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

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

1 - Department of Biology, University of Florida
2 - Department of Life Sciences, National Central University, Taiwan
3 - Department of Biology, Emory University
4 - University of Florida Genetics Institute

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

Abstract

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

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

Introduction

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

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

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

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

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

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

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

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

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

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

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

Methods and Materials

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

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

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

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

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

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

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

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

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

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

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

Data Analysis -

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

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

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

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

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

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

$$\pi_{ijkl} = \mu + t_k + b_j + c_{jk} + l_{l|jk} + \varepsilon_{il|jk}$$

where π_{ijkl} is a binary variate scored as 1 if the focal worm produced at least two offspring and 0 if it did not, μ 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 $\varepsilon_{il|jk}$ is the random residual, conditioned on block and treatment. Random effects were estimated separately for each block/treatment combination by means of the GROUP option in the RANDOM statement of the GLIMMIX procedure [26]. Significance of fixed effects was determined by F-test of Type III sums of squares, with degrees of freedom determined by the Kenward-Roger method [29].

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

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

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

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

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

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

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

Mutational Bias – The mutational bias is the per-generation rate of change in the trait mean. The slope of the regression of trait mean on generation of MA is often designated R_M [33]; the

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

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

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

meaningful to scale each group (G0 and MA) by its own mean [21]. Scaling by the group means is nearly equivalent to calculating ΔV_L from log-transformed data [37]. *CI* is on a logarithmic scale and cannot be mean-standardized. ΔV_L for CI_M and for π_M and π_H is not significantly different from 0 (see Results), so scaling is irrelevant.

Results

Rate of mutations with Zero Male Fitness - Of the 60 of the original 100 MA lines remaining in 2014, we were able to obtain fertile males from 53. We assume that the seven lines for which we were unable to obtain fertile males carry at least one mutation that leads to Zero Male Fitness (ZMF, i.e. inviable or sterile), and that these mutations follow a Poisson distribution - analogous to lethal equivalents [38]. With those assumptions, the expected proportion of lines with no mutations (p_0, i.e., the number of lines that produced fertile males) is: $p_0 = e^{-m}$, where m is the expected number of mutations carried by a line [39]. The expected number of mutations $m = \mu t$, where μ is the per-generation rate of ZMF mutations and t is the number of generations of MA. Thus, $(53/60) = e^{-250\mu}$, so $\mu_{ZMF} \approx 5\times10^{-4}$/generation, about half the lower-bound estimate on the frequency of males. If the ZMF mutation rate is greater than the frequency of males, the average chromosome will take a male-sterilizing hit before the next time it winds up in a male, leading to the loss of males by "error catastrophe" [40]. The same calculation from data reported in [41] gives an estimate of $\mu_{ZMF} \approx 3\times10^{-4}$/generation.

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

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

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

As noted, *CI* cannot be directly mean-standardized. However, *CI* is related to relative fitness by equation [1] above. ΔM_W can be calculated from $\frac{p_t}{q_t} = \frac{p_0}{q_0}\left(\frac{W_{foc}}{W_C}\right)^t$, where $CI = \left(\frac{p_t}{q_t}\right)$, $p_0 = q_0 = 0.5$, and $t = 1$, thus $CI = \left(\frac{W_{foc}}{W_C}\right)$. The ratio of the fitnesses relative to the competitor (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 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 x 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 ΔVL 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 (π_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 (ΔM = -1.9 x 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 π_H among lines is highly left-skewed (median π_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 π_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 (π_H > 97% for both G0 and MA; ΔM = -

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

(ii) Competitive Index (CI_H) – Mean CI_H declined significantly over the course of 250 generations 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 1; distributions of line means are shown in Supplementary Figure S3]. $R_{M,CI,H}$ calculated from the slope of the regression of $\log(CI_H)$ on generation of MA is $-2.54 \pm 0.49 \times 10^{-3}$ per-generation. From equation [1], $\frac{p_t}{q_t} = \frac{p_0}{q_0}\left(\frac{W_{foc}}{W_C}\right)^t$, where $CI = (\frac{p_t}{q_t})$, $p_0 = q_0 = 0.5$ and here t is equal to the number of generations of reproduction the population underwent over the course of the eight-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 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 that of the G0 ancestor, $\left(\frac{w_{MA}}{w_{G0}}\right)$.

