SUPPLEMENTARY MATERIAL

1. Testing mutation-selection balance using measures on a log scale, with multiplicative effects of mutations.

Assuming deleterious mutations reduce the value of a life history trait in a multiplicative fashion, the value of trait z in a given MA line will be

$$z = k\prod_{i}^{L}(1 - X_i a_i)$$

where k is the ancestral trait value, L is the number of loci at which mutations could occur, X_i is an indicator of the presence (1) or absence (0) of a mutation at locus i, and a_i is the heterozygous effect of a mutation at locus i on the trait.

It is useful to consider traits on the log scale, and we define

$$y = \log(z) = \log(k) + \sum_{i}^{L}\log\,(1 - X_i a_i)$$

Note that using a Taylor series approximation shows that when x is small,

$$\log(1 - x) \approx -x - 0.5x^2$$

and we can write

$$y \approx \log(k) - \sum_{i}^{L} X_i a_i - 0.5\sum_{i}^{L} X_i^2 a_i^2$$

Assuming no covariance between the mutation rate and mutational effect at a given locus, the expected log trait value is

$$E[y] \approx \log(k) - \sum_{i}^{L}\overline{X_i}a_i - 0.5\sum_{i}^{L} X_i^2 a_i^2$$

$$= \log(k) - \sum_{i}^{L}\overline{X_i}a_i - 0.5\sum_{i}^{L}(V[X_i] + \overline{X}_i^2])a_i^2$$

For the purposes of estimating $E[y]$ we ignore terms in a^2, which are small. Noting that $E[X]$ summed over all loci can be written Ut, where U is the total rate of mutation across all loci and t is the number of generations of MA, we have

$$E[y] \approx \log(k) - UtE[a]$$

When the trait value of controls represents the ancestral trait value (which is likely not the case in our study, but applies to systems where static controls are available), the rate of change in the trait value per generation, ΔM, can be estimated as

$$\Delta M = (E[y_{control}] - E[y_{MA}])/t$$

An approximation for the control mean in our experiment can be found below. To find the variance in trait values we need approximations for $E[y^2]$ and $E[y]^2$.

$$E[y^2] = E\left[\left(\log(k) + \sum_i^L \log(1 - X_i a_i)\right)^2\right]$$

$$= E\left[\log(k)^2 + 2\log(k)\sum_i^L \log(1 - X_i a_i) + \left(\sum_i^L \log(1 - X_i a_i)\right)^2\right]$$

$$= E\left[\log(k)^2 - 2\log(k)\sum_i^L X_i a_i - \log(k)\sum_i^L X_i^2 a_i^2 + \sum_i^L X_i^2 a_i^2 + 2\sum_i^{L-1}\sum_{j=i+1}^L X_i X_j a_i a_j + O(a^3)\right]$$

$$= \log(k)^2 - 2\log(k)\sum_i^L \overline{X_i} a_i - \log(k)\sum_i^L \left(V[X_i] + \overline{X_i}^2\right) a_i^2 + \sum_i^L \left(V[X_i] + \overline{X_i}^2\right) a_i^2 + 2\sum_i^{L-1}\sum_{j=i+1}^L \overline{X_i X_j} a_i a_j + O(a^3)$$

$$E[y]^2 = \log(k)^2 - 2\log(k)\sum_i^L \overline{X_i} a_i - \log(k)\sum_i^L \left(V[X_i] + \overline{X_i}^2\right) a_i^2 + \left(\sum_i^L \overline{X_i} a_i\right)^2 + O(a^3)$$

$$V[y] = E[y^2] - E[y]^2$$

$$= \sum_i^L \left(V[X_i] + \overline{X_i}^2\right) a_i^2 + 2\sum_i^{L-1}\sum_{j=i+1}^L \overline{X_i X_j} a_i a_j - \left(\sum_i^L \overline{X_i} a_i\right)^2 + O(a^3)$$

$$= \sum_i^L \left(V[X_i] + \overline{X_i}^2\right) a_i^2 + 2\sum_i^{L-1}\sum_{j=i+1}^L \overline{X_i X_j} a_i a_j - \sum_i^L \overline{X_i}^2 a_i^2 - 2\sum_i^{L-1}\sum_{j=i+1}^L \overline{X_i}\,\overline{X_j} a_i a_j + O(a^3)$$

$$= \sum_i^L \left(V[X_i] + \overline{X_i}^2\right) a_i^2 - \sum_i^L \overline{X_i}^2 a_i^2 + 2\sum_i^{L-1}\sum_{j=i+1}^L a_i a_j C(X_i, X_j) + O(a^3)$$

