Pervasive Antagonistic Interactions Among Hybrid Incompatibility Loci

Rafael F. Guerrero[1], Takuya Nakazato[1], Chris Muir[2], Sarah Josway[3], and Leonie C. Moyle[2,4]

1. Department of Biology, Indiana University, Bloomington IN U.S.A. 47405
2. Center for Biodiversity, University of British Columbia, Vancouver BC, Canada.
3. Environmental Science and Policy program, George Mason University, Washington D.C.
4. Corresponding author

ABSTRACT

Species barriers, expressed as hybrid inviability and sterility, are often due to epistatic interactions between divergent loci from two lineages. Theoretical models indicate that the strength, direction, and complexity of these genetic interactions can strongly affect the expression of interspecific reproductive isolation and the rates at which new species evolve. Nonetheless, empirical analyses have not quantified the frequency with which loci are involved in interactions affecting hybrid fitness, and whether these loci predominantly interact synergistically or antagonistically, or preferentially involve loci that have strong individual effects on hybrid fitness. We systematically examined the prevalence of interactions between pairs of short chromosomal regions from one species (*Solanum habrochaites*) co-introgressed into a heterospecific genetic background (*Solanum lycopersicum*). We used a diallel crossing design to generate most pairwise combinations of 15 chromosomal segments from *S. habrochaites* in the background of *S. lycopersicum* (*i.e.*, 95 double introgression lines). We compared the strength of hybrid incompatibility (either pollen sterility or seed sterility) expressed in each double introgression line to the expected additive effect of its two component single introgressions. We found that: epistasis was common among co-introgressed regions; epistastic effects were overwhelmingly antagonistic (*i.e.*, double hybrids were less unfit than expected from additive single introgression effects); and, epistasis was substantially more prevalent in pollen fertility compared to seed fertility phenotypes. Together, these results indicate that high-order interactions frequently contribute to postzygotic sterility barriers in these species. This pervasive epistasis leads to the decoupling of the patterns of accumulation of isolation loci and isolation phenotypes, and is expected to attenuate the rate of accumulation of hybrid infertility among lineages over time (*i.e.*, giving diminishing returns as more reproductive isolation loci accumulate). This decoupling effect might also explain observed differences between pollen and seed fertility in their fit to theoretical predictions of the accumulation of isolation loci, including the 'snowball' effect.

INTRODUCTION

Intrinsic postzygotic isolation (hybrid inviability and sterility that occurs independently of the environment) is often due to deleterious genetic interactions between loci that have functionally diverged during the evolution of new species (*i.e.*, Dobzhansky-Muller incompatibilities, or DMIs; [1]). Several models of this process assume that individual DMIs are due to epistasis between one locus in each diverging lineage (*i.e.*, pairwise genetic interactions), and that each DMI contributes independently to the expression of hybrid incompatibility between diverging species (although some models can relax the first assumption; see [2-4]. If hybrid dysfunction is due to more complex interactions among loci, however, there can be important consequences for the temporal accumulation of species barriers and the number of loci required to complete speciation [4-6]. In the simplest case, if epistasis between different hybrid incompatibility loci is antagonistic (*i.e.*, if the combined effect of two different DMIs is less than expected based on their individual effects on hybrid incompatibility), then a greater time to speciation is expected and correspondingly more loci are required. Conversely, if epistasis between different conspecific loci is typically synergistic (*i.e.*, the combined effect on hybrid incompatibility is greater than expected based on individual effects), fewer DMIs will be required for the expression of complete reproductive isolation, with a correspondingly shorter time to speciation. In the latter case, although the initial observation of incompatibility phenotypes requires more than one substitution per lineage, hybrid incompatibility can be expressed rapidly between two diverging species once substitutions have begun to accumulate. The prevalence of epistasis, and whether this epistasis is synergistic or antagonistic, can therefore be critical in determining patterns

and rates of evolution of isolation between diverging lineages by governing the accumulation dynamics of alleles that contribute to species barriers (*e.g.*, [4, 5]; and see below).

Empirically, however, very little is known about the nature of epistasis among different loci contributing to hybrid incompatibility ('complex epistasis', c.f. [7]) and whether these interactions typically act to enhance or retard the expression of hybrid incompatibility between species. Some evidence suggests that non-additive interactions affecting the fitness of hybrids might be common. In *Drosophila*, for example, quantitative trait loci (QTL) for male sterility show evidence of complex epistasis. In particular, several individual genomic regions are simultaneously required for the expression of some male sterility phenotypes (*e.g.*, [8]; other evidence is reviewed in [1]), consistent with synergism (*i.e.*, greater-than-additive effects) between conspecific loci for the expression of hybrid sterility. However, epistasis has been difficult to assess in many genetic analyses of quantitative trait loci because early recombinant generations have low power to detect these interactions [9]. The most promising method of directly estimating epistasis among loci is to subtract background effects by isolating two individual QTL in an otherwise isogenic background (*e.g.*, [10]). For example, this approach revealed less-than-additive effects for quantitative and ecophysiological traits in tomato, consistent with antagonistic interactions between different conspecific loci [11, 12]. Similarly, studies in microorganisms have used serial pairwise combinations of target loci to examine the size and direction of pairwise epistasis on the phenotypic effects of, for example, double deletion strains in yeast [13], and metabolic flux mutants in yeast and *E.*

coli [14]. Nonetheless, the serial combination of many pairs of conspecific loci has yet to be used to assess prevalence and direction of interactions influencing hybrid incompatibility [15].

The goal of this study was to assess the prevalence and direction of genetic interactions between different chromosomal regions from one species when combined pairwise in the background of a second species (Figure 1), focusing specifically on the consequences of pairwise conspecific epistasis for the expression of hybrid incompatibility. We selected fifteen chromosomal regions for inclusion in the study (Table 1), drawing from a set of fifteen introgression lines (ILs) between two plant species in the genus *Solanum* Section *Lycopersicon* (the tomato clade). Each IL contains a unique short chromosomal region from the wild species *S. habrochaites* ('*hab*') introgressed into the otherwise isogenic genetic background of the domesticated tomato, *S. lycopersicum* ('*lyc*') ([16]; see [17] for a previous summary). Several of the chosen introgressions have effects on either pollen or seed sterility—that is, they carry a QTL for partial interspecific sterility ([17], and see Results). We crossed 15 unique ILs to generate a collection of nearly 100 double introgression lines (DILs), which allow us to study the interactions among conspecific introgressed regions. We compared fertility phenotypes in DILs to those of their corresponding parental ILs to infer frequency, magnitude and direction of epistasis among conspecific introgressions.

We found that complex epistasis was common in both pollen and seed fertility phenotypes, although significantly more frequent for introgressions affecting pollen

sterility. These observed interactions were overwhelmingly antagonistic, whereby the combined effect of pairwise introgressions produced a less severe effect on fitness than predicted from individual effects. For pollen sterility phenotypes, we found that some chromosomal regions were considerably more prone to interaction than others: most complex epistasis occurred among introgressions with significant individual effects. In contrast, we found no evidence of a similar pattern for seed sterility.

This pervasive antagonism has critical implications for the predicted rate and pattern of accumulation of hybrid incompatibility between these lineages, and for future empirical studies of the evolution of post-zygotic isolation. We infer that these patterns of interdependency result from unexpectedly abundant interactions among genes underlying the phenotypes studied. This implies a violation of key assumptions—regarding the independence among different pairs of incompatibility loci—made by many speciation models of this process. Specifically, we surmise that high connectivity among loci that contribute to pollen sterility explains the observed flattened accumulation of (pairwise) pollen incompatibilities in this clade [18]. In addition to explaining this apparent lack of "snowball" effect, our results also imply that standard QTL mapping approaches can underestimate the number of incompatibility-causing interactions present in systems with high levels of epistasis.