We cannot be certain about the exact number of generations of reproduction, except that three is the upper bound (based on the worm's life cycle), and the true number is probably close to two. The basis for that judgment is this: if the average worm produces 200 offspring in its lifetime, after one generation there will be 2x200=400 worms on the plate and after two generations there will be 400x200=80,000 worms on the plate (density-dependence notwithstanding); after three generations there will be 80,000x200=16 million. There were many more than 400 and certainly fewer than 80,000, so we assume two generations is probably close to the true number of generations.

Assuming that $t = 2$, we find $\left(\frac{w_{MA}}{w_{G0}}\right) = 0.728$, or in other words, relative competitive fitness declined by about 27% over 250 generations of MA. Scaled relative to the G0 ancestor, $\Delta M_w \approx$ 1.09 x 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.

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

Discussion

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

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

encompassed by the ~0.11% per generation decline in relative competitive fitness. In males, both the probability of siring offspring and the fraction of offspring sired given that the male did mate successfully declined. A calculation based on the point estimates of the ΔMs (Table 1) shows that, after one generation of MA, hermaphrodite relative fitness will have declined by $1 - [(1 - 1.95\times10^{-5})(1 - 1.09\times10^{-3})] \approx 0.11\%$. By the same reasoning, male relative fitness will have declined by $1 - [(1 - 6.82\times10^{-4})(1 - 1.04\times10^{-3}) \approx 0.17\%$. Thus, male fitness is certainly no less sensitive, and perhaps slightly more sensitive to the cumulative deleterious effects of mutation than is hermaphrodite fitness. This result is quantitatively nearly identical to the finding that male *Drosophila melanogaster* decline in fitness ~1.5X faster from mutation accumulation than do females [20].

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

N_e of *C. elegans* has been estimated from the standing nucleotide diversity as being on the order of 10^4 [10]. If a mutant allele is under selection in males but is neutral in 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) = 4.4×10^{-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 (π_M and CI_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.

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

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

Acknowledgments

We thank Dustin Blanton, Whitney Bour, Myrnelle Damas, Thomas Keller, Laura Levy, Gigi Ostrow and Naomi Phillips for assistance in the lab. *C. elegans* strains were graciously 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

References

[1] Rockman, M.V. & Kruglyak, L. 2009 Recombinational Landscape and Population Genomics of Caenorhabditis elegans. *PLoS Genetics* **5**. (doi:e1000419 10.1371/journal.pgen.1000419).

[2] Kiontke, K. & Fitch, D.H.A. 2005 The phylogenetic relationships of Caenorhabditis and other rhabditids. *Wormbook*, 1-11.

[3] Cutter, A.D., Wasmuth, J.D. & Washington, N.L. 2008 Patterns of molecular evolution in Caenorhabditis preclude ancient origins of selfing. *Genetics* **178**, 2093-2104. (doi:10.1534/genetics.107.085787).

[4] Herman, R.K. 2005 Introduction to sex determination. In *Wormbook* (ed. T.*C.e.*R. Community).

[5] Stewart, A.D. & Phillips, P.C. 2002 Selection and maintenance of androdioecy in *Caenorhabditis elegans*. *Genetics* **160**, 975-982.

[6] Hodgkin, J. & Doniach, T. 1997 Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**, 149-164.

[7] Anderson, J.L., Morran, L.T. & Phillips, P.C. 2010 Outcrossing and the maintenance of males within *C. elegans* populations. *Journal of Heredity* **101**, S62-S74. (doi:10.1093/jhered/esq003).

[8] Fisher, R.A. 1930 *The Genetical Theory of Natural Selection*. Oxford, Clarendon Press.

[9] Charnov, E.L. 1982 *The Theory of Sex Allocation*. Princeton, NJ, USA, Princeton University Press.

[10] Andersen, E.C., Gerke, J.P., Shapiro, J.A., Crissman, J.R., Ghosh, R., Bloom, J.S., Felix, M.-A. & Kruglyak, L. 2012 Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nature Genetics* **44**, 285-U283. (doi:10.1038/ng.1050).

