doi:10.1111/evo.13876

Severe inbreeding depression is predicted by the "rare allele load" in *Mimulus guttatus**

Keely E. Brown[1,2] **and John K. Kelly**[1]

[1]*Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045*

[2]*E-mail: k034b363@ku.edu*

Received April 16, 2019

Accepted October 9, 2019

Most flowering plants are hermaphroditic and experience strong pressures to evolve self-pollination (automatic selection and reproductive assurance). Inbreeding depression (ID) can oppose selection for selfing, but it remains unclear if ID is typically strong enough to maintain outcrossing. To measure the full cost of sustained inbreeding on fitness, and its genomic basis, we planted highly homozygous, fully genome-sequenced inbred lines of yellow monkeyflower (*Mimulus guttatus)* in the field next to outbred plants from crosses between the same lines. The cost of full homozygosity is severe: 65% for survival and 86% for lifetime seed production. Accounting for the unmeasured effect of lethal and sterile mutations, we estimate that the average fitness of fully inbred genotypes is only 3–4% that of outbred competitors. The genome sequence data provide no indication of simple overdominance, but the number of rare alleles carried by a line, especially within rare allele clusters nonrandomly distributed across the genome, is a significant negative predictor of fitness measurements. These findings are consistent with a deleterious allele model for ID. High variance in rare allele load among lines and the genomic distribution of rare alleles both suggest that migration might be an important source of deleterious alleles to local populations.

KEY WORDS: **Field fitness, inbreeding depression, monkeyflower, rare alleles.**

Over 160 years since Darwin (1876, 1877) identified the problem, it remains a paradox that hermaphroditic species maintain outcrossing despite strong selection for self-fertilization (Fisher 1941; Lloyd 1979). Over 90% of flowering plants are hermaphroditic (Renner and Ricklefs 1995) and most are substantially or predominantly outcrossing (Goodwillie 2005; Igic et al. 2006). Plants have evolved complex mechanisms to prevent self-fertilization, such as molecular self-incompatibility (Takayama and Isogai 2005), herkogamy including flexistyly and heterostyly (Ganders 1979; Li et al. 2001a; Opedal 2018), and dichogamy (Bertin and Newman 1993). Classical theory predicts that a population should maintain outcrossing if inbreeding depression (ID) is strong enough, specifically that $\delta > 0.5$ where δ equals one minus the fitness of selfed relative to outcrossed progeny (Kimura 1959; Lande and Schemske 1985; Charlesworth and Charlesworth 1987).

The "$\delta > 0.5$ rule" has motivated experimental estimation of ID in many species. Winn et al. (2011) recently reviewed plant estimates; the mean δ for lifetime fitness was slightly greater than 0.5 for both highly outcrossing and mixed mating species. Unfortunately, interpretation of δ estimates near 0.5 is problematic because, for a number of reasons, the $\delta > 0.5$ rule underestimates the necessary strength of ID to halt selfing. Lloyd (1979) showed that "delayed selfing," where a plant self-fertilizes ovules after the opportunity to outcross has passed, can evolve even with very high ID. The reproductive assurance provided by selfing is also advantageous for the colonization of new habitats (Baker 1955) and for range expansion (Grossenbacher et al. 2015). In addition, delayed selfing can purge deleterious mutations from a population and thus increase the likelihood that "competing selfing" (self- and cross-fertilization compete for the same ovules; Lloyd 1979) might be favored.

Because it treats all inbred individuals as equivalent, δ is an imperfect statistic for ID. In fact, fitness varies with the level of inbreeding and should decline monotonically with

*This article corresponds to Hudson, A. I. 2020. Digest: Exposing the role of rare alleles in inbreeding depression in monkeyflower. Evolution. https://doi.org/10.1111/evo.13931.

F, the individual inbreeding coefficient (Morton et al. 1956; Kimura and Maruyama 1966; Charlesworth and Charlesworth 1987; Kondrashov 1988; Charlesworth et al. 1991). δ predicts the increase of a small-effect mutation that increases selfing within an outcrossing population, because in this case, nearly all inbred individuals will be the selfed progeny of outbred parents ($F = 0.5$). The fitness of inbred individuals carrying the selfing modifier will be approximately equal to, and unchanged from, the mean ID in the original population. ID is effectively characterized by δ because all inbred individuals are essentially equivalent in terms of expected genome-wide homozygosity.

This simplification no longer applies with large-effect selfing mutations, which can increase in certain situations even when $\delta >> 0.5$ (Lande and Schemske 1985; Holsinger 1988; Uyenoyama et al. 1993). Individuals carrying a large-effect selfing mutation become a distinct subpopulation, really a collection of distinct lineages, with varying levels of inbred fitness. If the selfing mutation fortuitously fixes within a lineage of low mutational load, this lineage can expand to exclude outcrossing genotypes from a population. The selfing mutation can "find" the right lineage through a combination of chance, if it occurs within a plant carrying fewer deleterious mutations, and subsequent intra-family purging of deleterious alleles (Kelly and Tourtellot 2006). Most importantly, the fate of the selfing mutation will be determined by the fitness of highly inbred (homozygous) individuals. At present, most estimates for ID are based on fitness measured in first-generation selfed progeny, not highly homozygous individuals.

The first objective of this paper is to estimate the relative fitness of highly homozygous genotypes ($F \approx 1$) under field conditions within a large outcrossing population. Most experimental estimates of ID are based on first-generation selfs ($F = 0.5$) assayed under benevolent conditions (Winn et al. 2011). Here, we compare inbred lines to F1 crosses between lines of yellow monkeyflower (*Mimulus guttatus*). The lines were derived from randomly sampled, field-collected individuals and allele frequencies in the lines match estimates from direct field collections (Troth et al. 2018).

The second objective of the paper is to use the full genome sequences of the lines in combination with the fitness estimates to further characterize the genetic basis of ID. ID may result from rare (partially) recessive deleterious alleles at low frequency in the population (the "dominance hypothesis," Charlesworth and Willis 2009), in which case fitness of inbred individuals should decline with the load of deleterious alleles. ID could also result from increased homozygosity at overdominant loci (Li et al. 2001b; Luo et al. 2001), and/or from loci with partially dominant alleles maintained at intermediate frequency by some form of balancing selection (Lee et al. 2016). Under the "overdominance hypothesis," fitness may positively correlate with genome-wide heterozygosity, although this prediction depends on how many loci exhibit heterozygote advantage.

