). (**A**) Shift in *mpt1* frequency per generation, obtained by subtracting *mpt1* frequency of mature
1098 adult progeny in Figure 4A from *mpt1* frequency of the respective parent, plotted as a function of parental
1099 *mpt1* frequency. (**B**) Shift in *mpt2* frequency per generation, obtained by subtracting *mpt2* frequency of
1100 mature adult progeny in Figure 4C from *mpt2* frequency of the respective parent, plotted as a function of
1101 parental *mpt2* frequency. (**C**) Shift in *mptDf3* frequency per generation, obtained by subtracting *mptDf3*
1102 frequency of mature adult progeny in Figure 4E from *mptDf3* frequency of the respective parent, plotted
1103 as a function of parental *mptDf3* frequency. (**D**) Shift in *mptDf2* frequency per generation, obtained by
1104 subtracting *mptDf2* frequency of mature adult progeny in Figure 4G from *mptDf2* frequency of the
1105 respective parent, plotted as a function of parental *mptDf2* frequency. (**E**) Shift in *mpt4* frequency per
1106 generation, obtained by subtracting *mpt4* frequency of mature adult progeny in Figure 4I from *mpt4*
1107 frequency of the respective parent, plotted as a function of parental *mpt4* frequency.

1108

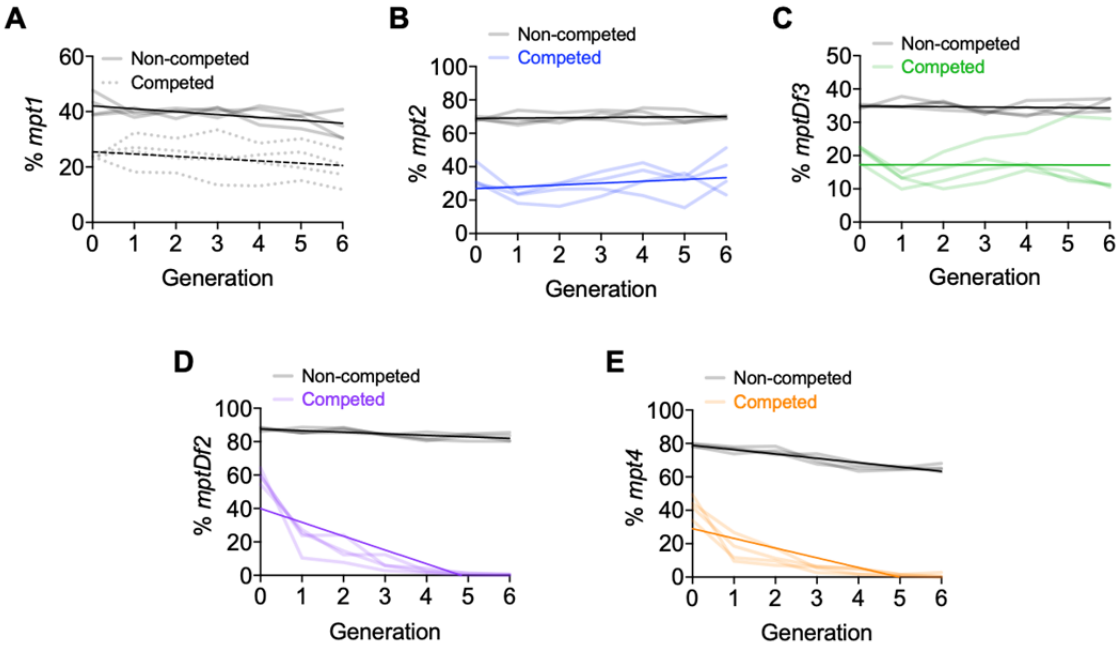


Figure 4—Figure Supplement 2. Raw measurements of decline in mutant frequency upon competition with homoplasmic (wildtype) counterparts, across the collection of heteroplasmies featured in Figure 3. (**A**) Same experiment as shown in Figure 4B but with y-axis expressed in raw (non-normalized) *mpt1* frequency. (**B**) Same experiment as shown in Figure 4D but with y-axis expressed in raw *mpt2* frequency. (**C**) Same experiment as shown in Figure 4F but with y-axis expressed in raw *mptDf3* frequency. (**D**) Same experiment as shown in Figure 4H but with y-axis expressed in raw *mptDf2* frequency. (**E**) Same experiment as shown in Figure 4J but with y-axis expressed in raw *mpt4* frequency. Straight lines represent best-fit regression across all replicates.