[11] Ohta, T. 1973 Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96-98.

[12] Murray, R.L., Kozlowska, J.L. & Cutter, A.D. 2011 Heritable determinants of male fertilization success in the nematode Caenorhabditis elegans. *Bmc Evolutionary Biology* **11**. (doi:10.1186/1471-2148-11-99).

[13] Charlesworth, B. 2015 Causes of natural variation in fitness: Evidence from studies of Drosophila populations. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 1662-1669. (doi:10.1073/pnas.1423275112).

[14] Davies, S.K., Leroi, A.M., Burt, A., Bundy, J. & Baer, C.F. 2016 The mutational structure of metabolism in *Caenorhabditis elegans*. *Evolution*. (doi:10.1111/evo.13020).

[15] Mukai, T., Chigusa, S.I., Crow, J.F. & Mettler, L.E. 1972 Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* **72**, 335-355.

[16] Willis, J.H. 1999 The contribution of male-sterility mutations to inbreeding depression in *Mimulus guttatus*. *Heredity* **83**, 337-346. (doi:10.1038/sj.hdy.6885790).

[17] Rowe, L. & Houle, D. 1996 The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society B-Biological Sciences* **263**, 1415-1421. (doi:10.1098/rspb.1996.0207).

[18] Mallet, M.A., Kimber, C.M. & Chippindale, A.K. 2012 Susceptibility of the male fitness phenotype to spontaneous mutation. *Biology Letters* **8**, 426-429. (doi:10.1098/rsbl.2011.0977).

[19] Mallet, M.A., Bouchard, J.M., Kimber, C.M. & Chippindale, A.K. 2011 Experimental mutation-accumulation on the X chromosome of *Drosophila melanogaster* reveals stronger selection on males than females. *BMC Evolutionary Biology* **11**. (doi:10.1186/1471-2148-11-156).

[20] Sharp, N.P. & Agrawal, A.F. 2013 Male-biased fitness effects of spontaneous mutations in *Drosophila melanogaster*. *Evolution* **67**, 1189-1195. (doi:10.1111/j.1558-5646.2012.01834.x).

[21] Baer, C.F., Shaw, F., Steding, C., Baumgartner, M., Hawkins, A., Houppert, A., Mason, N., Reed, M., Simonelic, K., Woodard, W., et al. 2005 Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 5785-5790.

[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

	π_M	π_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)	**-0.622** (0.159)	-0.0192 (0.050)	**-1.26** (0.48)	**-2.54** (0.49)
ΔM (x 10^3/gen)	-0.682	-0.0195	-1.04†	-1.09*,†

Table 1. Evolution of trait means, averaged across blocks; standard errors in parentheses. Sample sizes are given in Table 2. Values of R_M significantly different from zero (P<0.05) shown in bold type. See Methods for abbreviations and details of calculations. ΔM of CI is calculated for back-transformed data.

* - Assumes t = 2 generations of reproduction

† - ΔM calculated for *CI*, not log(*CI*)

Variance Component	***Block***	π_M $n_{MA,Blk1}$ = 46/51 $n_{MA,Blk2}$ = 53/53	π_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[1] (0, 4.299)	0	0.571 (0.128)
VM (x 10^3)	Ave	0	6.104[1] (0, 8.598)	0	**1.142** (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[2] (0.173)	0.822[2] (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[2] (0.298)	1.796[2] (0.136)
h_M^2 (x 10^3)	Ave	0	20.0	0	0.841

Table 2. Variances. Standard errors in parentheses. See Methods for abbreviations and details of calculations. For male traits, $n_{MA,Blk.}$ is the fraction of the 53 MA lines that were included in that block. For hermaphrodite traits, each block had a different set of MA lines, out of 80 total lines. VM significantly greater than 0 ($P<0.05$) shown in bold type.

[1] - 95% bootstrap confidence interval in parentheses

[2] - Best-fit model includes residual variances for each treatment (G0, MA) pooled over blocks.