In this study, we find no evidence for overdominance at individual single-nucleotide polymorphisms (SNPs), although we have limited power to detect such loci. However, there is a significant negative correlation between the number of rare alleles carried by a given line (the "rare allele load") and fitness measurements. This result is consistent with the dominance hypothesis insofar as the rare allele load is correlated with the deleterious mutation load. We find that rare alleles are distributed nonrandomly (in clumps) within individual plant genomes; the locations of these rare allele clusters (RACs) vary among plants. Importantly, the number and composition of RACs is a better predictor of fitness in both field and greenhouse than the overall rare allele load, which suggests that migration/introgression could be an important source of deleterious alleles in the population.

Methods

STUDY SYSTEM AND LINE DEVELOPMENT

Mimulus guttatus (Phrymaceae, syn. *Erythranthe guttata*) grows in western North America, from northern Mexico to Alaska. Populations are annual or short-lived perennials, reliant on bees for pollination. Estimated selfing rates vary among *M. guttatus* populations, and between years within a population, but most populations are predominantly outcrossing (Ritland and Ganders 1987; Awadalla and Ritland 1997; Sweigart et al. 1999). Here, we investigate the Iron Mountain (IM) population of *M. guttatus* (Oregon, U.S.A.; 44.402217 N, –122.153317 W), an annual population with an outcrossing rate of over 90% (Willis 1993b).

Each of the inbred lines used in the present study was initiated from a single seed sampled from IM, each seed from a distinct maternal plant. Note that 1200 lineages from one collection of wild plants were started in 1995 (those with a prefix "IM"), but after six generations of single seed descent (selfing with random selection of a single seedling per family for the next generation), only 300 remained (Willis 1999b; Kelly 2003). The purpose of this experiment was to allow lethal and sterile mutations to fix randomly within lines, in proportion to their frequency in the natural population. After extinction of these lineages, the resulting "purged" population ($f > 0.98$) carried only sublethal and mildly deleterious mutations. Some lines might also have died owing to the aggregate effect of many sublethal mutations, but the critical point for the current study is that a component of genetic load segregating in the natural population has been removed from the lines. Importantly, the line creation experiment differs from the hypothesized line formation process associated with large-effect selfing mutations (described above) because the experiment minimized the opportunity for intralineage purging.

The survival of 300 of 1200 lineages yields our estimate of 75% genetic death in line formation. Of course, plants can fail to reproduce for nongenetic reasons, but outbred IM plants have nearly 100% survival under greenhouse conditions. Also,

Kelly (2003) attempted to regrow seed from the fourth and fifth generation lines that had failed by generation 6, but very few of these lineages were resurrected. Since 2003, we have periodically germinated and selfed these lines and they are now 6–12 generations inbred. A second collection of inbred lines from the Zia-1 base population are distinguished by the prefix "Z." Each of the Z lines is derived from a single IM seed (Kelly 2008). However, single-seed descent for the Z lines was performed with cold stratification of seed (1 week) before germinating in the greenhouse. The IM lines were formed without cold treatment.

In total, 187 lines have been whole-genome sequenced and confirmed to be highly homozygous; population genomic results such as the allele frequency spectrum are reported in Puzey et al. (2017) and Troth et al. (2018). Whole-genome sequencing also revealed that several of the lines are more closely related to each other either due to splitting of a single founder into multiple inbred lines or by random sampling of close relatives in the field. For the field component of the present study, we selected 37 IM lines established to be wholly unrelated (not any more or less closely related in any of the other 36) by the kinship matrix of Troth et al. (2018). For the greenhouse study, we planted 165 of the sequenced lines (116 IM and 49 Z).

FIELD EXPERIMENTAL DESIGN

We generated both self-fertilized progeny and outcrossed progeny by growing three to four plants from each line to maturity in the U. Kansas greenhouses. We performed two different types of crosses to obtain outcrossed progeny: fertilization with pollen from a single non-self IM line (random mate-pair crosses) and fertilization with a mixture of pollen from seven other non-self IM lines (group crosses). The latter were done to test if producing more diverse progeny increases the average fitness of progeny. We germinated seed from each cross/self in the University of Oregon greenhouses, with seed to soil on May 7, 2018. On May 21–23, 2018, we transplanted 1176 greenhouse-germinated seedlings (37 lines, 4–40 seedlings per line, average 32) into absorbent peat/wood fiber pots (Jiffy Strip, Blue Ridge Greenhouses), one seedling per pot (548 selfed, 304 group outcrossed, 324 single outcrossed). One day after transplanting, we settled pot strips into the soil/moss matrix of the Browder Ridge Trailhead site (Oregon, USA; 44.373238 N, –122.130675 W) in two cohorts, one day apart. Browder Ridge is our "transplant site," geographically close and ecologically similar to IM (Mojica et al. 2012). We arranged the transplants such that in each flat of eight plants, there was a row of four outcrossed plants next to a row of four selfed plants all from the same maternal line when possible. This was meant to minimize the random effect of spatial variation across the field plot. Each maternal line group was planted in many different places across the plot. The timing of transplant ensured that wild individuals were at the same developmental stage as our experimental transplants (cotyledon or two leaf stage). On July 21, 2018, we harvested experimental plants to estimate fitness. All but five of the transplants were fully desiccated, and only one still had an open flower. We noted which individuals had survived to flower and collected all fruits. We scored all individuals for survival to flower (0/1). For all survivors, we determined the number of flowers, number of fruits, seed set per fruit, and total seed set.

GREENHOUSE FOLLOW-UP

We grew 134 of the whole-genome sequenced IM lines (on January 11, 2019) under the same conditions in which the lines were initiated (1–10 per line, average five plants). Z lines were stratified at 4°C for one week prior to being transferred to the U. Kansas greenhouse, whereas IM lines were not stratified. We had previously established an interaction between stratification treatment and line type (Z vs. IM) on germination success in the greenhouse (Fig. S1), so we have only grown lines in their initiation conditions for this experiment. We transplanted germinants to 2.25 in pots after 14 days in the greenhouse, randomizing them between flats on the greenhouse bench. Two weeks after transplanting, we cut each plant at the hypocotyl/root junction and collected the above-ground mass into coin envelopes. We dried the tissue in an oven at 300°F for 1 hour and weighed/recorded the dry mass, accurate to one-tenth of 1 mg.

