Genetic loci with parent of origin effects cause hybrid seed lethality between *Mimulus* species

Austin G. Garner[*1], Amanda M. Kenney[*2], Lila Fishman[§], Andrea L. Sweigart[*]

[*] Department of Genetics, University of Georgia, Athens, GA, USA 30602

[§] Department of Biological Sciences, University of Montana, Missoula, MT, USA 59812

[1]Current Address: Department of Biology, University of Virginia, Charlottesville, VA, USA 22904

[2]Current Address: School of Natural Sciences, St. Edwards University, Austin, TX, USA 78704

Running title: Genetics of hybrid seed lethality in Mimulus

Key words: speciation, hybrid lethality, parent of origin effects, *Mimulus*

Corresponding author: Andrea Sweigart

Mailing address: Department of Genetics, 120 East Green Street, Davison Life Sciences

Building, University of Georgia, Athens, GA 30602

Phone: 706-542-7001

Email: sweigart@uga.edu

ABSTRACT

The classic finding in both flowering plants and mammals that hybrid lethality often depends on parent of origin effects suggests that divergence in the underlying loci might be an important source of hybrid incompatibilities between species. In flowering plants, there is now good evidence from diverse taxa that seed lethality arising from interploidy crosses is often caused by endosperm defects associated with deregulated imprinted genes. A similar seed lethality phenotype occurs in many crosses between closely related diploid species, but the genetic basis of this form of early-acting F_1 postzygotic reproductive isolation is largely unknown. Here, we show that F_1 hybrid seed lethality is an exceptionally strong isolating barrier between two closely related *Mimulus* species, *M. guttatus* and *M. tilingii*, with reciprocal crosses producing less than 1% viable seeds. Using a powerful crossing design and high-resolution genetic mapping, we identify both maternally- and paternally-derived loci that contribute to hybrid seed incompatibility. Strikingly, these two sets of loci are largely non-overlapping, providing strong evidence that genes with parent of origin effects are the primary driver of F_1 hybrid seed lethality between *M. guttatus* and *M. tilingii*. We find a highly polygenic basis for both parental components of hybrid seed lethality suggesting that multiple incompatibility loci have accumulated to cause strong postzygotic isolation between these closely related species. Our genetic mapping experiment also reveals hybrid transmission ratio distortion and chromosomal differentiation, two additional correlates of functional and genomic divergence between species.

INTRODUCTION

Understanding the genetics of interspecific incompatibility provides insight both into the origins of species barriers and the processes driving genomic divergence within

species. In plants, postmating incompatibilities are often significant, and in particular, many early plant geneticists noted difficulties in generating F_1 hybrids from experimental crosses between closely related species (Thompson 1930, Stebbins 1957, Valentine and Woodell 1963, Vickery 1978). Some of these crossing problems are likely caused by pollen-pistil interactions that prevent fertilization altogether, but for others, successful fertilization occurs only to end in a high rate of F_1 hybrid seed lethality. This failure of F_1 seeds is also a common outcome of interploidy crosses (known as triploid block, Ramsey and Schemske 1998), and represents a form of postzygotic reproductive isolation that acts early in the life cycle. Relative to other forms of hybrid incompatibility (*e.g.* hybrid necrosis or hybrid sterility), which may act in older plants or only in one sex, F_1 hybrid seed lethality might be a particularly strong barrier to interspecific gene flow. Moreover, its prevalence in plant crosses from diverse taxa suggests that F_1 hybrid seed failure might represent a major class of reproductive isolation in plants – albeit one that has gone largely unmentioned in modern accounts of speciation (*e.g.*, Coyne and Orr 2004, but see Tiffin and Moyle 2001, Turelli and Moyle 2007).

Hybrid seed inviability has long been thought to result from developmental defects in the endosperm, an important nutritive tissue for the developing embryo (Brink and Cooper 1947, Woodell and Valentine 1961). In the seeds of flowering plants, a triploid endosperm is one of the two products of double fertilization. The male gametophyte (pollen tube) releases two sperm into the ovule: one fuses with the egg nucleus to produce the zygote, and the other fuses with the two nuclei of the central cell to form the endosperm. Classic work showed that deviations from the ratio of two maternal to one paternal genomes (2m:1p) disrupts endosperm development (Johnston *et al.* 1980), explaining the failure of many interploidy crosses, which double the maternal or paternal contribution. A similar disturbance to endosperm "balance" also

4

seems to cause hybrid seed lethality in many diploid crosses between species (Valentine and Woodell 1963, Johnston and Hanneman 1982, Josefsson *et al.* 2006, Ishikawa *et al.* 2011).

The sensitivity of the endosperm to mismatches in parental genome dosage led to the hypothesis that genomic imprinting is the molecular cause of both interploidy and interspecific seed lethality (Haig and Westoby 1991, Birchler 1993, Gutierrez-Marcos 2003). The expression of an imprinted gene is dependent on its parent of origin due to differential epigenetic modifications established during male and female gametogenesis (Kohler *et al.* 2012). If imprinted genes encode dosage-sensitive regulators, quantitative changes in parental contributions could cause a stoichiometric imbalance with downstream targets (Birchler *et al.* 2001, Birchler and Veitia 2012). This mechanism provides an explanation for the common observation that reciprocal interploidy crosses differ in seed phenotypes (*e.g.*, Thompson 1930, Brink and Cooper 1947, Scott *et al.* 1998, Sekine *et al.* 2013); disruptions to maternally versus paternally derived alleles have different consequences for endosperm development. Recently, several studies in *Arabidopsis* have shown that imprinted genes are misexpressed in interploidy crosses (Erilova *et al.* 2009, Jullien and Berger 2010, Wolff *et al.* 2011, Kradolfer *et al.* 2013). In principle, similar types of deregulation involving imprinted genes might occur in the endosperm of F_1 hybrid seeds between diploid species (Kohler *et al.* 2010), but aside from two studies in *Arabidopsis* (Josefsson *et al.* 2006, Burkart-Waco *et al.* 2012) and *Capsella* (Rebernig *et al.* 2015), this idea remains largely untested.

Intriguingly, parent of origin effects seem to contribute to hybrid incompatibilities not only in plants, but also in some animal systems (Wolf *et al.* 2014). As with the endosperm of flowering plants, genomic imprinting is an important feature of the mammalian placenta (Piedrahita 2011), and disrupted interactions between imprinted

genes cause placental defects that result in abnormal hybrid growth in mice (Vrana *et al.* 2000). Although there has been much debate about the evolutionary mechanisms that account for the origin and maintenance of genomic imprinting in angiosperms and mammals (Haig and Westoby 1989, Kohler *et al.* 2012, Spencer and Clark 2014), most explanations involve some form of genomic coevolution. For example, the parental conflict theory posits that imprinting is shaped by conflict between males and females over maternal investment (Haig and Westoby 1989), and the coadaptation theory holds that integration of offspring and maternal genomes drives imprinted gene patterns (Wolf and Hager 2006). In plants, it has also been suggested that imprinted genes in the endosperm might evolve as an incidental by-product of a defensive silencing mechanism that targets invading foreign DNA in the embryo (Gehring *et al.* 2009, Hsieh *et al.* 2009). Whatever its evolutionary cause, genomic imprinting in the endosperm is highly variable among (Luo *et al.* 2011, Waters *et al.* 2011, Jiang and Kohler 2012) and even within (Waters *et al.* 2013, Pignatta *et al.* 2014) plant species. Thus, if divergence in patterns of imprinting does contribute to postzygotic reproductive isolation, it might be particularly important in the early stages of speciation.

Nevertheless, it is still unknown if parent of origin effects represent a major source of hybrid incompatibilities in flowering plants. Because F_1 hybrid seeds combine divergent gene sequences, in addition to divergent patterns of imprinting, interactions among heterospecific alleles (*i.e.*, Dobzhansky-Muller incompatibilities; Dobzhansky 1937, Muller 1942) at non-imprinted genes might also contribute to hybrid lethality. Indeed, in taxa without genomic imprinting (*e.g.*, *Drosophila* and fish), there are many well-known cases of hybrid lethality due to incompatibilities between genes with bi-parental expression (*e.g.* Wittbrodt *et al.* 1989, Presgraves *et al.* 2003, Brideau *et al.* 2006, Tang *et al.* 2009). A key question, then, is whether F_1 hybrid seed lethality in

6

flowering plants is caused primarily by incompatibilities between imprinted or non-imprinted genes. The effects of the former are likely restricted to the endosperm, the site of most imprinting in plants, whereas the latter class of genes might act in the endosperm or in the embryo itself. Moreover, distinguishing between these two causes of F_1 seed lethality is fundamentally important for understanding its evolution: imprinted genes represent a special class of loci that almost certainly reflect explicit genomic coevolution within species, rather than the independent fixation of alleles that only interact in hybrids.

Here we explore the mechanisms and genetics of reproductive isolation between two diploid species of *Mimulus*, *M. guttatus* and *M. tilingii*. In nature, these species are mostly allopatric, but they occasionally co-occur in high alpine areas. Because both species are primarily outcrossing with large, bee-pollinated flowers, sympatric populations might be expected to experience interspecific gene flow. However, early crossing studies reported that F_1 hybrids between *M. guttatus* and *M. tilingii* are often difficult to generate (Vickery 1978), suggesting there might be some degree of postmating, prezygotic isolation that prevents fertilization and/or postzygotic isolation in the form of hybrid seed lethality. To investigate these possibilities, we began our study by intercrossing the two *Mimulus* species, characterizing both the mechanism and strength of reproductive isolation. In reciprocal crosses, we show there is only a modest reduction in hybrid seed number, but that almost all of these hybrid seeds are small, misshapen, and inviable.

This closely related *Mimulus* species pair thus presents a rare opportunity to test whether or not F_1 hybrid seed lethality involves parent of origin effects. The finding that hybrid seed lethality is severe in *both* cross directions does not necessarily rule out reciprocal differences in its underlying genetic basis. That is, loci for *Mimulus* hybrid

7

seed lethality might differ depending on their parent of origin. In this study, we use a powerful breeding design – backcrossing F_2 hybrids reciprocally to each parent – to assess both the maternal and paternal contributions to *Mimulus* hybrid seed inviability. At the outset, our genetic analyses reveal two correlates of genomic divergence, transmission ratio distortion in F_2 hybrids and chromosomal differentiation, which both provide insights into *Mimulus* speciation. Additionally, by performing high-resolution genetic mapping, we show that postzygotic reproductive isolation is primarily caused by loci with parent of origin effects, strongly suggesting a role for imprinted genes in the evolution of *Mimulus* hybrid seed lethality.

