

# SUPPLEMENTARY RESULTS

## Maximum likelihood estimation of fitness components in experimental evolution

Jingxian Liu<sup>1,2\*</sup>, Jackson Champer<sup>1,2\*+</sup>, Anna Maria Langmüller<sup>1,3,4</sup>, Chen Liu<sup>1,2</sup>, Joan Chung<sup>1,2</sup>, Riona Reeves<sup>1,2</sup>, Anisha Luthra<sup>1,2</sup>, Yoo Lim Lee<sup>1,2</sup>, Andrew H. Vaughn<sup>1</sup>, Andrew G. Clark<sup>1,2</sup>, Philipp W. Messer<sup>1+</sup>

<sup>1</sup>Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853

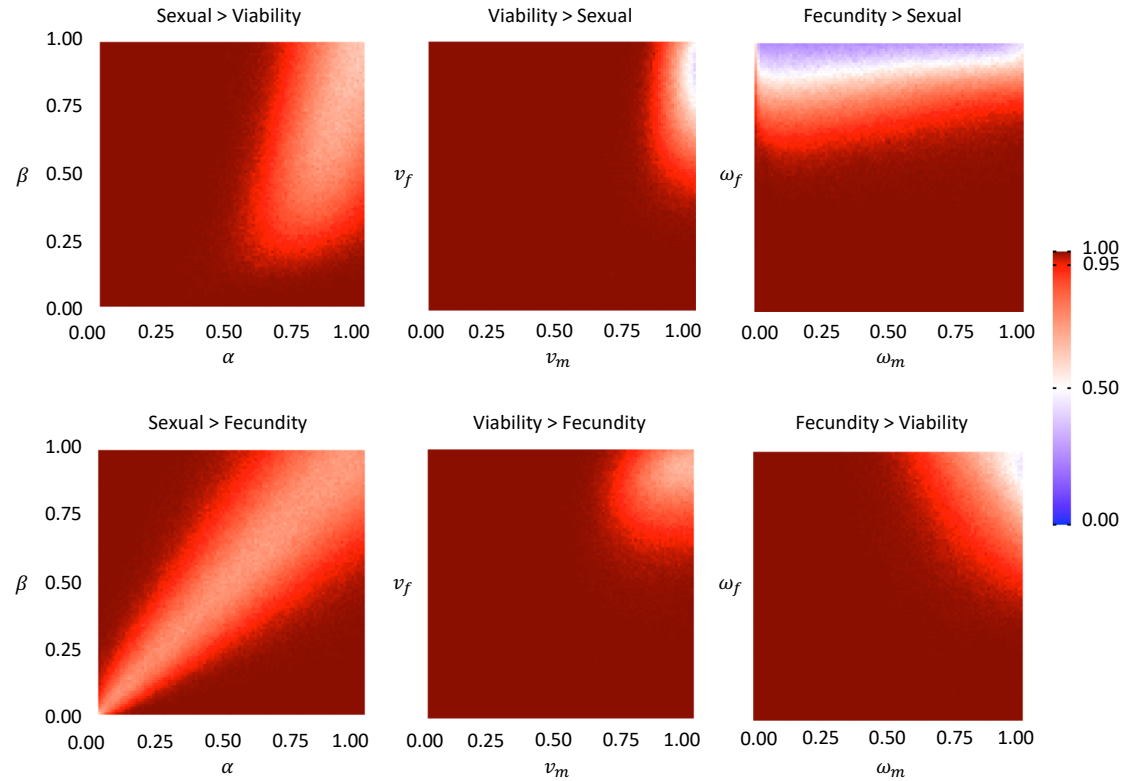
<sup>2</sup>Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

<sup>3</sup>Institut für Populationsgenetik, Vetmeduni Vienna, Veterinärplatz 1, 1210 Wien 1210, Austria

<sup>4</sup>Vienna Graduate School of Population Genetics, 1210, Austria

\*equal contribution

+corresponding authors: jc3248@cornell.edu (J.C.) or messer@cornell.edu (P.W.M.)



**Figure S1.** Power of the inference method to distinguish different types of selection. The analyses are similar to Figure 4 of the main manuscript, but here we show comparisons between two inference models. The first term in the title of each panel specifies the type of selection used in the simulations, while the second term specifies the type of selection used in the inference model to which the correct model is compared. The heatmaps show for each grid point the fraction of simulation runs in which the correct inference model yielded a higher maximum likelihood than the comparison model.