ANALYSIS

For the field data, we first tested the distinct outcross treatments to determine if progeny of group crosses were different from single matings. There was no evidence of difference for any fitness component (Fig. S2). Thus, for subsequent whole-plant analyses, we combine all outcrossed progeny of a maternal line into a single category. We cannot use group crosses for genomic analyses because the progeny genome sequence cannot be inferred (specific father unknown). Sample sizes are not sufficient for a meaningful SNP-level testing (Supporting Information Appendix 1). To calculate the rare allele load of each line, we determined allele frequencies in the full (unfiltered) variant call file obtained by Troth et al. (2018). We suppressed one line (IM764) from the frequency calculations because it is excessively divergent across multiple chromosomes. Table S1 reports all 5,018,997 SNPs where the minor allele is $\leq$5% in the inbred lines ("rare" for this analysis). We scored each line for number of rare alleles carried at these loci. The rare allele load is the count of rare alleles divided by the total scored sites for each line. For each family consisting of F1 progeny from reciprocal single-crosses planted in the field (17 in total), we also calculated the genome-wide heterozygosity as a fraction of scored loci.

To characterize rare allele distribution across the genome, we calculated rare allele load within 500 SNP windows in each line. To delineate RACs, we used the cpt.meanvar function in *changepoint* in R with the PELT method and a BIC penalty

(Killick and Eckley 2014; Fig. 2). The program distinguishes intervals in sequence where the mean and variance of the variable (rare allele load here) changes. We identified RACs as peaks of multiple adjacent windows in which rare allele load exceeds either 5%, 7.5%, or 10%. The overall rare allele load can then be partitioned into two components, within RACs and background (remainder of genome).

STATISTICAL TESTING

We performed all model fits using R (R Core Team 2013). We tested for the effect of cross-type, rare allele load, load inside/outside of RACs, and heterozygosity on field survival using generalized linear mixed-effect models. Here, survival was categorized as a binomial response (logit link function) estimated using the *glmer* function in the *lme4* package (Bates et al. 2007), which uses restricted maximum likelihood. The mixed-effects models included maternal line or family as a random effect and interactions when necessary. The mixed model was compared to a generalized linear model without random effects using the built-in R function *glm*.

Total fitness in the field, measured by seed set, is overdispersed relative to the Poisson distribution (Fig. S3). For this reason, we used the *reaster* function from the R package *aster* to fit an exponential family regression model (Geyer et al. 2007, 2013). Aster models test for an effect on total fitness as a cumulative trait composed of different stages of life history by allowing each stage to have a different response type. *Reaster* fits an aster model with random effects. We treated cross-type, rare allele load, load inside or outside of RACs, and heterozygosity as fixed effects and maternal line (and its interaction with the fixed effect) as a random effect. We structured the life history model into three stages: survival to flower (binomial), any seed set (binomial), number of seeds (zero-truncated Poisson), and compared these models to the corresponding model without random effects using *aster*. For plants grown in the greenhouse, the natural log of above ground dry mass yields normal residuals (Fig. S4), so we performed multiple linear regression on the average ln dry mass for every line. We used the built-in R function *lm* to test for an effect of rare allele load (and load inside/outside of RACs) on line means of ln(mass) using line collection (IM vs. Z) as an additional factor.

Results

Inbreeding depression: There was a sevenfold decrease in the total fitness of selfed individuals compared to outcrossed individuals in the field (7.6 vs. 1.1, $\delta = 0.86$, $z = 4.63$ for out vs. self, $P = 3.75 \times 10^{-6}$). This high level of ID is caused by an almost threefold decrease in survival to flowering (28.5% vs. 9.85%, $\delta = 0.65$, $z = -6.044$, $P = 1.5 \times 10^{-9}$), and among survivors, the mean outcrossed seed set was 26.8 seeds versus only 11 seeds for inbred plants. Seed counts are within the range reported in previous field experiments using IM genotypes (Mojica and Kelly 2010; Mojica et al. 2012; Monnahan and Kelly 2015). These estimates are from the full model including both maternal line and line by cross-interaction as random effects (See Table S2 for all model fits). We find a significant effect of line ($z = 3.32$, $P = 0.000454$) and line by type of cross-interaction ($z = 3.26$, $P = 0.0056$) on total fitness in the *reaster* model that includes both as random effects (see Fig. S5 for a visual representation of the interaction).

The variance in average phenotype among families was higher for outcrossed families than inbred families for both survival (0.050 vs. 0.026) and total seed (105.01 vs. 5.32). Given that our field fitness δ does not include lethal or sterile mutations, which segregate in the natural population but not in the inbred lines, our cumulative estimate for the relative fitness of homozygous plants is $(0.25)(0.14) = 0.035$. The resulting δ (0.965) accounts for the 75% of lines that perished owing to fixation of deleterious mutations over generation 5 of selfing (Willis 1999a,b). One inbred line, IM922, had a higher mean total fitness (9.05) than the outbred mean (7.6), although the confidence band on the former estimate is very large (including values far below the outbred mean).

Genomic predictors of inbred and outbred fitness: The 31 lines included in the field experiment have a rare allele load between 0.38% and 2.1% (mean 1.5%). Among selfed progeny, rare allele load is a highly significant predictor of survival (Fig. 1A; $z = -3.64$, $P = 0.00027$), but not lifetime seeds (Fig. 1B; $z = -1.45$, $P = 0.15$). If line is included as a random effect, tests become nonsignificant (Table S3) because line ID and rare allele load are strongly confounded (co-linear predictors). The lines with lower rare allele load have higher survival (left portion of Fig. 1A), but the limited replication at this end of the scale limits inference. The distribution of rare alleles is very nonrandom across the genome in a line-specific manner (Fig. 2). Only a small subset of lines is elevated within any window (otherwise the alleles would not be rare). We considered different thresholds to define RACs as predictors of field fitness (Table S4): 5% provided the greatest predictive power. The mean of load in 5% RACs is 0.58% across lines (about one-third of the total rare allele load). Even when line is included as a random effect, load within 5% RACs is a highly significant predictor of field survival (Fig. 1C, $z = -4.944$, $P = 7.67 \times 10^{-07}$). Ideally, we predict fitness using load within and outside of RACs as simultaneous predictors (multiple linear regression). Unfortunately, the inside/outside numbers are strongly positively correlated across lines ($r = 0.65$ for 5% RACs, 0.57 for 7.5% RACs, and 0.53 for 10% RACs), which makes the partial regression estimates unstable.

