

1 The mutational decay of male and hermaphrodite competitive fitness in the androdioecious  
2 nematode *C. elegans*, in which males are naturally rare

3

4 Shu-Dan Yeh<sup>1,2</sup>, Ayush Saxena<sup>1</sup>, Timothy Crombie<sup>1</sup>, Dorian Feistel<sup>1,3</sup>, Lindsay M. Johnson<sup>1</sup>,  
5 Isabel Lam<sup>1</sup>, Jennifer Lam<sup>1</sup>, Sayran Saber<sup>1</sup>, and Charles F. Baer<sup>1,4</sup>

6

7 1 - Department of Biology, University of Florida

8 2 - Department of Life Sciences, National Central University, Taiwan

9 3 - Department of Biology, Emory University

10 4 - University of Florida Genetics Institute

11

12 Keywords: *Caenorhabditis elegans*, Mutation accumulation, Mutational bias, Mutational  
13 variance, Spontaneous mutation

14

15 **Abstract**

16 Androdioecious *Caenorhabditis* have a high frequency of self-compatible hermaphrodites and a  
17 low frequency of males. The effects of mutations on male fitness are of interest for two reasons.  
18 First, when males are rare, selection on male-specific mutations is less efficient than in  
19 hermaphrodites. Second, males may present a larger mutational target than hermaphrodites  
20 because of the different ways in which fitness accrues in the two sexes.

21 We report the first estimates of male-specific mutational effects in an androdioecious  
22 organism. The rate of male-specific inviable or sterile mutations is  $\leq 5 \times 10^{-4}$ /generation, below  
23 the rate at which males would be lost solely due to those kinds of mutations. The rate of  
24 mutational decay of male competitive fitness is  $\sim 0.17\%$ /generation; that of hermaphrodite  
25 competitive fitness is  $\sim 0.11\%$ /generation. The point estimate of  $\sim 1.5X$  faster rate of mutational  
26 decay of male fitness is nearly identical to the same ratio in *Drosophila*. Estimates of mutational  
27 variance (VM) for male mating success and competitive fitness are not significantly different  
28 from zero, whereas VM for hermaphrodite competitive fitness is similar to that of non-  
29 competitive fitness. The discrepancy between the two sexes is probably due to the greater  
30 inherent variability of mating relative to internal self-fertilization.

31

32 **Introduction**

33

34 Several species of nematodes in the genus *Caenorhabditis*, among them the well-known *C.*  
35 *elegans*, have evolved an androdioecious mating system in which self-fertilizing hermaphrodites  
36 are very common and males are very rare. In *C. elegans*, for example, the frequency of  
37 outcrossing (= male-female mating, because hermaphrodites cannot mate with each other) is  
38 thought to be on the order of 1% or less, perhaps much less [1]. Dioecy is ancestral in the  
39 genus and most species in the genus are dioecious, although androdioecy has evolved  
40 independently at least three times [2]. Moreover, at least in *C. elegans*, androdioecy appears to  
41 have evolved quite recently [3].

42 Sex determination in *Caenorhabditis* is an XO type chromosomal system, with  
43 females/hermaphrodites having two copies of the X chromosome and males having a single X  
44 chromosome [4]. Laboratory populations of *C. elegans* kept under constant conditions in which  
45 the frequency of males is initially elevated above the background consistently lose males, until  
46 the frequency of males equilibrates at the frequency of non-disjunction of the X [5]. The  
47 frequency of males varies among strains [6] and depends on environmental conditions,  
48 averaging about 0.1% under standard husbandry conditions in the N2 strain. However,  
49 laboratory populations exposed to variable selection can keep males at significantly higher  
50 frequencies [7], consistent with the idea that recombination facilitates adaptive evolution.

51 In an androdioecious population, selection on male function is (i) necessarily weaker  
52 than selection on hermaphrodite function and (ii) weaker than selection on male or female  
53 function in a dioecious population, because in the absence of males (or females) a dioecious  
54 population immediately goes extinct whereas an androdioecious population plods on even in the  
55 complete absence of males. Moreover, although selection typically favors an equal sex-ratio in  
56 a randomly mating dioecious population [8], that is not generally true in a partially selfing  
57 population [9].

58           Importantly, the rarer males become, the less efficient is selection against genes with  
59 male-specific effects on fitness. For example, if males represent 0.1% of the population, as in a  
60 typical *C. elegans* lab population, an autosomal gene with effects only on male fitness will find  
61 itself in a male - and thus under selection - 0.1% of the time and in a hermaphrodite - and thus  
62 free from selection - 99.9% of the time. If the selection coefficient on that gene is 10% in males  
63 and 0 in females, the average selection coefficient at the gene will be 0.01%. The effective  
64 population size,  $N_e$ , of *C. elegans* is thought to be on the order of 10,000 [10], so an allele with  
65 those sex-specific selection coefficients will have an average selection coefficient of  
66 approximately  $1/N_e$ , roughly the boundary of effective neutrality [11]. Thus, the rarer males  
67 become, the stronger selection must be on male function to overcome random genetic drift.

68           These features of selection in androdioecious populations lead to a chicken-and-egg  
69 question with respect to the rarity of males in androdioecious Caenorhabditis: are males rare  
70 because the sex-ratio is near an evolutionary optimum, or are males on their way out, doomed  
71 to ultimately succumb to the ravages of deleterious mutation? Or perhaps both. A quantitative  
72 answer to that question requires an estimate of the distribution of effects of mutations affecting  
73 male fitness, both with respect to mutations that render males non-viable or sterile, and the  
74 effects on the ability of males to mate and for male sperm to compete with hermaphrodite  
75 sperm. Several elements of male fitness have been shown to be genetically variable among  
76 wild isolates of *C. elegans* [12].

77           Unfortunately, the distribution of fitness effects (DFE) on individual traits is very difficult  
78 to quantify reliably [13]. More tractable measures of the vulnerability of a trait to the cumulative  
79 effects of mutation are (i) the rate of change of the trait mean due to the accumulation of  
80 spontaneous mutations (the "mutational bias",  $\Delta M$ ) and (ii) the rate of increase in genetic  
81 variance (the "mutational variance",  $VM$ ). These quantities can be used to quantify the relative  
82 mutability of traits and populations.

83           Here we report on an experiment designed to estimate the cumulative effects of  
84 spontaneous mutations on male-male competitive fitness in a set of *C. elegans* mutation  
85 accumulation (MA) lines which were propagated by single hermaphrodite descent for  
86 approximately 250 generations. In this context, male function constitutes a truly neutral trait,  
87 since chromosomes were (almost) never passed through males. Cumulative effects of  
88 mutations on many hermaphrodite traits have been previously reported for these and other *C.*  
89 *elegans* MA lines (summarized in [14]), providing a robust baseline against which to compare  
90 male mutational properties. In addition, we report new results on the cumulative mutational  
91 effects on hermaphrodite-hermaphrodite competitive fitness.

92           The results will shed light on two questions of interest. First, the frequency of MA lines  
93 for which fertile males can be obtained provides a rough upper bound on the class of mutations  
94 resulting in "zero male fitness"; this class comprises male-specific lethal and male-sterile  
95 mutations. There are surprisingly few published estimates of the rate of mutation to alleles of  
96 zero male fitness. Mukai et al. [15] reported the frequency of matings of *Drosophila*  
97 *melanogaster* MA lines resulting in sterility was "below 2%". Willis [16] reported that a large  
98 fraction of inbreeding depression in *Mimulus guttatus* (~30%) could be attributed to male-sterile  
99 mutations, but actual rates could not be quantified.

100           Second, male components of fitness may be particularly vulnerable to the effects of  
101 deleterious mutations, for two reasons. Male mating success is much closer to a winner-take-all  
102 game than other components of fitness such as female fecundity or egg-to-adult viability (unless  
103 predation is an important cause of mortality). Small differences in performance may be  
104 magnified into a win-lose outcome, with the winner mating and the loser not mating. Also, the  
105 "genetic capture" hypothesis [17] predicts that sexual selection acts at least indirectly on male  
106 condition, in which case the mutational target of male fitness is potentially very large. Given the  
107 infrequency of outcrossing in *C. elegans*, it is arguable whether sexual selection is important to  
108 the evolution of the species. However, given that (*i*) outcrossing is the ancestral state in the

109 genus, (ii) androdioecy seems to have evolved relatively recently in *C. elegans*, and (iii) the  
110 basic biology of mating and fertilization appears similar throughout the genus, it seems  
111 reasonable that the cumulative mutational effects on male fitness in *C. elegans* would at least  
112 approximately reflect those in dioecious species.

113 To date, the only published estimates of cumulative mutational effects on male fitness in  
114 any animal come from *Drosophila melanogaster* [18-20], in which effects on male fitness are  
115 generally somewhat greater than those on female fitness.

116

## 117 **Methods and Materials**

118

119 *Mutation accumulation (MA)* - The details of the construction and maintenance of the MA lines  
120 have been reported elsewhere [21]. Briefly, 100 replicate lines were initiated from a highly  
121 homozygous population of the N2 strain of *C. elegans* and maintained by serial transfer of a  
122 single immature hermaphrodite every generation for approximately 250 generations, at which  
123 point each MA line was cryopreserved. The common ancestor (G0) of each set of MA lines was  
124 cryopreserved at the outset of the experiment.

125

126 *Recovery of males from MA lines* - Beginning in the winter of 2015, cryopreserved 250-  
127 generation (G250) MA lines were thawed and replicate populations collected on standard 60  
128 mm NGM agar plates. For lines in which males were not present on the thawed plate, we  
129 attempted to generate males using a standard heat shock protocol to induce non-disjunction of  
130 the X [22]. If males were obtained but pairings with hermaphrodites failed to produce male  
131 progeny, after three heat shock attempts the MA line was characterized as producing sterile  
132 males. If no males were obtained after three heat shock attempts, the MA line was  
133 characterized as incapable of producing males. Once males were obtained, a single male was  
134 paired with three young L4-stage hermaphrodites on a 35 mm NGM agar plate seeded with

135 OP50 strain *E. coli*, and the progeny split into two 100 mm NGM agar plates seeded with OP50,  
136 grown until food was just exhausted, and cryopreserved. A set of 46 male-segregating G0  
137 "pseudolines" was constructed in the same way and cryopreserved at the same time.

138

139 *Male competitive fitness assay* – Male-male competitive fitness was assayed by pairing a focal  
140 male (G0 ancestor or MA) with a marked competitor male of the ST-2 strain (homozygous for  
141 the dominant *ncls2* pH20::GFP reporter allele on an N2 genetic background) and a male-sterile  
142 hermaphrodite homozygous for a recessive null allele at the *fog-2* locus [*fog-2(q71)*]. *fog-2* is a  
143 recessive mutation that destroys spermatogenesis in hermaphrodites, thereby rendering  
144 hermaphrodites functionally female [23]. To minimize segregating variance in the maternal  
145 stock, we backcrossed the *fog-2(q71)* mutant allele into the ancestral N2 genetic background for  
146 ten generations prior to initiating the competitor population from a cross of a ST-2 male with a  
147 *fog-2* female.

