Supplementary Information

Materials and Methods:

MDS plot of all samples

Fig. S1: Multidimensional scaling plot of all 89 samples (one sample was removed due to low quality of the RNAseq library).

Multidimensional scaling (MDS) plots were constructed for unsupervised clustering of samples. MDS plots cluster samples based on similarity, in this case, based on pairwise $log₂$ fold changes. Two samples with larger differences in gene expression (larger absolute fold changes), are further apart in the plot. Samples that have more similar gene expression patterns (smaller absolute fold changes), cluster more closely together.

This plot demonstrates that most samples cluster by female genotype (blue: Beijing, black: Tasmania, red: Netherlands, green: Ithaca, grey: Zimbabwe). Six samples do not cluster by female genotype (I x Z-1, N x B-2, I x I-3, I x T-3, B x Z-1 and I x N-1). Further analysis demonstrated that these six samples contained

male-specific mRNAs (among which were transcripts from male accessory gland-specific genes). Based on genotype-specific SNPs, we found that these samples were probably contaminated with male mRNA (data not shown). When analyzing SNPs, evidence for male contamination was found in I x Z-3 as well (data not shown). We decided to remove I x Z-1, I x Z-3, N x B-2, I x I-3, I x T-3, B x Z- 1 and I x N-1 from the dataset.

MDS plot - 7 samples removed

Fig. S2: Multidimensional scaling plot after removal of seven samples.

Unsupervised clustering after removal of seven outliers demonstrates that samples cluster by female genotype. One sample, N x I-2, does not cluster with other Netherlands samples. Based on genotypespecific SNPs, we found that N x I-2 likely contained a mix of Netherlands and Ithaca females (data not shown). This sample was also removed from our analysis.

Fig. S3: Multidimensional scaling plot after removal of eight samples.

Unsupervised clustering after the removal of eight samples demonstrates clustering of samples by female genotype.

Fig. S4: Multidimensional scaling plots per female genotype.

Unsupervised clustering per female genotype identified additional outlier samples: three samples containing Beijing females clustered away from other Beijing samples. (B x I-1, B x I-2 and B x N-2). We further investigated these samples using MA plots (fig. 5).

In these MDS plots, samples were colored by replicate (replicate 1: red, replicate 2: blue, replicate 3: black). Samples tended to separate by replicate, indicating the presence of batch effects.

Fig. S5: MA plots were constructed to compare fold change differences between biological replicates for B x I and B x N combinations.

In these plots, each dot represents a gene. The y-axis (M) represents the log_2 fold change between two replicates. The x-axis (A) represents the average expression level of each gene in both replicates. Genes with a low average expression level usually have a larger inter-replicate variability (larger fold changes). When comparing biological replicates, we expect to see small fold changes (less than 2-fold) for genes with a high average expression. In the figures above, red dots represent genes with a ≥ 2-fold change.

We found that the expression of over 2000 genes differed 2-fold or more in B x I-2, relative to B x I-1 and B x I-3. On the other hand, the expression of only 418 genes differed 2-fold or more in B x I-1 relative to B x I-3.

Similarly, the expression of over 1000 genes differed 2-fold or more in B x N-2, relative to B x N-1 and B x N-3, while the expression of only 212 genes differed 2-fold or more in B x N-1 relative to B x N-3.

Based on this information, we decided to remove B x I-2 and B x N-2 from the dataset.

We decided to remove lowly expressed genes from our dataset, because genes with low cpm values (count per million) generally have a larger inter-replicate variability. The higher noise/signal ratio makes statistical inference more difficult. In addition, limiting the number of genes in the dataset also limits the number of tests that need to be performed during differential expression analysis (see edgeR user guide a[t www.bioconductor.org,](http://www.bioconductor.org/) and Law et al. 2016).

To set a filtering cutoff for lowly expressed genes, we examined the distribution of cpm values across all genes and samples in our dataset. In plots A and B, each line represents one of our 79 samples. Plot A shows that a large set of genes (high density on y-axis) is expressed at low cpm values (x-axis). After

filtering, using a cutoff of cpm> 3, the majority of genes in the dataset is expressed at higher cpm values. The code to generate these plots was used from Law et al. 2016.