## Supplementary Results: Evolutionary model for an autosomal locus

The evolutionary model developed in the main manuscript focused on an X-linked locus with two segregating alleles: wild-type (+) and yellow (*y*). Here, we will extend this model to an autosomal locus. We again assume a recessive phenotype and will use the following notation for the different genotypes:

$++\sigma^{\circ}$	male homozygous for wild-type allele
$++\varphi$	female homozygous for wild-type allele
$y+\sigma^{\circ}$	heterozygous male
$y+\varphi$	heterozygous female
$yy\sigma^{\circ}$	male homozygous for yellow allele
$yy\varphi$	female homozygous for yellow allele

The notation for phenotypes will be:

$+\sigma^{\circ}$	male with wild-type phenotype (thus either $++\sigma^{\circ}$ or $y+\sigma^{\circ}$ genotype)
$y\sigma^{\circ}$	male with yellow phenotype (thus $yy\sigma^{\circ}$ genotype)
$+\varphi$	female with wild-type phenotype (thus either $++\varphi$ or $y+\varphi$ genotype)
$y\varphi$	female with yellow phenotype (thus $yy\varphi$ genotype)

The definitions of the viability and fecundity parameters in the X-linked model translate directly to the autosomal model. Specifically, the relative viabilities of yellow-phenotype males vs. wildtype males, and yellow-phenotype females vs. wildtype females, are still defined as

$$\begin{aligned}\frac{\nu(y\sigma^{\circ})}{\nu(+\sigma^{\circ})} &= v_m \\ \frac{\nu(y\varphi)}{\nu(+\varphi)} &= v_f,\end{aligned}\tag{1}$$

and the relative fecundities of matings involving parents of different phenotypes are still defined by

	$+\varphi$	$y\varphi$
$+\sigma^{\circ}$	1	$\omega_f$
$y\sigma^{\circ}$	$\omega_m$	$\omega_b$

Note though that for the autosomal locus, the  $+\sigma^{\circ}$  individuals now include both  $++\sigma^{\circ}$  and  $y+\sigma^{\circ}$  genotypes. Due to the way in which we will set up the equations for calculating expected genotype frequencies in the next generation, it will be helpful to explicitly distinguish the three possible male genotypes when defining sexual selection in the autosomal model. Specifically, we will define the probabilities that a female of a given *phenotype* will mate with a male of a given *genotype* as

$$\begin{aligned}P(++\sigma^{\circ}|+\varphi) &= \frac{M_{++}}{M_{++} + M_{y+} + \alpha M_{yy}} \\ P(y+\sigma^{\circ}|+\varphi) &= \frac{M_{y+}}{M_{++} + M_{y+} + \alpha M_{yy}} \\ P(yy\sigma^{\circ}|+\varphi) &= \frac{\alpha M_{yy}}{M_{++} + M_{y+} + \alpha M_{yy}} \\ P(++\sigma^{\circ}|y\varphi) &= \frac{M_{++}}{M_{++} + M_{y+} + \beta M_{yy}} \\ P(y+\sigma^{\circ}|y\varphi) &= \frac{M_{y+}}{M_{++} + M_{y+} + \beta M_{yy}} \\ P(yy\sigma^{\circ}|y\varphi) &= \frac{\beta M_{yy}}{M_{++} + M_{y+} + \beta M_{yy}}\end{aligned}\tag{2}$$

With these definitions we can now calculate the expected genotype frequencies in the next generation, applying an analogous approach as used for deriving equation (4) in the main manuscript. Briefly, for each

genotype, we will sum over all possible mating pairs that can produce offspring of that genotype in the next generation, and then weight each such pair by its respective mating probability, fecundity, and the Mendelian ratio at which it is expected to produce the particular genotype. Genotype frequencies are then scaled by their relative viabilities. Finally, we normalize the frequencies (by the factors  $c_1$  and  $c_2$ ) such that male and females frequencies each sum up to 1. This yields:

$$\begin{aligned}
p_{++\sigma} &= c_1 \left( P(++\sigma|+\varnothing) + \frac{P(y+\sigma|+\varnothing)}{2} \right) \left( N_{++} + \frac{N_{y+}}{2} \right) \\
p_{++\varnothing} &= c_2 \left( P(++\sigma|+\varnothing) + \frac{P(y+\sigma|+\varnothing)}{2} \right) \left( N_{++} + \frac{N_{y+}}{2} \right) \\
p_{y+\sigma} &= c_1 \left[ \omega_f \left( P(++\sigma|y\varnothing) + \frac{P(y+\sigma|y\varnothing)}{2} \right) N_{yy} + \omega_m P(yy\sigma|+\varnothing) \left( \frac{N_{y+}}{2} + N_{++} \right) + P(y+\sigma|+\varnothing) \left( \frac{N_{y+}}{2} + \frac{N_{++}}{2} \right) \right. \\
&\quad \left. + P(++\sigma|+\varnothing) \frac{N_{y+}}{2} \right] \\
p_{y+\varnothing} &= c_2 \left[ \omega_f \left( P(++\sigma|y\varnothing) + \frac{P(y+\sigma|y\varnothing)}{2} \right) N_{yy} + \omega_m P(yy\sigma|+\varnothing) \left( \frac{N_{y+}}{2} + N_{++} \right) + P(y+\sigma|+\varnothing) \left( \frac{N_{y+}}{2} + \frac{N_{++}}{2} \right) \right. \\
&\quad \left. + P(++\sigma|+\varnothing) \frac{N_{y+}}{2} \right] \\
p_{yy\sigma} &= c_1 v_m \left[ \left( \frac{P(y+\sigma|+\varnothing)}{2} + \omega_m P(yy\sigma|+\varnothing) \right) \frac{N_{y+}}{2} + \omega_f \left( \frac{P(y+\sigma|y\varnothing)}{2} + \omega_m P(yy\sigma|y\varnothing) \right) N_{yy} \right] \\
p_{yy\varnothing} &= c_2 v_f \left[ \left( \frac{P(y+\sigma|+\varnothing)}{2} + \omega_m P(yy\sigma|+\varnothing) \right) \frac{N_{y+}}{2} + \omega_f \left( \frac{P(y+\sigma|y\varnothing)}{2} + \omega_m P(yy\sigma|y\varnothing) \right) N_{yy} \right]
\end{aligned} \tag{3}$$

The probabilistic model we will use to specify the probabilities of observing a given set of genotypes is the same as in the main manuscript. We will again assume  $N_m = N_f = N_e/2$  in the autosomal model. Because of the three distinct male genotypes, compared to only two in the X-linked case, the normalization coefficient  $c_m$  now changes for the  $d$ -dimensional standard simplex from  $d = 2$  to  $d = 3$ . Thus, we have  $c_m = c_f$ , as specified by equation (8) in the main manuscript.