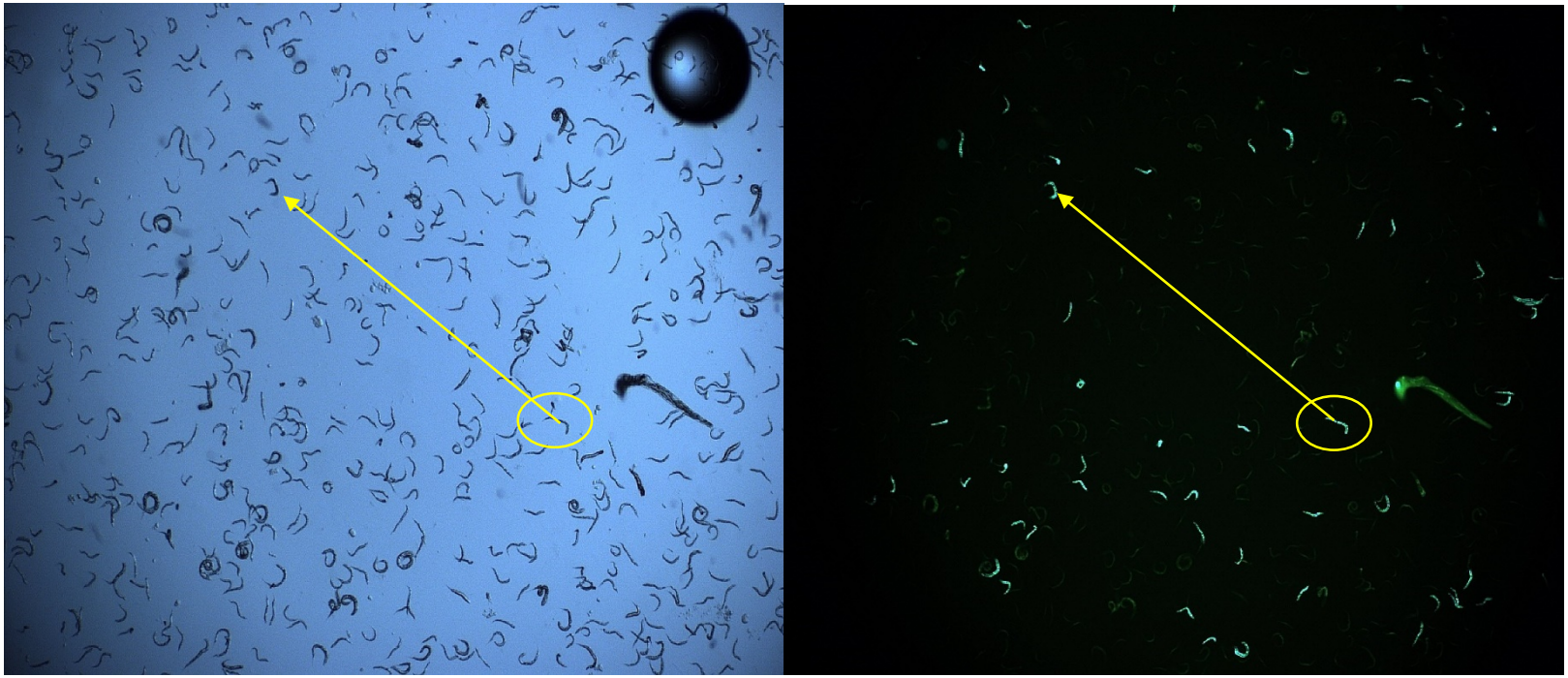
Model	Difference from preceding model	AICc
$y_{ijkl} = \mu + t_k + b_{j\|k} + c_{jk} + l_{l\|jk} + \varepsilon_{il\|jk}$	-	2524.3
$y_{ijkl} = \mu + t_k + b_{j\|k} + \varepsilon_{il\|jk}$	**No treatment by block interaction, no line term**	**2524.3**
$y_{ijkl} = \mu + t_k + b_j + \varepsilon_{il\|jk}$	Among-block variance pooled over treatments	2527.6
$y_{ijkl} = \mu + t_k + \varepsilon_{il\|jk}$	No among-block variance	2527.5
$y_{ijkl} = \mu + t_k + \varepsilon_{il\|k}$	Residual variance estimated for each treatment, pooled over blocks	2537.1
$y_{ijkl} = \mu + t_k + \varepsilon_{il\|j}$	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). μ 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 $\varepsilon_{il|jk}$ 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.