Plot C is a representative MA plot which shows that at cpm= 3 (A= $log₂(3)= 1.6$), inter-replicate variability in gene expression decreases.

ď Q	B		N	Τ	Z	
B	BxB	Bxl	BxN	BxT	BxZ	B
	IxB	lxl	IxN	l xT		
N	NxB	Nxl	NxN	NxT	NxZ	N
т	TxB	Txl	TxN	TxT	TxZ	т
Z	ZxB	ZxI	ZxN	ZxT	ZxZ	Z

Question 1: Which genes respond to mating, regardless of female or male genotype?

Question 2: Which genes respond differently to mating in females from a particular line mated to males from a particular line?

ď ò	B		N		Z	
B	BxB	Bxl	BxN	BxT	BxZ	B
	IxB	x	IxN	l xT		
N	NxB	Nxl	NxN	NxT	NxZ	Ν
т	TxB	TxI	TxN	TxT	TxZ	т
Z	ZxB	ZxI	ZxN	ZxT	ZxZ	Z

Question 3: Which genes respond differently to mating in females from a particular line, regardless of male genotype?

ď Q	B		N	Τ	Z	
B	BxB	Bxl	BxN	BxT	BxZ	B
	IxB	lxl	IxN	l xT		
Ν	NxB	Nxl	NxN	NxT	NxZ	N
T	TxB	TxI	TxN	TxT	TxZ	T
Ζ	ZxB	Zxl	ZxN	ZxT	ZxZ	Z

Question 4: Which genes respond differently to mating in females mated to a male from a particular line, regardless of female genotype?

Fig. S7: Overview of the four questions that were asked regarding the RNAseq dataset, and the general approach used to answer these questions.

To answer four different questions regarding the post-mating transcriptional response, four differential expression analyses were conducted, each with its own linear model. To calculate the transcriptional response to mating, we compared gene expression between mated females (dark green or red boxes) and resp. virgin females (light green or red boxes). For questions 2, 3 and 4, we compared the transcriptional response to mating in the genotype of interest (green) with the average response to mating across all genotypes (red). Questions 3 and 4 were investigated for the Beijing genotype, as shown in the figure, but were also repeated for the remaining 4 genotypes. Question 2 was investigated for females from the Beijing line mated to a male from the Beijing line, as shown in the figure, but was repeated for the remaining 23 combinations. Samples for the I x Z combination were excluded from our analysis.

Fig. S8: Quantile quantile-plots of p-values from differential expression tests conducted using edgeR. Shown here are qq-plots for question 1 (Which genes respond to mating regardless of female or male genotype?), question 3 (Which genes respond to mating in a female genotype-dependent manner?) and question 4 (Which genes respond to mating in a male genotype-dependent manner?).

The solid diagonal represents the distribution of p-values under the null hypothesis (no differentially expressed genes). Circles represent observed p-values for each gene in our dataset. Circles that lie above the diagonal have smaller p-values compared to what is expected under the null hypothesis.

Fig. S9: Quantile quantile-plots of p-values from differential expression tests conducted using edgeR. Shown here are qq-plots for question 2 (Which genes respond to mating in a female x male genotype interaction-dependent manner?), for the B and I female genotypes.

The solid diagonal represents the distribution of p-values under the null hypothesis (no differentially expressed genes). Circles represent observed p-values for each gene in our dataset. Circles that lie above the diagonal have smaller p-values compared to what is expected under the null hypothesis.

The solid diagonal represents the distribution of p-values under the null hypothesis (no differentially expressed genes). Circles represent observed p-values for each gene in our dataset. Circles that lie above the diagonal have smaller p-values compared to what is expected under the null hypothesis.

Fig. S11: Quantile quantile-plots of p-values from differential expression tests conducted using edgeR. Shown here are qq-plots for question 2 (Which genes respond to mating in a female x male genotype interaction-dependent manner?), for the Z female genotype. The solid diagonal represents the distribution of p-values under the null hypothesis (no differentially expressed genes). Circles represent observed pvalues for each gene in our dataset. Circles that lie above the diagonal have smaller p-values compared to what is expected under the null hypothesis.