METHODS

Experimental design

Using a library of 71 ILs that covered 85% of the *hab* genome in the background of *lyc*, we had previously detected 8 and 5 QTL for hybrid pollen and seed sterility, respectively, between these species [17]. Here we chose 15 of these 71 introgression lines (or ILs) to use in the construction of double introgression genotypes. Each line contained a single, marker delimited chromosomal region from *hab* introgressed into the background of *lyc*. In a previous study [17], we inferred that ten of these lines carried QTL for hybrid incompatibility (five with significant pollen, and five with significant seed, sterility effects) when compared to pure parental genotypes. The remaining five lines contain *hab* introgressions with no previously detected phenotypic effect on pollen or seed sterility; these additional lines were chosen based on their distribution in the genome (each was located on a different chromosome), and to be roughly comparable in terms of introgression length (as a percent of *hab* genome introgressed) as the partially sterile lines (Table 1). Nonetheless, because the detection of sterility effects can be dependent on experimental power and sensitive to environmental conditions, all 15 lines were reassessed for sterility effects in the current experiment (see Results).

Generation of Double Introgression Lines (DILs)

Generation of lines with two introgressed regions (DILs) has been previously described [15]. Briefly, a half diallel cross was performed among our 15 lines to combine each introgression with each other introgression (Figure 1), for a total of 105 unique pairwise combinations of the 15 regions. Heterozygote F_1 DILs from each of these 105 pairwise IL combinations were selfed to generate F_2 seeds ('DIL family'), and up to 120 F_2s in each DIL family were grown and genotyped for two markers, one located in each introgressed

region (Moyle and Nakazato 2009). Species-specific co-dominant PCR-based markers were used to diagnose the presence of each introgression (*hab* versus *lyc* alleles), for a total experimental size of 105 DIL families x 120F_2s x 2 markers = 25,200. From these families, individuals that were confirmed to carry two *hab* alleles at both introgression locations (*i.e.*, double *hab* homozygotes) were retained. For the current experiment, these lines were selfed to produce double *hab* homozygous seed (Figure 1, Table 2). In seven DIL families, plants failed to produce seed from self-fertilizations, despite repeated hand-pollinations; these genotypes were necessarily excluded from the experiment (Supplemental Table 1). Finally, introgression lines were selfed to produce new homozygous IL seeds. In total, for this experiment we generated and assessed 106 unique genotypes: the isogenic parental *lyc*, 15 homozygous ILs, and 95 (of 105 possible) homozygous DILs.

Assessment of fertility phenotypes

Each unique genotype was grown and assessed in replicate (total experiment size = 721 individuals; Supplemental Table 2) in a fully randomized common garden (greenhouse) experiment at the IU Biology greenhouse facility. Cultivation methods have been previously described (Muir and Moyle 2009). Briefly, all experimental seeds were germinated under artificial lights on moist filter paper, transferred to soil post-germination (at the cotyledon stage), and repotted into 1gallon pots at 4-6 weeks post-germination. Due to mortality and/or fecundity variation among individuals, total experimental size of individuals that reached reproductive maturity was N=1456. Because of these inviability or fertility effects, not all DIL genotypes had sufficient

individuals (*i.e.*, three or more) to be included in analyses (see Supplemental Table 1, Results and Discussion).

Seed fertility (seeds per fruit) was determined by quantifying seed production from self-pollinated fruits ('self-seed set'), as with previous QTL analyses [19]. At least two flowers per plant were allowed to produce fruit via selfing; when fruits did not develop automatically, flowers were self-pollinated manually to ensure that floral morphology was not responsible for lack of seed set. Upon maturation, individual fruits were harvested, seeds extracted by hand, and seed fertility determined by counting the number of visible seeds from each fruit. Average seed per fruit for each plant was used to generate self-seed set estimates for each introgression genotype and the *lyc* control parent.

Pollen fertility (quantified as the proportion of fertile pollen) was estimated on two unopened flowers on each plant, as previously described [17]. Briefly, all pollen (the entire anther cone) from each target flower was collected into lactophenol-aniline blue histochemical stain, homogenized, and a known sub-sample of homogenate used to count inviable and viable pollen grains using microscopy. Pollen inviability was indicated by the absence of a stained cytoplasm, a conservative measure of pollen infertility [20].

Seed set phenotypes may be affected by three components: ovule viability, pollen fertility, and gamete compatibility in the zygote. Therefore, pollen sterility in our lines may influence estimates of seed sterility. In fact, we found a weak but significant

correlation between pollen fertility and self-seed set across our experimental lines (p =0.0014; negative-binomial GLM). However, pollen fertility explained only about 5% of variation in seed fertility (pseudo-R^2 = 0.046), which is consistent with previous studies [17, 19]. We removed this small pollen effect by carrying out our analyses of self-seed set on their residuals from the regression on pollen fertility. For completeness, we carried out the same analyses on the uncorrected self-seed set values and found no qualitative differences (results not shown).

Analysis

Estimation of sterility in ILs

To quantify the individual effect of introgressions on hybrid fitness we tested for differences between the phenotypes of the ILs and those of the background parent, *lyc*, using generalized linear models (GLMs) with a single fixed effect. For pollen fertility, we carried out an arcsine square root transformation to improve normality of residual distributions. Seed fertility data were analyzed assuming a negative binomial error distribution (as appropriate for count data). Fitness of each IL was tested independently against *lyc*, allowing slightly different significance thresholds for each fitness component. For corrected self-seed set data, we took 'sterile' ILs as those with $P < 0.01$ (which implied a false discovery rate of 4.5%; Table 1). We were slightly more stringent with significance for pollen fertility, for which we took 'sterile' ILs at a false discovery rate (FDR) of 1% (which corresponded to $P < 10^{-4}$; Table 1). The significance threshold of individual IL effects on sterility does not affect our key findings (see *Results*, Suppl Fig 1).

Inference of epistatic effects in DILs

We used a maximum likelihood approach to infer the presence and strength of epistasis from DIL phenotypic data. This approach allowed us to compare two alternative fitness models, choose the one that best fit the data for each DIL, and infer the amount of epistasis, within a single analysis. The two fitness models we compared are additive (where fitness effects in DILs are the sum of the fitness effects of the two parental ILs), and multiplicative (in which DIL fitness is the product of parental IL fitness).

First, we define fitness in DILs and their IL parents as the phenotype (pollen or seed fertility) relative to the mean phenotype in *lyc*. For each DIL carrying introgressions i and j, we assume that relative fitness, w_{ij}, is normally distributed with mean and variance, $\mu_{(\psi)}$ and $\sigma^2_{(\psi)}$, determined by the fitness model ψ (=$\{a, p\}$), relative fitness of the parental ILs (w_i and w_j) and ε_{ij}, the epistasis parameter. Under the additive fitness model, $\mu_{(a)}$ and $\sigma^2_{(a)}$, are the mean and variance of the vector $\mathbf{v}^{(a)}_{ij} = \mathbf{w}_i + \mathbf{w}_j - 1 + \varepsilon_{ij}$, where $\mathbf{w}_i$ and $\mathbf{w}_j$ are vectors holding k observed values for w_i and w_j (that is, there are k biological replicates of each parental IL). Similarly, under the multiplicative fitness model the parameters $\mu_{(p)}$ and $\sigma^2_{(p)}$ are the mean and variance of the vector $\mathbf{v}^{(p)}_{ij} = (1 + \varepsilon_{ij})(\mathbf{w}_i \cdot \mathbf{w}_j)$. The likelihood function for a single DIL under model ψ is $L_\psi = \prod_{n=1}^{N} \mathbb{P}(w_{ij}^n = x \mid \mu_{(\psi)}, \sigma^2_{(\psi)})$, the product of the normal probability density over N observations of relative fitness. In both models, the likelihood function has the epistasis term, ε_{ij}, as a single free parameter.