MATERIALS AND METHODS

Study system and plant material

The yellow monkeyflower *Mimulus guttatus* is highly polymorphic with natural populations distributed across much of western North America. The species occupies diverse environments, ranging from sand dunes along the Pacific coast to high alpine habitats. *M. tilingii*, a mat-forming perennial, occurs throughout much of the same geographic area, but is largely restricted to high elevations (2000+ m). Both species are self-compatible, but predominantly outcrossing with large, bee-pollinated flowers. *M. guttatus* and *M. tilingii* are closely related (Beardsley and Olmstead 2002) and belong to the same *Simiolus* section of primarily yellow-flowered taxa. Still, the two species have been classified as members of different species complexes (Vickery 1978) and variation at 16 nuclear loci suggest they form distinct genetic groups (Oneal *et al.* 2014). In areas of sympatry, there have been a few reports of putative hybridization (Lindsay and Vickery 1967, Carrie Wu pers comm.), but classic crossing experiments have shown that

8

F_1 hybrids are difficult to generate (Vickery 1978), suggesting reproductive isolation between *M. guttatus* and *M. tilingii* is strong.

In this study, we use one inbred line for each of the two focal species. The *M. guttatus* parental line (DUN10) is derived from a population located in the Oregon Dunes National Recreation Area along the Pacific coast. The *M. tilingii* parental line (LVR) originated from a high-alpine population in California's Yosemite Valley (at 2751 m). Both of these inbred lines were formed by more than six generations of self-fertilization with single-seed descent.

Measuring genomic divergence among *Mimulus* species

To determine the extent of genetic differentiation between *M. tilingii* and the well-studied *M. guttatus* complex (Brandvain *et al.* 2014), we measured genome-wide divergence (*i.e.*, average pairwise nucleotide differences) among *M. tilingii*, *M. guttatus*, and *M. nasutus*. For *M. tilingii*, we generated new sequence data for the parental line LVR (see below) that yielded 68.4 million 150-bp paired-end reads. For *M. guttatus* and *M. nasutus*, we used six previously published lines (Table S1), including the parental *M. guttatus* line DUN10. Before alignment of *M. guttatus* and *M. nasutus* reads, we removed adapter contamination using Trimmomatic (Bolger *et al.* 2014) and confirmed removal using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). We checked LVR for adapter contamination using FastQC but did not detect any. We aligned the paired-end reads from each line to the *M. guttatus* v2.0 reference genome (Hellsten *et al.* 2013) using the Burrows-Wheeler Aligner (BWA, BWA-MEM used for LVR, BWA-backtrack used for all others; Li and Durbin 2009) with a minimum quality threshold of Q29 (filtering done using SAMtools, Li *et al.* 2009). Following alignment, we removed PCR duplicates using Picard (http://picard.sourceforge.net/) and performed

9

local realignment around indels using the Genome Analysis Toolkit (GATK, Depristo *et al.* 2011, McKenna *et al.* 2010).

We produced a high quality set of invariant sites and SNPs simultaneously using the GATK Unified Genotyper (Depristo *et al.* 2011, McKenna *et al.* 2010) with a site quality threshold of Q40. We excluded sites with more than two alleles and set cutoffs for coverage depth to a minimum of 10 and a maximum of 75 reads per site. To assign genotypes at heterozygous sites, we randomly selected one of two alternate alleles. We then quantified patterns of sequence variation for each pair of *Mimulus* samples using a custom python script (L. Flagel, pers. comm.). For each pairwise comparison, we counted the number of pairwise sequence differences and number of sites for which both samples have data above our quality and depth thresholds.

Measurement of reproductive isolation and genetic crosses

To characterize postmating reproductive isolation between *M. guttatus* and *M. tilingii*, we performed crosses within and between the two parental lines ($N = 20$ each for DUN10 x DUN10, LVR x LVR, DUN x LVR, and LVR x DUN). For each of these crosses, we dissected one ripened fruit and measured total seed set. For each fruit, we also assessed seed viability by determining the proportion of viable seeds (number of viable seeds/total seeds). Seed viability was straightforward to score by eye: viable seeds were plump and tan in color, whereas inviable seeds were shriveled, darker in color, and often adhering to each other (Figure S1). To determine if our measure of seed viability is correlated with germination rate, we planted the seeds from a single fruit for crosses within parental lines ($N = 2$ each) and for a subset of our F_2 x parental crosses ($N = 20$, see below). Because we found that our visual assessment of seed viability correlates

strongly with germination rate (Spearman's Correlation, *rho* = 0.92, *P* < 0.0001), all further measurements of seed viability were performed by eye.

To study the genetics of reproductive isolation between *M. guttatus* and *M. tilingii*, we intercrossed DUN10 (maternal parent) and LVR (paternal parent) to form F_1 hybrids. We then backcrossed F_1 hybrids (*N* = 20) reciprocally to each parental line and assessed seed viability as described above. To generate a recombinant population for genetic mapping, we self-fertilized a single F_1 to form an F_2 generation (*N* = 240). For each of these F_2 hybrids, we performed reciprocal backcrosses to each parental line; seed viability was measured from a single fruit for each of the four cross treatments (Figure 1). Note that because the DUN10 line was used as the original maternal parent, all F_1 and F_2 hybrids carried an *M. guttatus* cytoplasm.

All plants were grown using similar conditions at the University of Georgia. Seeds were planted into 2.5" pots containing Fafard 3b potting mix, chilled for 7 days at 4°C to promote germination, and then placed in a Conviron growth chamber with lights set to 16-hr days. Plants were bottom watered daily and temperatures were maintained at 22°C during the day and 16°C at night.

DNA extraction, library preparation, and sequencing

For each of the 240 F_2 individuals, we collected bud tissue into 96-well plates, immediately placed samples on dry ice, and stored them at -80°C. We isolated genomic DNA using a standard CTAB/chloroform extraction protocol as described in Holeski *et al.* (2014). Following extraction, we quantified DNA using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen) and diluted each sample to 5 ng/μL.

11

We generated a genomic library for genotyping the 240 F_2 hybrids by using the multiplexed shotgun genotyping (MSG) method of Andolfatto *et al.* (2011). In 96-well plates, we digested 50 ng of genomic DNA (15 U *Mse*I; New England Biolabs) for each sample in 20-µL reactions for 3 hr at 37°C followed by heat inactivation at 65°C for 20 min. Next, using 48 unique bar-coded adapters, we formed five sets of F_2 individuals (each set with 48 F_2s carrying bar-codes 1-48). We added 5 pmol of the bar-coded adapters to each well and performed ligation reactions using 1200 U of T4 DNA ligase (New England Biolabs) in total volumes of 50 µL at 16°C for 3 hr followed by heat inactivation at 65°C for 10 min. We precipitated DNA in each well by adding 5 µL of 3M sodium acetate and 50 µL of isopropanol. We then pooled each set of 48 F_2 individuals into a single tube, added 1 µL of glycogen, and incubated overnight at 4°C. After resuspending each of the five samples in 100 µL TE, we purified them using a 1.5 volume of Agencourt AMPure beads. Following purification, the Georgia Genomics Facility size-selected fragments of 375-425 bp using a QIAquick Gel Extraction Kit (QIAGEN). For each of the five sets of 48 F_2 hybrids, we then performed 18 cycles of PCR using a Phusion High-Fidelity PCR master mix (New England Biolabs) and one of five unique "index" primers that bind to common regions in the bar-coded adapters. Following this PCR amplification, each of the 240 F_2 individuals carried a unique bar-code/index identifier. We performed cleanup on each of the five sets using a 0.8 volume of Agencourt AMPure beads, resuspended in 40 µL QIAGEN EB buffer, and quantified the DNA using a Qubit fluorometer. We combined the five indexed sets into one library of 240 F_2 individuals at a final concentration of 15 µM.

We sent our library of 240 F_2 individuals to the Genome Sequencing Facility at Duke University for sequencing. We also included a sample of genomic DNA from the LVR *M. tilingii* line (the library was prepared by the Duke Genome Sequencing Facility).

Because LVR is a highly inbred line, this genomic DNA should be nearly identical to that of the maternal line used to generate the F_2 mapping population. Sequencing was performed across two lanes on an Illumina Hiseq 2500 for paired-end, 150-bp reads.

Determining F_2 genotypes

We used the MSG software pipeline (Andolfatto *et al.* 2011) to genotype 240 F_2 hybrids between *M. guttatus* (DUN10) and *M. tilingii* (LVR). The MSG HMM algorithm requires reference genomes for both parents of the mapping population. To generate pseudo-reference sequences for DUN10 and LVR, we used the processed alignment files described above. We used SAMtools mpileup to call bases with a minimum base quality threshold of Q15 and depth capped at 250. We then used SAMtools bcftools followed by SAMtools vcfutils vcf2fq to output a fastq file for each line (with an indel filtering window of 20 bases, and minimum and maximum base depth thresholds set to five and the mean plus two times the standard deviation, respectively, for each line). Finally, we used Heng Li's seqtk toolset (https://github.com/lh3/seqtk) to convert each fastq file to a fasta file, with a base quality threshold of Q15.

For all 240 *M.guttatus-M. tilingii* F_2 hybrids, we assigned ancestry probabilities across chromosomes using the MSG v0.4.3 pipeline (Andolfatto *et al.* 2011). Because MSG does not consider reverse reads, we only used forward reads from the F_2 hybrids. We removed low quality reads (less than an average Q6 across any 100-bp window) and reads with potential adapter contamination using Trimmomatic (Bolger *et al.* 2014). We then parsed reads for all 240 F_2s by unique 6-bp barcode (within each of the five indices) using Stacks (Catchen *et al.* 2011). The parsed reads were aligned to each of the two parental reference genomes using bwa with default settings, followed by base calling with SAMtools. Next, a Hidden Markov Model (HMM) estimated a posterior

13

probability of each possible ancestry (*i.e.*, homozygous *M. guttatus*, heterozygous, or homozygous *M. tilingii*) across the genomes of each F_2 individual. Finally, genotype probabilities were imputed for all F_2 individuals at each informative genomic position that was typed in at least one individual. This process resulted in 151,669 SNPs across the 14 assembled chromosomes (SNPs per chromosome: mean = 10,834; min = 5,886; max = 19,217). The MSG parameters were set as follows: the priors on ancestry homozygosity for *M. guttatus* = 0.25, heterozygosity = 0.5, and homozygosity for *M. tilingii* = 0.25; the ancestry error probabilities for *M. guttatus* = 0.03 and for *A. tilingii* = 0.03; the genome-wide recombination rate = 28 crossovers per genome per generation (1 per chromosome for each F_1 parent); and the arbitrary recombination scaling factor, rfac = 0.01 (qualitatively similar results were achieved at lower and greater values of rfac).