148 The assay was performed in two blocks, in the same conditions as the MA phase of the  
149 experiment (plates seeded with 100  $\mu$ l of an overnight culture of the OP50 strain of *E. coli* as  
150 food, incubated at 20°), with the exception that the assay plates were 40% agarose (NGMA)  
151 rather than 100% agar, to prevent worms from burying in the substrate. Each MA line was  
152 included in each block. Assay blocks were initiated by thawing MA lines and G0 pseudolines,  
153 followed by one generation of male recovery in ten replicate 35 mm plates. Each replicate plate  
154 contained two or three males and three young hermaphrodites. All lines were thawed and  
155 replicated on the same day. Replicate plates were assigned random numbers and were  
156 subsequently handled in order by random number. On the third day after the replicate plates  
157 were initiated, competition assay plates were initiated by transfer of a single young focal male  
158 from each replicate plate and a single similarly-staged competitor male from a stock plate. The  
159 two males were allowed to acclimate to the plate for one day, at which time a female was  
160 introduced to the plate at a location approximately intermediate between the two males. Two

161 days after the introduction of the female, the three adult worms were removed from the plate  
162 and their progeny allowed to grow for another two days. In Block 1, after the two-day growth  
163 period, plates were stored at 4° C for four days prior to counting. In Block 2, worms were  
164 counted directly after the two-day growth period without refrigeration.

165 Worms were counted with the aid of a Union Biometrica BioSorter™ large-particle flow  
166 cytometer (aka, a "worm sorter") equipped with the LP Sampler™ microtiter plate sampler. The  
167 detailed counting protocol is given in Supplementary Appendix A1. Worms were washed from  
168 the competition plates in approximately 1.5 ml of M9 buffer into 1.5 ml microcentrifuge tubes.  
169 Tubes were centrifuged at ~2K x g for 1 minute, the supernatant decanted, and the pelleted  
170 worms resuspended in 100 µl of M9 and pipetted into a well in a 96-well microtiter plate, which  
171 was counted with the BioSorter as outlined in Supplementary Appendix A1.

172

173 *Hermaphrodite competitive fitness assay* - Competitive fitness of hermaphrodites was assayed  
174 in two blocks beginning in May, 2005. At the outset of each block, the cryopreserved G0  
175 ancestor of the MA lines was thawed and 20 replicate populations initiated from a single L3/L4  
176 stage worm placed on a standard 60 mm NGM agar plate seeded with 100 µl of an overnight  
177 culture of the OP50 strain of *E. coli*. These populations are referred to as "pseudolines" and  
178 designated the P0 generation. Seven L3/L4 stage offspring from each pseudoline were  
179 transferred singly to new plates, designated the P1 generation. From this point on, pseudolines  
180 were treated identically to the MA lines. G250 MA lines were thawed and seven revived L3/L4  
181 stage worms from each line were placed individually on standard 60 mm NGM plates, labeled  
182 P1. All P1 plates were assigned a unique random number and all subsequent experimental  
183 manipulations were performed in sequence by random number. All replicate populations were  
184 maintained for two more generations (P2-P3) by transfer of a single L3/L4 stage offspring at  
185 four-day intervals to control for parental and grandparental effects. At the same time, we



186 thawed a replicate of the GFP-marked competitor strain ST-2 and made several large replicate  
187 populations by transferring a chunk from the initial plate to a new 100 mm plate.

188 On the second day after the P3 worm began reproduction, a competition plate for each  
189 replicate was set up by transferring a single L1-stage larva from the P3 plate and a single L1-  
190 stage ST-2 competitor onto a 60 mm NGM agar plate seeded with 100  $\mu$ l of the HB101 strain of  
191 *E. coli* and supplemented with nystatin to retard fungal contamination. Competition plates were  
192 incubated at 20° C for eight days, at which point food was exhausted. Worms were washed  
193 from competition plates in cold M9 buffer, settled on ice and 100  $\mu$ l of the settled worms  
194 transferred into a drop of glycerol on the lid of an empty 60 mm agar dish and the bottom of the  
195 empty dish pressed into the lid. The glycerol immobilizes the worms and pressing them  
196 between halves of the plate puts them into the same focal plane. We took two pictures of each  
197 plate at 40X magnification through a Leica MZ75 dissecting microscope fitted with a 100 W  
198 mercury arc lamp and epifluorescence GFP filter cube (470/40 nm excitation filter, 525/50 nm  
199 emission filter) using a Leica DFC280 camera connected to a computer running the Leica IM50  
200 software (Leica Microsystems Imaging Solutions Ltd). The first picture used the arc lamp and  
201 GFP filter cube (called the "green" image) and the second, taken immediately afterwards, used  
202 transmitted white light (called the "white" image). All worms are visible in the white image,  
203 whereas wild-type (non-GFP) worms appear only very faintly in the green image  
204 (Supplementary Figure S1). The difference between the number of worms in a white image and  
205 in the matching green image is the number of focal worms in the sample.

206 Images were imported into ImageJ software (<http://rsb.info.nih.gov/ij/>) and worms were  
207 counted as follows. If there appeared to be fewer than 200 worms visible in a white image, we  
208 first counted every worm in the white image and then each worm visible in the accompanying  
209 green image. If there appeared to be > 200 worms in the white image we drew a rectangle  
210 around approximately 200 worms and counted them. We then pasted the same rectangle in the  
211 green image and counted the worms visible within the rectangle.

212

213 *Data Analysis -*

214 i) Measures of competitive fitness - Competitive fitness has two components: (1) did the focal  
215 individual reproduce at all? If not, relative fitness is zero regardless of the number of offspring of  
216 the competitor, and (2) given that the focal individual did reproduce, what fraction of the  
217 offspring belong to the focal individual? These considerations apply both to male-male  
218 competitive fitness and to hermaphrodite-hermaphrodite competitive fitness. Given that a focal  
219 individual did reproduce, the ratio  $p/(1-p)$  is related to competitive fitness by the relationship

220 
$$\frac{p_t}{q_t} = \frac{p_0}{q_0} \left( \frac{W_{foc}}{W_C} \right)^t \quad \text{Equation 1}$$

221 [24, equation 17.2], where  $t$  represents the number of generations in the fitness assay (NOT the  
222 number of MA generations),  $p_0$  is the frequency of the focal type (G0 or control) at the beginning  
223 of the assay,  $p_t$  is the frequency of the focal type (G0 control or MA) at the conclusion of the  
224 assay,  $q = 1-p$ ,  $W_{foc}$  is the absolute fitness of the focal type, and  $W_C$  is the absolute fitness of the  
225 competitor. Each trial was started with one focal worm and one competitor, so the ratio  $\frac{p_0}{q_0} = 1$ .  
226 We refer to the ratio  $p/(1-p)$  as the "competitive index",  $CI$  [25].  $CI$  provides a measure of fitness  
227 of the focal type relative to the competitor, raised to the power  $t$ . All analyses of  $CI$  were  
228 performed on natural log-transformed data.

229 ii) Probability of reproduction,  $\pi$  - Probability of reproduction is a binary trait. If a focal worm  
230 reproduced the replicate is scored as a success ("event=1"); if the focal worm did not reproduce  
231 it is scored as a failure ("event=0"). Data were analyzed by Generalized Linear Mixed Model  
232 (GLMM) with estimation by Residual Subject-specific Pseudolikelihood (RSPL) as implemented  
233 in the GLIMMIX procedure of SAS v.9.4 with a logit link function and a random residual.  
234 Treatment (MA vs. Control) is a fixed effect and Line and Replicate (nested within Line) are  
235 random effects. Block is a random effect in principle. However, pseudolikelihoods are not  
236 appropriate criteria for model selection (e.g., by AIC; [26]), so rather than include or exclude

237 variance components including block on the basis of estimates for which there is little power  
238 (because  $n=2$ ), we chose to model block as a fixed effect for this analysis. It is common in the  
239 analysis of MA fitness assays to treat block as a fixed effect when the number of blocks is small  
240 (e.g., [27, 28]).

241 Each line (MA and G0 pseudoline) was assayed for male-male competitive fitness,  $\pi_M$ ,  
242 in each of the two assay blocks. The full model is written as:

$$243 \quad \pi_{ijkl} = \mu + t_k + b_j + c_{jk} + l_{ljk} + \varepsilon_{iljk}$$

244 where  $\pi_{ijkl}$  is a binary variate scored as 1 if the focal worm produced at least two offspring and 0  
245 if it did not,  $\mu$  is the overall mean,  $t_k$  is the fixed effect of treatment  $k$  (G0 or MA),  $b_j$  is the fixed  
246 effect of block  $j$ ,  $c_{jk}$  is the fixed effect of the treatment by block interaction,  $l_{ljk}$  is the random  
247 effect of line (or pseudoline)  $l$ , conditioned on block and treatment, and  $\varepsilon_{iljk}$  is the random  
248 residual, conditioned on block and treatment. Random effects were estimated separately for  
249 each block/treatment combination by means of the GROUP option in the RANDOM statement of  
250 the GLIMMIX procedure [26]. Significance of fixed effects was determined by F-test of Type III  
251 sums of squares, with degrees of freedom determined by the Kenward-Roger method [29].

252 Hermaphrodite probability of reproduction,  $\pi_H$ , was modelled similarly, with the exception  
253 that each line (or pseudoline) was represented in only one of the two assay blocks, so line is  
254 nested within block. The distribution of  $\pi_H$  was strongly left-skewed, so means and standard  
255 errors were calculated by an empirical bootstrap procedure [30, 31]. Resampled datasets were  
256 constructed by resampling lines within blocks, followed by estimation of means and variance  
257 components from the GLMM described above.

258 Competitive Index (CI) – (i) *Males*. Given that both male worms – the focal worm and the  
259 competitor – sired at least 2% of the offspring on the competition plate, we analyzed false-  
260 positive corrected male-male CI ( $CI_M$ ) using a standard general linear model (GLM) as  
261 implemented in the MIXED procedure of SAS v. 9.4. The correction for false positives is  
262 explained in Supplementary Appendix A1. Studentized residuals of natural log-transformed

263 data were scrutinized for outliers by eye against a Q-Q plot. After removal of three outliers ( $n =$   
264 671), the data were initially fit to the linear model:

$$265 \quad y_{ijkl} = \mu + t_k + b_{j|k} + c_{jk} + l_{ij|k} + \varepsilon_{il|jk}$$

266 where  $y_{ijkl}$  is the  $\log(CI)$  of the individual replicate and the independent variables are defined as  
267 in the previous section. Block and Line-by-block interaction were modelled as random effects in  
268 this analysis. Variance components of random effects were estimated by restricted maximum  
269 likelihood (REML). The among-block component of variance was estimated separately for each  
270 treatment and the among-line and among-replicate (nested within line) components were  
271 estimated separately for each treatment/block combination by means of the GROUP option in  
272 the RANDOM or REPEATED statement of the MIXED procedure [32].