We maximized likelihood for each model numerically (using the *mle* function from the stats4 package in R; [21]). Because we have different sample sizes for the phenotypes of parental ILs (*i.e.*, vectors $\mathbf{w}_i$ and $\mathbf{w}_j$ have different lengths), we subsampled without replacement each set of observations to the minimum parental pair of each DIL (that is, took $k_m = \min(k_i, k_j)$ elements from each vector). We repeated this subsampling 100 times, and took the average of $\mu_{(\psi)}$ and $\sigma^2_{(\psi)}$ from all of them in order to obtain estimates that use all our data and minimize the effect of pairing observations from parental ILs.

To choose the best model, we compared the resulting maxima using the Bayesian Information Criterion (BIC; [22]). Additionally, to confidently infer cases of epistasis, we required that our best-fit model have a better BIC score than that of two null models (additive and multiplicative fitness models with $\varepsilon_{ij} = 0$). Finally, we obtained statistical support for our inference of epistasis through a likelihood-ratio test of the best-fit model against the best null model (assuming a chi-squared distribution of the likelihood surface and 1 degree of freedom). We applied a correction for multiple testing within each phenotype, taking highly significant epistatic interactions at a 1% FDR. We visualized significant interactions between ILs as a network using the *igraph* [23] and *ggraph* packages [24]. In these networks, each node represents an IL (specifically, the introgressed chromosomal region within that IL), while edges represent an epistatic interaction between two ILs. All scripts used, written in R [21], are available in the Supplement.

Our approach is similar to the one described by Gao *et al.* [25] for the classification of epistasis. In contrast to that previous work, however, we have not considered a third fitness model (the 'minimum' model) in our reported estimations. The 'minimum' model aims to assess statistical independence among mutations, however it leads to difficulty detecting certain types of biological interactions (*e.g.*, 'masking' and 'co-equal' interactions; Mani et al 2008), and might not be adequate when testing epistasis among loci with known similar functions [25]. Moreover, when our data are re-analyzed including this model, there is no change in the direction and only a slight decline in the frequency of significant epistatic interactions (results not shown).

For simplicity, throughout this manuscript we refer to individual effects of QTL and interaction effects between loci –although strictly discussing the phenotypic consequences of entire introgressions. The estimated length of some of our introgressions (Table 1) suggests that they could carry up to several hundred genes each. Accordingly, our study more accurately is an analysis of the frequency and strength of *trans* interaction effects between short conspecific chromosomal regions and is unable to differentiate, for example, the single versus combined phenotypic effects of multiple genes within any given introgression. Depending upon genetic structure of these QTL at finer chromosomal scales, the specific genetic interpretation of types and dynamics of substitutions would change. We note, however, that most observations of complex epistasis outside of microbial studies have come from examining *trans* interactions between chromosomal regions rather than individual genes (Drosophila, [26]; Tigriopus, [27]; Mimulus, [28]). In addition, prior QTL mapping studies in this and other *Solanum*

species pairs indicate that the genetic basis of hybrid incompatibility is modestly, not highly, polygenic [17, 19]; therefore, unless individual loci underlying sterility QTL are unusually (and unexpectedly) highly clustered within the genome, these data suggest most introgressions do not harbor multiple different sterility-causing loci.

RESULTS

We found three prevalent patterns of conspecific epistasis for hybrid incompatibility. First, epistatic interactions between conspecific loci are common, especially among introgressions with larger individual effects on fertility. Second, these interactions are predominantly antagonistic (*i.e.*, less-than-additive) in their effects on hybrid incompatibility phenotypes. Third, both these patterns are significantly more pronounced for pollen than seed sterility effects.

For seed sterility, we found evidence of non-independence between introgressions for about one fifth of DILs analyzed (Figure 2). Out of 94 DILs, 18 showed departures from fitness models of independence ($P < 0.01$, FDR=4.4%). Interactions were qualitatively more common for pollen fertility phenotypes (Figure 3), where 25 out of 95 analyzed DILs were best fit by epistatic fitness models ($P < 0.01$, FDR=3.5%). Remarkably, very few of these interactions were synergistic: in seed phenotypes, only five DIL hybrids were more sterile than expected by their IL parents' individual effects. This pattern was considerably more dramatic for pollen fertility, where all detected epistatic interactions were antagonistic.

We found no differences across phenotypes in the magnitude of epistasis or the predicted fitness model. Epistatic interactions caused, on average, a 37% excess in pollen fitness when compared to the fitness predicted from the joint effect of parental ILs. In seed sterility, we found similar effect sizes in synergistic and antagonistic interactions (34% and 45%, respectively). For both studied phenotypes, fitness in most DILs was better described by an additive model of epistasis than a multiplicative one (15 out of 20 in seed; 21 out of 26 in pollen). Across DILs with no significant epistatic effects, we found that an additive model was favored over multiplicative more frequently in seed (44 out of 76 DILs) but not in pollen phenotypes (16 out of 70).

Epistatic interactions for fitness were common across the genome: most parental ILs were involved in at least one interaction with another conspecific introgression (all except *LA3975* for seed set, and *LA3931* for pollen fertility). In pollen phenotypes, however, interactions more frequently involved introgressions with individual sterility effects. This bias becomes more apparent in highly significant interactions (14 DILs at FDR=1%; bordered panels in Figure 3), which always happened in DILs with at least one pollen-sterile IL parent genotype. Moreover, most of these highly significant interactions were between two parental ILs with significant individual effects on pollen fertility. This pattern was confirmed when we represented the distribution of highly significant interactions as a network (Figure 4). In pollen, ILs formed a single interconnected cluster in which each IL had 1.87 highly significant interactions on average. Sterile ILs for pollen had more interactions (*i.e.*, higher node degree; 3.5 on average) than non-sterile introgressions (0.5 on average). In contrast, the overall network of interactions in seed

phenotypes appeared sparser than in pollen phenotypes, with three separate clusters of up to 4 ILs. The seed sterility network showed no direct connections between two sterile ILs, and sterile ILs had similar number of interactions than non-sterile ILs (0.8 on average for both groups). The number of interactions of ILs was also correlated with their mean effect size in pollen but not seed phenotypes (Suppl Fig 1), suggesting that these patterns are robust to our arbitrary cut-off for significance of individual IL sterility (FDR of 1% for pollen and 4.5% for seed).

Our results suggest that the hybrid sterility effects of individual introgressions are frequently non-independent, especially for loci affecting hybrid pollen fertility. Individually, each of the observed interactions could be consistent with one of three scenarios. First, each parental IL could carry one half of the same incompatibility (*i.e.*, ILs carry both loci involved in a pairwise DMI), resulting in phenotype rescue (*i.e.,* complementation) in the DIL when both loci are introgressed. This is a parsimonious interpretation for antagonistic interactions between sterile and non-sterile ILs (found in 4 cases for seed, and 4 for pollen phenotypes), but not for antagonistic epistasis between two sterile ILs (11 cases in pollen phenotypes) because it would not be consistent with the expected pattern of asymmetry of DMIs [29]. Second, the parental ILs could carry loci involved in separate pairwise DMIs, and the interaction could be the result of complex epistasis between these two incompatibilities. This seems likely to be the case for interactions between ILs with individual sterility effects (which are carriers of putative incompatibility factors). Third, the introgressions are part of a third-order DMI in which a specific combination of three alleles is necessary for sterility. The presence of

a higher-order DMI is consistent with all antagonistic interactions, and is a likely explanation for synergistic interactions between two ILs with no observed individual effects on fitness (2 cases in seed phenotypes).