Assuming loci mutate independently, $C(X_i, X_j) = 0$, leaving

$$V[y] \approx \sum_i^L V[X_i]a_i^2$$

Assuming mutations arise at random among lines, i.e., following a Poisson distribution where $V[X] = E[X]$, we can write

$$V[y_{MA}] = UtE[a^2]$$

The change in trait variance per generation can therefore be estimated as

$$\Delta V = V[y_{MA}]/t$$

When the variance in a standing population at equilibrium is due only to segregating deleterious mutations, the above approach can be applied analogously, where $V[X] = 2pq$, giving

$$V[y_{standing}] = \sum_i^L 2p_i q_i a_i^2$$

$$\approx 2\sum_i^L \frac{\mu_i a_i^2}{h_i s_i}$$

Because $a = chs$, we can write

$$V[y_{standing}] \approx 2\sum_i^L \mu_i a_i c_i = UE[ac]$$

This relationship can also be found by considering standing variance in the relative untransformed trait value z

$$\frac{V[z_{standing}]}{\left(E[z_{standing}]\right)^2} \approx UE[ac]$$

and noting that $V[\log(x)] \approx V[x]/(E[x])^2$ in a Taylor-series approximation.

2. Stocks and crosses

The outbred lab population was collected in 1970 in Dahomey (now Benin) West Africa, and maintained in the current lab for more than >3 years (>75 generations) before this experiment in a population of several thousand adults, with overlapping generations. All flies were maintained under standard conditions (25°C, 70% RH, 12 h light). Visible markers and balancer chromosomes , which suppress recombination on the homologous chromosome, were obtained from the Bloomington Drosophila

Stock Center (Bloomington, IN). Except where noted, all crosses took place in 37 mL vials containing 7 mL of yeast-sugar-agar food seeded with live yeast, using virgin females where appropriate. All flies used to initiate male and female fitness assays were virgins. Experimental flies were typically 2-6 days post eclosion at the time of crossing. For both mutational decline and standing variance, all traits were measured for heterozygous focal second chromosomes situated on a common isogenic background derived from the outbred Dahomey population using standard balancer chromosome techniques. Additional stocks of this isogenic background with isogenic markers and balancer chromosomes were also created as required.

3. Trait measurements

(a) Viability

In each assay replicate, two males heterozygous for the focal chromosome and an isogenic chromosome bearing the dominant phenotypic marker *L* were crossed to two females heterozygous for an isogenic standard second chromosome and an isogenic balancer chromosome, *CyO*. After 4 d these adults were removed and offspring were scored up to 15 d following vial initiation. We scored the number of offspring bearing the focal chromosome (heterozygous with the standard isogenic chromosome) compared with the number of *L*/*CyO* offspring (Figure S2C).

(b) Male mating success

Male mating success was measured under male-biased sex ratio conditions in competition with males homozygous for a ubiquitously-expressed phenotypically dominant red fluorescent protein allele (*DsRed*). In each assay replicate 5 virgin focal males (heterozygous for the focal chromosome), 5 virgin competitor males, and 6 virgin isogenic females were allowed to interact for 3 ± 0.25 h (Figure S2D) providing sufficient time for mating to take place while limiting the opportunity for multiple mating (Manning 1967). The males were then discarded; 1-2 d later the females were placed individually into 16 mL oviposition tubes containing ~2 mL of food without live yeast. After a further 8-10 d, each group of six oviposition tubes was examined to determine the number of females that produced offspring, and

examined under fluorescent light to determine the number of females who produced offspring sired by *DsRed* competitor males (*DsRed* is visible in larvae, pupae, and adults). There was no evidence of multiple mating (fluorescent and non-fluorescent offspring in a single tube) except in one case, which was removed from the dataset. Some females either died or did not produce any offspring; on average, >5 females produced offspring per replicate.