273 We first analyzed the full model above, then sequentially simplified the model by first  
274 pooling the random effects across grouping levels (e.g., estimating a single among-line variance  
275 rather than estimating it separately for each block) and then removing the effect entirely. The  
276 model with the smallest corrected AIC (AICc) was chosen as the best model, and significance of  
277 the fixed effect of treatment (MA or G0) in that model was determined by F-test of Type III sums  
278 of squares, with degrees of freedom determined by the Kenward-Roger method [29]. If two  
279 models had equal AICc, the simpler model was chosen as the best model. AICc's of the models  
280 tested are given in Supplementary Table S1. In addition, we calculated empirical bootstrap  
281 estimates of the mean and standard error of  $CI_M$ , resampling over lines within blocks followed by  
282 estimation of means and variance components from the GLM described above [30, 31].

283 Hermaphrodite  $CI$ ,  $CI_H$  was modelled analogously to  $CI_M$ , with the exception that each  
284 line (or pseudoline) was represented in only one of the two assay blocks, therefore line is  
285 nested within block and there is no line-by-block interaction term. Outliers were identified as for  
286 males; five outliers ( $n = 727$ ) were removed prior to further analysis.

287 Mutational Bias – The mutational bias is the per-generation rate of change in the trait mean.

288 The slope of the regression of trait mean on generation of MA is often designated  $R_M$  [33]; the

289 per-generation change scaled as a fraction of the ancestral (G0) trait mean is often referred to  
290 as  $\Delta M_z = \frac{\bar{z}_{MA} - \bar{z}_0}{t\bar{z}_0}$ , where  $\bar{z}_{MA}$  and  $\bar{z}_0$  represent the MA and ancestral trait means and  $t$  is the  
291 number of generations of MA. For  $\pi_M$  and  $\pi_H$ , MA and G0 means were estimated by least  
292 squares, given the general linear mixed model, and  $\Delta M_\pi$  calculated directly from the least-  
293 squares means.  $CI$  is on a logarithmic scale so mean-standardization of the data is not  
294 appropriate because  $CI$  can be negative or positive. For  $CI$ ,  $R_{M,CI} = \frac{\bar{z}_{MA} - \bar{z}_0}{t}$  and the mutational  
295 bias represents the per-generation change in competitive fitness of the focal genotype relative  
296 to the competitor strain.  $R_{M,CI}$  can be related to the per-generation mutational change in relative  
297 fitness *per se*,  $\Delta M_w$ , from Equation 1, as explained below in the Results.

298

299 Mutational Variance (VM) – The per-generation increase in genetic variance resulting from new  
300 mutations,  $VM$ , is equal to the product of the per-genome, per-generation mutation rate ( $U$ ) and  
301 the square of the average effect of a mutation on the trait of interest,  $\alpha$ , i.e.,  $VM = U\alpha^2$  [34]. In  
302 this experiment, MA lines are assumed to be homozygous, in which case  $VM$  is equal to half the  
303 increase in the among-line component of phenotypic variance, divided by the number of  
304 generations of MA, i.e.,  $VM = \Delta V_L = \frac{V_{L,MA} - V_{L,G0}}{2t}$ , where  $V_{L,MA}$  is the variance among MA lines,  
305  $V_{L,G0}$  is the variance among the G0 pseudolines, and  $t$  is the number of generations of MA ([35],  
306 p. 330). For all traits ( $\pi$  and  $C$ ),  $V_L$  is the among-line component of variance of the treatment  
307 group (MA or G0) extracted from the relevant GLM or GLMM.

308  $VM$  is typically scaled either relative to the residual (environmental) variance,  $VE$  or to  
309 the square of the trait mean,  $\bar{z}$ . The ratio  $VM/VE$  is the mutational heritability ( $h_M^2$ ), and the ratio  
310  $\frac{VM}{\bar{z}^2}$  is sometimes called the mutational evolvability ( $I_M$ ) and is equivalent to the squared  
311 mutational coefficient of variation [36]. Usually,  $VM$  is scaled relative to the ancestral (G0)  
312 mean, but if the mean changes substantially over the course of MA (i.e.,  $\Delta M \neq 0$ ), it is more

313 meaningful to scale each group (G0 and MA) by its own mean [21]. Scaling by the group  
314 means is nearly equivalent to calculating  $\Delta V_L$  from log-transformed data [37].  $CI$  is on a  
315 logarithmic scale and cannot be mean-standardized.  $\Delta V_L$  for  $CI_M$  and for  $\pi_M$  and  $\pi_H$  is not  
316 significantly different from 0 (see Results), so scaling is irrelevant.

317

## 318 **Results**

319 Rate of mutations with Zero Male Fitness - Of the 60 of the original 100 MA lines remaining in  
320 2014, we were able to obtain fertile males from 53. We assume that the seven lines for which  
321 we were unable to obtain fertile males carry at least one mutation that leads to Zero Male  
322 Fitness (ZMF, i.e. inviable or sterile), and that these mutations follow a Poisson distribution -  
323 analogous to lethal equivalents [38]. With those assumptions, the expected proportion of lines  
324 with no mutations ( $p_0$ , i.e., the number of lines that produced fertile males) is:  $p_0 = e^{-m}$ , where  $m$   
325 is the expected number of mutations carried by a line [39]. The expected number of mutations  
326  $m = \mu t$ , where  $\mu$  is the per-generation rate of ZMF mutations and  $t$  is the number of generations  
327 of MA. Thus,  $(53/60) = e^{-250\mu}$ , so  $\mu_{ZMF} \approx 5 \times 10^{-4}$ /generation, about half the lower-bound estimate  
328 on the frequency of males. If the ZMF mutation rate is greater than the frequency of males, the  
329 average chromosome will take a male-sterilizing hit before the next time it winds up in a male,  
330 leading to the loss of males by "error catastrophe" [40]. The same calculation from data  
331 reported in [41] gives an estimate of  $\mu_{ZMF} \approx 3 \times 10^{-4}$ /generation.

332 Male-Male Competitive Fitness – (i) *Probability of mating* ( $\pi_M$ ). After 250 generations of  
333 completely relaxed selection, MA males are significantly less likely to successfully mate under  
334 competitive conditions than are their unmutated G0 ancestors ( $F_{1,131.8} = 12.51$ ,  $P < 0.0001$ ).  
335 Averaged over the two blocks, the probability that a G0 male mated successfully (defined as an  
336 estimated frequency of offspring sired  $> 2\%$ ) was  $\sim 90\%$  ( $\bar{\pi}_M = 0.912 \pm 0.019$ ) whereas the  
337 probability that a MA male mated significantly declined to  $\sim 75\%$  ( $\bar{\pi}_M = 0.757 \pm 0.023$ ). Scaled  
338 relative to the G0 mean, the probability of a male successfully mating under the assay

339 conditions decreased by about 0.06% per generation ( $\Delta M_{\pi} = -0.622 \pm 0.159 \times 10^{-3}$ /generation;  
340 Table 1; distributions of line means are shown in Supplementary Figure S2).

341 In neither block did VM of  $\pi_M$  differ significantly from 0. In the first block, the RSPL  
342 estimate of VL among G0 pseudolines was greater than VL among MA lines; in the second  
343 block VL was greater in the MA lines but the difference was not significant (Table 2). (ii)  
344 *Competitive Index (CI<sub>M</sub>)*. When the estimated frequency of offspring sired by the focal male was  
345 at least 2%, MA males sired a smaller fraction of offspring than did their unmutated G0  
346 ancestors [ $\log(CI_{M,G0}) = -0.145 \pm 0.074$ ;  $\log(CI_{M,MA}) = -0.461 \pm 0.095$ ; standard errors  
347 represented by the standard deviation of the empirical bootstrap distribution]. The best-fit linear  
348 model includes a separate among-block component of variance for each treatment  
349 (Supplementary Table S1), and under that model the change in the trait mean is not significantly  
350 different from zero ( $F_{1,1.38} = 1.67$ ,  $P > 0.37$ ). However, when the among-block variance is pooled  
351 over the two treatments, the change in the mean  $CI_M$  becomes significant ( $F_{1,603} = 5.57$ ,  $P <$   
352  $0.02$ ). The lack of significance in the best-fit model potentially represents Type II error resulting  
353 from having to estimate a variance component with  $n=2$ . To test that possibility, we estimated  
354 the change in the trait mean from the mean of 1000 bootstrap replicates, resulting in an  
355 empirical  $P < 0.007$ . Averaged over the two blocks,  $\log(CI_M)$  declined by slightly more than  
356 0.1% per generation ( $R_{M,CI,M} = -1.26 \pm 0.48 \times 10^{-3}$  /generation; Table 1; distributions of line  
357 means are shown in Supplementary Figure S3).

358 As noted,  $CI$  cannot be directly mean-standardized. However,  $CI$  is related to relative  
359 fitness by equation [1] above.  $\Delta M_W$  can be calculated from  $\frac{p_t}{q_t} = \frac{p_0}{q_0} \left( \frac{W_{foc}}{W_C} \right)^t$ , where  $CI = \left( \frac{p_t}{q_t} \right)$ ,  $p_0 =$   
360  $q_0 = 0.5$ , and  $t = 1$ , thus  $CI = \left( \frac{W_{foc}}{W_C} \right)$ . The ratio of the fitnesses relative to the competitor  
361 (designated by a capital  $W$ ),  $\left( \frac{W_{MA}}{W_C} \right) \left( \frac{W_{G0}}{W_C} \right)^{-1}$  gives the fitness of the MA lines relative to that of the  
362 G0 ancestor (designated by a lower-case  $w$ ),  $\left( \frac{w_{MA}}{w_{G0}} \right)$ . The ratio  $\left( \frac{W_{G0}}{W_C} \right) = \exp(-0.145) = 0.865$  and

363  $\left(\frac{W_{MA}}{W_C}\right) = \exp(-0.461) = 0.631$ , so  $\left(\frac{W_{MA}}{W_{G0}}\right) \approx 0.73$ . Thus, male competitive fitness relative to the  
364 ST-2 competitor declined by about 27% over the course of 250 generations of MA, or by about -  
365  $1.04 \times 10^{-3}$ /generation when scaled as a fraction of the G0 mean.