While our approach does not allow us to distinguish between these three alternatives for most individual DILs, taken together, the structure of these observed interactions implies a complex landscape of interdependence among sterility loci. Interactions in seed suggest at least two higher-order incompatibilities, composed of three and four loci (if we focus on highly significant edges in Figure 4). In the pollen network, ILs form a single cluster of highly significant interactions, and most of these interacting ILs are epistatic with at least two other introgressions (8 out of 10 ILs; Figure 4). Speculating on the genetic details underlying these patterns of interdependence is not straightforward. One extreme interpretation is that this large module represents a single incompatibility that involves ten loci (*i.e.*, a tenth-order DMI). Alternatively, each of these antagonistic interactions represents a pairwise hybrid incompatibility, which implies that pollen-sterility loci are frequently involved in more than one DMI (since many ILs show more than one interaction). This latter interpretation is also consistent with our observation that sterile ILs show more interactions than non-sterile ones, since each interaction in which they engage could contribute to their observed fitness (*i.e.*, ILs involved in more DMIs would, on average, show larger fitness effects). However, that a locus can be involved in more than one hybrid incompatibility supposes a violation of a central assumption in most speciation models of this process (an issue that we examine below).

DISCUSSION

By examining multiple pairwise combinations of conspecific loci, co-introgressed into a heterospecific background, we have shown evidence for pervasive non-independence among introgressions that is likely due to a combination of pairwise, higher-order incompatibilities, and complex epistatic effects. This general observation suggests that hybrid phenotypes might frequently be the product of epistatic interactions among several factors. Complex epistasis was common in both pollen and seed fertility phenotypes, although significantly more frequent for introgressions affecting pollen sterility. For pollen sterility phenotypes, we found that some chromosomal regions were considerably more prone to interaction than others: most complex epistasis occurred among introgressions with significant individual effects. In contrast, we found no evidence of a similar pattern for seed sterility. Finally, the observed interactions were overwhelmingly antagonistic, whereby the combined effect of pairwise introgressions produced a less severe effect on fitness than predicted from individual effects. Overall, our results indicate that the hybrid sterility effects of individual heterospecific introgressions are frequently non-independent, especially for loci affecting hybrid pollen fertility, and that the resulting interactions are largely antagonistic (less-than–additive). This observed pervasive antagonism could have important implications for understanding both the predicted patterns and the mechanism of accumulation of hybrid incompatibility between lineages, as well as for the detection of such loci in QTL mapping efforts.

First, these observations suggest an important role for evolutionary dynamics that elevate the opportunity for complex epistatic interactions, in the fixation of postmating

reproductive isolation alleles. The non-independence observed among incompatibility-affecting loci indicates that the 'genetics of isolation' (especially pollen sterility) in this species pair is enriched for loci engaged in complex genetic interactions. This high connectivity implies that the fixation of underlying substitutions may have been driven by evolutionary processes, such as antagonistic coevolution, that enrich for these interactions. For reproductive traits, this coevolution is likely to arise from antagonistic male-female interactions during reproduction—a well-studied phenomenon in animals [30, 31]. Although sexual conflict has been less prominently assessed in plants, some studies indicate it can be an important force in pollen-pistil interactions [32] and models suggest that sexual antagonism over resource investment in seeds could contribute to the evolution of seed sterility among diverging lineages [33] and the evolutionary accumulation of pollen-sterility loci via cytoplasmic male sterility (reviewed in [34]). Thus, our results contribute—albeit indirectly—to a budding view of sexual conflict as a potentially powerful driver of plant postmating reproductive barriers.

Second, the differences in connectivity between seed and pollen phenotypes could also explain previously observed differences in the pattern of accumulation of hybrid incompatibilities (or 'snowball' effect) in this group. Moyle & Nakazato [18] showed that, while the accumulation of self-seed set QTL among increasingly divergent *Solanum* lineages follows the predicted snowball (*i.e.*, it is faster than linear), this does not seem to be the case for QTL involved in pollen fertility. At first, this linear accumulation of pollen fertility QTLs may seem at odds with our observation of high connectivity for the phenotype, because theory predicts a faster increase in the number of incompatibilities

with increasing connectivity among loci. In a simple version of the snowball model, the number of incompatibilities is roughly proportional to the square of K, number of genetic differences between two species, and p, the probability that any pair of these is incompatible. Therefore, factors that increase K or p should increase the speed of the snowball. For instance, antagonistic coevolution may result in an accelerated rate of substitution (*i.e.*, increased K; [35, 36]). Similarly, the probability that any pair of genetic differences is incompatible (p) is elevated when there are more opportunities for incompatibilities to arise, such as when the phenotype is controlled by a highly connected network of genes. Given these predictions, our results would suggest that the number of detected pollen sterility loci should snowball faster than seed sterility loci (yet, they do not).

A possible explanation for this discordance lies in the assumption that the number of loci involved in incompatibilities represents an adequate proxy for the total number of incompatibilities. Because of the challenges of identifying underlying molecular loci, empirical studies have used the number of sterility QTL as an indicator of the number of DMIs, with the underlying assumption that each QTL is involved in only one (pairwise) incompatibility. Under a pairwise snowball model, each locus is involved in $(K\text{-}1)p$ incompatibilities on average [3]. Since p (the probability of an incompatible pair) is assumed to be very small compared to K, such that $(K\text{-}1)p << 1$, each QTL is therefore likely to be involved in just one incompatibility. In cases with increased p, however, enough genetic differences could accumulate such that $(K\text{-}1)p \geq 1$ (*i.e.*, loci are, on average, involved in one or more incompatibilities). After this "saturation" point, each

new genetic difference will continue to generate (K-1)p new incompatible interactions, but the number of loci involved in hybrid incompatibility increases just by one. We illustrate this behavior by simulation of the accumulation of pairwise incompatibilities (Figure 5). These simulations (carried out in R; [21]) simply count the number of pairwise incompatibilities between two lineages as substitutions fix, allowing each new substitution to be incompatible with any of the existing ones at that time (with probability p). As Figure 5 shows, incompatibilities (*i.e.*, interactions that affect hybrid fitness) in saturated systems continue to snowball while the number of individual loci contributing to hybrid incompatibility starts to increase linearly, breaking the relationship between sterility-affecting loci and DMIs.

Our results on pollen fertility, together with the lack of snowball previously found [18], suggest saturation for this fitness component. Several pollen-sterile ILs showed multiple significant epistatic interactions, suggesting that at least some of these sterility loci are involved in more than one hybrid incompatibility; moreover, the number of loci involved in pollen sterility appears to be accumulating linearly in this clade [18]. These two observations are consistent with a scenario in which most genetic differences at loci underlying pollen fertility are already involved in at least one incompatible interaction, and the accumulation of new DMIs does not result in the increase of new unique sterility loci at a similar rate.

At first pass, our inference of saturation does not seem consistent with rough estimates of p for this clade. Previously, Moyle & Nakazato [18] estimated $p \sim 10^{-7}$ for tomatoes,

which is roughly equivalent to *p* estimates reported for *Drosophila* (10^{-7}– 10^{-8}, [3]). If *p* is not particularly large, then how could saturation occur in this system? Three factors may contribute to explaining this apparent discordance. First, the number of relevant genetic differences – *K* at pollen loci only rather than across the whole genome– could be elevated due to antagonistic coevolution (see above). Second, high modularity in the genome may allow for high *p* within a module even when there is a low genome-wide mean probability of incompatibility. Our results suggest saturation for the pollen module, which implies large *p* among pollen loci only – not among the genome as a whole. Third, high variance in the probability of incompatibility across loci could result in apparent saturation. For instance, some loci could be highly connected (i.e., be network 'hubs'), therefore having a much higher probability of being involved in a DMI than others. Systems with high variance in their gene networks may appear saturated if hubs are involved in two or more DMIs. While modularity itself is expected to have little effect on genome-wide *p* and the rate of accumulation of DMIs [3], the importance of the structure of networks for the absolute size and the variance of *p* requires further theoretical work.