The outcross progeny between single lines (F1s) have known whole genome sequences. Among the F1s grown in the field, the overall heterozygosity of plants varied from 15.7% to 21.2%. Heterozygosity is not a significant predictor of survival either as

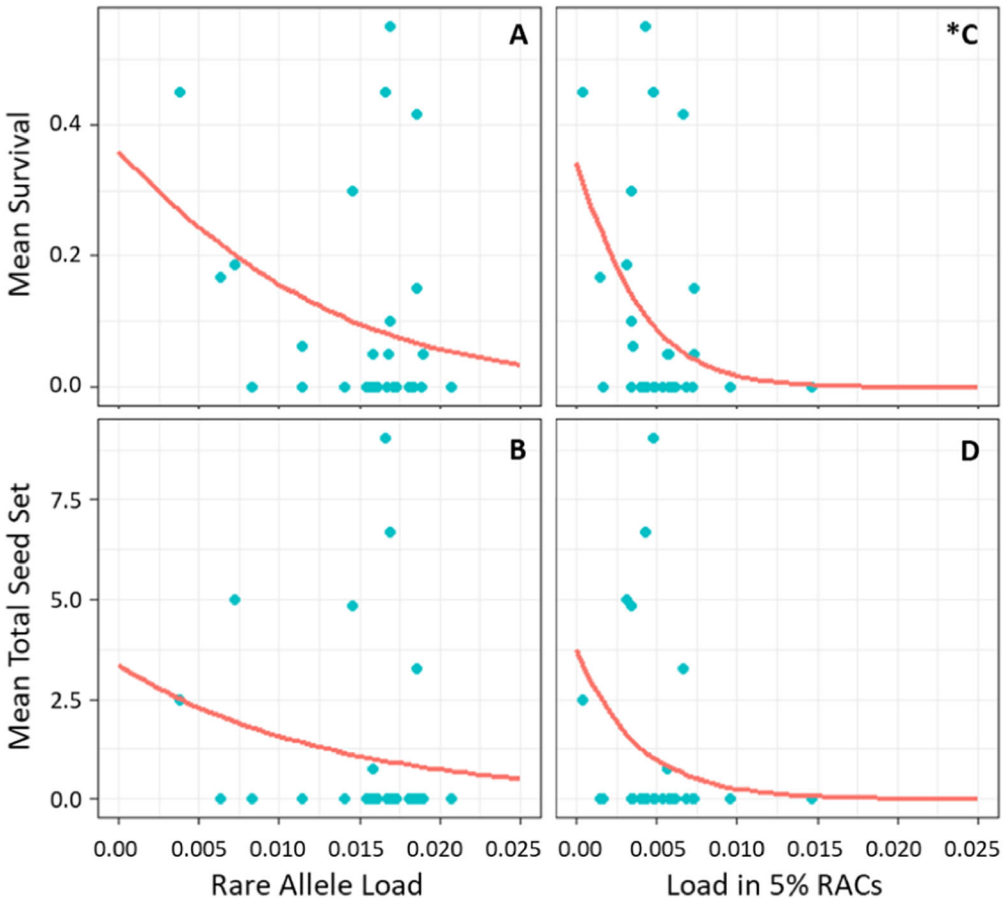

Figure 1. Effect of rare allele load (A, B) and load within RACs (C, D) on inbred line survival and total seed set in the field experiment. Curves are predicted from simple model fits in *glm*. Asterisk for (C) indicates the only relationship that remains significant in the full model with line included as a random factor. Predictions are given by the following: (A) survival = (e^(−0.577 − 110.6*x))/(1 + e^(−0.577 − 110.6*x)); (B) total fitness = e^(1.206 − 75.44*x); (C) survival = (e^(−0.6577 − 336.9103*x))/(1 + e^(−0.6577 − 336.9103*x)); (D) total fitness = e^(1.32007 − 268.58975*x).

a single predictor or with line included as a random factor (Table S7). It has an apparently positive effect on lifetime seeds, but this effect is marginally nonsignificant ($P = 0.075$) without line and entirely nonsignificant ($P = 0.46$) when line is included as a random factor.

We performed the greenhouse experiment on a much larger collection of lines to more clearly distinguish the effect of rare allele load. This experiment reveals a highly significant effect of both rare allele load and load in 7.5% RACs (the best fit model, Table S6) on above-ground biomass at day 28 (Fig. 3, Table S5; $F = 13.55$, $P = 0.00034$ for rare allele load, $F = 15.22$, $P = 0.00015$ for load in peaks). The 5% RAC cut-off used in the field experiment analysis is also highly significant (Table S6). Line set has a marginal effect (IM vs. Z: $F = 4.01$, $P = 0.047$). One line (Z12) is an outlier for rare allele load (0.033) and load in 7.5% peaks (0.013), but after dropping that point, both predictors remain highly significant ($F = 10.53$, $P = 0.0015$ for rare allele load, $F = 11.85$, $P = 0.00078$ for load in peaks). Dropping set as a factor changes the estimated effect of rare allele load minimally (Table S5).

Discussion

Severity of ID: Hermaphroditic populations should evolve self-fertilization unless ID is sufficiently strong (Kimura 1959; Charlesworth and Charlesworth 1987). The prediction that δ must be greater than 0.5 to maintain outcrossing is burdened with many caveats (Lloyd 1979; Uyenoyama et al. 1993; Johnston et al. 2009), but most of these exceptions favor selfing and thus increase the necessary severity of ID to maintain outcrossing. In this experiment, we find that inbreeding to (nearly) full homozygosity has an enormous fitness cost. In the field, the lifetime seed production of inbred plants was only 14% that of their outbred competitors ($\delta = 0.86$). Accounting for lethal and sterile mutations that segregate in the natural population but not in our experimental lines, the δ estimate increases to 0.965. This value is much higher