366 The REML estimate of VL for  $\log(CI_M)$  for both G0 pseudolines and MA lines is zero in  
367 each block (Table 2), and the distributions of line means are similar in the two groups  
368 (Supplementary Figure S3). Taken at face value, a change in the mean coupled with no change  
369 in the among-line variance implies that each line changed at the same rate, or at least at rates  
370 that were indistinguishable. A more plausible explanation is that the true genetic variance is  
371 small relative to the environmental variance (which includes experimental error) and the sample  
372 sizes employed here were not large enough to provide power to detect small differences. In the  
373 male fitness assay, each line was initially replicated tenfold, five replicates per block. In the  
374 hermaphrodite competitive fitness assay, in which  $\Delta VL$  for  $\log(CI_H)$  is highly significant  
375 ( $P < 0.0001$ ; see next section), each line was replicated sevenfold, but in only one of the two  
376 blocks.

377 Hermaphrodite-Hermaphrodite competitive fitness - (i) *Probability of reproducing* ( $\pi_H$ ). The  
378 probability of a hermaphrodite reproducing was high (>98% for both G0 and MA), and changed  
379 little over 250 generations of MA ( $\Delta M = -1.9 \times 10^{-5}$ /generation; Table 1; distributions of line  
380 means are shown in Supplementary Figure S2). The mutational heritability is large ( $h_M^2 \approx 0.02$ ;  
381 Table 2), but the distribution of  $\pi_H$  among lines is highly left-skewed (median  $\pi_H = 1$ ;  
382 Supplementary Figure S2) so the point estimate of  $h_M^2$  is probably not very meaningful.

383 It is likely that some GFP-marked competitors were misidentified as focal types (non-  
384 GFP) in our image analysis, which could potentially inflate the apparent probability of  
385 reproduction. However, these values of  $\pi_H$  are nearly identical to the probability of reproduction  
386 of hermaphrodites in a different experiment in which hermaphrodites of the same set of MA lines  
387 were allowed to reproduce in non-competitive conditions ( $\pi_H > 97\%$  for both G0 and MA;  $\Delta M = -$



388  $3.3 \times 10^{-5}$ /generation; reanalysis of data in [30]). Thus, the very high rate of reproduction does  
389 not appear to be an experimental artifact.

390 (ii) *Competitive Index ( $CI_H$ )* – Mean  $CI_H$  declined significantly over the course of 250 generations  
391 of MA [ $\log(CI_{H,G0}) = 0.862 \pm 0.331$ ;  $\log(CI_{H,MA}) = 0.228 \pm 0.338$ ;  $F_{1,169} = 26.97$ ,  $P < 0.0001$ ; Table  
392 1; distributions of line means are shown in Supplementary Figure S3].  $R_{M,CI,H}$  calculated from  
393 the slope of the regression of  $\log(CI_H)$  on generation of MA is  $-2.54 \pm 0.49 \times 10^{-3}$  per-generation.

394 From equation [1],  $\frac{p_t}{q_t} = \frac{p_0}{q_0} \left( \frac{W_{foc}}{W_C} \right)^t$ , where  $CI = \left( \frac{p_t}{q_t} \right)$ ,  $p_0 = q_0 = 0.5$  and here  $t$  is equal to the  
395 number of generations of reproduction the population underwent over the course of the eight-  
396 day assay. Therefore,  $\log(CI) = t \times \log\left(\frac{W_{foc}}{W_C}\right)$ , so  $[\exp(\log(CI))]^{1/t} = \left(\frac{W_{foc}}{W_C}\right)$ , and the ratio of the  
397 fitnesses relative to the competitor,  $\left(\frac{W_{MA}}{W_C}\right) \left(\frac{W_{G0}}{W_C}\right)^{-1}$  gives the fitness of the MA lines relative to  
398 that of the G0 ancestor,  $\left(\frac{W_{MA}}{W_{G0}}\right)$ .

399 We cannot be certain about the exact number of generations of reproduction, except that  
400 three is the upper bound (based on the worm's life cycle), and the true number is probably close  
401 to two. The basis for that judgment is this: if the average worm produces 200 offspring in its  
402 lifetime, after one generation there will be  $2 \times 200 = 400$  worms on the plate and after two  
403 generations there will be  $400 \times 200 = 80,000$  worms on the plate (density-dependence  
404 notwithstanding); after three generations there will be  $80,000 \times 200 = 16$  million. There were many  
405 more than 400 and certainly fewer than 80,000, so we assume two generations is probably  
406 close to the true number of generations.

407 Assuming that  $t = 2$ , we find  $\left(\frac{W_{MA}}{W_{G0}}\right) = 0.728$ , or in other words, relative competitive fitness  
408 declined by about 27% over 250 generations of MA. Scaled relative to the G0 ancestor,  $\Delta M_w \approx$   
409  $1.09 \times 10^{-3}$ /generation. If  $t$  is closer to 3,  $\left(\frac{W_{MA}}{W_{G0}}\right) \approx 0.81$  and  $\Delta M_w \approx 0.76 \times 10^{-3}$ /generation.

410 Averaged over the two blocks, the REML estimate of the among-line variance in  $\log(CI_H)$   
411 increased from zero in the G0 ancestor to 0.534 in the MA lines, giving  $VM = 1.09 \times 10^{-3}$   
412  $^3/\text{generation}$  (Table 2; Likelihood Ratio Chi-square = 23.7,  $df=2$ ,  $P<0.0001$ ). Scaled as a  
413 fraction of VE, the mutational heritability  $h_m^2 = 0.83 \times 10^{-3}/\text{generation}$ . By way of comparison, the  
414 point-estimate of  $h_m^2$  for non-competitive fitness in these lines under the same conditions is ~  
415  $1.29 \times 10^{-3}/\text{generation}$  [30].

416

## 417 Discussion

418 A simple but important finding is the close quantitative agreement between our estimate of the  
419 Zero Male Fitness mutation rate and that gleaned from the results of a previous MA experiment  
420 on the N2 strain [41]. Those estimates are subject to several sources of uncertainty, both  
421 experimental (e.g., perhaps if we had tried harder, we could have gotten fertile males from the  
422 lines that did not produce them) and biological (the distribution of mutational effects). The  
423 experimental uncertainty in this case leads to an overestimate of the ZMF rate. By way of  
424 comparison, the lethal recessive mutation rate in N2 has been estimated, from very limited data,  
425 to be on the order of  $\sim 0.01/\text{generation}$  [42, 43]. If we assume that each genetic death (lethal  
426 mutation) or male dysfunction (ZMF) is due to one and only one mutation, and that the estimate  
427 of 80 mutations per MA genome is not far off, then the fraction of ZMF mutations is  $7/(80 \times 60)$ ,  
428 about 0.15%. In every 100 genomes there will be  $\sim 32$  new mutations, of which about one will  
429 be lethal, so the fraction of mutations that are lethal is  $1/32$ , or about 3%. Thus, we infer that  
430 the fraction of ZMF mutations is around 5% of the lethal fraction.

431 What of the idea that male fitness is more susceptible to the cumulative effects of  
432 mutation than hermaphrodite fitness? Generally speaking, fitness is a function of survival,  
433 fecundity, and timing of reproduction. In the hermaphrodite assay, there was no discernible  
434 effect of mutation accumulation on the probability of reproducing (a finding which recapitulates  
435 the result from a previous non-competitive assay), so the decline in relative fitness with MA is

436 encompassed by the  $\sim 0.11\%$  per generation decline in relative competitive fitness. In males,  
437 both the probability of siring offspring and the fraction of offspring sired given that the male did  
438 mate successfully declined. A calculation based on the point estimates of the  $\Delta M$ s (Table 1)  
439 shows that, after one generation of MA, hermaphrodite relative fitness will have declined by  $1 -$   
440  $[(1 - 1.95 \times 10^{-5})(1 - 1.09 \times 10^{-3})] \approx 0.11\%$ . By the same reasoning, male relative fitness will have  
441 declined by  $1 - [(1 - 6.82 \times 10^{-4})(1 - 1.04 \times 10^{-3})] \approx 0.17\%$ . Thus, male fitness is certainly no less  
442 sensitive, and perhaps slightly more sensitive to the cumulative deleterious effects of mutation  
443 than is hermaphrodite fitness. This result is quantitatively nearly identical to the finding that  
444 male *Drosophila melanogaster* decline in fitness  $\sim 1.5X$  faster from mutation accumulation than  
445 do females [20].

446 The cumulative effects of selection depend on both the effects of an allele on fitness (the  
447 selection coefficient,  $s$ ) and the effective population size,  $N_e$ . Based on whole-genome  
448 sequencing of a subset of the MA lines in this experiment [44], the per-genome mutation rate is  
449 not less than 0.2 (one mutation every five generations) and unlikely to be more than about one  
450 mutation per generation, so the average MA line carries somewhere between 50 and 250  
451 mutations; we think the true average is likely to be about 80 (AS and CFB, unpublished results).  
452 The cumulative decline in hermaphrodite relative fitness is about 27%. Thus, we can bracket the  
453 (arithmetic) mean homozygous effect on relative competitive fitness of new mutations,  $\bar{s}$ , as  
454 lying somewhere between  $t\Delta M/50$  and  $t\Delta M/250$ , where  $t$  is the number of generations of MA.  
455 For hermaphrodite competitive fitness,  $t\Delta M_w \approx 0.27$ , so the average selective effect  $\bar{s}$  is  
456 bracketed between about  $0.27/250 \approx 0.1\%$  and  $0.27/50 \approx 0.6\%$ ; if our estimate of 80 new  
457 mutations is correct,  $\bar{s} \approx 0.33\%$ . For male relative fitness,  $t\Delta M \approx 0.35$ , so  $\bar{s}$  is bracketed  
458 between about 0.14% and 0.7%, with a best-estimate value  $\bar{s} \approx 0.44\%$ .

459  $N_e$  of *C. elegans* has been estimated from the standing nucleotide diversity as being on  
460 the order of  $10^4$  [10]. If a mutant allele is under selection in males but is neutral in  
461 hermaphrodites and males represent 1% of the population, the average selection coefficient on

462 a mutant autosomal allele would be  $(0.99)(0) + (0.01)(0.0044) = 4.4 \times 10^{-4}$ , on the cusp of  
463 effective neutrality. Deleterious alleles with selection coefficients  $s \approx 1/N_e$  are the most  
464 pernicious with respect to population mean fitness [45]. On the face of it, it would appear that  
465 males in *C. elegans* are in peril of mutating their way out of existence. However, that conclusion  
466 is based on the strong assumption that mutations that affect male fitness have no pleiotropic  
467 effects on hermaphrodite fitness.