To facilitate linkage map construction, we converted ancestry probabilities from the MSG HMM to "hard ancestry calls," assigning each individual a confident genotype at each marker. We first culled ancestry calls with less than 95% certainty by converting those calls to missing data. We then converted the remaining high confidence calls to a probability of 1.0 (or 0.0) for each parent. To remove markers with redundant information and convert the probabilities to hard calls, we used the script pull_thin.py (available on github at https://github.com/dstern/pull_thin) to thin markers with 100% identical neighbors (including identical missing data). Then, we thinned this dataset further using R/qtl (Broman *et al.* 2003) to include markers with at least one unique genotype across all 240 individuals (keeping markers with the least amount of missing data). The final, thinned dataset containing 2960 markers was then used for linkage map construction and MQM mapping, which requires hard genotype calls.

14

Linkage mapping, transmission ration distortion, and QTL mapping

We performed linkage mapping in JoinMap 4.0 (Van Ooijen 2011) using a LOD threshold of 10.0, the Haldane mapping function, and the default Maximum Likelihood settings. In some cases, marker positions disagreed with their predicted locations from the *M. guttatus* v2.0 reference genome (often in regions known to be misassembled). For each linkage group, we assessed map quality using the nearest neighbor stress parameter in JoinMap, rlod scores output from the MSG pipeline, and visual inspection of genotypes to minimize double recombinants. Note that many of the regions that conflict between our genetic map and the genome assembly were also found in Holeski *et al.* (2014). Among the 240 F_2 hybrids, six individuals were discovered to be pure *M. guttatus* (all markers were homozygous for DUN10 alleles); these six individuals were removed from all subsequent analyses. We also used JoinMap to test markers for significant non-Mendelian genotype frequencies.

To identify putative inversions between *M. guttatus* and *M. tilingii*, we compared genetic distances between the markers in our linkage map to physical distances from the *M. guttatus* v2.0 assembly. Because this final set of markers included only those with one or more unique genotype (among the 240 F_2 hybrids), adjacent markers spanning large physical regions might represent areas of suppressed recombination. Alternatively, large physical jumps between adjacent markers could reflect low initial marker density due to a local paucity of SNPs. To exclude this second possibility, we measured initial marker densities in 500-kb windows using the raw, un-thinned marker outputs from the MSG pipeline. Regions identified as carrying putative inversions had initial marker densities similar to that of the genome-wide average.

We performed QTL mapping using MapQTL ®6 (Van Ooijen 2009). As a first step, we performed interval mapping (IM) using a 1-cM step size. Significance of QTL

15

detected by IM was evaluated at the 5% significance level by permutation tests ($N =$ 1000 permutations) at both the chromosome level and genome-wide level. Markers surrounding the QTL (i.e., with LOD scores exceeding the empirical threshold) were selected as cofactors for restricted Multiple QTL Mapping (MQM; analogous to composite interval mapping). To narrow the set of cofactors, we used the automatic cofactor selection tool in MapQTL with a threshold of $P = 0.005$. Automatic cofactor and restricted MQM analyses were repeated until a stable set of significant cofactors remained (note that restricted MQM excludes linked cofactors). At the end of this process, selected cofactors were used for a final round of restricted MQM to estimate for each QTL the maximum LOD, the additive genotypic effect (a), and the proportion of the F_2 variance explained. As before, QTL were considered significant if LOD peaks reached the genome-wide and/or chromosome significance threshold of 5% ($N = 1000$ permutations). Finally, we calculated 1.5-LOD intervals for QTL found through MQM.

RESULTS

Nucleotide divergence among *Mimulus* species

Several recent studies have characterized population genomic variation within and between closely related species of the *M. guttatus* complex (Brandvain *et al.* 2014, Puzey and Vallejo-Marin 2014, Twyford and Friedman 2015), but samples from *M. tilingii*, which classic crossing studies suggest is more distantly related (Vickery 1978), were not included in these analyses. As a first step toward understanding the extent of genetic differentiation between these species, we examined genome-wide nucleotide variation within and between *M. guttatus*, *M. nasutus*, and *M. tilingii* (Table S2). Consistent with classic taxonomic groupings (Vickery 1978) and previous phylogenic

16

work (Beardsley and Olmstead 2002), genome-wide divergence between *M. guttatus* and *M. tilingii* (d = 6.94%, *SE* = 0.28%) is substantially higher than divergence between *M. guttatus* and *M. nasutus* (d = 4.38%, *SE* = 0.09%). The latter pair is largely interfertile with ongoing introgression (Brandvain *et al.* 2014), and species divergence is comparable to levels of diversity within *M. guttatus* (π = 4.02%, *SE* = 0.19%). Higher genomic divergence between *M. guttatus* and *M. tilingii* suggests an earlier split for this pair and, potentially, stronger interspecific isolating barriers.

Pattern of hybrid seed production and viability between *M. guttatus* and *M. tilingii*

To examine the strength of postmating reproductive isolation between *M. guttatus* and *M. tilingii*, we compared seed set from reciprocal interspecific crosses to that from crosses within each parental line (Figure S2). The DUN10 line of *M. guttatus* is a large flowered ecotype and produces significantly more seeds per fruit (mean = 313, *SE* = 16.6, *N* = 20) than the smaller flowered LVR line of *M. tilingii* (mean = 188, *SE* = 16.6, *N* = 20). For both sets of interspecifc crosses, seed set was lower than parental levels, but this reduction was only significant with *M. tilingii* as the maternal parent (*M. guttatus* x *M. tilingii*: mean = 167, *SE* = 16.6, *N* = 20; *M. tilingii* x *M. guttatus*: mean = 113, *SE* = 16.2, *N* = 20). Backcross seed set with *M. tilingii* as the maternal parent was similarly low (*M. tilingii* x F_1: mean = 103, *SE* = 19.2, *N* = 15). In contrast, seed sets from the other backcross classes were similar to that of *M. tilingii* (F_1 x *M. guttatus*: mean = 166, *SE* = 18.6, *N* = 16; *M. guttatus* x F_1: mean = 214, *SE* = 17.5, *N* = 18; F_1 x *M. tilingii*: mean = 204, *SE* = 18.0, *N* = 16). Note that because all F_1 hybrids were generated using the DUN10 line of *M. guttatus* as the maternal parent, differences in seed set between reciprocal backcrosses to *M. tilingii* might involve cytonuclear interactions. Overall, there is a modest reduction in interspecific seed set, suggesting that reproductive isolating

17

barriers (*e.g.* pollen-pistil incompatibilities) might partially interfere with the production of F_1 hybrid seeds.

A much more dramatic isolating barrier was observed when we compared the proportion of viable seeds produced within and between species (Figure 2). Although seed viability differed significantly among parental lines (*M. guttatus*: mean = 0.95, *SE* = 0.02, *N* =20; *M. tilingii*: mean = 0.56, *SE* = 0.02, *N* = 20), both had much higher proportions than either interspecific cross. Indeed, F_1 hybrid seed viability – in both directions of the cross – was <2% of the mid-parent value and not significantly different from zero (*M. guttatus* x *M. tilingii*: mean = 0.004, *SE* = 0.02, *N* = 20; *M. tilingii* x *M. guttatus*: mean = 0.01, *SE* = 0.02, *N* = 20). This F_1 seed lethality represents an extremely strong barrier to interspecific reproduction, allowing less than 1% of F_1 hybrid seeds to survive.

Moreover, this strong F_1 hybrid lethality appears to be driven, at least in part, by loci with parent of origin effects: in backcrosses to either of the recurrent parents, hybrid lethality is much more severe when the F_1 acts as the paternal parent. For the backcross to *M. guttatus*, this effect cannot be explained by cytonuclear interactions because the reciprocal backcross progeny inherit the same DUN10 cytoplasm. In contrast, the progeny of reciprocal backcrosses to *M. tilingii* do carry different cytoplasms; this fact, along with the possible maternal effects from *M. tilngii*, might explain the particularly low levels of seed viability observed when *M. tilingii* acts as the seed parent (in Figure 2, compare seed viability for T x F_1 versus F_1 x T). For the other backcross classes, seed viability is generally additive, with levels of seed failure among the progeny intermediate to that from parental and interspecific crosses (*e.g.*, in Figure 2, compare seed viability for F_1 x G to G x G and T x G).

Transmission ratio distortion in the F$_2$ hybrid mapping population

Our genetic map is based on 2960 markers and has a total length of 1318 cM with an average spacing of 0.4 cM. Two-thirds (1,987) of the markers genotyped in our F$_2$ mapping population deviate from the expected 1:2:1 genotype ratios at $\alpha = 0.05$, and nearly one-third (940) show significant transmission ratio distortion at a higher threshold ($\alpha = 0.001$). The bias we observe is highly directional. Of the 940 markers distorted at $\alpha = 0.001$, 692 have an excess of *M. guttatus* (GG) homozygous genotypes and a deficit of *M. tilingii* (TT) genotypes, whereas only 70 markers show the opposite pattern (excess of TT, deficit of GG). The vast majority of these markers (749 of 762) also show a significant bias in allele frequency from the expected 1:1 ratio ($\alpha = 0.05$). The remaining markers with significant genotypic distortion (178) show an excess of heterozygotes, whereas no markers show a deficit of heterozygous genotypes.

Transmission ratio distortion is highly variable across the genome and affects 12 of the 14 linkage groups (Figure 3). By examining the genome-wide distribution of transmission bias, we identified 16 regions that contain genetically linked clusters of distorted markers. Ten of these regions show an excess of *M. guttatus* genotypes, three show an excess of *M. tilingii* genotypes, and three have an excess of heterozygotes (red, blue, and purple horizontal bars, respectively, in Figure 3). Two linkage groups (LG2 and LG6) have highly distorted markers across their entire lengths – both are overrepresented for *M. guttatus*.

Evidence for chromosomal inversions

Each of the 14 linkage groups in our genetic map corresponds to a chromosome from the *M. guttatus* v2.0 sequence assembly (http://www.phytozome.org), which

19

represents ~294 Mb of the genome (total haploid genome size is ~450 Mb, http://www.mimulusevolution.org). Using genetic distances from our map and physical lengths from the genome sequence, the average recombination rate per chromosome is 4.6 cM/Mb (range = 3.4-8.1 cM/Mb, σ = 1.2 cM/Mb). The genome-wide estimate of (1318 cM/450 Mb) ~2.9 cM/Mb is substantially lower because it includes portions of the genome not included in the sequence assembly. These results are broadly consistent with previous genetic maps from *M. guttatus* (*et al.* 2014, Holeski *et al.* 2014).