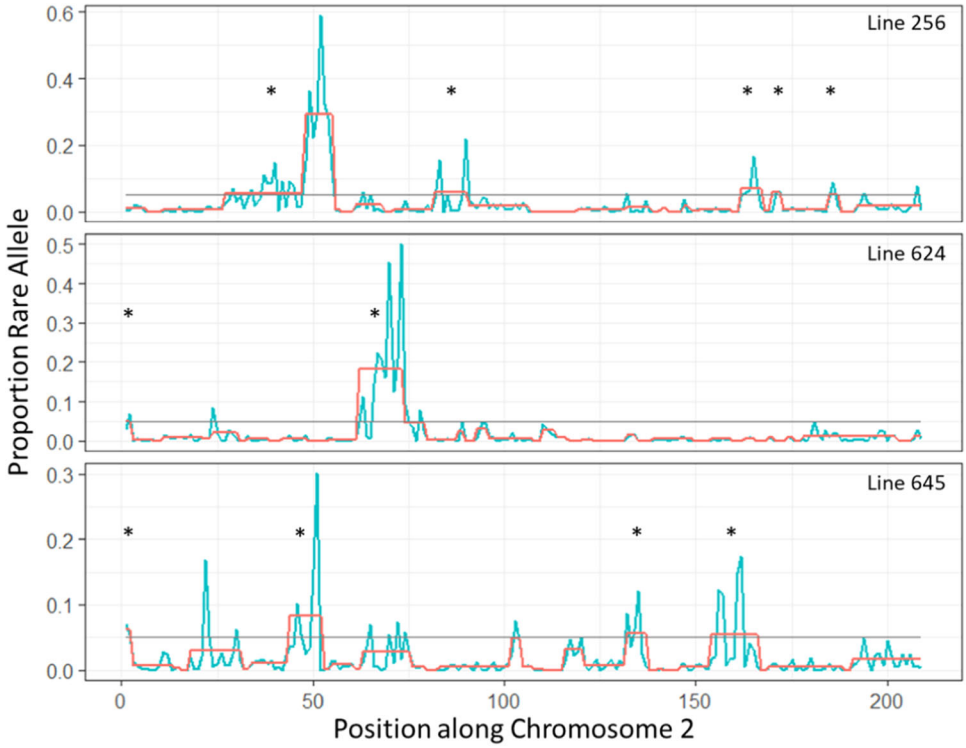

Figure 2. Rare allele proportion, within windows of 500 SNPs, is clustered into localized regions that vary among lines. Shown as an example is a section of chromosome 2 from ~10Mb to 15Mb. Red overlay indicates where *change point* has contiguous windows with similar mean and variance. The gray line at 0.05 is our cutoff for RACs in the field analysis and * denotes RACs.

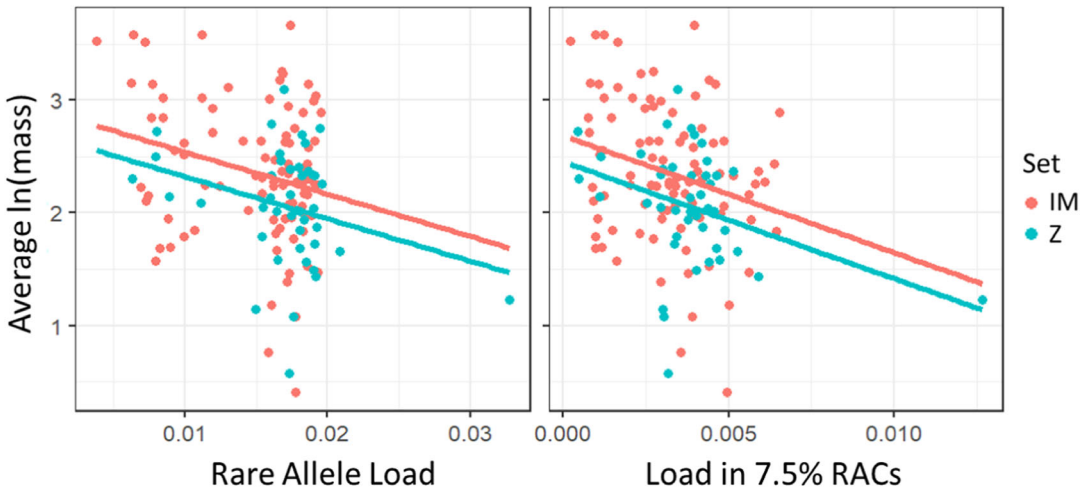

Figure 3. Regression of log mass onto (A) Rare allele load or (B) load in 7.5% RACs for greenhouse experiment. Model included line set (IM vs. Z) as a factor.

than previously obtained from populations with similar selfing rates (5−15%, Winn et al. 2011), which range from 0.21 in *Campanula americana* (Galloway et al. 2003) up to 0.53 in *Yucca filamentosa* (Huth and Pellmyr 2000) . However, ID in these studies was measured using greenhouse-grown first-generation selfed progeny from field collected plants. Here, we evaluate the fitness consequences of high homozygosity and find the cost is great enough to maintain predominant outcrossing. It may be strong enough even to resist large-effect selfing mutations that can invade when $\delta >> 0.5$, although this conclusion remains tentative given the relatively strong performance of a few lines (e.g., IM922).

Our cumulative δ (0.965) is at least roughly consistent with previous estimates from *M. guttatus*. Using first-generation selfs

($F = 0.5$), Willis (1993b) obtained $\delta = 0.69$ for the IM population in Oregon, whereas Carr and Dudash (1995) estimated even higher values (0.70−0.73) for their S and T populations in California. These δ are lower than 0.965, but this is not surprising given that the fitness decline is predicted to double when progressing from first-generation selfs to full homozygosity (Morton et al. 1956; Charlesworth and Charlesworth 1987). In stark contrast, the DUN population of *M. guttatus* (Oregon) exhibits minimal ID, a few percent depending on fitness component (Marriage and Kelly 2009). However, this extensively asexual population exhibits minimal heterozygosity/variation at both microsatellite loci and for quantitative traits. There is not expected to be much intrapopulation ID in populations where all plants have very nearly the same highly homozygous genotype. In contrast to DUN, the IM population studied here has one the highest levels of sequence variation yet documented in a plant population with synonymous nucleotide diversity of approximately 3.3% genome wide (Puzey et al. 2017).

Two caveats require attention. First, adaptation to the greenhouse environment likely occurred while making the inbred lines. Homozygous genotypes cannot adapt (except by de novo mutation), but differential germination of seeds during line formation could have favored some genotypes over others (see, for example, Fig. S1). For the field study, we germinated plants in the greenhouse and then transplanted seedlings into the field, but germination relevant loci could have pleiotropic effects on field performance. Our experimental design is insulated from this potential bias. The outbred plants in the experiments are derived from crosses between the inbred lines and thus carry the same (putatively) lab-adapted alleles in the same frequency. Any shift in allele frequencies between lines and the field population is shared equally by inbred and outbred plants, which differ only in heterozygosity. Also, given that germination is a component of fitness (not considered our experiments), ID may actually be greater than we estimate.