The distribution of interactions across sterility loci suggests that the pollen fertility phenotype might be close to saturation for hybrid incompatibility loci (the nodes in our network; Figure 4), even while the DMIs in which they are involved (the edges in our network; Figure 4) continue to snowball. If pollen fertility is indeed saturated, it follows that pollen DMIs have been accumulating in tomatoes at a rate much faster than previously thought. This inference is similar to that of recent simulation work on the accumulation of DMIs in RNA folding, which showed that the number of pairwise DMIs

will not snowball in systems where these interactions are converted into high-order DMIs (*i.e.*, involving alleles at three or more loci) as a result of later substitutions [37].

Conversely, our observations imply that seed sterility is farther from saturation, and that seed sterility loci in these species might be more likely to engage in DMIs that involve 'simple' pairwise interactions (rather than higher-order epistasis). Each additional seed sterility-causing substitution is likely to originate a new DMI with a unique interacting partner, rather than engaging in fitness-affecting interactions with existing sterility loci. In this case, seed sterility loci themselves will still appear to 'snowball' with divergence between lineages; indeed, seed sterility loci have been shown to snowball between Solanum species [18]. While this speculation regarding the saturation of different sterility phenotypes needs to be confirmed with further work, a more general implication of our results is that QTL mapping (or any count of individual hybrid incompatibility loci) might not always provide a suitable estimate of the number and complexity of accumulating DMIs. Despite the considerable additional experimental burden this implies, future empirical tests of the snowball effect should take into account the possibility of highly connected networks and saturated traits.

Finally, we note that the high frequency of interactions with antagonistic effects on sterility has implications for the pattern of accumulation of reproductive isolation phenotypes (rather than loci or DMIs) between these species. The amount of reproductive isolation is expected to snowball if DMIs are independent and have, on average, similar fitness effect sizes. If antagonistic epistasis among sterility loci is common, however, the

accumulation of reproductive isolation is expected to be less-than-linear, because new DMIs will tend to contribute 'diminishing returns' on fitness. This lack of a snowball in incompatibility phenotypes has been previously noted [38, 39], and pervasive phenotypic antagonism could contribute to explaining this common observation {Kalirad, #41}}. In contrast to its phenotypic consequences, however, the mechanistic basis of pervasive antagonism is unclear. This pattern of diminishing deleterious effects is consistent with observations in *E. coli* [40] and yeast [13], but the underlying basis is unresolved. For Drosophila, Yamamoto et al. [41] proposed that similar patterns of antagonism might be expected among loci that are involved in stabilizing selection within populations; however, this is unlikely to be relevant here both because our interacting loci are not normally segregating within populations, and because fertility traits are less likely to be subject to stabilizing selection at some intermediate phenotypic optimum (compared to other phenotypic traits). Alternatively, less-than-additive fitness effects can result when different mutations each have deleterious individual effects within the same developmental pathway, but their pairwise combination suppresses or ameliorates these deleterious effects (as has been observed, for example, with double deletion strains in yeast; [13]). Our inference here is that, in saturated systems, new substitutions are likely to interact with loci that are already involved in hybrid incompatibility, which would be consistent with sterility-affecting mutations accumulating in a tightly-connected network (such as a specific developmental pathway). Of course, additional empirical work is needed to evaluate this possibility. Moreover, the frequency of antagonism and the potential for phenotypic diminishing returns are likely to be dependent on the properties of the diverging gene network. Further theoretical work is needed to determine how

specific network properties, such as its connectivity and modularity, can affect the patterns of accumulation of incompatibilities and reproductive isolation.

Literature Cited

1. Coyne JA, Orr HA. Speciation. Sunderland, MA: Sinauer Associates, Inc; 2004.
2. Turelli M, Orr HA. Dominance, epistasis and the genetics of postzygotic isolation. Genetics. 2000;154(4):1663-79.
3. Orr HA, Turelli M. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution. 2001;55(6):1085-94.
4. Welch JJ. Accumulating Dobzhansky-Muller incompatibilities: reconciling theory and data. Evolution. 2004;58(6):1145-56.
5. Kondrashov AS. ACCUMULATION OF DOBZHANSKY-MULLER INCOMPATIBILITIES WITHIN A SPATIALLY STRUCTURED POPULATION. Evolution. 2003;57(1):151-3.
6. Turelli M, Barton NH, Coyne JA. Theory and speciation. Trends in Ecology & Evolution. 2001;16(7):330-43.
7. Cabot EL, Davis AW, Johnson NA, Wu CI. Genetics of reproductive isolation in the Drosophila simulans clade: complex epistasis underlying hybrid male sterility. Genetics. 1994;137(1):175-89.
8. Wu C-I, Johnson NA, Palopoli MF. Haldane's rule and its legacy: Why are there so many sterile males? Trends in Ecology & Evolution. 1996;11(7):281-4.
9. Kao C-H, Zeng Z-B, Teasdale RD. Multiple interval mapping for quantitative trait loci. Genetics. 1999;152(3):1203-16.
10. Doebley J, Stec A, Gustus C. teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics. 1995;141(1):333-46.
11. Eshed Y, Zamir D. Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics. 1996;143(4):1807-17.
12. Muir CD, Moyle LC. Antagonistic epistasis for ecophysiological trait differences between Solanum species. New Phytologist. 2009;183(3):789-802.
13. Jasnos L, Korona R. Epistatic buffering of fitness loss in yeast double deletion strains. Nature genetics. 2007;39(4):550-4.
14. He X, Qian W, Wang Z, Li Y, Zhang J. Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks. Nature genetics. 2010;42(3):272-6.
15. Moyle LC, Nakazato T. Complex epistasis for Dobzhansky–Muller hybrid incompatibility in Solanum. Genetics. 2009;181(1):347-51.
16. Monforte A, Friedman E, Zamir D, Tanksley S. Comparison of a set of allelic QTL-NILs for chromosome 4 of tomato: deductions about natural variation and implications for germplasm utilization. Theoretical and Applied Genetics. 2001;102(4):572-90.
17. Moyle LC, Graham EB. Genetics of hybrid incompatibility between Lycopersicon esculentum and L. hirsutum. Genetics. 2005;169(1):355-73.
18. Moyle LC, Nakazato T. Hybrid incompatibility “snowballs” between Solanum species. Science. 2010;329(5998):1521-3.
19. Moyle LC, Nakazato T. Comparative genetics of hybrid incompatibility: sterility in two Solanum species crosses. Genetics. 2008;179(3):1437-53.
20. Kearns CA, Inouye DW. Techniques for pollination biologists: University Press of Colorado; 1993.
21. R Core Team. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
22. Schwarz G. Estimating the dimension of a model. The annals of statistics. 1978;6(2):461-4.
23. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal, Complex Systems. 2006;1695(5):1-9.
24. Pedersen TL. ggraph: An implementation of grammar of graphics for graphs and networks. R package version 0.1.1 ed2016.
25. Gao H, Granka JM, Feldman MW. On the classification of epistatic interactions. Genetics. 2010;184(3):827-37.
26. Chang AS, Noor MA. The genetics of hybrid male sterility between the allopatric species pair Drosophila persimilis and D. pseudoobscura bogotana: dominant sterility alleles in collinear autosomal regions. Genetics. 2007;176(1):343-9.