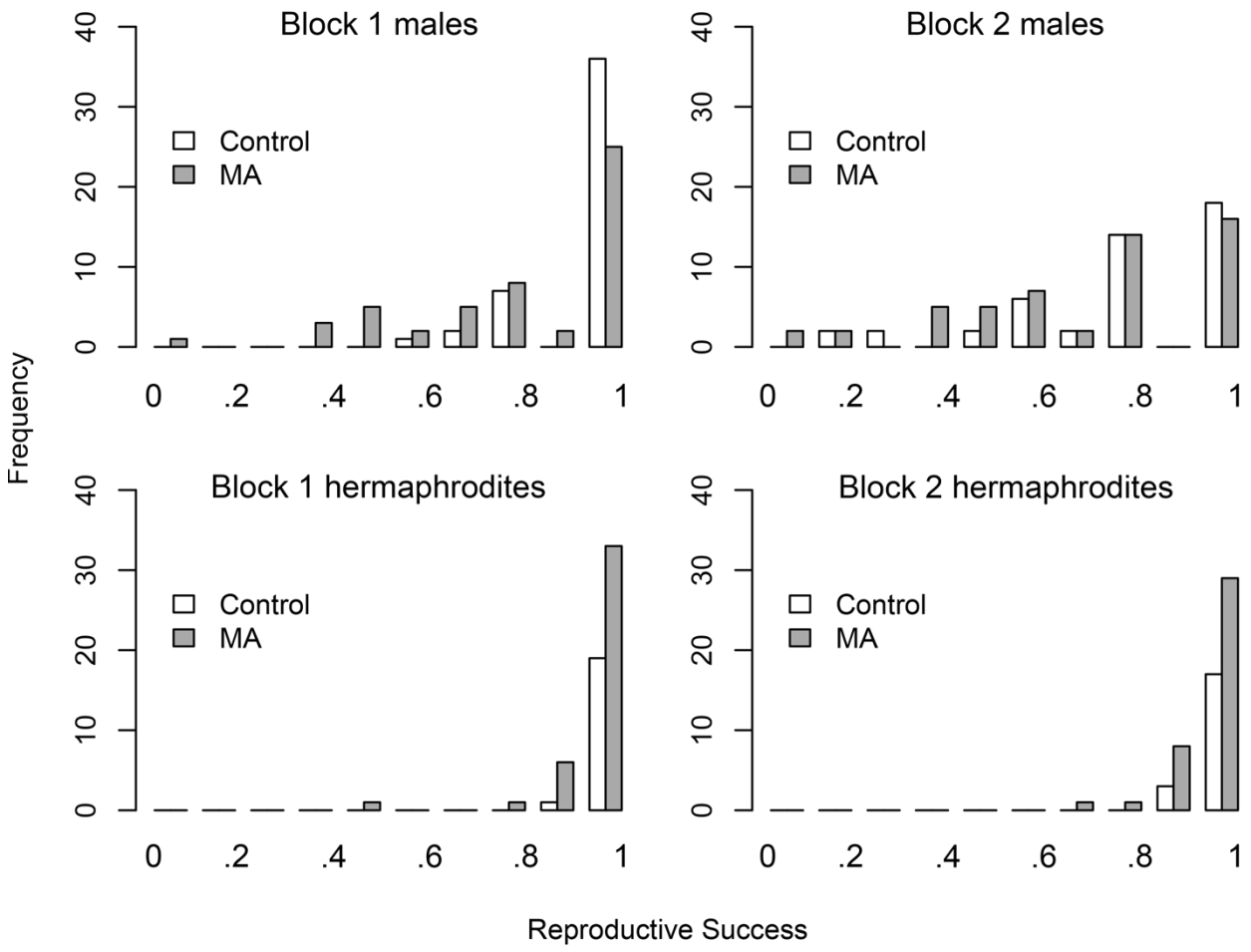
Model	Difference from preceding model	AICc
$y_{ijkl} = \mu + t_k + b_{j\|k} + l_{l\|jk} + \varepsilon_{il\|jk}$	Full model	2353.5
$y_{ijkl} = \mu + t_k + b_j + l_{l\|jk} + \varepsilon_{il\|jk}$	Among-block variance pooled over treatments	2352.6
$y_{ijkl} = \mu + t_k + l_{l\|jk} + \varepsilon_{il\|jk}$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

Supplementary Table S1b. AICc scores of linear models tested for hermaphrodite CI (CI_H). 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.