Across our genetic map, we observed three regions where a large number of physically dispersed markers map to the same location (Figure 4, regions highlighted in yellow), suggesting the presence of inversions between *M. guttatus* and *M. tilingii*. One of these regions extends from ~0.9-7.6 Mb on chromosome 8 and overlaps with a previously discovered inversion known as *DIV1* that differentiates *M. guttatus* annual and perennial ecotypes (Lowry and Willis 2010, Oneal *et al.* 2014, Twyford and Friedman 2015). Because the DUN10 *M. guttatus* parent is known to carry the "perennial" *DIV1* arrangement (Lowry and Willis 2010), our results strongly suggest that *M. tilingii* is collinear with annual forms of *M. guttatus*. The two other putative inversions are from ~13.4-18.8 Mb on chromosome 5 and ~15.9-20.5 Mb on chromosome 13. This region on chromosome 5 overlaps with a putative inversion segregating within *M. guttatus* (Holeski *et al.* 2014).

QTL mapping of hybrid seed lethality and characterization of parent of origin effects

The distribution of hybrid seed viability varies substantially among the four different F_2 cross types (Figure 4). Intriguingly, the distinct patterns of phenotypic variation in reciprocal crosses (F_2 x G vs. G x F_2; F_2 x T vs. T x F_2) suggest parent of

20

origin effects on hybrid seed viability (but note that in cross types involving *M. tilingii* we cannot rule out an effect of cytoplasm). In three of the four cross types, seed viability appears generally consistent with additivity, with the means for F_1 and F_2 hybrids similar to the midparent values (see the first three panels of Figure 4). In contrast, seed viability from the *M. tilingii* x F_2 crosses is skewed strongly to the left, potentially suggesting a role for cytonuclear interactions and/or maternal effects. Additionally, we observed transgressive inheritance for seed viability in the F_2 x *M. tilingii* crosses (with a substantial number of F_2 hybrids producing more viable seeds than pure *M. tilingii*).

Our QTL analyses show a polygenic basis for hybrid seed lethality between *M. guttatus* and *M. tilingii*. For three of the four cross types, we detected multiple QTLs (18 total; Figure 5, Table 1). In the *M. tilingii* x F_2 cross, we found only a single QTL, but our power of detection was likely limited by a smaller sample size (only 158 F_2 hybrids were phenotyped for this cross type) and non-normal distribution of seed viability (Figure 4; Beavis 1998). For 17 of the 18 QTLs, additive effects were in the expected directions based on parental values. Individual QTLs often had strong effects on seed lethality: reductions in backcross seed viability for the "lethal" versus alternative homozygote ranged from 20-61% (Table 1). On average, the seed viability of F_2 hybrids carrying a lethal genotype at one of the QTLs was reduced by 39%.

In crosses to both species, we discovered distinct sets of QTLs for hybrid seed lethality in reciprocal F_2 crosses, implying that the allelic effects of these loci differ depending on whether they are inherited from the maternal or paternal parent. Indeed, there is little overlap between QTLs mapped in F_2 x *M. guttatus* crosses and those mapped in *M. guttatus* x F_2 crosses (orange vs. red bars in Figure 5); the same is true for QTLs mapped in the two cross types involving *M. tilingii* (blue vs. green bars in Figure 5). Additionally, two QTLs show opposite allelic effects – in the predicted

21

directions – when crossed to different species. One of these QTLs (on chromosome 4) affects hybrid seed lethality when it segregates in the paternal parent. The other QTL (on chromosome 14 at ~72-78 cM) affects seed lethality through the maternal parent. Taken together, these results provide strong evidence that QTLs with parent of origin effects play a major role in *M. guttatus-M. tilingii* hybrid seed lethality.

One exception to this general pattern is the QTL on chromosome 2. At this locus, the *M. guttatus* allele increases seed fertility in three of the four cross types irrespective of direction or species. This result suggests that the chromosome 2 QTL is part of a hybrid incompatibility or simply reflects the segregation of deleterious alleles from the *M. tilingii* parent (in Figure 2, note the lower seed viability for this inbred line).

DISCUSSION

In this study, we have shown that hybrid seed lethality is a highly effective isolating barrier between *M. guttatus* and *M. tilingii* with F_1 seed viability being less than 1%. Additionally, we used a high-resolution mapping experiment to characterize the genetic basis of this widespread and exceptionally strong form of postzygotic reproductive isolation. Along with identifying hybrid lethality QTLs, our genetic analyses revealed both transmission ratio distortion (TRD) in F_2 hybrids and chromosomal differentiation resulting in suppression of hybrid recombination. These phenomena are common in other plant mapping populations, and may provide insights into functional and genomic divergence between species. Furthermore, our reciprocal cross design allowed for direct tests of parent of origin effects on hybrid seed lethality at the QTL-

22

level. Strikingly, most of the QTLs we identified contribute to hybrid seed lethality when inherited from only one of the parents (maternal or paternal), providing strong evidence for the involvement of imprinted genes. This finding is consistent with recent empirical work in *Arabidopsis* suggesting that imprinting plays a central role in triploid block (seed lethality in interploid hybrids), but our study is the first to identify both maternal- and paternal-effect loci for hybrid seed lethality between diploid species. Further work will be needed to identify the causal genes, but the results presented here suggest that divergence in genomic imprinting can generate strong postzygotic isolation between plant species in the early stages of divergence.

Transmission ratio distortion and chromosomal differentiation between *Mimulus* species

In addition to QTLs for hybrid lethality, our mapping experiment revealed two common correlates of species divergence, transmission ratio distortion and evidence for chromosomal rearrangements. Just within the *M. guttatus* complex, TRD has been instrumental in the discovery of centromere-associated female meiotic drive (Fishman and Willis 2005, Fishman and Saunders 2008), loci underlying gamete competition and conspecific pollen precedence (Fishman *et al*. 2008), and potential cytoplasm-dependent hybrid incompatibilities (Lowry *et al*. 2009). In addition to these mechanisms, TRD in hybrids may also reflect inbreeding depression, barriers to fertilization, and postzygotic hybrid seed lethality. Similarly, suppression of recombination can reveal inversions or other rearrangements distinguishing species, which may be important in the development of both premating (Kirkpatrick and Barton 2006) and postzygotic barriers (Noor *et al*. 2001, Navarro and Barton 2003).

23

A large number of markers had distorted genotypic frequencies in our F_2 mapping population, similar to other crosses within (Hall and Willis 2005) and between *Mimulus* species (Fishman *et al.* 2001, Fishman *et al.* 2015). The proportion of distorted markers in our F_2 population (67% at $\alpha = 0.05$, 32% at $\alpha = 0.001$) is somewhat higher than the proportion found in an interspecific cross between *M. guttatus* and *M. nasutus* (49%, Fishman *et al.* 2001), consistent with the higher divergence between our focal species (*d* for *M. guttatus-M. nasutus* = 4.38%, *d* for *M. guttatus-M. tilingii* = 6.94%). In general, the distorted markers cluster in particular regions, suggesting that patterns of transmission distortion are caused by underlying loci, rather than by chance or error.

F_2 mapping alone does not allow differentiation of the many mechanisms of TRD; however, our seedset data provide some clues about likely contributors. Least interesting, but a possibility in our cross, is inbreeding depression. Both parental lines were highly inbred, but they did not have equivalent fertility. Whereas the *M. guttatus* DUN10 line has high fitness, the *M. tilingii* LVR line produces a large fraction of inviable selfed seed (Figure 2). Segregation of deleterious recessive alleles fixed in the LVR parent could account, in part, to the genome-wide excess of *M. guttatus* genotypes and alleles we observe throughout the genome. Alternatively, TRD could reflect genetic interactions unique to hybrids. These heterospecific interactions can arise in gametes, where they have the potential to influence viability or fertilization success, or in F_2 zygotes, where they might cause differential survival. Consistent with the action of gametic incompatibilities, in both reciprocal crosses of *M. guttatus* and *M. tilingii*, we observed a modest reduction in seed set (Figure S2), suggesting that certain loci might interfere with interspecific fertilization. Finally, the presence of strong F_1 seed lethality, as well as the common observation of postzygotic hybrid incompatibilities between less divergent *Mimulus* species (*e.g.*, Christie and MacNair 1987, Fishman and Willis 2006,

24

Sweigart *et al.* 2007, Sweigart and Flagel 2015), suggests that Dobzhansky-Muller interactions causing F_2 hybrid seed lethality might also contribute to TRD.

One intriguing possibility is that some of the same genetic loci might cause both hybrid seed lethality in our F_2 crossing experiments and transmission ratio distortion in the F_2 mapping population. At the LG2 QTL, F_2 hybrids that carry *M. guttatus* alleles produce a higher proportion of viable seeds in three of four cross treatments. Similarly, if F_2 seeds with *M. guttatus* alleles at LG2 are more likely to be viable themselves, it would lead to an overrepresentation of *M. guttatus* alleles in the F_2 adults. The same effect might also occur at the distorted region on LG13; when the LG13 QTL carries *M. guttatus* alleles, it increases seed viability in reciprocal F_2 crosses to *M. guttatus*. Interestingly, both the QTL and transmission ratio distortion on LG13 map to a putative inversion (Figure 5), suggesting that this region might contain multiple loci contributing to these phenotypes.

Importantly, for understanding the history of divergence between *M. guttatus* and *M. tilingii*, and for further studies of adaptation and speciation in this system, our genetic mapping revealed strong suppression of recombination in the *DIV1* region on chromosome 8 (Figure 5). The *DIV1* inversion defines widespread annual and perennial ecotypes of *M. guttatus* and contains QTLs for flowering time and growth-related traits that differentiate them (Hall *et al.* 2006, Lowry and Willis 2010, Twyford and Friedman 2015). Our finding that the LVR strain of *M. tilingii* is not collinear with the perennial DUN10 parent suggests that the perennial arrangement might be evolutionarily derived, consistent with an observed reduction in genetic diversity in the *DIV1* region in *M. guttatus* perennials (Oneal *et al.* 2014, Twyford and Friedman 2015).