27. Willett CS. Complex deleterious interactions associated with malic enzyme may contribute to reproductive isolation in the copepod Tigriopus californicus. PLoS One. 2011;6(6):e21177.
28. Kelly JK, Mojica JP. Interactions among flower-size QTL of Mimulus guttatus are abundant but highly variable in nature. Genetics. 2011;189(4):1461-71.
29. Orr HA. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics. 1995;139(4):1805-13.
30. Arnqvist G, Rowe L. Sexual conflict: Princeton University Press; 2013.
31. Holland B, Rice WR. Perspective: chase-away sexual selection: antagonistic seduction versus resistance. Evolution. 1998:1-7.
32. Madjidian JA, Lankinen Å. Sexual conflict and sexually antagonistic coevolution in an annual plant. PLoS One. 2009;4(5):e5477.
33. Brandvain Y, Haig D. Divergent mating systems and parental conflict as a barrier to hybridization in flowering plants. The American Naturalist. 2005;166(3):330-8.
34. Chase CD. Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. Trends in Genetics. 2007;23(2):81-90. doi: http://dx.doi.org/10.1016/j.tig.2006.12.004.
35. Gavrilets S. Rapid evolution of reproductive barriers driven by sexual conflict. Nature. 2000;403(6772):886-9.
36. Rice WR, Holland B. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. Behavioral Ecology and Sociobiology. 1997;41(1):1-10.
37. Kalirad A, Azevedo R. The Melting Snowball: A Test of the Snowball Model Using RNA. bioRxiv. 2016:076232.
38. Gourbiere S, Mallet J. Are species real? The shape of the species boundary with exponential failure, reinforcement, and the “missing snowball”. Evolution. 2010;64(1):1-24.
39. Johnson N. Patterns and processes of speciation: the evolution of reproductive isolating barriers. In: Fox CW, Wolf JB, editors. Evolutionary genetics concepts and case studies New York, NY, USA: Oxford University Press; 2006. p. 374-86.
40. Wiser MJ, Ribeck N, Lenski RE. Long-term dynamics of adaptation in asexual populations. Science. 2013;342(6164):1364-7.
41. Yamamoto A, Anholt RR, Mackay TF. Epistatic interactions attenuate mutations affecting startle behaviour in Drosophila melanogaster. Genetics research. 2009;91(06):373-82.

Table 1: Introgression lines (ILs) phenotyped and used as parents to generate double-introgression lines. Accession identifiers from the Tomato Genome Resource Center (tgrc.ucdavis.edu). QTL: Prior evidence of IL carrying QTL for pollen sterility (P), seed sterility (S) or none (N). Size: length of introgressed region, as a percentage of the *lyc* genome. Seed: Mean (and standard deviation) for self-seed count. Pollen: Mean (and standard deviation) for proportion of fertile pollen. P-values are from negative-binomial (for seed) and quasi-likelihood binomial (for pollen) general linear models. (*)IL has a significant individual effect on the corresponding phenotype.

Accession	QTL	Chromosome	Percent of Genome	Seed	P-value	Pollen	P-value
LA3948	P	7	4	56.2 (16)	0.03	0.546 (0.33)	7e-06*
LA3935	P	4	4.2	15.1 (13)	0.001*	0.573 (0.29)	5e-06*
LA3950	P	7	2.7	13.2 (10)	3e-04*	0.583 (0.07)	1e-07*
LA3956	P	9	4.6	18.3 (12)	0.001*	0.723 (0.11)	1e-04*
LA3963	P	10	2.4	27.6 (20)	0.6	0.289 (0.2)	3e-10*
LA3915	S	1	2.8	30.6 (9.5)	0.3	0.565 (0.24)	2e-06*
LA3931	S	4	1.5	22.1 (13)	0.006*	0.749 (0.12)	0.002
LA3939	S	5	2	26.5 (13)	0.02	0.783 (0.17)	0.01
LA3943	S	5	2.7	17.6 (13)	0.002*	0.667 (0.25)	7e-04
LA3977	S	4	1.5	15.4 (13)	0.01	0.402 (0.4)	2e-06*
LA3947	N	6	0.68	48.1 (15)	0.5	0.722 (0.27)	0.005
LA3957	N	9	3.6	26.8 (17)	0.03	0.724 (0.3)	0.005
LA3964	N	10	1.8	44.7 (21)	1	0.771 (0.28)	0.03
LA3968	N	12	1.1	34.5 (16)	0.3	0.71 (0.16)	3e-04
LA3975	N	3	0.96	44.2 (25)	0.9	0.785 (0.19)	0.02
LA4024	-	-	-	42.1 (16)	-	0.899 (0.071)	-

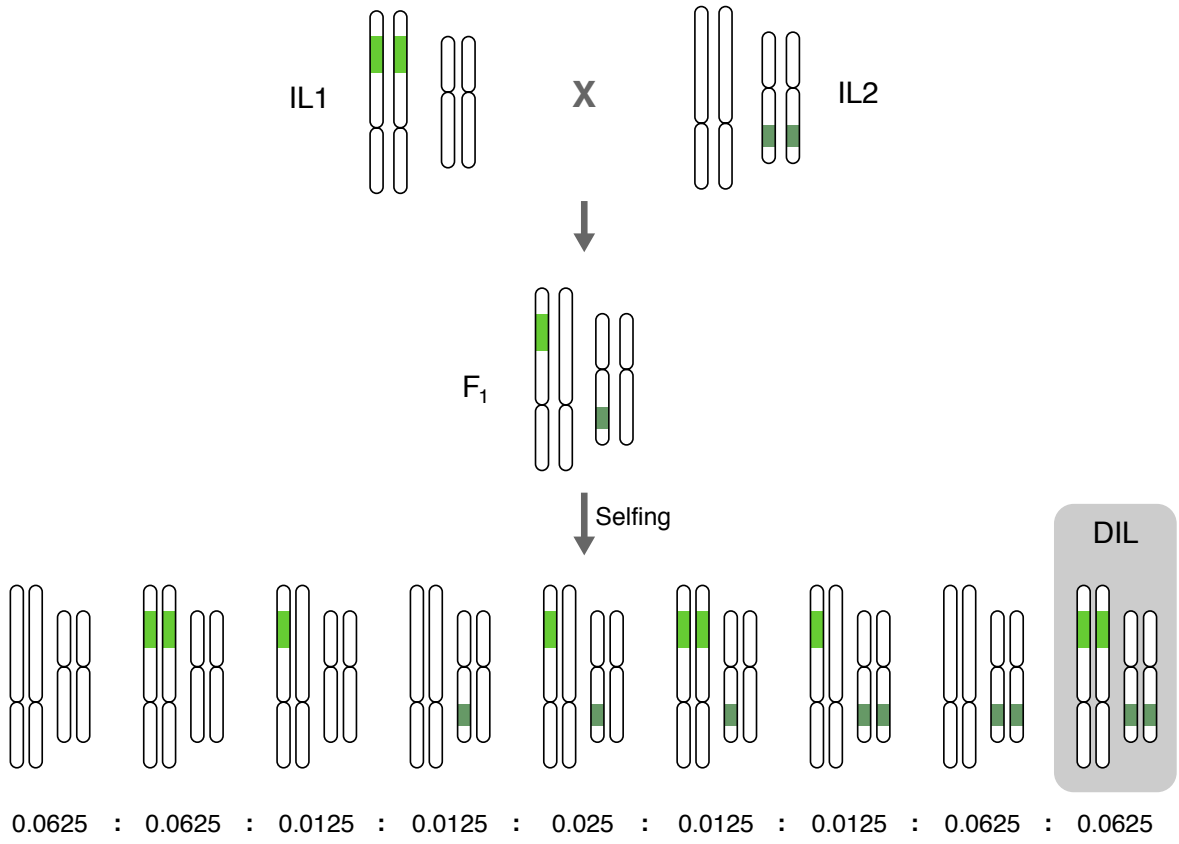

Figure 1: Diagram of cross made to generate the double-introgression lines (DILs). Green shaded areas within a chromosome represent introgressed *hab* regions in an isogenic *lyc* background. Included are the expected genotypic ratios in an F2 population for each DIL family, if marker transmission is Mendelian.