Table S1: Post-mating log₂ fold changes of *Rp49* in each of the 24 mating combinations, based on the RNAseq dataset. Q-values are shown in parentheses. In all combinations, *Rp49* undergoes fold changes after mating that are smaller than 2-fold, and none of these fold changes are significant when comparing mated and virgin females. These data demonstrate that *Rp49* transcript levels do not change after mating, making *Rp49* a suitable gene for qPCR normalization.

Table S2: qPCR primer sequences.

Fig. S12: Overview of the design used to collect flies for immune gene qRT-PCR.

Two bottles containing flies with the female genotype of interest were set up (female bottle A and female bottle B), together with two bottles containing flies with the male genotype of interest (male bottle A and male bottle B). From each bottle, we collected two biological replicates on two different days, giving us a total of four replicates: (1) Females from bottle A, replicate 1 were mated to males from bottle A, replicate 1. (2) Females from bottle B, replicate 1 were mated to males from bottle B, replicate 1. These matings were performed on the same day. Two days later, (3) females from bottle A, replicate 2 were mated to males from bottle A, replicate 2. (4) Females from bottle B, replicate 2 were mated to males from bottle B, replicate 2. For each of the four replicates, virgin females were sampled as well.

Keeping track of which bottle the flies were collected from allowed us to take variation in the microenvironment of each bottle into account. For example, if virgin females from bottle A were found to have higher than expected expression of antimicrobial peptides, this suggests that pathogens might have been present in bottle A. In this case, expression levels of antimicrobial peptides in mated females from bottle A might have been influenced by other factors (pathogens) besides mating.

RNA-seq analysis – alternative method:

A 5 by 5 factorial ANOVA was used as a different method to address the roles of female and male genotype on post-mating gene expression changes. In particular, an ANOVA was used to set up a combined model to test the overall contribution to variance in post-mating fold changes due to female effects, male effects and their interaction (questions 3, 4 and 2). We were unable to obtain adequate fits of this model to the data, and qq-plots displayed significant inflation of p-values. Despite this caveat, the trend of genes showing significant expression differences was essentially a subset of those found by edgeR (results not shown).

Effect of *Wolbachia* on post-mating gene expression changes

Of the five Global Diversity Lines used in this study, four lines carried the bacterial endosymbiont *Wolbachia pipientis* (only the Netherlands line is uninfected). This raised the concern that observed female x male genotypic effects on post-mating transcriptional changes might have been caused by the presence or absence of *Wolbachia*, rather than true genotype interaction effects.

To control for *Wolbachia* effects, we divided the 24 female x male combinations into four groups, based on the presence or absence of *Wolbachia* in female and male (table 1). We used edgeR to find genes that were differentially expressed after mating in each of the four groups. We set up four contrasts to compare mated females of each group with their respective virgin females (fig. 1).

No genes were found to be differentially expressed after mating in females that were not infected with *Wolbachia* (Netherlands females).

We found 170 genes that were differentially expressed after mating if both female and male carried *Wolbachia* (q< 0.05; this includes all crosses except crosses with Netherlands flies). Of these 170 genes, one gene was among the genes that were found to be involved in female x male mating interactions:

- Ect3 was down-regulated in IxB (log² fold change= -1.2). It was also down-regulated in *Wolbachia* infected females mated to *Wolbachia* infected males, but this downregulation was not as strong as in IxB (average log_2 fold change= -0.4). This suggests that the strong down-regulation of *Ect3* in IxB is due to the I and B genotypes, and not solely due to the presence of *Wolbachia*.

In addition to this, we found 200 genes that were differentially expressed after mating a *Wolbachia* infected female with an uninfected male (q< 0.05). Of these 200 genes, only one gene was also found among the genes involved in female x male genotype interactions (IxN). *Obp49a* was up-regulated in *Wolbachia* infected females mated to uninfected males (log₂ fold change= 1, this includes BxN, IxN, TxN and ZxN). However, *Obp49a* was up-regulated more strongly in IxN

($log₂$ fold change = 1.9). This suggests that the stronger up-regulation observed in IxN is mediated by genotype interactions, rather than the presence of *Wolbachia*.

In short, the variation in presence/absence of *Wolbachia* across the five *Drosophila* lines used in this study did not appear to generate false positive calls for genes whose expression levels are robustly impacted by mating.

Table 2: The 24 female x male mating combinations can be subdivided into four groups, based on the presence or absence of *Wolbachia*. In parentheses are the numbers of genes that are DE after mating in each of the four groups.