468 This study has two limitations. First, we would like to have an estimate of the mutational  
469 correlation between male fitness and hermaphrodite fitness, because those data would  
470 illuminate the extent to which intersex pleiotropy ("intralocus conflict" if effects are antagonistic  
471 between the sexes, [46]) is an inherent feature of genomic architecture, without the confounding  
472 influence of natural selection. However, the lack of significant mutational variance in either of  
473 the male fitness traits ( $\pi_M$  and  $Cl_M$ ) obviously means the estimate of any covariance with those  
474 traits is zero. The male-fitness assay included fewer MA lines (53) than did the full  
475 hermaphrodite assay (80), although each block of the hermaphrodite competitive fitness assay  
476 included fewer lines (~40 vs ~50) and the estimates of VM were highly significant in each block  
477 (LRT,  $P < 0.001$ ). We have assayed many hermaphrodite traits with ~50 250-generation MA  
478 lines and detected significant VM (e.g., [14, 47]). Mating is inherently subject to many more  
479 sources of variation than is internal self-fertilization, and the results reflect that greater  
480 variability.

481 The second limitation is that males compete for fitness not only with other males, but  
482 also with the hermaphrodite itself. Measuring male-hermaphrodite competitive fitness in our  
483 context requires a recessive marker in the hermaphrodite competitor, so that the offspring of a  
484 cross can be distinguished from the hermaphrodite's self-progeny. Unfortunately, we were  
485 unable to find a recessive marker that had reasonably high fitness and could also be reliably  
486 scored at sufficiently high throughput to enable a full assay, either with the worm sorter, by  
487 image analysis, or by eye.

488           Male-male competitive index includes both a behavioral component and a sperm-  
489 competition component, which cannot be discriminated with our assay. Both features could  
490 potentially affect male-hermaphrodite competitive fitness. A hermaphrodite paired with a male  
491 that is a poor mater may sire a larger fraction of offspring prior to exhaustion of its sperm than  
492 will a hermaphrodite paired with a good mater. Similarly, a hermaphrodite mated to a male with  
493 poor sperm will presumably sire a larger fraction of offspring than a hermaphrodite mated to a  
494 male with good sperm. Male sperm generally outcompete hermaphrodite sperm [48], so if it is  
495 assumed that the entire difference in male-male competitive fitness is due to reduction in sperm-  
496 competitive ability and that wild-type hermaphrodite sperm would be no better competitors than  
497 wild-type male sperm (which seems reasonable), then the strength of selection against male-  
498 male competitive fitness provides an upper bound on the strength of selection acting on the  
499 competitive ability of male sperm relative to hermaphrodite sperm. Alas, no such simple  
500 approximation is possible with respect to male mating behavior, because the fitness  
501 consequences of even the simplest aspect of male behavior, time to mating, depend on the  
502 distribution of timing of hermaphrodite self-fertilization.

503           To conclude, the results of this study indicate that selection acting on mutations affecting  
504 male function is similar to, or perhaps slightly stronger than, selection on mutations affecting  
505 hermaphrodite function. However, a full accounting of mutations affecting the full spectrum of  
506 components of male fitness remains incomplete.

507

## 508 **Acknowledgments**

509

510 We thank Dustin Blanton, Whitney Bour, Myrnelle Damas, Thomas Keller, Laura Levy, Gigi  
511 Ostrow and Naomi Phillips for assistance in the lab. *C. elegans* strains were graciously  
512 provided by Keith Choe (UF) and the Caenorhabditis Genetics Center. Support was provided

513 by NIH grants R01GM107227 to CFB, E. C. Andersen and J. M. Ponciano, and

514 S1010OD012006 to CFB, L. Bianchi, K. Choe, and A. S. Edison.

515

516 **References**

- 517 [1] Rockman, M.V. & Kruglyak, L. 2009 Recombinational Landscape and Population Genomics of  
518 *Caenorhabditis elegans*. *PLoS Genetics* **5**. (doi:e1000419  
519 10.1371/journal.pgen.1000419).
- 520 [2] Kiontke, K. & Fitch, D.H.A. 2005 The phylogenetic relationships of *Caenorhabditis* and other  
521 rhabditids. *Wormbook*, 1-11.
- 522 [3] Cutter, A.D., Wasmuth, J.D. & Washington, N.L. 2008 Patterns of molecular evolution in  
523 *Caenorhabditis* preclude ancient origins of selfing. *Genetics* **178**, 2093-2104.  
524 (doi:10.1534/genetics.107.085787).
- 525 [4] Herman, R.K. 2005 Introduction to sex determination. In *Wormbook* (ed. T.C.e.R. Community).
- 526 [5] Stewart, A.D. & Phillips, P.C. 2002 Selection and maintenance of androdioecy in *Caenorhabditis*  
527 *elegans*. *Genetics* **160**, 975-982.
- 528 [6] Hodgkin, J. & Doniach, T. 1997 Natural variation and copulatory plug formation in *Caenorhabditis*  
529 *elegans*. *Genetics* **146**, 149-164.
- 530 [7] Anderson, J.L., Morran, L.T. & Phillips, P.C. 2010 Outcrossing and the maintenance of males within *C.*  
531 *elegans* populations. *Journal of Heredity* **101**, S62-S74. (doi:10.1093/jhered/esq003).
- 532 [8] Fisher, R.A. 1930 *The Genetical Theory of Natural Selection*. Oxford, Clarendon Press.
- 533 [9] Charnov, E.L. 1982 *The Theory of Sex Allocation*. Princeton, NJ, USA, Princeton University Press.
- 534 [10] Andersen, E.C., Gerke, J.P., Shapiro, J.A., Crissman, J.R., Ghosh, R., Bloom, J.S., Felix, M.-A. &  
535 Kruglyak, L. 2012 Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity.  
536 *Nature Genetics* **44**, 285-U283. (doi:10.1038/ng.1050).
- 537 [11] Ohta, T. 1973 Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96-98.
- 538 [12] Murray, R.L., Kozłowska, J.L. & Cutter, A.D. 2011 Heritable determinants of male fertilization success  
539 in the nematode *Caenorhabditis elegans*. *Bmc Evolutionary Biology* **11**. (doi:10.1186/1471-2148-11-99).

- 540 [13] Charlesworth, B. 2015 Causes of natural variation in fitness: Evidence from studies of *Drosophila*  
541 populations. *Proceedings of the National Academy of Sciences of the United States of America* **112**,  
542 1662-1669. (doi:10.1073/pnas.1423275112).
- 543 [14] Davies, S.K., Leroi, A.M., Burt, A., Bundy, J. & Baer, C.F. 2016 The mutational structure of  
544 metabolism in *Caenorhabditis elegans*. *Evolution*. (doi:10.1111/evo.13020).
- 545 [15] Mukai, T., Chigusa, S.I., Crow, J.F. & Mettler, L.E. 1972 Mutation rate and dominance of genes  
546 affecting viability in *Drosophila melanogaster*. *Genetics* **72**, 335-355.
- 547 [16] Willis, J.H. 1999 The contribution of male-sterility mutations to inbreeding depression in *Mimulus*  
548 *guttatus*. *Heredity* **83**, 337-346. (doi:10.1038/sj.hdy.6885790).
- 549 [17] Rowe, L. & Houle, D. 1996 The lek paradox and the capture of genetic variance by condition  
550 dependent traits. *Proceedings of the Royal Society B-Biological Sciences* **263**, 1415-1421.  
551 (doi:10.1098/rspb.1996.0207).
- 552 [18] Mallet, M.A., Kimber, C.M. & Chippindale, A.K. 2012 Susceptibility of the male fitness phenotype to  
553 spontaneous mutation. *Biology Letters* **8**, 426-429. (doi:10.1098/rsbl.2011.0977).
- 554 [19] Mallet, M.A., Bouchard, J.M., Kimber, C.M. & Chippindale, A.K. 2011 Experimental mutation-  
555 accumulation on the X chromosome of *Drosophila melanogaster* reveals stronger selection on males  
556 than females. *BMC Evolutionary Biology* **11**. (doi:10.1186/1471-2148-11-156).
- 557 [20] Sharp, N.P. & Agrawal, A.F. 2013 Male-biased fitness effects of spontaneous mutations in  
558 *Drosophila melanogaster*. *Evolution* **67**, 1189-1195. (doi:10.1111/j.1558-5646.2012.01834.x).
- 559 [21] Baer, C.F., Shaw, F., Steding, C., Baumgartner, M., Hawkins, A., Houppert, A., Mason, N., Reed, M.,  
560 Simonelic, K., Woodard, W., et al. 2005 Comparative evolutionary genetics of spontaneous mutations  
561 affecting fitness in rhabditid nematodes. *Proceedings of the National Academy of Sciences of the United*  
562 *States of America* **102**, 5785-5790.
- 563 [22] He, F. 2011 Making males of *C. elegans*. *Bio-protocol Bio101* **e58**. (doi:DOI: 10.21769/BioProtoc.58).



- 564 [23] Schedl, T. & Kimble, J. 1988 Fog-2, a germ-line-specific sex determination gene required for  
565 hermaphrodite spermatogenesis in *Caenorhabditis elegans*. *Genetics* **119**, 43-61.
- 566 [24] Barton, N.H., Briggs, D.E.G., Eisen, J.A., Goldstein, D.B. & Patel, N.H. 2007 *Evolution*. Cold Spring  
567 Harbor, NY, USA, Cold Spring Harbor Laboratories Press.
- 568 [25] Shabalina, S.A., Yampolsky, L.Y. & Kondrashov, A.S. 1997 Rapid decline of fitness in panmictic  
569 populations of *Drosophila melanogaster* maintained under relaxed natural selection. *Proceedings of the*  
570 *National Academy of Sciences of the United States of America* **94**, 13034-13039.
- 571 [26] 2013 The GLIMMIX Procedure. In *SAS/STAT 13.1 User's Guide* (Cary, NC, SAS Institute, Inc.
- 572 [27] Houle, D., Hughes, K.A., Hoffmaster, D.K., Ihara, J., Assimacopoulos, S., Canada, D. & Charlesworth,  
573 B. 1994 The effects of spontaneous mutation on quantitative traits. 1. Variances and covariances of life-  
574 history traits. *Genetics* **138**, 773-785.
- 575 [28] Shaw, R.G., Byers, D.L. & Darmo, E. 2000 Spontaneous mutational effects on reproductive traits of  
576 *Arabidopsis thaliana*. *Genetics* **155**, 369-378.
- 577 [29] Kenward, M.G. & Roger, J.H. 1997 Small sample inference for fixed effects from restricted maximum  
578 likelihood. *Biometrics* **53**, 983-997. (doi:10.2307/2533558).
- 579 [30] Baer, C.F., Phillips, N., Ostrow, D., Avalos, A., Blanton, D., Boggs, A., Keller, T., Levy, L. & Mezerhane,  
580 E. 2006 Cumulative effects of spontaneous mutations for fitness in *Caenorhabditis*: Role of genotype,  
581 environment and stress. *Genetics* **174**, 1387-1395.
- 582 [31] Efron, B. & Tibshirani, R.J. 1993 *An Introduction to the Bootstrap*. New York, Chapman and Hall.
- 583 [32] Fry, J.D. 2004 Estimation of genetic variances and covariances by restricted maximum likelihood  
584 using PROC MIXED. In *Genetic Analysis of Complex Traits Using SAS* (ed. A.M. Saxton), pp. 11-34. Cary,  
585 NC, SAS Institute, Inc.
- 586 [33] Vassilieva, L.L. & Lynch, M. 1999 The rate of spontaneous mutation for life-history traits in  
587 *Caenorhabditis elegans*. *Genetics* **151**, 119-129.