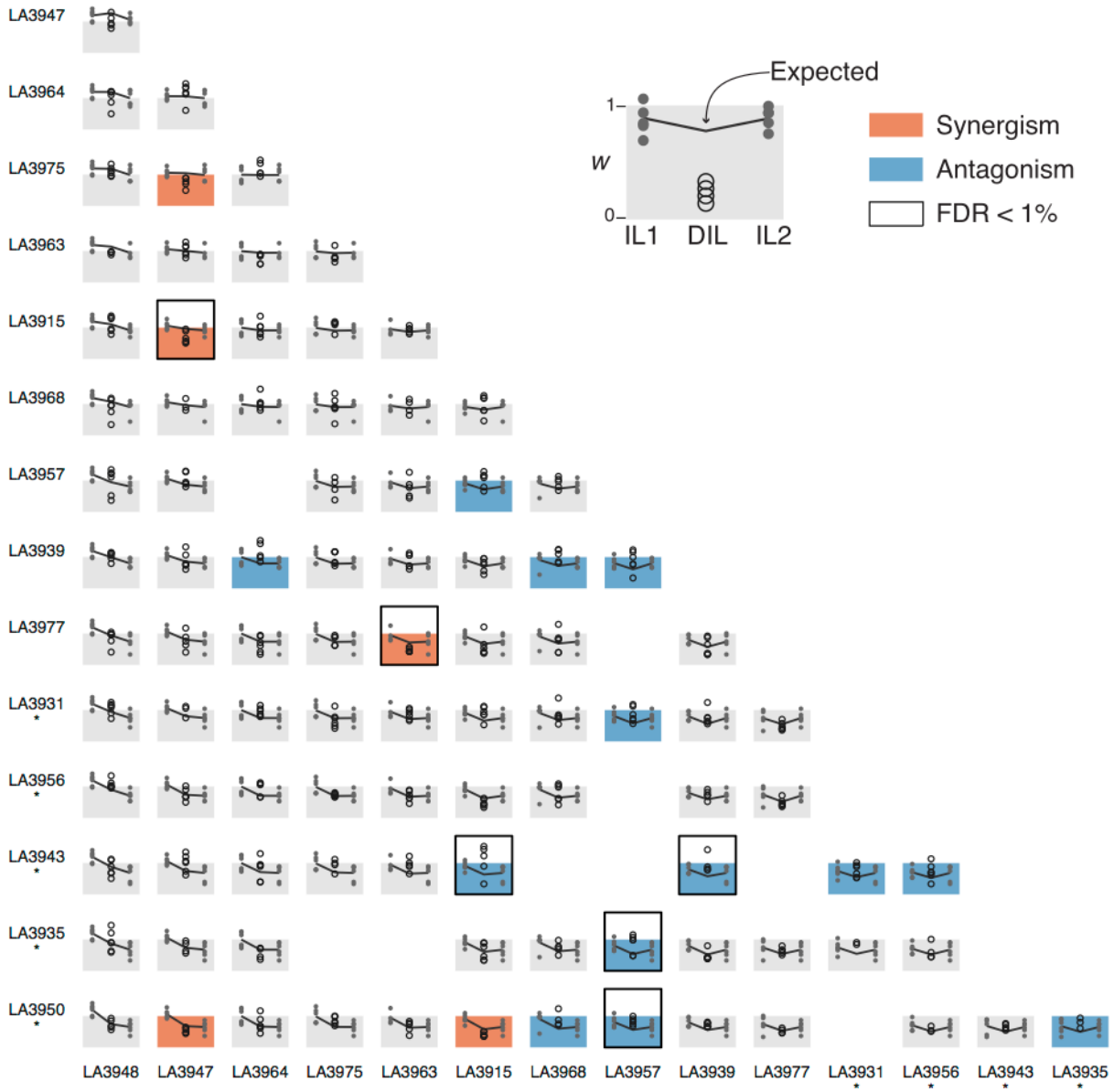

Figure 2: Self-seed set, relative to *S. lycopersicum*, for all parental introgression lines (ILs) and double-introgression lines (DILs). Parental ILs (across rows and columns) are arranged by their individual effect on the phenotype (larger effects down and to the right). Asterisks below IL accession names denote significant individual effects in ILs. Each panel shows the observed values for two parents (small circles) and their corresponding DIL (large circles, rightmost in each panel). Lines connect the observed mean phenotypes for each parental IL and the expected mean phenotype for the DIL. The background of each panel spans from 0 to 1 along the vertical axis (1 corresponds to the mean fitness value of *S. lycopersicum*). Accented panels show significant support ($P < 0.01$) for epistasis (antagonistic interactions = blue panels; synergistic = red panels). Black bordered panels indicate highly significant interactions (FDR=1%).

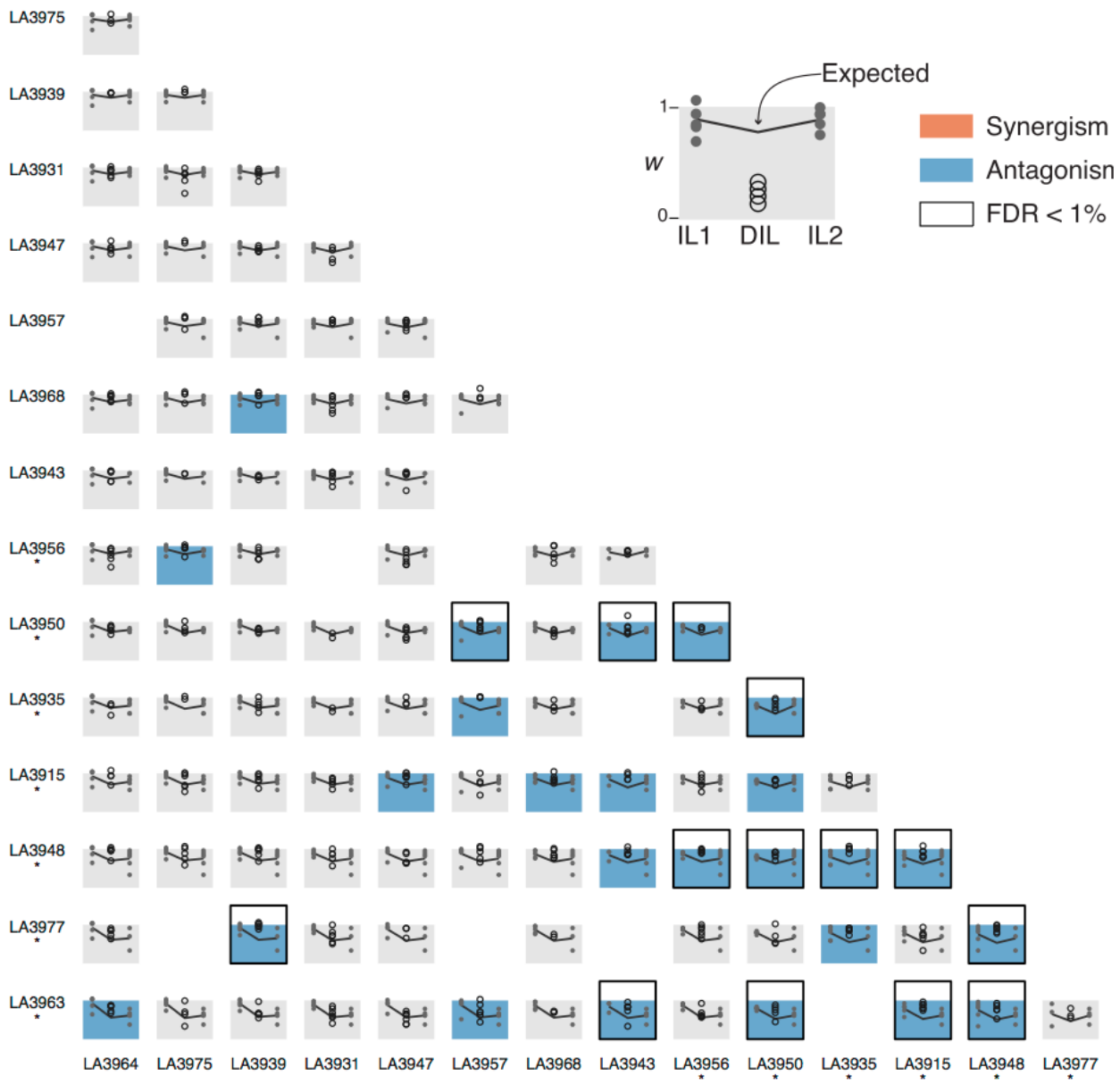

Figure 3: Proportion of fertile pollen, relative to *S. lycopersicum*, for all parental introgression lines (ILs) and double-introgression lines (DILs). All conventions as in Figure 2.