Fig. 1: QQ-plots of p-values for differential expression tests conducted using edgeR. The solid diagonal represents the distribution of p-values under the null hypothesis (no DE genes). Circles represent observed p-values for each gene in our dataset. Circles that lie above the diagonal have smaller p-values compared to what is expected under the null hypothesis. Shown here are qq-plots for all four mating combinations of *Wolbachia* infected and uninfected females and males. Contrasts were set up in edgeR to find genes that were differentially regulated post-mating, for each of the four groups.

Fig. S13: Diagnostic plots to check assumptions of the lmer model used to analyze fecundity data. A: Residuals plot for homogeneity of variance. B: Residuals follow a normal distribution.

Fig. S14: Diagnostic plots to check assumptions of the lmer model used to analyze hatchability data. A: Residuals plot for homogeneity of variance. B: Residuals follow a normal distribution. When using a GLM with binomial error distribution, instead of lmer with normal distribution, our data was over-dispersed. Adding an observation level random effect (Harrison 2014) did not sufficiently reduce the over-dispersion. An alternative was to use a quasi-binomial model, which accounts for over-dispersion, but this model did not allow us to add random effects (which meant we would have to analyze each block separately). Because of this, and because the diagnostic plots for the normal lmer model looked good, we decided to use lmer instead.

Fig. S15: Diagnostic plots to check assumptions of the lmer model used to analyze receptivity data.

 $20 \right.$ A: Residuals plot for homogeneity of variance, for female refractoriness at day 1. B: Residuals follow a normal distribution for female refractoriness at day 1. C: Residuals plot for homogeneity of variance, for female refractoriness at day 4. D: Residuals follow a normal distribution for female refractoriness at day 4. To analyze female refractoriness to re-mating, we added three random effects to the lmer models: (1) block, (2) the interaction between block and female genotype, and (3) the interaction between block and male genotype. Adding these random interaction terms to the model did not affect the results for refractoriness to re-mating on day 4. However, he results changed for refractoriness to re-mating on day 1. A model with "block" as the only random effect, found significant effects of female and male genotype on re-mating rate on day 1. Adding additional random interaction terms removed these significant effects. These random interaction terms consider how the overall effect of female genotype, or the overall effect of male genotype, changes across blocks. We did no longer find significant effects of female or male genotype due to large variability between the blocks for 1 day refractoriness to re-mating.

Results:

Table S3: Permutation tests were performed to determine the likelihood to find the observed number of differentially expressed (DE) genes by chance. This table contains the observed number of differentially expressed genes in the original RNAseq dataset for each of the 24 combinations. The proportion of how often this number, or a larger number, of differentially expressed genes was observed in any of the 500 randomized datasets is shown in parentheses. This proportion can be considered as a permutation pvalue. All permutation p-values were < 0.05. Intra-population crosses are highlighted in yellow.

Fig. S16: Six genes for which the post-mating fold changes were impacted by female x male genotype interactions. Boxplots represent normalized read counts (cpm) across all replicates, for virgin and mated females of all genotypes. A: *Obp49a* transcript levels were up-regulated higher than average in I x N (q= 0.015). B: *Snmp1* transcript levels were down-regulated more than average in T x B (q= 0.048). C: *Def* transcript levels were up-regulated more than average in T x T (q= 0.19) and T x Z (q= 6 x 10⁻⁶). D: *Cyp4p2* mRNA levels were down-regulated more in B x N (q= 1.5 x 10⁻⁵). E: *AttB* transcripts were up-regulated more than average in B x I (q= 0.041). F: *CG32277* mRNA was down-regulated more than average in $B \times B$ (q= 0.006).

Fig. S17: Post-mating $log₂$ fold changes for immune transcripts in all 24 female x male mating combinations. A significant upregulation of immune gene transcripts was seen in B x I (*AttB*, *Dro*, *Drs*, *PGRP-SC2*, *edin*, *Tep2*), T x Z (*Tep1*, *Def*) and T x T (*Def*, *Spn28Dc*). A significant down-regulation of *CG11313* was observed in T x I, while *Jon65Aii* was significantly down-regulated in T x B, T x I, T x N and T x T.