- 588 [34] Barton, N.H. 1990 Pleiotropic models of quantitative variation. *Genetics* **124**, 773-782.
- 589 [35] Lynch, M. & Walsh, B. 1998 *Genetics and Analysis of Quantitative Traits*. Sunderland, MA., Sinauer;  
590 xvi, 980 p.
- 591 [36] Houle, D., Morikawa, B. & Lynch, M. 1996 Comparing mutational variabilities. *Genetics* **143**, 1467-  
592 1483.
- 593 [37] Fry, J.D. & Heinsohn, S.L. 2002 Environment dependence of mutational parameters for viability in  
594 *Drosophila melanogaster*. *Genetics* **161**, 1155-1167.
- 595 [38] Morton, N.E., Crow, J.F. & Muller, H.J. 1956 An estimate of the mutational damage in man from  
596 data on consanguinous marriages. *Proceedings of the National Academy of Sciences of the United States*  
597 *of America* **42**, 855-863. (doi:10.1073/pnas.42.11.855).
- 598 [39] Luria, S.E. & Delbrück, M. 1943 Mutations of bacteria from virus sensitivity to virus resistance.  
599 *Genetics* **28**, 491-511.
- 600 [40] Eigen, M. 1971 Self-organization of matter and evolution of biological macromolecules.  
601 *Naturwissenschaften* **58**, 465-&.
- 602 [41] Vassilieva, L.L., Hook, A.M. & Lynch, M. 2000 The fitness effects of spontaneous mutations in  
603 *Caenorhabditis elegans*. *Evolution* **54**, 1234-1246.
- 604 [42] Clark, D.V., Rogalski, T.M., Donati, L.M. & Baillie, D.L. 1988 The Unc-22(IV) region of *Caenorhabditis*  
605 *elegans*: Genetic analysis of lethal mutations. *Genetics* **119**, 345-353.
- 606 [43] Rosenbluth, R.E., Cuddeford, C. & Baillie, D.L. 1983 Mutagenesis in *Caenorhabditis elegans*. 1. A  
607 rapid eukaryotic mutagen test system using the reciprocal translocation *et1*(III-V). *Mutation Research*  
608 **110**, 39-48.
- 609 [44] Denver, D.R., Wilhelm, L.J., Howe, D.K., Gafner, K., Dolan, P.C. & Baer, C.F. 2012 Variation in base-  
610 substitution mutation in experimental and natural lineages of *Caenorhabditis* nematodes. *Genome*  
611 *Biology and Evolution* **4**, 513-522. (doi:10.1093/gbe/evs028).

- 612 [45] Schultz, S.T. & Lynch, M. 1997 Mutation and extinction: The role of variable mutational effects,  
613 synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution* **51**, 1363-1371.  
614 (doi:10.2307/2411188).
- 615 [46] Rice, W.R. & Chippindale, A.K. 2001 Intersexual ontogenetic conflict. *Journal of Evolutionary Biology*  
616 **14**, 685-693. (doi:10.1046/j.1420-9101.2001.00319.x).
- 617 [47] Farhadifar, R., Ponciano, J.M., Andersen, E.C., Needleman, D.J. & Baer, C.F. 2016 Mutation is a  
618 sufficient and robust predictor of genetic variation for mitotic spindle traits in *Caenorhabditis elegans*.  
619 *Genetics*. (doi:10.1534/genetics.115.185736).
- 620 [48] Lamunyon, C.W. & Ward, S. 1995 Sperm precedence in a hermaphroditic nematode (*Caenorhabditis*  
621 *elegans*) is due to competitive ability of male sperm. *Experientia* **51**, 817-823. (doi:10.1007/bf01922436).  
622  
623  
624

625

	$\pi_M$	$\pi_H$	$\log(CI_M)$	$\log(CI_H)$
G0	0.912 (0.016)	0.987 (0.008)	-0.145 (0.074)	0.862 (0.331)
MA	0.757 (0.023)	0.982 (0.005)	-0.461 (0.095)	0.228 (0.338)
$R_M$ (x $10^3$ /gen)	<b>-0.622</b> (0.159)	-0.0192 (0.050)	<b>-1.26</b> (0.48)	<b>-2.54</b> (0.49)
$\Delta M$ (x $10^3$ /gen)	-0.682	-0.0195	-1.04 <sup>†</sup>	-1.09 <sup>*,†</sup>

626

627 **Table 1.** Evolution of trait means, averaged across blocks; standard errors in parentheses.

628 Sample sizes are given in Table 2. Values of  $R_M$  significantly different from zero ( $P < 0.05$ )

629 shown in bold type. See Methods for abbreviations and details of calculations.  $\Delta M$  of CI is

630 calculated for back-transformed data.

631 \* - Assumes  $t = 2$  generations of reproduction

632 † -  $\Delta M$  calculated for  $CI$ , not  $\log(CI)$

<b>Variance Component</b>	<b>Block</b>	$\pi_M$ $n_{MA,Blk1} = 46/51$ $n_{MA,Blk2} = 53/53$	$\pi_H$ $n_{MA,Blk1} = 41$ $n_{MA,Blk2} = 39$	$\log(CI_M)$ $n_{MA,Blk1} = 46/53$ $n_{MA,Blk2} = 53/53$	$\log(CI_H)$ $n_{MA,Blk1} = 41$ $n_{MA,Blk2} = 39$
VL (G0)	1	2.662 (0.953)	0		0
	2	0.705 (0.319)	0		0
	Ave	0.574 (0.288)	0	0	0
VL (MA)	1	0.256 (0.406)	3.547 (1.213)		0.743 (0.212)
	2	0.412 (0.304)	2.546 (0.952)		0.343 (0.147)
	Ave	0.266 (0.187)	3.052 <sup>1</sup> (0, 4.299)	0	0.571 (0.128)
VM (x 10 <sup>3</sup> )	Ave	0	6.104 <sup>1</sup> (0, 8.598)	0	<b>1.142</b> (0.256)
VE (G0)	1	0.328 (0.037)	1.008 (0.130)		
	2	0.823 (0.093)	1.007 (0.120)		
	Ave	0.644 (0.052)	1.008 (0.088)	2.211 <sup>2</sup> (0.173)	0.822 <sup>2</sup> (0.068)
VE (MA)	1	0.900 (0.105)	0.261 (0.025)		
	2	0.903 (0.099)	0.341 (0.030)		
	Ave	0.902 (0.072)	0.301 (0.020)	3.408 <sup>2</sup> (0.298)	1.796 <sup>2</sup> (0.136)
$h_M^2$ (x 10 <sup>3</sup> )	Ave	0	20.0	0	0.841

0 **Table 2.** Variances. Standard errors in parentheses. See Methods for abbreviations and  
1 details of calculations. For male traits,  $n_{MA,Blk}$  is the fraction of the 53 MA lines that were  
2 included in that block. For hermaphrodite traits, each block had a different set of MA lines, out  
3 of 80 total lines. VM significantly greater than 0 ( $P < 0.05$ ) shown in bold type.  
4 <sup>1</sup> - 95% bootstrap confidence interval in parentheses  
5 <sup>2</sup> - Best-fit model includes residual variances for each treatment (G0, MA) pooled over blocks.

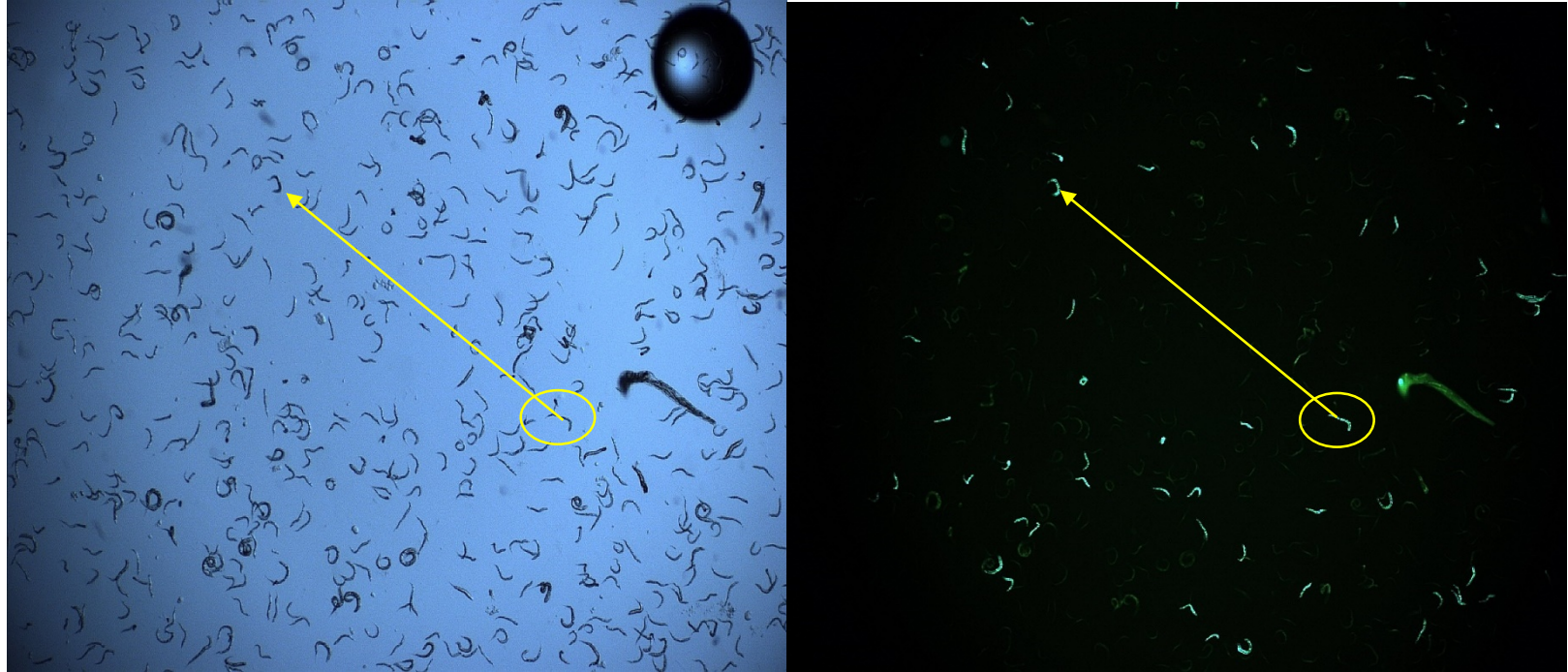
Model	Difference from preceding model	AICc
$y_{ijkl} = \mu + t_k + b_{j k} + c_{jk} + l_{l jk} + \varepsilon_{iljk}$	-	2524.3
<b><math>y_{ijkl} = \mu + t_k + b_{j k} + \varepsilon_{iljk}</math></b>	<b>No treatment by block interaction, no line term</b>	<b>2524.3</b>
$y_{ijkl} = \mu + t_k + b_j + \varepsilon_{iljk}$	Among-block variance pooled over treatments	2527.6
$y_{ijkl} = \mu + t_k + \varepsilon_{iljk}$	No among-block variance	2527.5
$y_{ijkl} = \mu + t_k + \varepsilon_{iljk}$	Residual variance estimated for each treatment, pooled over blocks	2537.1
$y_{ijkl} = \mu + t_k + \varepsilon_{ilj}$	Residual variance estimated for each block, pooled over treatments	2540.2