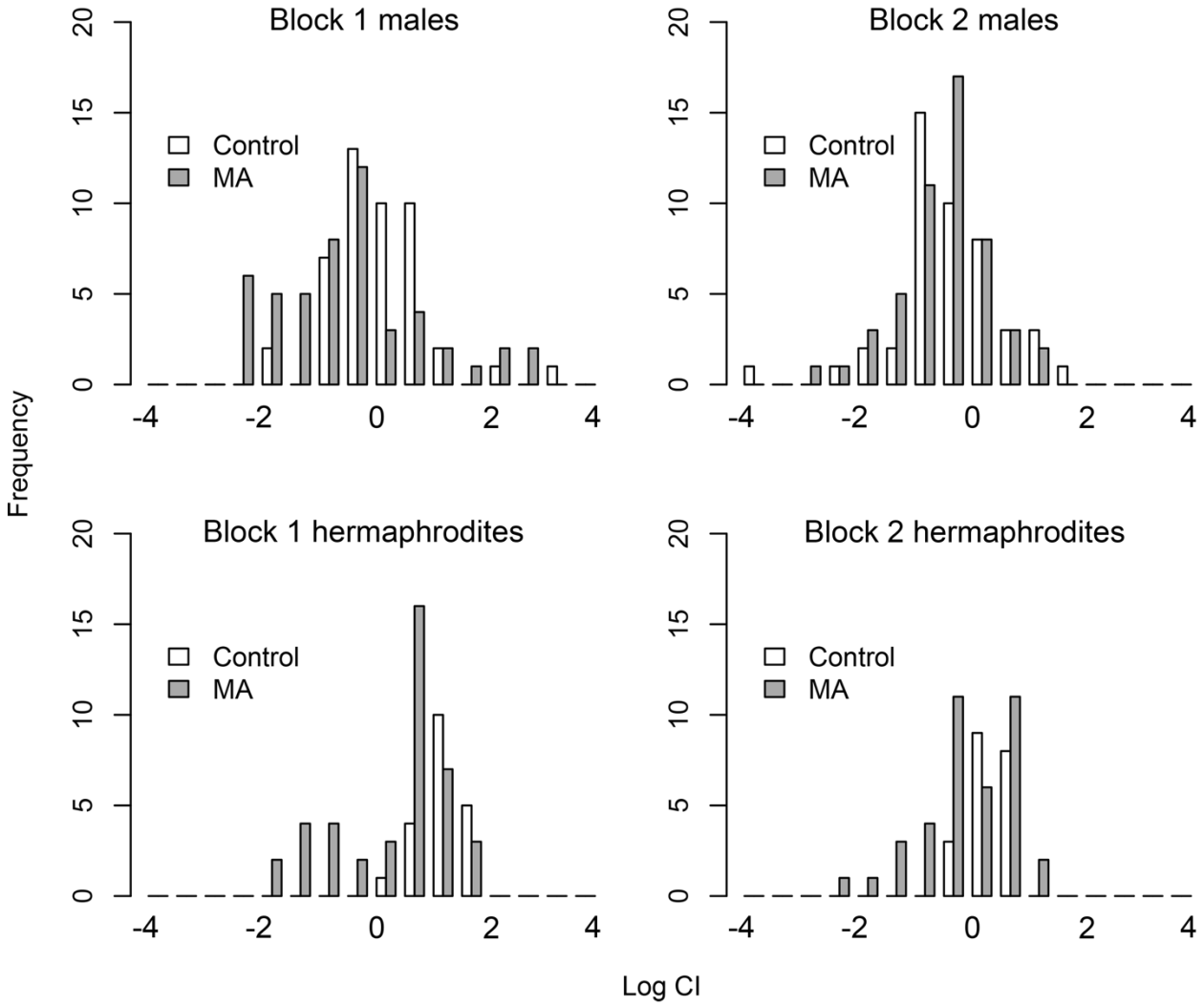
x



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 (π_M); lower panels, hermaphrodite reproductive success (π_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).