(c) Female fecundity

Virgin focal females (heterozygous for the focal chromosome) and outbred virgin brown-eyed females (*bw*/*bw*) were held in individual vials with food but without live yeast for 4 d. In each assay replicate, a single focal female was then placed in a vial with a single *bw*/*bw* female and two isogenic males (Figure S2E). Each vial was supplemented with ~10 µl of 0.02 g/mL live yeast in solution. This amount of yeast can be consumed by a single female in < 24 h [25], and therefore represents a "limited" resource for females in this context, where two females are present. After 24 ± 0.5 h focal females were placed in individual oviposition vials without live yeast, containing standard media with added food coloring to facilitate egg counting. Females were allowed to oviposit for 18 ± 0.5 h, after which they were discarded, and the number of eggs was scored. Note that the adult fitness components we measured do not involve the number of offspring produced, and are therefore independent of larval viability.

3. Maximum likelihood model

This model was implemented on the log scale, where $y = \log(z)$. Building from the material presented in the main text, we first describe how observed means and variances are related to underlying parameters.

Genetic means and variances of MA lines:

$$E[y_{v,MA}] = \log(k_v) - 52\ U\ E[a_v]$$

$$E[y_{f,MA}] = \log(k_f) - 52\ U\ E[a_f]$$

$$E[y_{m,MA}] = \log(k_m) - 52\ U\ E[a_m]$$

$$V[y_{v,MA}] = 52U(E[a_v]^2 + V[a_v])$$
$$V[y_{f,MA}] = 52U(E[a_f]^2 + V[a_f])$$
$$V[y_{m,MA}] = 52U(E[a_m]^2 + V[a_m])$$

Mean and variance of selection on heterozygous mutations:

$$E[hs] = E[a_v] + 0.5E[a_f] + 0.5E[a_m]$$

$$V[hs] = V[a_v] + 0.25V[a_f] + 0.25V[a_m] + \sqrt{V[a_v]V[a_f]} + \sqrt{V[a_v]V[a_m]} + 0.5\sqrt{V[a_m]V[a_f]}$$

$$E\left[\frac{1}{hs}\right] \approx \frac{1}{E[hs]} + \frac{V[hs]}{E[hs]^3}$$

Trait means of the control lines for the MA experiment:

$$E[y_{v,control}] \approx \log(k_v) - \sum_j E[X_j]a_{v,j}$$
$$E[y_{f,control}] \approx \log(k_f) - \sum_j E[X_j]a_{f,j}$$
$$E[y_{m,control}] \approx \log(k_m) - \sum_j E[X_j]a_{m,j}$$

where $E[X_j]$ is the expected number of alleles on a control chromosome in deleterious size class j (defining the original copy of chromosome 2 used in this experiment as wild-type at all sites).

Ideally, the control would be mutation-free (i.e., an exact replica of the original copy of chromosome 2, so that $E[X_j] = 0$ for all j). In reality, deleterious alleles will segregate at low frequencies within the control populations so $E[X_j] > 0$. Under a deterministic model, the average number of mutations of size class j in the controls after t (=52) generations is

$$E[X_j] \approx \sum_{g=0}^{t} Uf_j(1 - h_j s_j)^g$$

where f_j is the fraction of all new mutations that are in class j. We assume mutational fitness effects follow a gamma distribution (Keightley 1994) with mean $E[hs]$ and

variance $V[hs]$ (shape $E[hs]^2/V[hs]$ and scale $V[hs]/E[hs]$), and divide the distribution into 100 discrete size classes, which we use to assign values to f_j and $h_j s_j$. We assume the effect of a mutation on total fitness can be approximated by its effect on the fitness components we studied, such that

$\mathrm{E}[hs] = E[a_v] + \mathrm{E}[a_f]/2 + \mathrm{E}[a_m]/2$ and

$$V[hs] = V[a_v] + \frac{1}{4}V[a_m] + \frac{1}{4}V[a_f] + Cov[a_v, a_m] + Cov[a_v, a_f] + \frac{1}{2}Cov[a_m, a_f]$$

We assume the traits are perfectly positively correlated such that $Cov[a_1, a_2] = SD[a_1]SD[a_2]$ because this is conservative with respect to the test for excess genetic variance.

The remaining parameters are not of primary interest, and are free to vary, namely block effects on male mating success (2 parameters), control variances (3 parameters), and overdispersion for female fecundity and larval viability (4 parameters). The 19 total parameters are listed in Table S3.