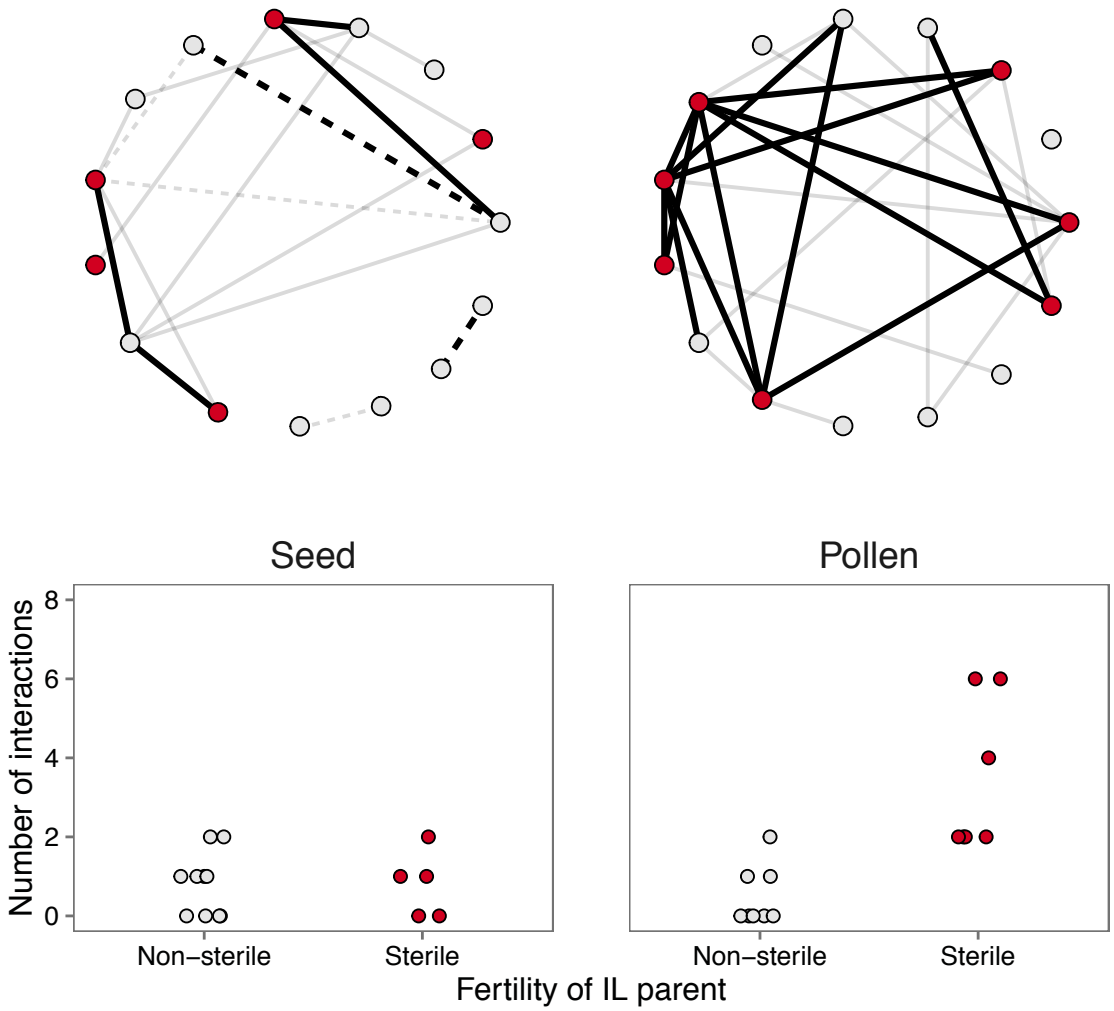

Figure 4: Above, the network representation of interactions in self-seed set (left) and proportion of fertile pollen (right). Parental ILs (nodes; red = sterile IL) are connected if they showed a highly significant interaction (FDR=1%). Solid lines denote antagonistic interactions; dashed lines mean synergistic interactions. Below, the relationship between parental IL fertility on the number of interactions (node degree) showed.

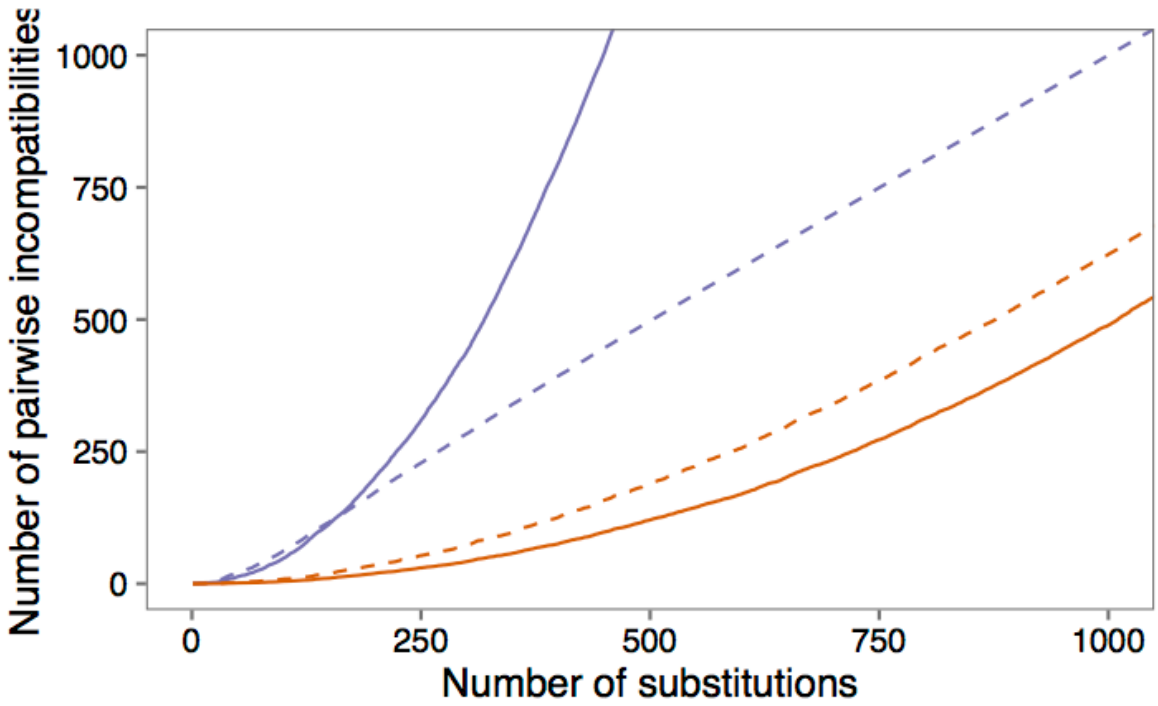

Figure 5: The number of pairwise incompatibilities (showing the "snowball" effect, solid lines), and the number of loci involved in hybrid incompatibilities (dashed lines), for two cases that differ in the per-substitution probability of pairwise incompatibility (orange, $p = 0.0001$; purple, $p = 0.01$). After saturation (roughly around 200 substitutions for the case in purple lines) the number of loci involved in incompatibilities no longer tracks the number of DMIs. In each case, values are averaged over 1000 simulations.