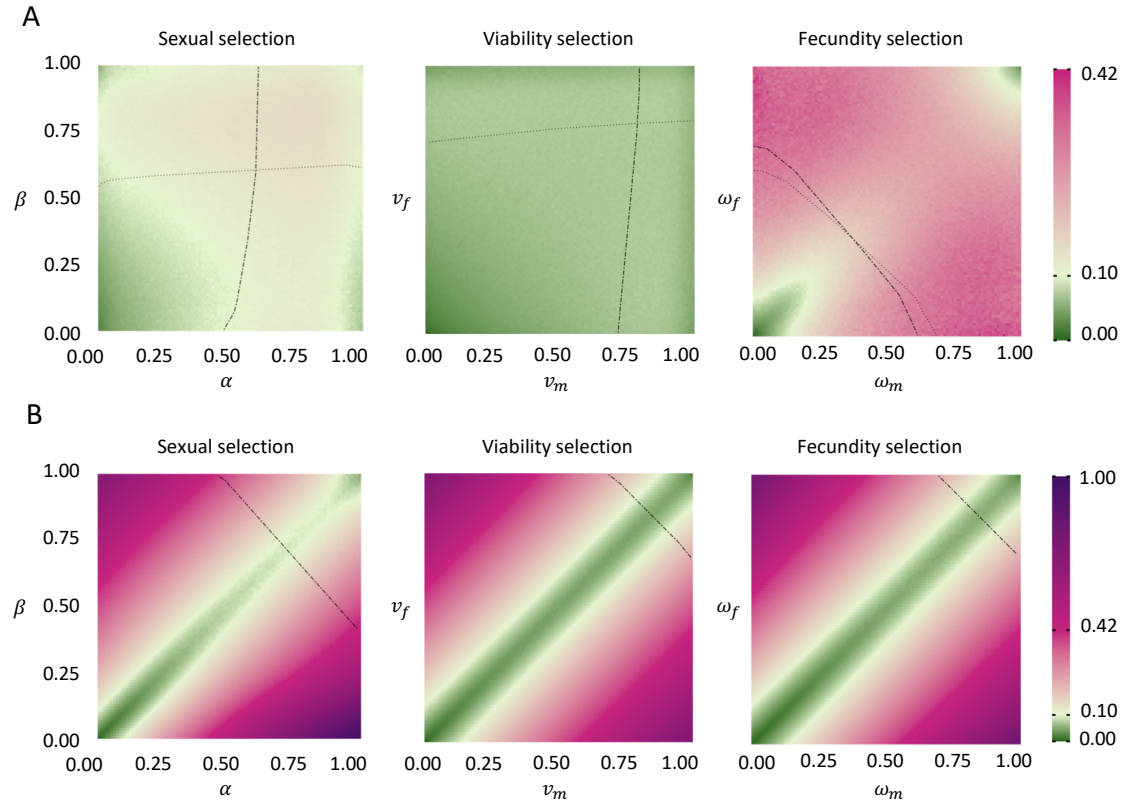
## Power evaluation

To test the performance of our inference method for the autosomal model, we performed a power analysis similar to the one for the X-linked model in the main manuscript. Specifically, we simulated evolution experiments lasting 10 generations with an effective population size of  $N_e = 200$  and assuming a female:male ratio of 1. The simulations were initialized with a  $y$ -allele frequency of 70% in both males and females in generation zero. Only homozygotes were assumed to be present at the start of the experiment. For the fecundity scenario, we again assumed  $\omega_b = \omega_m \omega_f$ . Each of the three selection scenarios (sexual selection, viability selection, and fecundity selection) was therefore defined by two parameters. 1000 simulations were run at each grid point of the parameter space for each the three models.

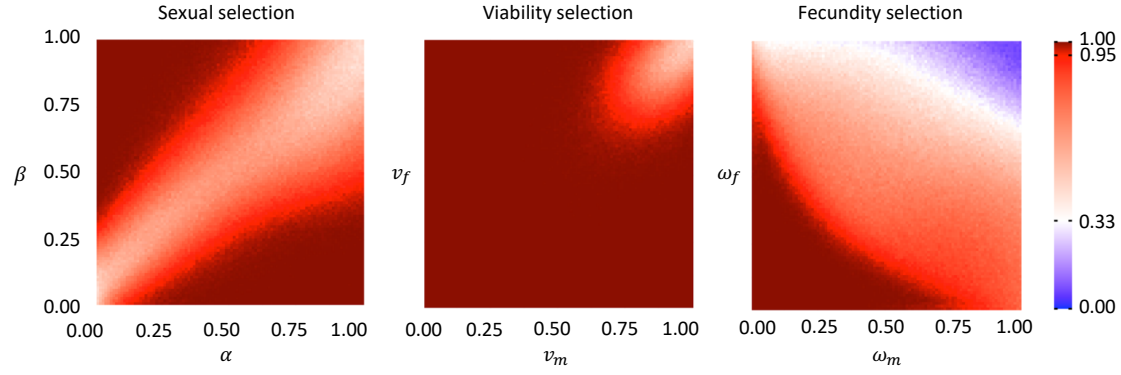
We first evaluated the accuracy of parameter MLE for the autosomal model, analogous to Figure 3D for the X-linked model. The resulting heatmaps are shown in Figure S2A. This analysis demonstrates that inference accuracy is still good for viability and sexual selection in the autosomal model. Parameter MLEs for fecundity selection, however, are less accurate compared to the X-linked model. Figure S2A also depicts the selection thresholds below which we can reject a value of 1 for a given selection parameter (this analysis corresponds to Figure 3C in the main manuscript for the X-linked model). For viability and sexual

selection scenarios, these thresholds are again quite similar between the X-linked and autosomal models, but for fecundity selection, the autosomal model has distinctly lower sensitivity than the X-linked model. Interestingly, the thresholds between the male and female parameters are highly anti-correlated for the autosomal model. This hints at the possibility that our inference method could still have good sensitivity to detect selection overall, but may just no longer be able to distinguish between fitness costs in males and females. To test this hypothesis, we evaluated inference models where we forced the male and female selection parameters to be equal, so that there was only one free inference parameter for a given selection type (e.g.  $\nu = \nu_m = \nu_f$  for the viability model). The results of this analysis are shown in Figure S3B. Indeed, parameter MLEs are now very accurate along the diagonal of the parameter space in this one-parameter inference model (note that the diagonal of the parameter space is now the only area where the inference model is the “correct” model). This one-parameter model also recovers sensitivity for detecting selection, as indicated by the fact that the selection thresholds where we can reject a value of 1 for a selection parameter are now again much higher.