Loci with parent of origin effects cause strong reproductive isolation between *Mimulus* species

At first glance, the finding that severe F_1 hybrid seed lethality occurs in both reciprocal crosses between *M. guttatus* and *M. tilingii* seems contrary to the idea that imprinted genes are involved. Unlike many interploidy and interspecies crosses, which often show pronounced reciprocal differences in seed lethality (see Thompson 1930, Haig and Westoby 1991), we detected no parent of origin effects in these first-generation *Mimulus* hybrid seeds. So, if imprinted genes *do* cause F_1 hybrid seed lethality, the lack of reciprocal differences points to an independent genetic basis for the phenotype in each cross direction, implying genetic changes have evolved in both *Mimulus* lineages. Indeed, this is exactly what our mapping of distinct maternally- and paternally-contributed QTLs reveals: many loci with parent of origin effects contribute to *Mimulus* hybrid seed lethality. Of course, a more detailed phenotypic characterization of *Mimulus* F_1 hybrid seed lethality might also reveal reciprocal differences at the level of endosperm growth and/or development that are associated with these distinct sets of genetic loci.

We found a polygenic basis for hybrid seed lethality between *M. guttatus* and *M. tilingii*. Across the four F_2 backcross treatments, we detected 18 QTLs, although several of these QTLs seem not to be independent. On chromosome 14, for example, two maternally contributed QTLs map to overlapping regions and have opposite allelic effects in backcrosses to *M. guttatus* and *M. tilingii*. The simplest explanation for this pattern is that one QTL causes both effects: F_2 hybrids carrying *M. guttatus* alleles at the causal locus promote seed compatibility in crosses to *M. guttatus*, whereas *M. tilingii* alleles improve seed viability in crosses to *M. tilingii*. Similarly, on chromosome 4, overlapping paternal QTLs have opposite phenotypic effects in the two cross treatments

26

(in the predicted directions), suggesting a common genetic basis. For two additional genomic regions – on chromosomes 2 and 13 – we detected overlapping QTLs from reciprocal crosses, implying that allelic effects at the underlying loci do not depend on parent of origin. For all remaining, non-overlapping QTLs, dominance relations and/or genetic background effects might make detection more likely in one backcross than in the other. Fortunately, our four-way crossing design maximizes the chances of mapping such QTLs despite these complicating effects. In total, we found eight QTLs with seed lethality effects only through the maternal parent, three QTLs with effects only through the paternal parent, and two QTLs with effects through both parents. Similar to these results in *Mimulus*, postzygotic barriers affecting hybrid seed survival between *Arabidopsis* species (Burkart-Waco *et al.* 2012) and between *Capsella* species (Rebernig *et al.* 2015) are also determined by many genetic loci. In *Arabidopsis*, epistasis among the causal loci indicates that hybrid seed incompatibility is controlled by a complex genetic network (Burkart-Waco *et al.* 2012). Thus, in three diverse systems, strong F_1 postzygotic isolation seems to have evolved between closely related species due to the accumulation of multiple incompatibilities that combine to cause severe defects in hybrid seed development.

Given the high variability in patterns of genomic imprinting even among strains of *Arabidopsis* (Pignatta *et al.* 2014), information from other species may have limited predictive value for identifying candidate genes in our *Mimulus* seed lethality QTLs. However, as a first step toward identifying candidate genes for parent of origin effects in *Mimulus* hybrid seed lethality, we performed a blast search of the *M. guttatus* reference genome using genes previously identified as imprinted in the *A. thaliana* endosperm (Hsieh *et al.* 2009, Wolff *et al.* 2011, Pignatta *et al.* 2014). We blasted 438 maternally expressed and 150 paternally expressed *Arabidopsis* genes and recovered 303 and 92

27

best hits, respectively, from the *M. guttatus* genome (using an e-value cutoff of 1E-06).

Among these 395 *M. guttatus* genes, 26 co-localize with maternal QTLs and six co-localize with paternal QTLs. Based on their annotations, none of these genes has an obvious functional role in endosperm development, and of course, there is no guarantee that any of them are imprinted in *Mimulus*. Thus, further fine-mapping and functional/expression analyses will be necessary to identify the molecular basis of parent of origin effects in Mimulus hybrids; such analyses will also enlarge our understanding of this highly dynamic phenomenon beyond a few model systems.

A key question is which evolutionary forces might have led to divergence at hybrid lethality loci between *M. guttatus* and *M. tilingii*. One intriguing possibility is that F_1 hybrid lethality in both reciprocal crosses is the outcome of unique coevolutionary histories between imprinted genes and their targets in each of the two *Mimulus* lineages. According to the parental conflict model, maternally expressed genes should be selected to restrict endosperm growth, whereas paternally expressed genes might function to promote growth (Haig and Westoby 1989). If patterns of genomic imprinting have evolved in *Mimulus* because of parental conflict over maternal investment, it is certainly possible that different species have resolved this conflict using distinct genetic routes. In nature, both *M. guttatus* and *M. tilingii* are predominantly outcrossing, so parental conflict has the potential to be strong within each species (Brandvain and Haig 2005). Of course, lineage-specific changes may also occur if imprinted genes have evolved by some other selective mechanism [*e.g.*, coadaptation between maternal and offspring traits (Wolf and Hager 2006)] or by incidental proximity to silenced TEs (Gehring *et al.* 2009, Hsieh *et al.* 2009). In the latter case, patterns of genomic imprinting might be expected to be particularly dynamic given the high degree of variation in TE position within and between plant species (Tenaillon *et al.* 2010, Cao *et al.* 2011). Interestingly, F_1 seed lethality also

28

occurs in some crosses between geographically distant populations of *M. tilingii* (Sweigart *et al.*, unpubl. results), suggesting that, as in *Arabidopsis* and maize (Waters *et al.* 2013, Pignatta *et al.* 2014), there may be variation within *M. tilingii* for patterns of genomic imprinting.

Conclusions

In this study, we have presented the first investigation of reproductive isolation and its genetic basis between *M. guttatus* and *M. tilingii*. Both species have abundant natural populations that occasionally co-occur, potentially providing the opportunity for interspecific gene flow. However, we have shown that there is incredibly strong postzygotic reproductive isolation between these species. Our quantitative genetic analysis of hybrid seed lethality – the most comprehensive to date between diploid plant species – confirms a central role for genetic loci with parent of origin effects. These findings set the stage for future fine-mapping, phenotypic characterizations of hybrid seed development, and gene expression studies to identify the underlying genes. This possibility, along with the potential for discovering natural variation in genomic imprinting, makes this *M. guttatus-M. tilingii* system a particularly rich one for investigating the evolutionary mechanisms of this early acting form of postzygotic reproductive isolation.

ACKNOWLEDGEMENTS

29

LITERATURE CITED

Andolfatto P, Davison D, Erezyilmaz D, Hu TT, Mast J, et al. (2011) Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Research 21: 610-617.

Beardsley PM, Yen A, Olmstead RG (2003) AFLP phylogeny of Mimulus section Erythranthe and the evolution of hummingbird pollination. Evolution 57: 1397-1410.

Beavis WD (1998) QTL analyses: power, precision, and accuracy. Molecular Dissection of Complex Traits, 145-162.

Birchler JA (1993) Dosage analysis of maize endosperm development. Annual Review of Genetics 27: 181-204.

Birchler JA, Veitia RA (2012) Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. Proceedings of the National Academy of Sciences 109: 14746-14753.

Birchler JA, Bhadra U, Bhadra MP, Auger DL (2001) Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. Developmental Biology 234: 275-288.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.

Brandvain Y, Haig D (2005) Divergent mating systems and parental conflict as a barrier to hybridization in flowering plants. The American Naturalist 166: 330-338.

Brandvain Y, Kenney AM, Flagel L, Coop G, Sweigart AL (2014) Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. PLoS Genetics 10(6): e1004410.

Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X *et al.* (2006) Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. Science 314: 1292-1295.

Brink RA, Cooper DC (1947) The endosperm in seed development. The Botanical Review 13: 479-541.

Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889-890.

Burkart-Waco D, Josefsson C, Dilkes B, Kozloff N, Torjek O, et al. (2012) Hybrid incompatibility in *Arabidopsis* is determined by a multiple-locus genetic network. Plant Physiology 158: 801-812.

Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, et al. (2011) Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nature Genetics 43: 956-963.

Catchen J, Amores A, Hohenlohe P, Cresko W, Postlethwait J (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics 1: 171-182.

Christie P, Macnair MR (1987) The distribution of postmating reproductive isolating genes in populations of the yellow monkey flower, *Mimulus guttatus*. Evolution 41: 571-578.

31

Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Sunderland, MA.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43: 491-498.

Dobzhansky TH (1937) Genetics and the Origin of Species, Columbia University Press, New York.

Erilova A, Brownfield L, Exner V, Rosa M, Twell D, et al. (2009) Imprinting of the Polycomb Group Gene MEDEA Serves as a Ploidy Sensor in Arabidopsis. PLoS Genet 5: e1000663.

Fishman L, Saunders A (2008) Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. Science 322: 1559-1562.

Fishman L, Willis JH (2005) A novel meiotic drive locus almost completely distorts segregation in *Mimulus* (monkeyflower) hybrids. Genetics 169: 347-353.

Fishman L, Willis JH (2006) A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. Evolution 60: 1372-1381.

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics 159: 1701-1716.

Fishman L, Willis JH, Wu CA, Lee YW (2014) Comparative linkage maps suggest that fission, not polyploidy, underlies near-doubling of chromosome number within monkeyflowers (*Mimulus*; Phrymaceae). Heredity 112: 562-568.

Fishman L, Beardsley PM, Stathos A, Williams CF, Hill JP (2015) The genetic architecture of traits associated with the evolution of self-pollination in *Mimulus*. New Phytologist 205: 907-917.

Gehring M, Bubb KL, Henikoff S (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science 324: 1447-1451.

Gutierrez-Marcos JF, Pennington PD, Costa LM, Dickinson HG (2003) Imprinting in the endosperm: a possible role in preventing wide hybridization. Philosophical Transactions of the Royal Society of London B: Biological Sciences 358: 1105-1111.

Haig D, Westoby M (1989) Parent-specific gene expression and the triploid endosperm. American Naturalist 134: 147-155.

Haig D, Westoby M (1991) Genomic imprinting in endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis. Philosophical Transactions of the Royal Society B: Biological Sciences 333: 1-13.

Hall MC, Willis JH (2005) Transmission ratio distortion in intraspecific hybrids of Mimulus guttatus implications for genomic divergence. Genetics 170: 375-386.

Hall MC, Basten CJ, Willis JH (2006) Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in *Mimulus guttatus*. Genetics 172: 1829-1844.