Table S4: Permutation tests were performed to determine the likelihood to find the observed number of differentially expressed (DE) genes by chance. This table contains the observed number of differentially expressed genes in the original RNAseq dataset for each of the five female genotypes. The proportion of how often this number, or a larger number, of differentially expressed genes was observed in any of the 500 randomized datasets is shown in parentheses. Of the 35 differentially expressed genes indicated below, eleven genes were also differentially expressed for the female x male interaction effects, leaving 24 genes differentially expressed depending on female genotype only.

Table S5: Permutation tests were performed to determine the likelihood to find the observed number of differentially expressed (DE) genes by chance. This table contains the observed number of differentially expressed genes in the original RNAseq dataset for each of the five male genotypes. The proportion of how often this number, or a larger number, of differentially expressed genes was observed in any of the 500 randomized datasets is shown in parentheses. Of the seven differentially expressed genes indicated below, five genes were also differentially expressed for the female x male interaction effects, leaving two genes differentially expressed depending on male genotype only: *CG16743* in females mated to Zimbabwe males, and *Diedel* in females mated to Tasmania males.

CG16743 transcripts were on average up-regulated after mating, confirming findings from earlier studies (McGraw et al. 2004; McGraw, Clark, and Wolfner 2008; Zhou, Mackay, and Anholt 2014; Hollis, Houle, and Kawecki 2016). However, in females that mated to a male from the Zimbabwe line, the mRNA levels of *CG16743* did not change. The post-mating up-regulation of *Diedel* transcripts was found to be significantly higher than average if a female mated to a male from the Tasmania line. Yet, permutation tests showed that the number of differentially regulated genes depending on male genotype, was not significantly different from the number of differentially regulated genes found by chance.

Fig. S18: Normalized read counts (cpm, counts per million) for virgin females of each genotype (each boxplot represents data from three biological replicates). Virgin females from the Tasmania and Zimbabwe lines had higher cpm values for a range of antimicrobial peptides.

Table S6: Summary statistics for qRT-PCR results. A t-test was performed to determine if post-mating fold changes (gene expression in mated females relative to gene expression in virgin females) differed significantly between female x male combinations. After removing outlier replicates, significant differences (p< 0.05) were found for all genes, except for *Dro*. For *Dro*, the trends did confirm the pattern observed in the RNAseq dataset. (SE=standard error)

Fig. S19: QRT-PCR results for post-mating gene expression changes in six genes.

Variation in post-mating gene expression changes was validated for six genes using qRT-PCR. For each of the six genes, two female x male combinations were compared, that showed the largest difference in postmating response based on the RNAseq data. Each bar represents one biological replicate (10 females pooled per replicate).

A-B: Based on the RNAseq data, *CG3088* and *Cyp4p2* were significantly down-regulated in B x N relative to the average response to mating in all female x male combinations. We ran qRT-PCR to compare post-mating transcript levels in B x N and B x T (*CG3088*) and B x N and I x T (*Cyp4p2*). QRT-PCR results confirmed the down-regulation of *CG3088* and *Cyp4p2* in the B x N combination. C: QRT-PCR confirmed an up-regulation of *Def* in T x T. D-E: qPCR results confirmed an up-regulation of antimicrobial peptides (AMPs) *AttB* and *Dro* in B x I for 3 out of 4 biological replicates. When investigating Ct values for the aberrant biological replicate, it became clear that the low fold change in this replicate was not due to the absence of AMP induction in mated females, but that it was due to higher AMP expression in virgin females (fig. S20). F: *Obp49a* was upregulated in I x N in the RNAseq data. Only one out of three biological replicates confirmed this up-regulation using qRT-PCR.

BA-1

BA-2

BB-1

B B-2

fig. S20: Ct values for mated and virgin females, for two antimicrobial peptides. Virgin females from replicates B-1 (bottle B, replicate 1) had lower Ct values for *Dro* and *AttB*, indicating an up-regulation of antimicrobial peptides in these virgin females.