**Supplementary Table S1a.** AICc scores of linear models tested for male CI ( $CI_M$ ).  $\mu$  is the overall mean,  $t_k$  is the fixed effect of treatment  $k$  (G0 or MA),  $b_j$  is the fixed effect of block  $j$ ,  $c_{jk}$  is the fixed effect of the treatment by block interaction,  $l_{l|jk}$  is the random effect of line (or pseudoline)  $l$ , conditioned on block and treatment, and  $\varepsilon_{iljk}$  is the random residual, conditioned on block and treatment. Terms with a REML estimate of zero were removed from the full model prior to testing simpler models. The best-fit model is shown in bold type.

Model	Difference from preceding model	AICc
$y_{ijkl} = \mu + t_k + b_{jk} + l_{ijk} + \varepsilon_{ijkl}$	Full model	2353.5
$y_{ijkl} = \mu + t_k + b_j + l_{ijk} + \varepsilon_{ijkl}$	Among-block variance pooled over treatments	2352.6
$y_{ijkl} = \mu + t_k + l_{ijk} + \varepsilon_{ijkl}$	No among-block variance	2374.6
$y_{ijkl} = \mu + t_k + b_j + l_{ijk} + \varepsilon_{ijkl}$	Among-block variance pooled over treatments, among-line variance pooled over blocks	2352.9
$y_{ijkl} = \mu + t_k + b_j + l_{ij} + \varepsilon_{ijkl}$	Among-line variance pooled over treatments	2376.7
$y_{ijkl} = \mu + t_k + b_j + l_i + \varepsilon_{ijkl}$	Among-line variance pooled over blocks and treatments	2378.0
$y_{ijkl} = \mu + t_k + b_j + l_{ijk} + \varepsilon_{ijkl}$	Among-line variance estimated separately for each block/treatment combination; residual variance pooled across blocks within each treatment	2394.3
<b><math>y_{ijkl} = \mu + t_k + b_j + l_{ijk} + \varepsilon_{ijl}</math></b>	<b>Residual variance pooled across treatments within each block</b>	<b>2349.4</b>
$y_{ijkl} = \mu + t_k + b_j + l_{ijk} + \varepsilon_{il}$	Residual variance pooled across treatments and blocks	2393.3

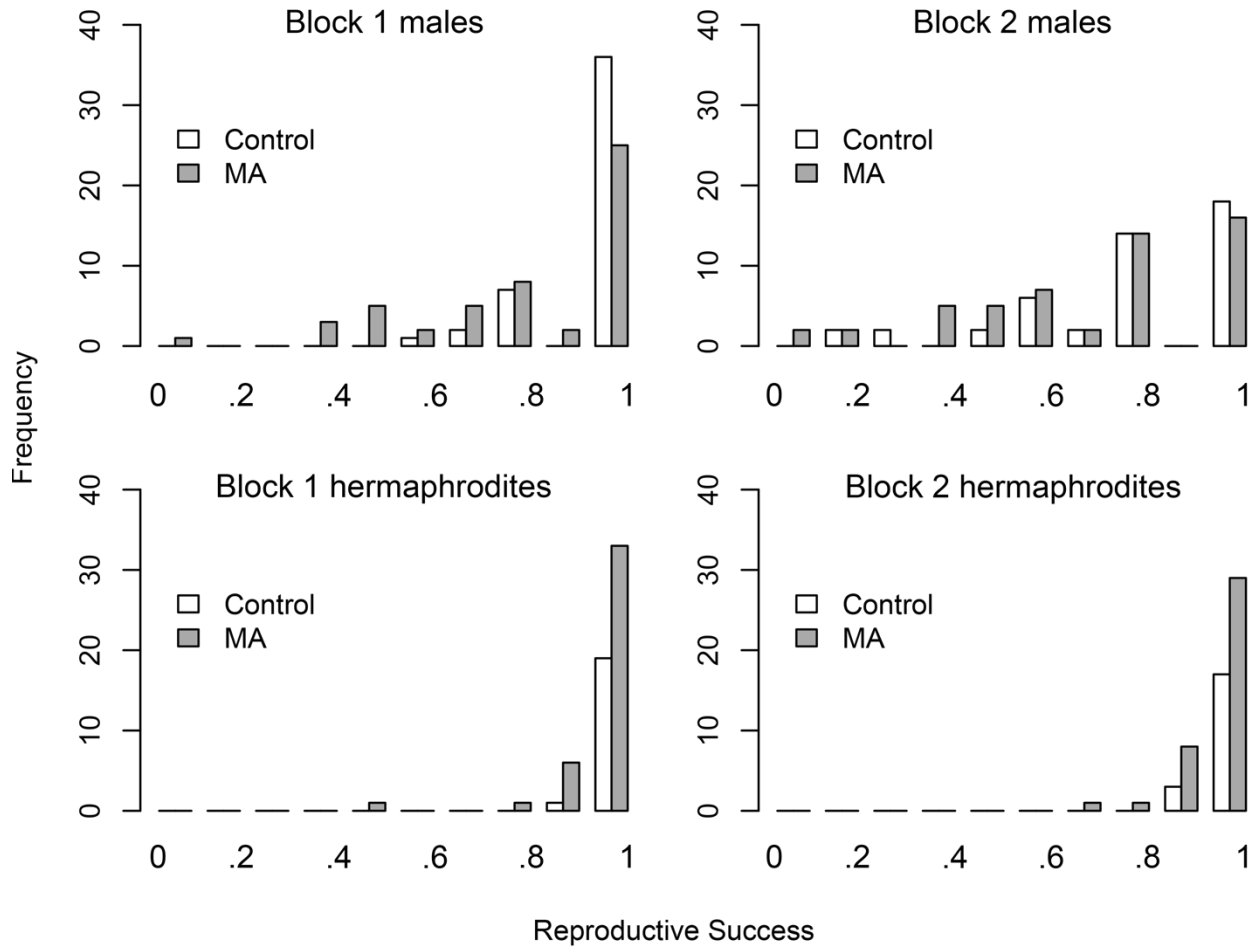
**Supplementary Table S1b.** AICc scores of linear models tested for hermaphrodite CI ( $C_{IH}$ ). Definitions of variables are the same as in Table S1a. Note that since each MA line was present in only one block, there is no treatment by block interaction in these models. The best-fit model is shown in bold type.

x

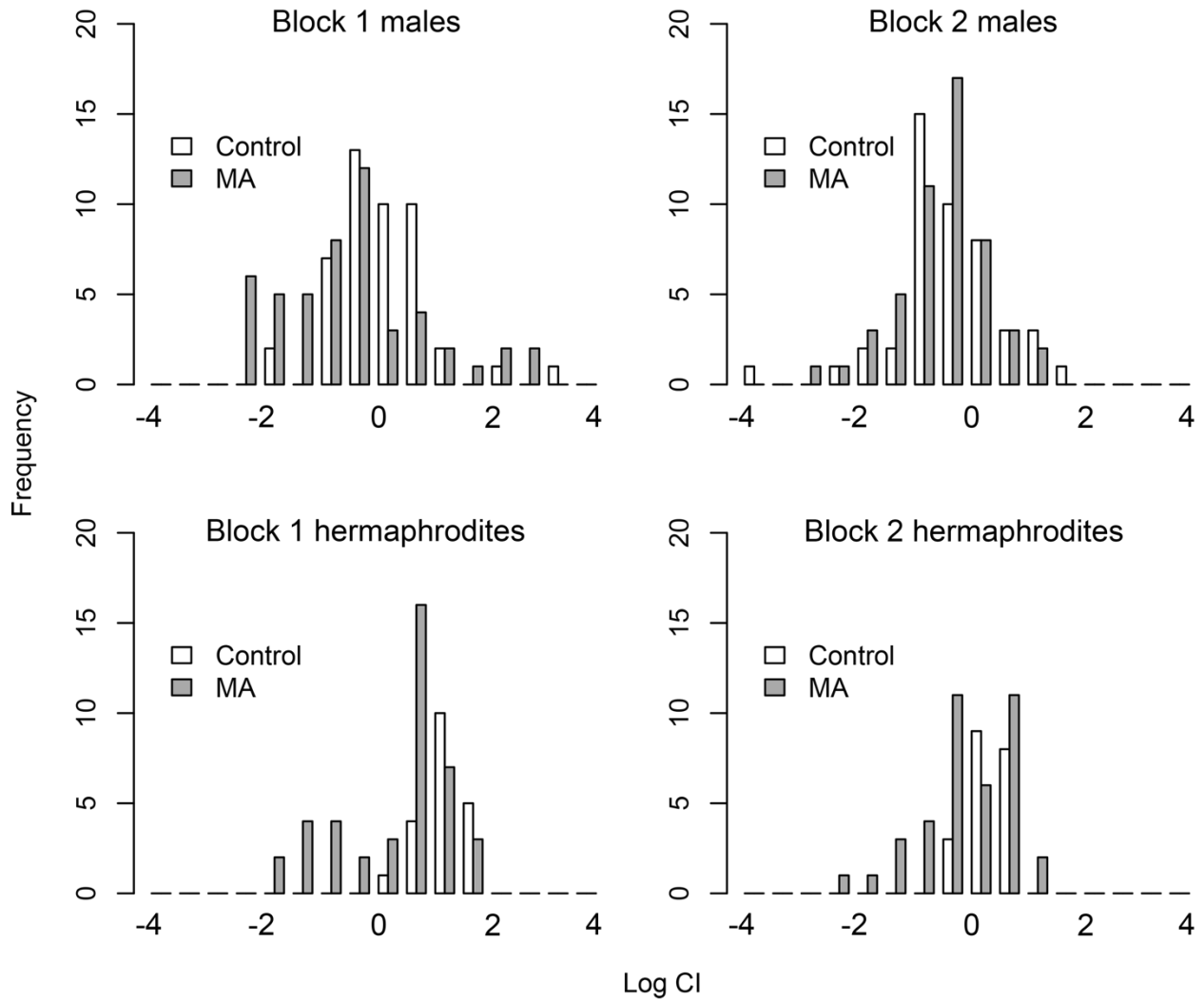


**Supplementary Figure S1.** Example of a side-by-side comparison of the same sample of worms taken under transmitted white light (left) and under 470 nm fluorescent light (right) in the hermaphrodite competitive fitness assay. Worms that appear green under fluorescent light are the GFP-marked competitor (ST-2); worms that do not fluoresce are the focal type, either MA or G0. The yellow oval and arrow highlight the same individual worms in the two images and are shown to emphasize that the images are an exact overlay of each other. The difference between the number of worms in the left image and the right image is the fraction  $p$  of the focal type.