Hellsten U, Wright KM, Jenkins J, Shu S, Yuan Y, et al. (2011) Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. Proceedings of the National Academy of Sciences of the United States of America. 110: 19478-19482.

Holeski LM, Monnahan P, Koseva B, McCool N, Lindroth RL, et al. (2014) A high-resolution genetic map of yellow monkeyflower identifies chemical defense QTLs and recombination rate variation. G3: Genes, Genomes, Genetics 4: 813-821.

Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, et al. (2009) Genome-wide demethylation of *Arabidopsis* endosperm. Science 324: 1451-1454.

Ishikawa R, Ohnishi T, Kinoshita Y, Eiguchi M, Kurata N, et al. (2011) Rice interspecies hybrids show precocious or delayed developmental transitions in the endosperm

without change to the rate of syncytial nuclear division. The Plant Journal 65: 798-806.

Jiang H, Köhler C (2012) Evolution, function, and regulation of genomic imprinting in plant seed development. Journal of Experimental Botany 63: 4713-4722.

Johnston SA, Hanneman RE (1982) Manipulations of endosperm balance number overcome crossing barriers between diploid *Solanum* species. Science 217: 446-448.

Johnston SA, Den Nijs TPM, Peloquin SJ, Hanneman RE (1980) The significance of genic balance to endosperm development in interspecific crosses. Theoretical and Applied Genetics 57: 5-9.

Josefsson C, Dilkes B, Comai L (2006) Parent-dependent loss of gene silencing during interspecies hybridization. Current Biology 16: 1322-1328.

Jullien PE, Berger F (2010) Parental genome dosage imbalance deregulates imprinting in *Arabidopsis.* PLoS Genetics 6: e1000885.

Kirkpatrick M, Barton N (2006) Chromosome inversions, local adaptation and speciation. Genetics 173: 419-434.

Köhler C, Scheid OM, Erilova A (2010) The impact of the triploid block on the origin and evolution of polyploid plants. Trends in Genetics 26: 142-148.

Köhler C, Wolff P, Spillane C (2012) Epigenetic mechanisms underlying genomic imprinting in plants. Annual Review of Plant Biology 63: 331-352.

Kradolfer D, Wolff P, Jiang H, Siretskiy A, Köhler C (2013) An imprinted gene underlies postzygotic reproductive isolation in *Arabidopsis thaliana*. Developmental Cell 26: 525-535.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence
Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Lindsay DW, Vickery RK (1967) Comparative evolution in *Mimulus guttatus* of the
Bonneville Basin. Evolution 21: 439-456.

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism
contributes to a major life-history transition, local adaptation, and reproductive
isolation. PLoS Biology 8: 2227.

Lowry DB, Hall MC, Salt DE, Willis JH (2009) Genetic and physiological basis of
adaptive salt tolerance divergence between coastal and inland *Mimulus guttatus*.
New Phytologist 183: 776-788.

Luo M, Taylor JM, Spriggs A, Zhang H, Wu X, et al. (2011) A genome-wide survey of
imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm.
PLoS Genetics 7: e1002125.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. (2010) The Genome
Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA
sequencing data. Genome Research 20: 1297-1303.

Muller HJ (1942) Isolating mechanisms, evolution, and temperature. Biol. Symp. 6: 71-
125.

Navarro A, Barton NH (2003) Accumulating postzygotic isolation genes in parapatry: a
new twist on chromosomal speciation. Evolution 57: 447-459.

Noor MA, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the
reproductive isolation of species. Proceedings of the National Academy of Sciences,
98: 12084-12088.

Oneal E, Lowry DB, Wright KM, Zhu Z, Willis JH (2014) Divergent population structure
and climate associations of a chromosomal inversion polymorphism across the
Mimulus guttatus species complex. Molecular Ecology 23: 2844-2860.

Piedrahita JA (2011) The role of imprinted genes in fetal growth abnormalities. Birth Defects Research Part A: Clinical and Molecular Teratology 91: 682-692.

Pignatta D, Erdmann RM, Scheer E, Picard CL, Bell GW, et al. (2014) Natural epigenetic polymorphisms lead to intraspecific variation in *Arabidopsis* gene imprinting. Elife 3: e03198.

Presgraves DC, Balagopalan L, Abmayr SM, Orr HA (2003) Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. Nature 423: 715-719.

Puzey J, Vallejo-Marín M (2014) Genomics of invasion: diversity and selection in introduced populations of monkeyflowers (*Mimulus guttatus*). Molecular Ecology 23: 4472-4485.

Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 29: 467-501.

Rebernig CA, Lafon-Placette C, Hatorangan MR, Slotte T, Köhler C (2015) Non-reciprocal interspecies hybridization barriers in the Capsella genus are established in the endosperm. PLoS Genetics 11: e1005295.

Sekine D, Ohnishi T, Furuumi H, Ono A, Yamada T, et al. (2013) Dissection of two major components of the post-zygotic hybridization barrier in rice endosperm. The Plant Journal 76: 792-799.

Spencer, H. G., & Clark, A. G. (2014). Non-conflict theories for the evolution of genomic imprinting. Heredity, 113(2), 112-118.

Stebbins GL (1957) The inviability, weakness, and sterility of interspecific hybrids. Advances in Genetics 9: 147-215.

Sweigart AL, Flagel LE (2015) Evidence of natural selection acting on a polymorphic hybrid incompatibility locus in *Mimulus*. Genetics 199: 543-554.

Sweigart AL, Mason AR, Willis JH (2007) Natural variation for a hybrid incompatibility between two species of *Mimulus*. Evolution 61: 141-151.

Tang S, Presgraves DC (2009) Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. Science 323: 779-782.

Tenaillon MI, Hollister JD, Gaut BS (2010) A triptych of the evolution of plant transposable elements. Trends in Plant Science 15: 471-478.

Thompson WP (1930) Causes of difference in success of reciprocal interspecific crosses. American Naturalist 64: 407-421.

Tiffin P, Olson S, Moyle LC (2001) Asymmetrical crossing barriers in angiosperms. Proceedings of the Royal Society of London B: Biological Sciences 268: 861-867.

Turelli M, Moyle LC (2007) Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. Genetics 176: 1059-1088.

Twyford AD, Friedman J (2015) Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. Evolution 69: 1476-1486.

Valentine DH, Woodell SRJ (1963) Studies in British Primulas. X. Seed incompatibility in intraspecific and interspecific crosses at diploid and tetraploid levels. New Phytologist 62: 125-143.

Van Ooijen JW, Kyazma BV (2009) MapQTL 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma BV: Wageningen, Netherlands.

Van Ooijen JW (2011) Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genetics Research 93: 343-349.

Vickery Jr RK (1978) Case studies in the evolution of species complexes in *Mimulus* (pp. 405-507). Springer US.

Vrana PB, Fossella JA, Matteson P, del Rio T, O'Neill MJ, et al. (2000) Genetic and epigenetic incompatibilities underlie hybrid dysgenesis in *Peromyscus*. Nature Genetics 25: 120-124.

Waters AJ, Bilinski P, Eichten SR, Vaughn MW, Ross-Ibarra J, et al. (2013). Comprehensive analysis of imprinted genes in maize reveals allelic variation for imprinting and limited conservation with other species. Proceedings of the National Academy of Sciences 110: 19639-19644.

Wolf JB, Hager R (2006) A maternal-offspring coadaptation theory for the evolution of genomic imprinting. PLoS Biology 4: e380.

Wolf JB, Oakey RJ, Feil R (2014) Imprinted gene expression in hybrids: perturbed mechanisms and evolutionary implications. Heredity 113: 167-175.

Wolff P, Weinhofer I, Seguin J, Roszak P, Beisel C, et al. (2011) High-resolution analysis of parent-of-origin allelic expression in the *Arabidopsis* endosperm. PLoS Genetics 7: e1002126.

Wittbrodt J, Adam D, Malitschek B, Mäueler W, Raulf F, Telling A *et al.* (1989) Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in *Xiphophorus*. Nature 341: 415-421.

Woodell SRJ, Valentine DH (1961) Studies in British Primulas IX. Seed incompatibility in diploid-autotetraploid crosses. New Phytologist 60: 282-294.

Table 1. Summary of seed lethality QTL peak locations, statistical significance (LOD score), additive effect estimates (*a*), and viability effects.

Cross	Linkage group	Position	LOD	a^a	Reduction in viability[b]
F_2 x *M. guttatus*	2	70.53	3.62	0.078	0.42
	5	33.88	2.74	0.066	0.35
	8	58.39	5.00**	0.080	0.42
	10	64.05	3.51	0.069	0.37
	13	58.35	5.59**	0.093	0.50
	14	21.23	5.14**	0.131	0.61
	14	72.46	6.74**	0.120	0.49
M. guttatus x F_2	2	57.95	3.06	0.063	0.20
	4	2.36	2.95	0.067	0.21
	6	13.41	2.63	0.082	0.25
	7	66.99	3.14	0.086	0.26
	13	51.32	3.31	0.089	0.27
F_2 x *M. tilingii*	2	72.73	2.93	0.092	0.41
	6	31.94	3.91	-0.094	0.41
	7	20.17	3.09	-0.073	0.33
	9	7.73	3.17	-0.156	0.45
	14	78.10	7.57**	-0.140	0.55
M. tilingii x F_2	4	18.59	2.92	-0.037	0.53

**QTL is significant at a genome-wide threshold of 5% (*N* = 1000). All other QTLs are significant at the chromosomal level with a threshold of 5% (*N* = 1000).

[a]Positive values indicate the *M. guttatus* allele increases seed viability, negative values indicate the *M. tilingii* allele increases seed viability.