Table S7: ANOVA output from the linear mixed effects model used to analyze fecundity data. Fecundity differed depending on interactions between female and male genotype, and depending on the day after mating (p= 7.6 x 10⁻⁶).

approximation for degrees of freedom					
female				5043 1260.7 4 515.87 10.435 3.911e-08 ***	
male		12549 3137.2		4 515.87 25.966 < 2.2e-16 ***	
day	65041 16260.4			4 2070.92 134.582 < 2.2e-16 ***	
female:male		6384 399.0		16 515.88 3.302 1.601e-05 ***	
female:day		60575 3785.9		16 2070.91 31.335 < 2.2e-16 ***	
male:day		22163 1385.2		16 2070.92 11.465 < 2.2e-16 ***	
female:male:day		15357 240.0		64 2070.91 1.986 7.547e-06 ***	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Analysis of Variance Table of type III with Satterthwaite approximation for degrees of freedom

Table S8: Pairwise comparisons were made between all 25 female x male combinations to evaluate differences in mean egg production per day. If egg production differed significantly between two female x male combinations, these combinations were assigned to groups with no numbers or letters in common. Comparisons across days were not made. (Lsmean= least square means calculated based on three replicates, with average n per female x male combination= 32; SE= standard error; df= degrees of f reedom; lower.CL and upper.CL= boundaries for 95% confidence interval)

 $day = 1:$

Degrees-of-freedom method: satterthwaite Confidence level used: 0.95 P value adjustment: tukey method for comparing a family of 25 estimates significance level used: alpha = 0.05

Fig. S21: Egg production over the course of five days differed depending on interactions between female and male genotype (p= 7.6 x 10⁻⁶). Shown here is mean egg production per day, for each of the 25 female x male combinations (mean based on three replicate experiments; average number of females (n) per combination across all replicates= 21.7). Error bars represent standard errors. Generally, egg production decreased over time. A high egg production persisted over the course of five days in Ithaca females. For other females, the decrease in egg production over time differed depending on the genotype of the female's mate. Specifically, males from the Netherlands and Zimbabwe lines were unable to stimulate a high egg production in females from Beijing, the Netherlands, Tasmania and Zimbabwe.

Table S9: ANOVA output from the linear mixed effects model used to analyze hatchability data. Hatchability differed depending on interactions between female and male genotype, and depending on the day after mating ($p= 0.01$).

approximation for degrees of reedom					
		Sum Sq Mean Sq NumDF DenDF F. value $Pr(>=F)$			
female				4.5365 1.1341 4 511.41 36.861 $\overline{2.2e-16}$ ***	
male				1.9811 0.4953 4 512.04 16.097 2.002 e-12 ***	
day				15.9630 3.9908 4 1919.17 129.706 < 2.2e-16 ***	
female:male	1.0755 0.0672			16 511.29 2.185 0.004993 **	
female:day	4.9412 0.3088			16 1915.87 10.037 < 2.2e-16 ***	
male:day 1.6658 0.1041				16 1918.16 3.384 6.060e-06 ***	
female:male:day 2.8701 0.0448				64 1914.02 1.458 0.011303 *	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Analysis of Variance Table of type III with Satterthwaite approximation for degrees of freedom

Table S10: Pairwise comparisons were made between all 25 female x male combinations, to evaluate differences in mean hatchability per day. If hatchability differed significantly between two mating combinations, these combinations were assigned to groups with no numbers or letters in common. Comparisons across days were not made. (Lsmean= least square means calculated based on three replicates, with average n per female x male combination= 32; SE= standard error; df= degrees of freedom; lower.CL and upper.CL= boundaries for 95% confidence interval)

 $day = 1$:

Degrees-of-freedom method: satterthwaite Confidence level used: 0.95 P value adjustment: tukey method for comparing a family of 25 estimates significance level used: alpha = 0.05

Fig. S22: Hatchability (number of pupae/ number of eggs) over the course of five days differed depending on interactions between female and male genotype (p= 0.01). Shown here is mean hatchability per day, for each of the 25 female x male combinations (based on three replicate experiments; average number of females(n) per combination across all replicates= 21.7). Error bars represent standard errors. Hatchability in Zimbabwe females was consistently high. In the four remaining female genotypes, hatchability decreased strongly over time. This decrease was more pronounced females mated to Ithaca or Tasmania males. Egg production was low in females mated to males from the Netherlands or Zimbabwe lines, but hatchability was high in these mating combinations. Low hatchability in intra-population crosses could be due to inbreeding effects. The Netherlands line is the only line not infected with *Wolbachia*. Low hatchability was observed in Netherlands females mated to *Wolbachia*-infected males, likely due to cytoplasmic incompatibility.