There are six 'sets' of data (control and MA lines for each of the three traits). For each set we evaluated the log-likelihood of the data for given values of the mean, genetic variance, and overdispersion using the *lme4* generalized linear mixed model (*glmm*) deviance evaluation function for that subset. For male sets there is also a block effect on the mean, and overdispersion is absent. The overall log-likelihood is then the sum of the log-likelihoods across the six sets. After optimization we obtained the most likely ancestral control value, $log(k)$, for each trait.

Table S1. Means and genetic variances for each trait on the original scale of measurement, determined by integration of the inverse link function using means and variances of each trait on the link scale.

Trait	Group	Mean (SE) [95% CI†]	Genetic variance (SE) [95% CI†]
Male mating success	Mutant	1.598 [1.412, 1.825]	0.142 [0.006, 0.445]
	Control	1.974 [1.738, 2.276]	0.293 [0.023, 0.866]
	Standing	2.040 [1.886, 2.219]	0.190 [0.018, 0.469]
Female fecundity	Mutant	28.157 [26.575, 29.885]	18.245 [5.386, 37.533]
	Control	30.008 [28.383, 31.775]	25.795 [11.307, 47.831]
	Standing	53.094 [51.394, 54.871]	22.660 [2.452, 51.151]
Viability	Mutant	2.745 [2.558, 2.956]	0.133 [0.007, 0.389]
	Control	3.341 [3.058, 3.675]	0.405 [0.039, 1.056]
	Standing	1.759 [1.712, 1.809]	0.014 [0.001, 0.034]

†Bayesian credible interval.

Table S2. Summary of quantitative genetic estimates for each trait, for haploid second chromosomes

Trait	Metric	Estimate × 10^3 [95% CI†]
Male mating success	ΔM	3.897 [0.538, 7.216]
	ΔV	0.988 [0.050, 2.740]
	σ^2	43.507 [4.570, 98.051]
Female fecundity	ΔM	1.172 [–0.353, 2.724]
	ΔV	0.435 [0.133, 0.863]
	σ^2	7.976 [0.879, 17.723]
Larval viability	ΔM	3.598 [1.433, 5.784]
	ΔV	0.332 [0.018, 0.933]
	σ^2	4.381 [0.277, 11.057]

†Bayesian credible interval.

Table S3. Maximum likelihood model estimates.

Parameter	Estimate
$\log(k_m)$	0.712
$\log(k_f)$	3.463
$\log(k_v)$	1.197
$E[a_m]$	0.069
$E[a_f]$	0.034
$E[a_v]$	0.049
$V[a_m]$	0.616×10^{-3}
$V[a_f]$	3.370×10^{-3}
$V[a_v]$	0.226×10^{-6}
U	0.077
$V[y_{m,\,control}]$	0.043
$V[y_{f,\,control}]$	0.024
$V[y_{v,\,control}]$	0.024
$\text{block}_{m,\,MA}$	0.107
$\text{block}_{m,\,control}$	0.151
$\text{error}_{f,\,MA}$	0.092
$\text{error}_{f,\,control}$	0.073
$\text{error}_{v,\,MA}$	0.118
$\text{error}_{v,\,control}$	0.205