[b]Relative seed viability of the alternative homozygotes. The value reported is the reduction in backcross seed viability for the "incompatible" genotype relative to the "compatible" genotype.
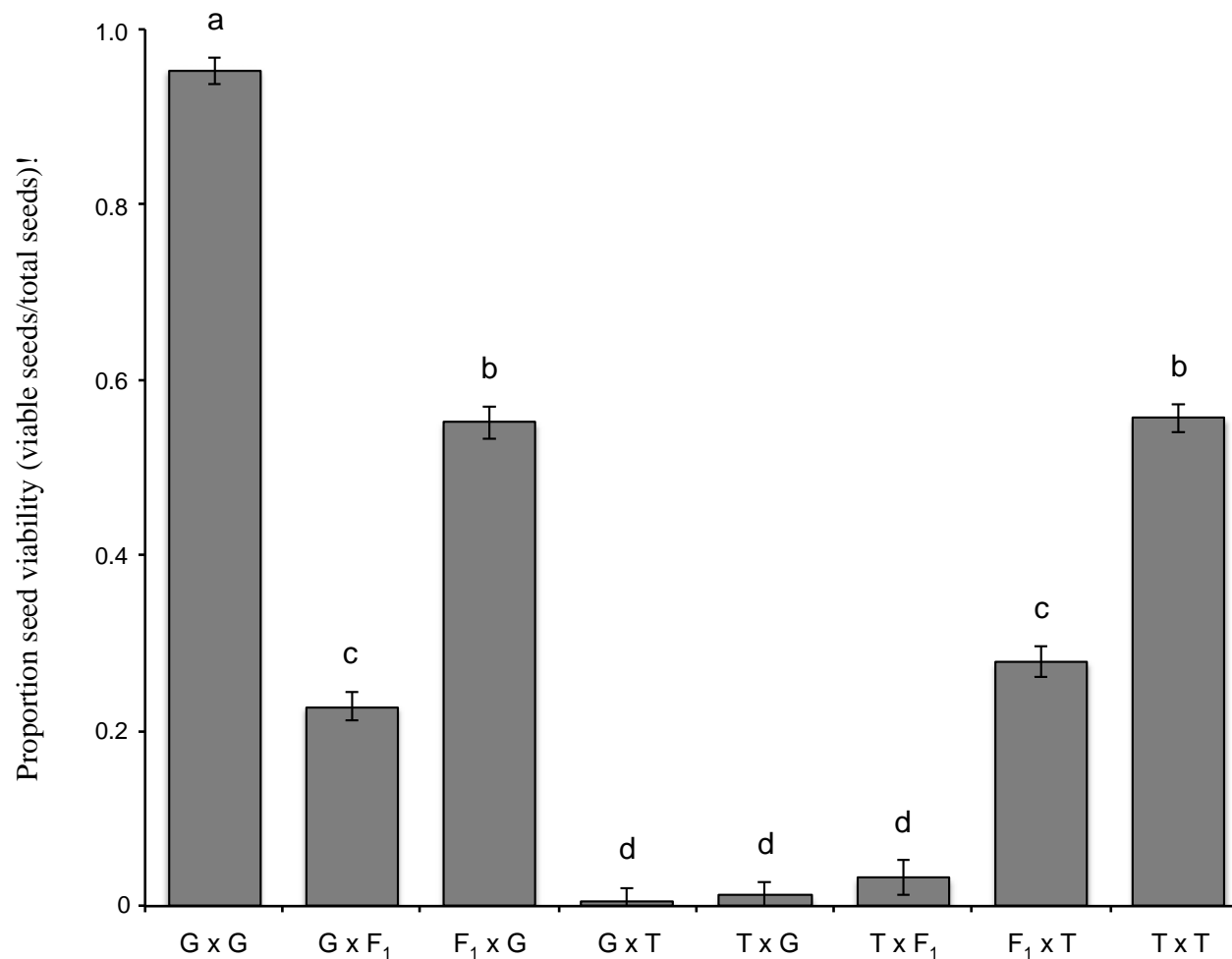
Figure 1
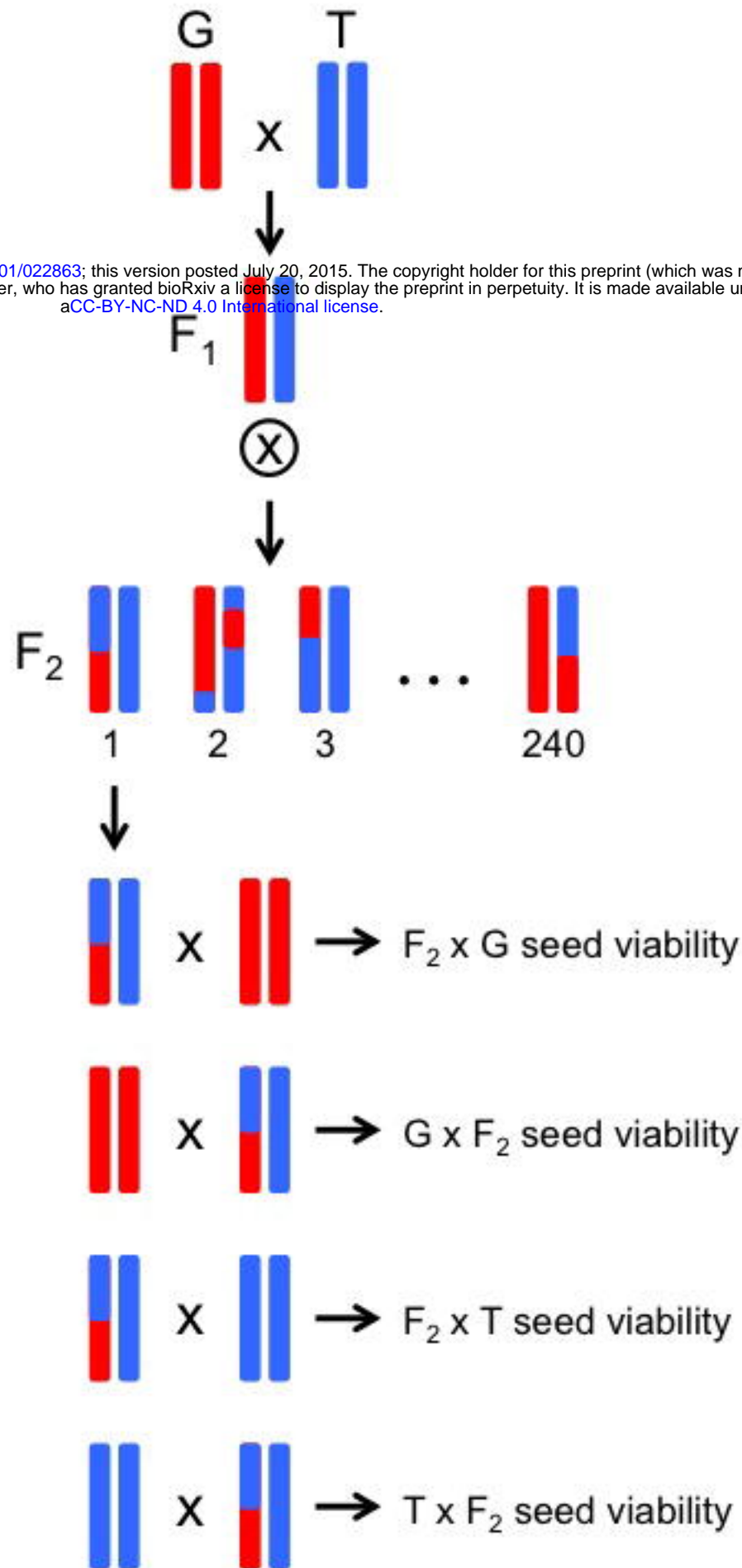
Figure 1. Crossing design to determine the genetic basis of seed lethality between *M. guttatus* and *M. tilingii.* We formed F_1 hybrids by intercrossing *M. guttatus* (G, red) and *M. tilingii* (T, blue), and self-fertilized a single F_1 to generate the F_2 generation. We then reciprocally crossed F_2 hybrids to each parental line and assessed seed viability. The diagram shows a single pair of chromosomes for each individual with the maternal parent listed first.
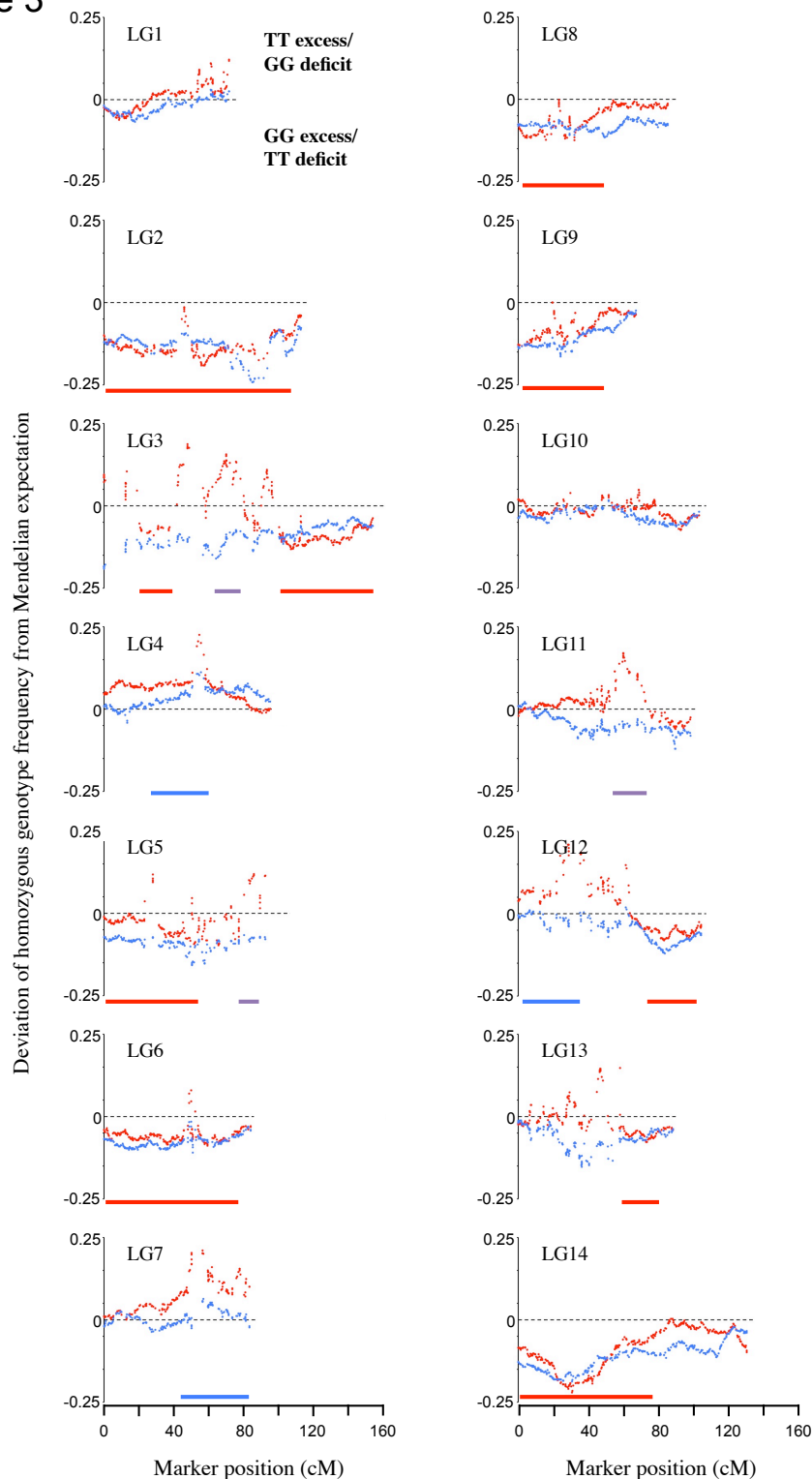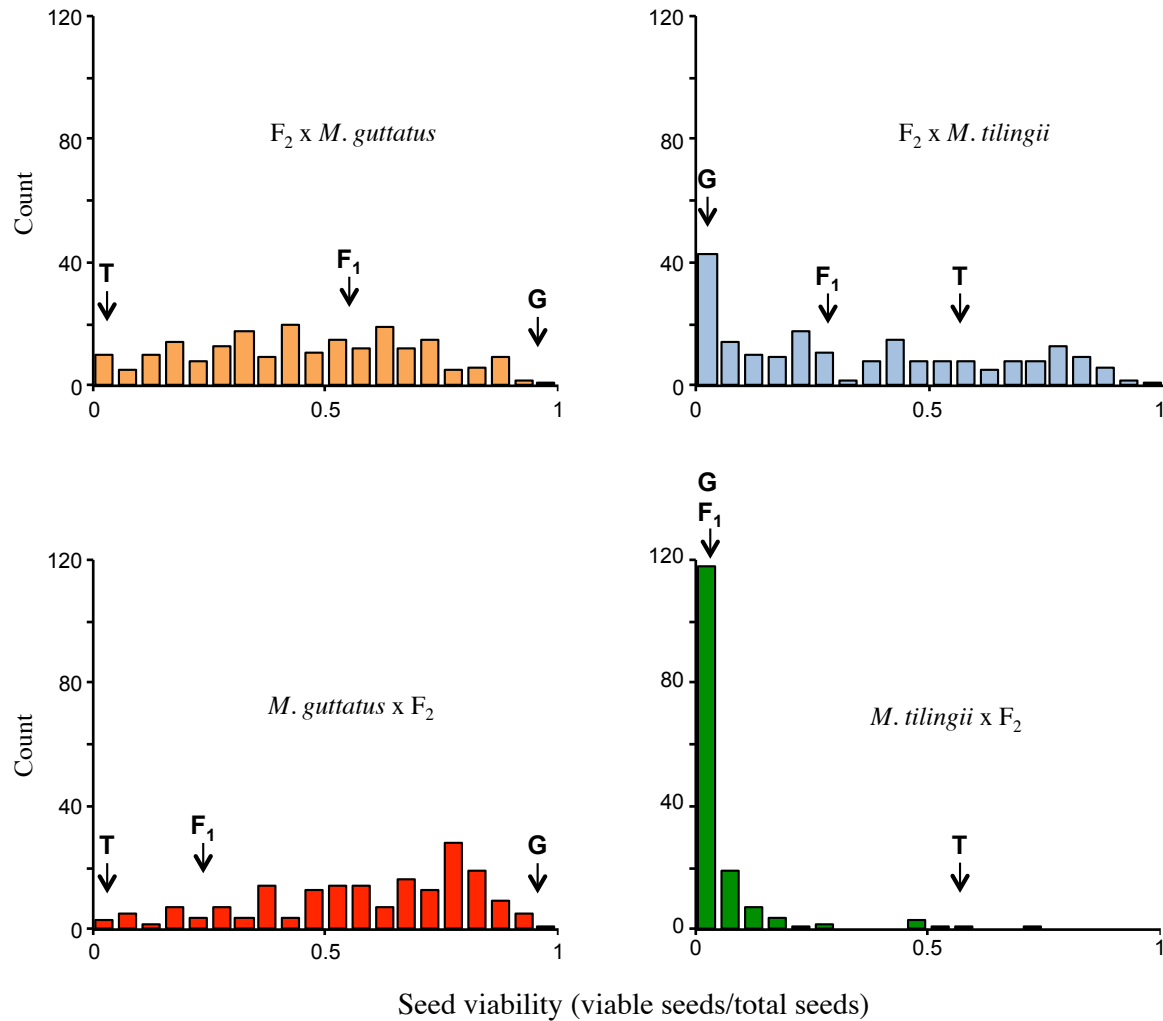
Figure 2