A more serious concern is that we measured fitness as lifetime seed production. We did not measure outcross siring success; half of adult fitness in the predominantly outcrossing IM population. However, several previous experiments in *M. guttatus* indicate a severe effect of inbreeding on male reproductive capacity. The number of viable pollen grains produced by a plant declines more substantially than other fitness components, with several populations showing an accelerating decline of pollen viability with increasing homozygosity (Willis 1993a; Carr and Dudash 1997; Carr et al. 1997; Kelly 2005). Thus, as with germination, our neglect of male fitness suggests that our already severe estimate of ID is likely underestimated.

Cause of ID: Genome sequencing provides a new opportunity to investigate mechanisms for ID and heterosis. Results vary significantly among species and even among crosses/phenotypes within species (Yu et al. 1997; Li et al. 2001b; Luo et al. 2001; Springer and Stupar 2007; Schnable and Springer 2013). In principle, sequencing allows direct evaluation of the models by attributing fitness effects to individual loci. Unfortunately, genome-wide association studies (GWAS) struggle to estimate the effects of rare alleles (Myles et al. 2009; Josephs et al. 2017), which are the cause of ID under the dominance model. Allele frequencies should be intermediate with overdominance, and thus it is noteworthy that we found no evidence of heterozygote superiority at individual SNPs (Supporting Information Appendix 1). However, our field experiment is underpowered to detect SNPs with slight overdominance, and the small variance among F_1s in heterozygosity impedes testing for an aggregate effect of (putative) overdominant loci.

We did obtain significant associations between fitness and genome-wide distillations of the number and location of rare alleles (Figs. 2, 3). It is noteworthy that a simple statistic like rare allele load predicts fitness given that the great majority of rare alleles might be neutral or nearly neutral (Kimura and Ohta 1971). The first remarkable feature of rare allele load is the extent to which it varies among lines. If minor alleles were randomly assigned to lines (linkage equilibrium), the variance in rare allele load would be approximately equal to the mean divided by the number of SNPs, that is, the Poisson expectation (Schultz and Willis 1995). The actual variance across lines (1.50×10^{-05}) is over 4000 times greater than the Poisson predicted variance (3.51×10^{-09} given an average rare allele frequency of 0.0154 and 4,379,000 SNPs scored (on average) per line). The inflation is due to positive linkage disequilibrium (LD)—rare alleles tend to co-occur within localized genomic windows of each line (Fig. 2) (Kelly et al. 2013). The rare allele load is a standardized sum of 0/1 values across SNPs, and thus its variance is the sum of single locus variances (binomial) plus twice the sum of covariances across all pairs of loci. Thus, although the rare allele load is not the deleterious allele load, the former may be an indicator of the latter due to LD between neutral and deleterious alleles. Recently, Kremling et al. (2018) showed that rare alleles per inbred line are correlated with extreme gene expression patterns in maize. They also found fitness consequences to this "dysregulation," but noted that maize breeding history somewhat confounds this effect.

We partitioned the rare allele load after noting the nonrandom genomic distribution of rare alleles within lines (Fig. 1). We found that rare alleles in genomic clusters (RACs) were a slightly better predictor of fitness decline in both field (Fig. 2) and greenhouse (Fig. 3) than either the overall rare allele load or the load outside of RACs. This observation is preliminary but intriguing. Migration of pollen and/or seed from other populations into IM might be an important source of deleterious alleles insofar as RACs are the remnants of immigrant genomes. Indeed, RACs resemble "introgressed segments" from an experimental intercross population (Fig. 1; Tanksley 1983; Patterson et al.

2004). Of course, if local environmental conditions make rare alleles deleterious, why is the relationship evident in the novel greenhouse environment? Our greenhouse experiment evaluated early growth rate. IM has a short growing season that exerts selection for fast growth and rapid progression to flower (Mojica et al. 2012). If the migrant alleles confer slower growth early in life (in terms of above ground biomass), introgressed alleles are likely to reduce fitness in this assay. Indeed, some populations close to Iron Mountain exhibit a slower growth life history (Troth et al. 2018).

RACs are defined based on allele frequency and SNP location and they do not likely provide a clean partitioning of mutations into different types (e.g., locally adaptive alleles vs. unconditionally deleterious mutations). In contrast, Genomic Evolutionary Rate Profiling (GERP) uses phylogenetic conservation across highly divergent taxa to identify sites that are constrained (Cooper et al. 2005). Scoring novel alleles at conserved sites as putatively deleterious, Yang et al. (2017) found a negative correlation between GERP score for a SNP and frequency in a collection of inbred maize lines, as well as a positive correlation between GERP score and effect on grain yield (fitness in the agricultural environment). GERP uses conservation across a broad taxonomic scale and the behavior of the statistic for within-species variation resulting from local adaptation is unclear. We propose that combining RAC analysis and GERP scoring might be fruitful in future studies to better classify rare alleles as neutral, conditionally deleterious, or unconditionally deleterious.

In summary, we here demonstrate that sustained inbreeding is severely detrimental in yellow monkeyflower, and further that ID is predicted by the load of rare alleles carried by an inbred line. Rare alleles are nonuniformly distributed across the genome, and partitioning them into load within clusters and background load better predicts both field and greenhouse fitness. The presence of fitness-determining clusters of rare alleles indicates that migration or introgression could be an important source of deleterious variants in the Iron Mountain population of yellow monkeyflower.

AUTHOR CONTRIBUTIONS

JKK and KEB conceived of the project and undertook field and greenhouse planting and data collection collaboratively. JKK performed all genomic tests, whereas KEB performed most of the statistical analyses. KEB wrote the manuscript, with substantial contributions and revisions by JKK.

ACKNOWLEDGMENTS

We thank J. Willis, R. Shaw, J. Colicchio, D. Charlesworth, C. Wessinger, D. Hall, M. Streisfeld, and two anonymous reviewers for input on the project and/or manuscript. We thank A. Arteaga, P. Gamero, A. Martin, and P. Sterkhova for assistance in data collection and C. Friesen (U.S. Forest Service) for site access. KEB thanks S. Brown and C. Gray for field work transportation. This work was supported by grants from the National Institutes of Health to JKK (R01 GM073990) and KU Botany Endowment to KEB.

DATA ARCHIVING

All sequence data is already available through NCBI. R and python code and data will be made available upon manuscript acceptance through Dryad https://doi.org/10.5061/dryad.00000000g.