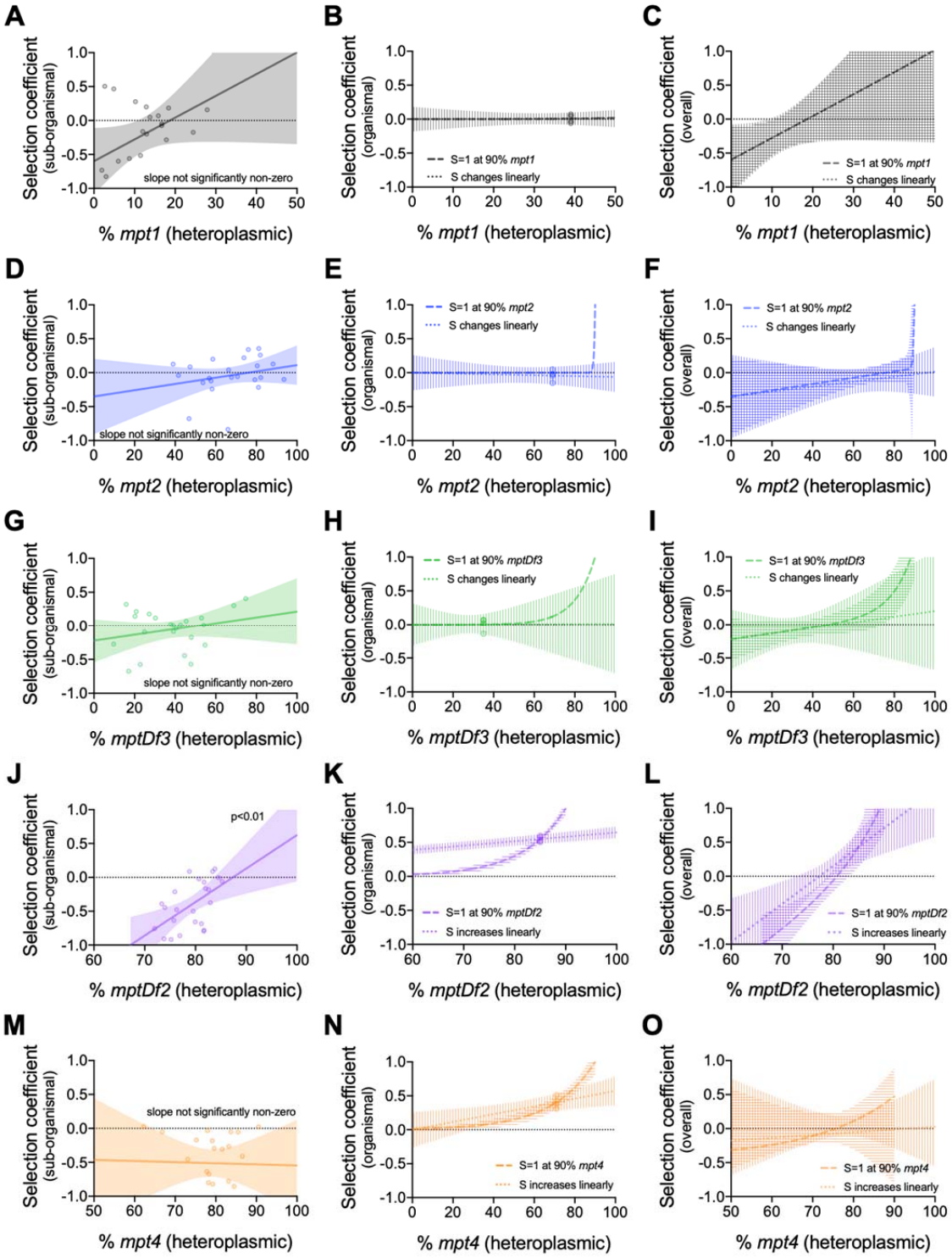


Figure 5—Figure Supplement 1. Integration of the multilevel selection dynamics of all five heteroplasmies using the Price Equation framework with alternate organismal covariance assumptions. The same empirical data and analysis as in Figure 5 but with the following model adjustments: dashed

1123 line and horizontally-striped region represents the best fit and 95% confidence interval, respectively,
1124 assuming the organismal selection coefficient reaches 1 when mutant mtDNA frequency is 90%; dotted
1125 line and vertically-striped region represents the best fit and 95% confidence interval, respectively,
1126 assuming the selection coefficient changes linearly with mutant mtDNA frequency. For each graph shown
1127 above, the horizontal dotted line represents no net selection. Above the dotted horizontal line: selection
1128 favors wildtype mtDNA. Below the dotted horizontal line: selection favors mutant mtDNA. Note that
1129 confidence ranges capable of accommodating a horizontal regression indicate a slope that is not
1130 significantly non-zero, implying no significant relationship between mutant mtDNA frequency and
1131 selection coefficient, at least within the mutant frequency range in which the data was collected.
1132