Figure 2. Mean seed viability varies among experimental crosses. Seed viability is highest from crosses within parental lines (*M. guttatus*: G x G, *M. tilingii*: T x T) and lowest in reciprocal interspecific crosses (G x T and T x G). Seed viability is generally intermediate in the four classes of backcross progeny (G x F_1, F_1 x G, T x F_1, F_1 x T). For all crosses, the maternal parent is listed first. Bars indicate standard errors and lower case letters above show significant differences by Tukey's HSD test.

Figure 3. Transmission ratio distortion across the *M. guttatus* x *M. tilingii* linkage map. For each of the 14 linkage groups, homozygous parental genotypes at marker loci are shown in red (*M. guttatus*, GG) and blue (*M. tilingii*, TT). For each marker, the vertical positions of the red and blue dots show the deviations of genotype frequencies from the Mendelian expectation of 0.25. Note that the *M. tilingii* homozygote (TT) deviations are plotted directly [deviation = (frequency of TT) – 0.25] and the *M. guttatus* (GG) deviations are plotted as negative [deviation = – ((frequency of GG) – 0.25)]. As a result, the values above the zero line show excesses of TT homozygotes or deficits of GG homozygotes and values below the zero line show excesses of GG homozygotes or deficits of TT homozygotes. Horizontal bars beneath linkage groups indicate regions with excess GG (red), TT (blue), or heterozygotes (purple).

Figure 4

Figure 4. Frequency distributions of F$_2$ hybrid seed viability phenotypes for the four experimental crosses. F$_2$ x *M. guttatus* (average = 0.45, *N* = 215), *M. guttatus* x F$_2$ (average = 0.58, *N* = 190), F$_2$ x *M. tilingii* (average = 0.36, *N* = 207), and *M. tilingii* x F$_2$ (average = 0.05, *N* = 158). Approximate values for parental (*M. guttatus*: G, *M. tilingii*: T) and F$_1$ means are indicated with arrows in each histogram.

Figure 5
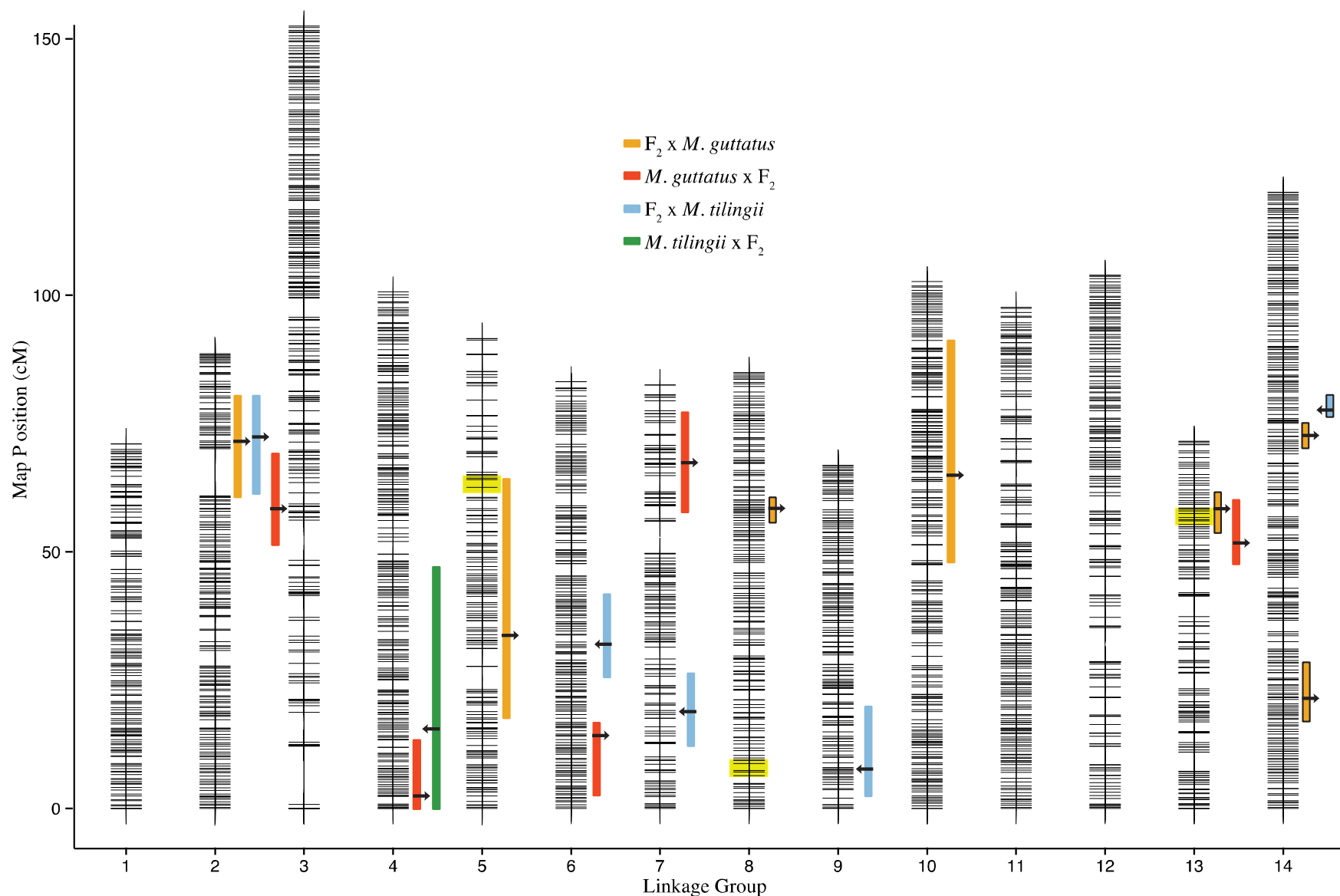
Figure 5. *M. guttatus* x *M. tilingii* genetic map and seed viability QTL. Each marker is shown as a horizontal line and the three hypothesized inversions are indicated by yellow shading. Arrows show the QTL LOD score peak locations and directions of additive effects (right pointing arrows indicate *M. guttatus* alleles increase the trait value, left pointing arrows indicate *M. tilingii* alleles increase the trait value). QTL bars show the 1.5 LOD drop confidence intervals and those outlined in black are significant at the genome-wide level (all others are significant at the chromosome level).