LITERATURE CITED

Awadalla, P., and K. Ritland. 1997. Microsatellite variation and evolution in the *Mimulus guttatus* species complex with contrasting mating systems. Mol. Biol. Evol. 14:1023–1034.

Baker, H. G. 1955. Self-compatibility and establishment after "long-distance" dispersal. Evolution 9:347–349.

Bates, D., D. Sarkar, M. D. Bates, and L. Matrix. 2007. The lme4 package. R package version 2, 74.

Bertin, R. I., and C. M. Newman. 1993. Dichogamy in angiosperms. Bot. Rev. 59:112–152.

Carr, D. E., and M. R. Dudash. 1995. Inbreeding depression under a competitive regime in *Mimulus guttatus*: consequences for potential male and female function. Heredity 75:437–445.

———. 1997. The effects of five generations of enforced selfing on potential male and female function in *Mimulus guttatus*. Evolution 51:1797–1807.

Carr, D. E., C. B. Fenster, and M. R. Dudash. 1997. The relationship between mating-system characters and inbreeding depression in *Mimulus guttatus*. Evolution 51:363–372.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1991. Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. Genet. Res. 57:177–194.

Charlesworth, D., and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. 18:237–268.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10:783.

Cooper, G. M., E. A. Stone, G. Asimenos, E. D. Green, S. Batzoglou, and A. Sidow. 2005. Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 15:901–913.

Darwin, C. R. 1876. The effects of cross- and self-fertilisation in the vegetable kingdom. John Murray, London.

———. 1877. The different forms of flowers on plants of the same species. John Murray, London.

Fisher, R. A. 1941. Average excess and average effect of a gene substitution. Ann. Eugen. 11:53–63.

Galloway, L., J. Etterson, and J. Hamrick. 2003. Outcrossing rate and inbreeding depression in the herbaceous autotetraploid, *Campanula americana*. Heredity 90:308.

Ganders, F. R. 1979. The biology of heterostyly. New Zealand J. Bot. 17:607–635.

Geyer, C. J., S. Wagenius, and R. G. Shaw. 2007. Aster models for life history analysis. Biometrika 94:415–426.

Geyer, C. J., C. E. Ridley, R. G. Latta, J. R. Etterson, and R. G. Shaw. 2013. Local adaptation and genetic effects on fitness: calculations for exponential family models with random effects. Ann. Appl. Statist. 7:1778–1795.

Goodwillie, C. 2005. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. Annu. Rev. Ecol. Evol. Syst. 36:47–79.

Grossenbacher, D., R. Briscoe Runquist, E. E. Goldberg, and Y. Brandvain. 2015. Geographic range size is predicted by plant mating system. Ecol. Lett. 18:706–713.

Holsinger, K. E. 1988. Inbreeding depression doesn't matter: the genetic basis of mating-system evolution. Evolution 42:1235–1244.

Huth, C. J., and O. Pellmyr. 2000. Pollen-mediated selective abortion in yuccas and its consequences for the plant–pollinator mutualism. Ecology 81:1100–1107.

Igic, B., L. Bohs, and J. R. Kohn. 2006. Ancient polymorphism reveals unidirectional breeding system shifts. Proc. Natl. Acad. Sci. 103: 1359–1363.

Johnston, M. O., E. Porcher, P.-O. Cheptou, C. G. Eckert, E. Elle, M. A. Geber, S. Kalisz, J. K. Kelly, D. A. Moeller, M. Vallejo-Marin et al. 2009. Correlations among fertility components can maintain mixed mating in plants. Am. Nat. 173:1–11.

Josephs, E. B., J. R. Stinchcombe, and S. I. Wright. 2017. What can genome-wide association studies tell us about the evolutionary forces maintaining genetic variation for quantitative traits? New Phytol. 214:21–33.

Kelly, J. K. 2003. Deleterious mutations and the genetic variance of male fitness components in *Mimulus guttatus*. Genetics 164:1071–1085.

———. 2005. Epistasis in monkeyflowers. Genetics 171:1917–1931.

———. 2008. Testing the rare alleles model of quantitative variation by artificial selection. Genetica 132:187–198.

Kelly, J. K., and M. K. Tourtellot. 2006. The genetic analysis of family structured inbreeding depression studies. Heredity 97:346–354.

Kelly, J. K., B. Koseva, and J. P. Mojica. 2013. The genomic signal of partial sweeps in *Mimulus guttatus*. Genome Biol. Evol. 5:1457–1469.

Killick, R., and I. Eckley. 2014. changepoint: an R package for changepoint analysis. J. Statist. Softw. 58:1–19.

Kimura, M. 1959. Conflict between self fertilization and outbreeding in plants. Ann. Rep. Natl. Inst. Genet. Japan 9:87–88.

Kimura, M., and T. Ohta. 1971. Theoretical aspects of population genetics. Princeton Univ. Press, Princeton, NJ.

Kimura, M., and T. Maruyama. 1966. The mutation load with epistatic gene interactions in fitness. Genetics 54:1337–1351.

Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336:435–440.

Kremling, K. A., S.-Y. Chen, M.-H. Su, N. K. Lepak, M. C. Romay, K. L. Swarts, F. Lu, A. Lorant, P. J. Bradbury, and E. S. Buckler. 2018. Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. Nature 555:520.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24–40.

Lee, Y. W., L. Fishman, J. K. Kelly, and J. H. Willis. 2016. A segregating inversion generates fitness variation in yellow monkeyflower (*Mimulus guttatus*). Genetics 202:1473–1484.

Li, Q.-J., Z.-F. Xu, W. J. Kress, Y.-M. Xia, L. Zhang, X.-B. Deng, J.-Y. Gao, and Z.-L. Bai. 2001a. Pollination: flexible style that encourages outcrossing. Nature 410:432.

Li, Z.-K., L. Luo, H. Mei, D. Wang, Q. Shu, R. Tabien, D. Zhong, C. Ying, J. Stansel, and G. Khush. 2001b. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. Genetics 158:1737–1753.

Lloyd, D. G. 1979. Some reproductive factors affecting the selection of self-fertilization in plants. Am. Nat. 113:67–79.

Luo, L., Z.-K. Li, H. Mei, Q. Shu, R. Tabien, D. Zhong, C. Ying, J. Stansel, G. Khush, and A. Paterson. 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. Genetics 158:1755–1771.