Supplementary Appendix A1. Counting worms with the BioSorter™

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

Worms were washed from the competition plates (see Methods in the main text) in approximately 1.5 ml of M9 buffer into 1.5 ml microcentrifuge tubes. Tubes were centrifuged at ~2K x g for one minute, the supernatant decanted, and the pelleted worms resuspended in 100 µl of M9 and pipetted into a well in a 96-well microtiter plate to be counted with the Biosorter. The contents of each well of the 96-well plate was counted using the LP Sampler™ microtiter plate sampler.

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

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

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

N_T = Total events recorded in the test sample

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

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

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

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

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

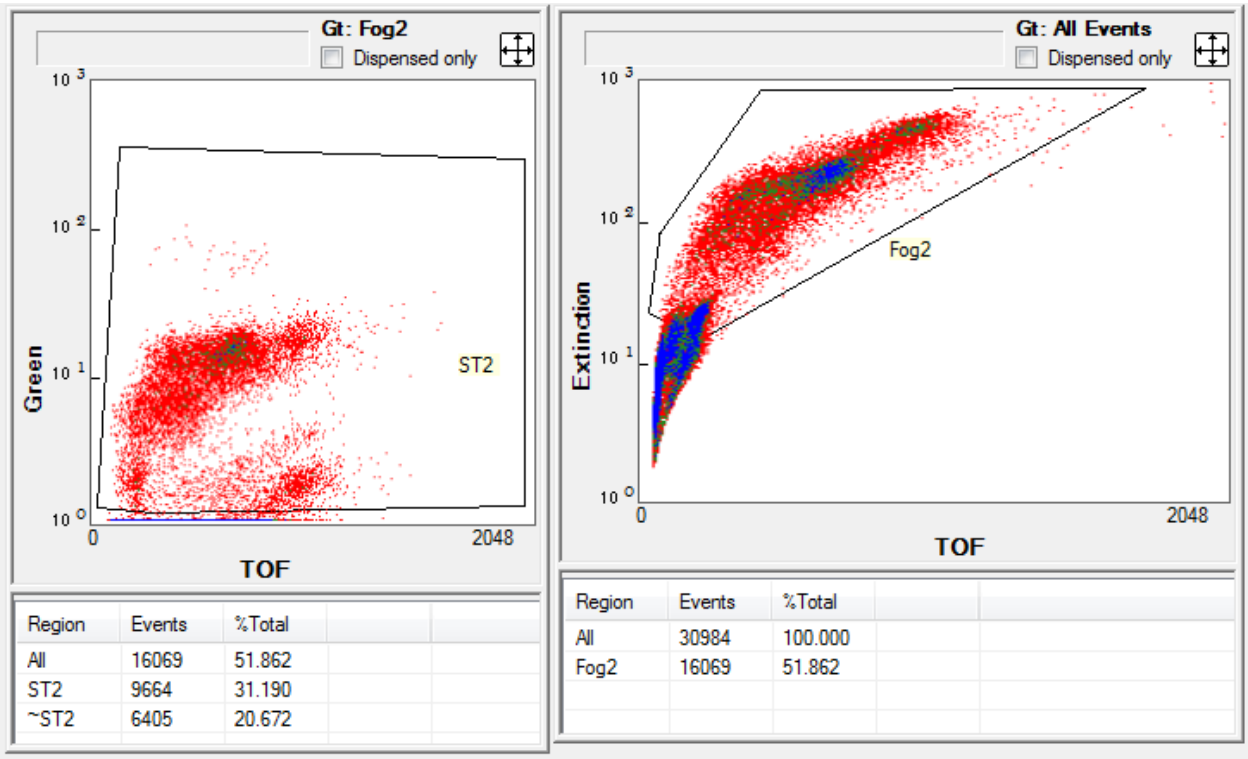
R = 1-L, Recovery rate

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

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

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

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



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