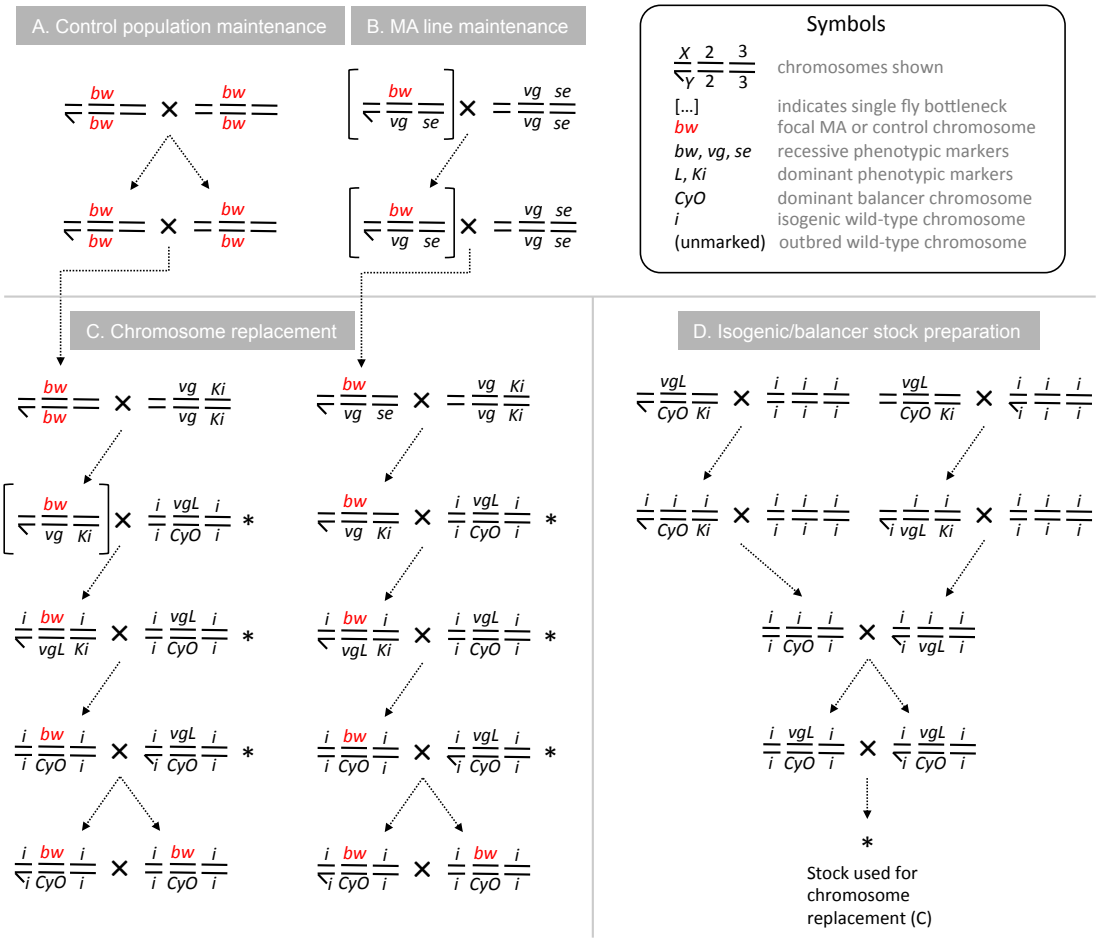

Figure S1. Details of mutation accumulation and control line maintenance and preparation. The first three chromosomes are shown for each genotype; the tiny fourth chromosome was not manipulated. Males generally lack recombination, and are identified here by the presence of a *Y* chromosome. Crosses took place using virgin females where appropriate. Chromosomes were identified using recessive phenotypic markers (*bw*, *vg*, *se*), dominant phenotypic markers (*L*, *Ki*), and a balancer chromosome (*CyO*), which suppresses recombination on the second chromosome. For mutation accumulation (MA), a single second chromosome marked with *bw* was used to initiate three control populations and numerous MA lines. These focal chromosomes are shown in red. (A) Control populations homozygous for the focal chromosome were maintained at a moderate size (450 adults) to prevent mutation accumulation. (B) MA chromosomes were propagated by bottlenecking to a single heterozygous male each generation, allowing new mutations to accumulate. (C) Following 52 generations of MA, crosses were performed to replace all non-focal chromosomes with an isogenic background. Within-line variation on the focal chromosome was eliminated by bottlenecking (square brackets). Each cross included 1-4 males and 4 females per line. These crosses involved several marker stocks, which were created using standard crossing methods (not shown). An isogenic stock with *vgL*/*CyO* was created as shown in (D), after creating a completely isogenic genotype using standard balancer chromosome methods (not shown).