Marriage, T. N., and J. K. Kelly. 2009. Inbreeding depression in an asexual population of *Mimulus guttatus*. J. Evol. Biol. 22:2320–2331.

Mojica, J. P., and J. K. Kelly. 2010. Viability selection prior to trait expression is an essential component of natural selection. Proc. R. Soc. B-Biol. Sci. 277:2945–2950.

Mojica, J. P., Y. W. Lee, J. H. Willis, and J. K. Kelly. 2012. Spatially and temporally varying selection on intrapopulation quantitative trait loci for a life history trade-off in *Mimulus guttatus*. Mol. Ecol. 21:3718–3728.

Monnahan, P. J., and J. K. Kelly. 2015. Naturally segregating loci exhibit epistasis for fitness. Biol. Lett. 11. https://doi.org/10.1098/rsbl.2015.0498

Morton, N. E., J. F. Crow, and H. J. Muller. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. Proc. Natl. Acad. Sci. USA 42:855–863.

Myles, S., J. Peiffer, P. J. Brown, E. S. Ersoz, Z. Zhang, D. E. Costich, and E. S. Buckler. 2009. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21:2194–2202.

Opedal, Ø. H. 2018. Herkogamy, a principal functional trait of plant reproductive biology. Int. J. Plant Sci. 179:677–687.

Patterson, N., N. Hattangadi, B. Lane, K. E. Lohmueller, D. A. Hafler, J. R. Oksenberg, S. L. Hauser, M. W. Smith, S. J. O'Brien, D. Altshuler et al. 2004. Methods for high-density admixture mapping of disease genes. Am. J. Hum. Genet. 74:979–1000.

Puzey, J. R., J. H. Willis, and J. K. Kelly. 2017. Population structure and local selection yield high genomic variation in *Mimulus guttatus*. Mol. Ecol. 26:519–535.

Renner, S. S., and R. E. Ricklefs. 1995. Dioecy and its correlates in the flowering plants. Am. J. Bot. 82:596–606.

Ritland, K., and F. R. Ganders. 1987. Covariation of selfing rates with parental gene fixation indices within populations of *Mimulus guttatus*. Evolution 41:760–771.

Schnable, P. S., and N. M. Springer. 2013. Progress toward understanding heterosis in crop plants. Annu. Rev. Plant Biol. 64:71–88.

Schultz, S. T., and J. H. Willis. 1995. Individual variation in inbreeding depression: the roles of inbreeding history and mutation. Genetics 141:1209–1223.

Springer, N. M., and R. M. Stupar. 2007. Allelic variation and heterosis in maize: how do two halves make more than a whole? Genome Res. 17:264–275.

Sweigart, A., K. Karoly, A. Jones, and J. H. Willis. 1999. The distribution of individual inbreeding coefficients and pairwise relatedness in a population of *Mimulus guttatus*. Heredity 83:625.

Takayama, S., and A. Isogai. 2005. Self-incompatibility in plants. Annu. Rev. Plant Biol. 56:467–489.

Tanksley, S. D. 1983. Molecular markers in plant breeding. Plant Mol. Biol. Rep. 1:3–8.

Team, R. C. 2013. R: A language and environment for statistical computing.

Troth, A., J. R. Puzey, R. S. Kim, J. H. Willis, and J. K. Kelly. 2018. Selective trade-offs maintain alleles underpinning complex trait variation in plants. Science 361:475–478.

Uyenoyama, M. K., K. E. Holsinger, and D. M. Waller. 1993. Ecological and genetic factors directing the evolution of self-fertilization. Oxford Surv. Evol. Biol. 9:327–381.

Willis, J. H. 1993a. Effects of different levels of inbreeding on fitness components in *Mimulus guttatus*. Evolution 47:864–876.

———. 1993b. Partial self fertilization and inbreeding depression in two populations of *Mimulus guttatus*. Heredity 71:145–154.

———. 1999a. Inbreeding load, average dominance, and the mutation rate for mildly deleterious alleles in *Mimulus guttatus*. Genetics 153:1885–1898.

———. 1999b. The role of genes of large effect on inbreeding depression in *Mimulus guttatus*. Evolution 53:1678–1691.

Winn, A. A., E. Elle, S. Kalisz, P. O. Cheptou, C. G. Eckert, C. Goodwillie, M. O. Johnston, D. A. Moeller, R. H. Ree, and R. D. Sargent. 2011. Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. Evolution 65:3339–3359.

Yang, J., S. Mezmouk, A. Baumgarten, E. S. Buckler, K. E. Guill, M. D. McMullen, R. H. Mumm, and J. Ross-Ibarra. 2017. Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. PLoS Genet. 13:e1007019.

Yu, S., J. Li, C. Xu, Y. Tan, Y. Gao, X. Li, Q. Zhang, and M. S. Maroof. 1997. Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc. Natl. Acad. Sci. 94:9226–9231.

Associate Editor: M. A. Streisfeld
Handling Editor: D. W. Hall

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1: Interaction between line set (IM vs Z) and stratification treatment on greenhouse germination rate.
Figure S2: We find no difference in either survival or total fitness in the field between outcrossed progeny from single vs group sires.
Figure S3: Histogram of seed set in the field.
Figure S4: Histogram of ln(mass) in greenhouse grown plants, and residual values for the multiple linear regression of average mass and rare allele load (Figure 2).
Figure S5: Examples of the differences in effect of cross type between lines on field survival.
Figure S6: Effect of rare allele load (RAL) of inbred lines (A, B) and heterozygosity (H) of F1s (C, D) on survival and total seed in the field experiment.
Figure S7: Regression of log mass onto (A) Rare allele load or (B) load in 5% RACs for greenhouse experiment.
Table S1 (separate file): Table of 5,018,997 variants for which one allele is considered rare (at a frequency of <= 5% in the inbred lines).
Table S2: Model fits for the effect of self vs. outcross on field fitness.
Table S3: Model test for the effect of rare allele load or load in 5% peaks on field fitness in the inbred lines.
Table S4: Model fits for different cutoff values to define "peaks" of high rare allele load on field survival and seed set (total fitness).
Table S5: Model fits for effect of rare allele load or load in 5% peaks on ln(mass) in the greenhouse implemented in the built-in R function *lm*.
Table S6: Model fits for different cutoff values to define "peaks" of high rare allele load on greenhouse growth rate.
Table S7: Model fits for the effect of proportion heterozygous loci on field fitness in the single outcrossed families.