**Supplementary Figure S2.** Distributions of line means of reproductive success (= mating success in males); X-axis values are probabilities of reproducing. Upper panels, male reproductive success ( $\pi_M$ ); lower panels, hermaphrodite reproductive success ( $\pi_H$ ).



**Supplementary Figure S3.** Distributions of line means of  $\log(CI)$ . See Methods for description of calculation of  $CI$ . Upper panels, male reproductive success ( $CI_M$ ); lower panels, hermaphrodite reproductive success ( $CI_H$ ).

1 **Supplementary Appendix A1.** Counting worms with the BioSorter™

2 The Union Biometrica BioSorter™ is a flow cytometer equipped with a flow cell of diameter large  
3 enough (250  $\mu\text{m}$ ) to permit worm-sized particles to be detected; the specifications are included  
4 in the manufacturer's documentation at <http://www.unionbio.com/biosorter/>. The Biosorter  
5 registers an event (a worm, piece of debris, etc.) when a laser detector detects a reduction in  
6 optical density ("extinction") relative to the background optical density. The extinction profile  
7 (the time of flight and area under the curve) is correlated with the length and shape of the object  
8 passing the detector, enabling worms to be distinguished from debris and other non-worm  
9 objects. The Biosorter is equipped with a 488 nm excitation laser and can detect GFP  
10 fluorescence. Events can be "gated" according to time of flight, extinction, and intensity of  
11 fluorescence. Worms at the L2 stage of development or larger can be distinguished from non-  
12 worm events such as debris with reasonably high confidence (details follow). Eggs and L1  
13 stage worms have a much lower signal to noise ratio, so we exclude those developmental  
14 stages from the analysis.

15 Worms were washed from the competition plates (see Methods in the main text) in  
16 approximately 1.5 ml of M9 buffer into 1.5 ml microcentrifuge tubes. Tubes were centrifuged at  
17  $\sim 2K \times g$  for one minute, the supernatant decanted, and the pelleted worms resuspended in 100  
18  $\mu\text{l}$  of M9 and pipetted into a well in a 96-well microtiter plate to be counted with the Biosorter.  
19 The contents of each well of the 96-well plate was counted using the LP Sampler™ microtiter  
20 plate sampler.

21 The sample may contain bacterial clumps, shed worm cuticles, and other non-worm  
22 debris that can register as false positive events in a "worm" gate. Because these false positives  
23 are not fluorescent (beyond a certain background level), and we count both fluorescent and  
24 non-fluorescent worms, false positives inflate the apparent fraction of non-fluorescent worms in  
25 a sample. Overestimation of the fraction of non-fluorescent worms,  $p$ , leads to an overestimate  
26 of the competitive index  $p/(1-p)$ . To quantify and correct for false positives, we set up 125 35

27 mm NGMA test plates with five different population sizes (50, 100, 150, 200, 250) at five  
28 fluorescent-to-non-fluorescent-worm ratios (1:4, 2:3, 1:1, 3:2, 4:1), each replicated five-fold.  
29 Worms were sorted from a mixed-stage mass culture onto the test plates using the BioSorter in  
30 "sort" mode using the same gates as in the fitness assay. Worms were washed from the plates  
31 into a well of a 96-well microtiter plate and counted as described above.

32 The sorting efficiency and rate of detection of fluorescent worms are both > 99%. Some  
33 worms are lost in the wash step, and not all worms present in a microtiter plate will be recorded  
34 as events based on extinction (as opposed to fluorescence). However, there is no reason to  
35 think that loss or failure to record are genotype-specific. The false-positive rate is calculated as  
36 follows. Terms in bold text are depicted in Supplementary Figure S4 below.

37  **$N_T$  = Total events recorded in the test sample**

38  $N_W$  = Total events recorded in the **worm gate**, gating on extinction.

39  $N_{FS}$  = Number of fluorescent worms sorted onto the test plate

40  $N_{FR}$  = Total fluorescent events recorded in the worm gate, gating on fluorescence.

41  $N_W - N_{FR}$  = [number of wild-type worms + non-worms in the worm gate]

42  $L = [N_{FS} - N_{FR}] / N_{FS}$ , Loss Rate, i.e., fraction of worms lost in washing and counting in LP

43 Sampler

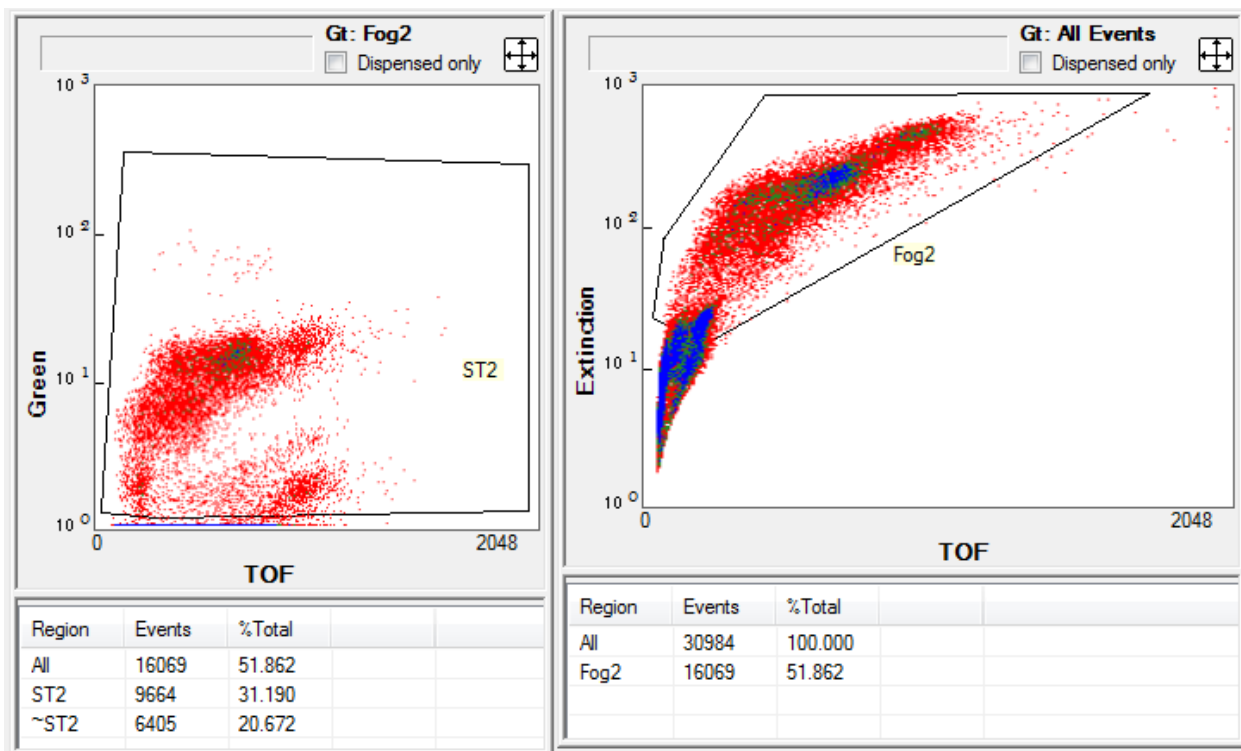
44  $R = 1 - L$ , Recovery rate

45  $FP = [(N_W - N_{FR}) - N_{FS} \times R] / (N_T - N_{FR})$  is the estimated false-positive rate.

46 These calculations were applied to 96 of the test plates, leading to an overall estimated false-  
47 positive rate of 9% ( $FP = 0.09$ ). The number of non-fluorescent worms is estimated as number of  
48 non-fluorescent events in the worm gate,  $(N_W - N_{FR})(1 - FP)$ .

49 In assay block 1, worms were stored at 4° C for four days prior to counting. To account  
50 for the potential effect of refrigeration and storage, 29 of the 125 test plates were stored at 4° C  
51 for four days prior to counting. On average, the number of fluorescent events recorded after four  
52 days of cold storage ( $N_{FR,C}$ ) declined, leading to an increase in the loss rate after cold storage

53 ( $L_C$ ). We assume the difference between L and  $L_C$  (and thus R and  $R_C$ ) is due to dead worms  
 54 that no longer express GFP but that still register as events when gated on extinction. In block 1,  
 55 the number of fluorescent worms in a sample was estimated as  $N_{FR,C} + N_{FR,CX}[(R - R_C)/R_C]$ .  
 56



57  
 58 **Supplementary Figure S4.** The plot on the right displays a density plot of all the events recorded in  
 59 one well of a 96-well plate, with log(extinction) on the y-axis plotted against time of flight on the  
 60 x-axis (TOF is a proxy for object length). The density of events increases from red to green to  
 61 blue. The polygon enclosed by the black lines is the "worm gate", designated "Fog2" in this  
 62 figure. Events characterized by extinction/TOF ratios within the gate are classified as worms;  
 63 events falling outside the gate are classified as not worms. In this example, the total number of  
 64 events recorded in this plate is 30,984, 51.86% of which fell in our worm gate. The plot on the  
 65 left is a subset of the worm gate (designated "ST2" in this example) and shows a plot of  
 66 log(intensity of green fluorescence) against TOF. In this example, 9664/16069 events are  
 67 classified as fluorescent (and thus as worms of the ST-2 strain).