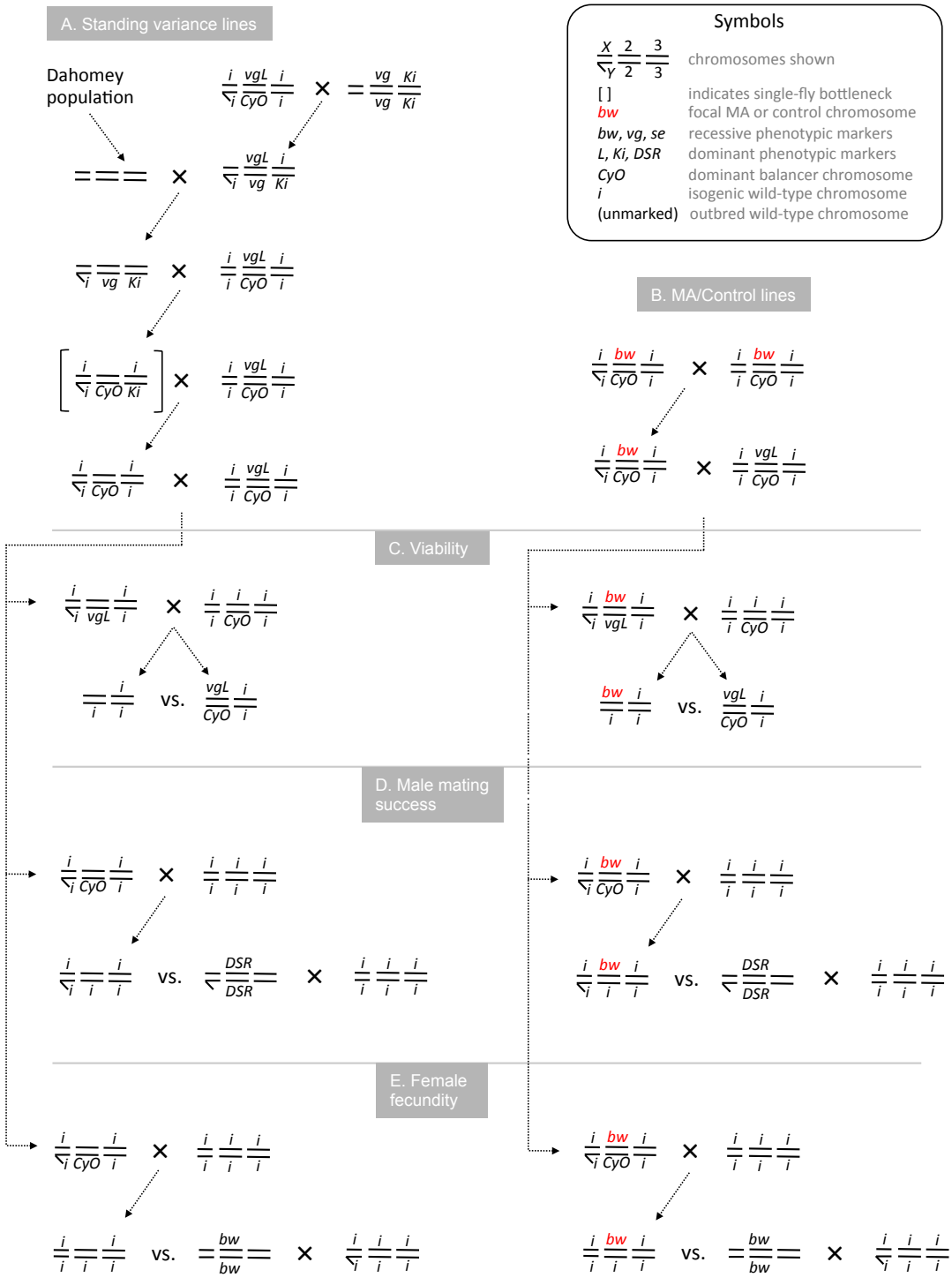

Figure S2. Details of standing variance line preparation and assay crosses for all lines. The first three chromosomes are shown for each genotype; the tiny fourth chromosome was not manipulated. Males lack recombination, and are identified here by the presence of a *Y* chromosome. Crosses took place using virgin females where appropriate. Chromosomes were identified using recessive phenotypic markers (*bw*, *vg*, *se*),

dominant phenotypic markers (*L*, *Ki*, *DSR*), and a balancer chromosome (*CyO*), which suppresses recombination on the second chromosome. After obtaining focal second chromosomes from either the outbred Dahomey laboratory population (A) or MA lines and control lines (B; generated as shown in Figure S1), crosses were performed to assess viability (C), male mating success (D), and female fecundity (E), as described in the text.