Table S11: ANOVA output from the linear mixed effects model used to analyze female refractoriness to re-mating, one day after the first mating. Refractoriness on day 1 after mating did not differ depending on female or male genotype.

Table S12: ANOVA output from the linear mixed effects model used to analyze female refractoriness to re-mating, four days after the first mating. Re-mating rate at day 4 after the first mating differed significantly depending on interactions between female and male genotype.

Analysis of Variance Table of type III with Satterthwaite approximation for degrees of freedom

	Sum Sq Mean Sq NumDF DenDF F.value			$Pr(>=)$	
female	0.83747 0.209368	4		60 10.9928 9.407e-07 ***	
male	1.12139 0.280347	-4		12 14.7196 0.0001409 ***	
female: male 1.26232 0.078895		16		60 4.1424 2.835e-05 ***	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table S13: Pairwise comparisons were made between all 25 female x male combinations, to evaluate differences in female refractoriness to re-mating, four days after the first mating. If two mating combinations differed significantly from each other, these combinations were assigned to groups with no numbers or letters in common. (Lsmean= least square means calculated based on four replicates, with average n per female x male combination= 32; SE= standard error; df= degrees of freedom; lower.CL and upper.CL= boundaries for 95% confidence interval)

Degrees-of-freedom method: satterthwaite

Fig S23: Using a Spearman correlation test, a negative correlation was found between (1) the average number of eggs produced per day, and (2) the proportion of females that re-mated with a standard male, four days after the first mating with a male from the Global Diversity Lines (p= 0.001). Generally, egg production was low, and receptivity to re-mating was high in females mated to males from the Netherlands or Zimbabwe lines. Colors indicate the genotype of the first male a female mated with (B= Beijing, I= Ithaca, N= Netherlands, T=Tasmania, Z= Zimbabwe).

Fig. S24: Using a Spearman correlation test and Benjamini-Hochberg correction for multiple testing, a significant correlation was found between the post-mating fold change of *CG32277* and the average number of eggs produced on day 1 after mating ($q= 0.03$; B= Beijing, I= Ithaca, N= Netherlands, T=Tasmania, Z=Zimbabwe). Beijing females, which produced a high number of eggs on day 1 after mating, underwent a strong post-mating down-regulation of *CG32277* transcripts.

References:

- Harrison, Xavier A. 2014. "Using Observation-Level Random Effects to Model Overdispersion in Count Data in Ecology and Evolution." *PeerJ* 2 (October): e616. doi:10.7717/peerj.616.
- Hollis, B., D. Houle, and T. J. Kawecki. 2016. "Evolution of Reduced Post-Copulatory Molecular Interactions in Drosophila Populations Lacking Sperm Competition." *Journal of Evolutionary Biology* 29 (1): 77–85. doi:10.1111/jeb.12763.
- Law, Charity W., Monther Alhamdoosh, Shian Su, Gordon K. Smyth, and Matthew E. Ritchie. 2016. "RNA-Seq Analysis Is Easy as 1-2-3 with Limma, Glimma and edgeR." *F1000Research* 5 (June): 1408. doi:10.12688/f1000research.9005.1.
- McGraw, Lisa A., Andrew G. Clark, and Mariana F. Wolfner. 2008. "Post-Mating Gene Expression Profiles of Female Drosophila Melanogaster in Response to Time and to Four Male Accessory Gland Proteins." *Genetics* 179 (3): 1395–1408. doi:10.1534/genetics.108.086934.
- McGraw, Lisa A., Greg Gibson, Andrew G. Clark, and Mariana F. Wolfner. 2004. "Genes Regulated by Mating, Sperm, or Seminal Proteins in Mated Female Drosophila Melanogaster." *Current Biology* 14 (16): 1509–14. doi:10.1016/j.cub.2004.08.028.

Zhou, Shanshan, Trudy FC Mackay, and Robert RH Anholt. 2014. "Transcriptional and Epigenetic Responses to Mating and Aging in Drosophila Melanogaster." *BMC Genomics* 15: 927. doi:10.1186/1471-2164-15-927.