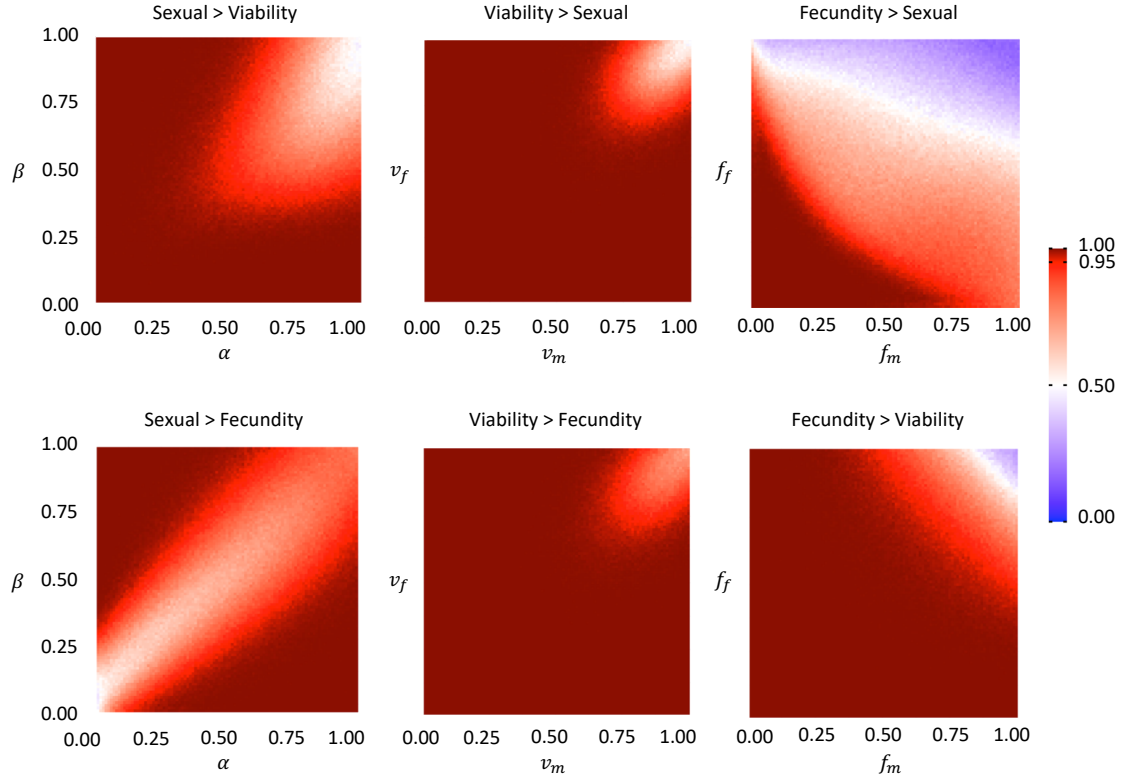
Figures S3 and S4 demonstrate that for the autosomal model, our method also retains reasonable power in distinguishing the different types of selection, except for fecundity, which tends to be more difficult to distinguish from sexual selection unless fitness costs are very high.



**Figure S2.** Accuracy of parameter MLEs of our method and sensitivity to detect selection for the autosomal model. (A) shows results for the inference models where male and female selection parameters can still vary freely, while (B) shows the results for inference models where male and females parameters are set to be equal. The heatmaps show the average Euclidean distance between pairs of parameter MLEs and the true values of the simulations over the whole parameter space. 1000 simulations were run for each grid point. Lines show the threshold below which the parameter can be distinguished from 1 with 95% confidence in 95% of simulations (by a likelihood ratio test). Dashed lines show the horizontal axis parameter and dotted lines show the vertical axis parameter.



**Figure S3.** Power of the inference method to distinguish the three different types of selection for the autosomal model (compare with the analogous plots shown in Figure 4 for the X-linked model). The heatmap shows for each grid point the fraction of simulation runs in which the correct inference model yielded the highest maximum likelihood among the three models.



**Figure S4.** Power of the inference method to distinguish different types of selection when explicitly testing between two inference models (compare with the analogous plots shown in Figure S1 for the X-linked model).