Original article: Heterosis in crosses between remnant populations of a rare prairie forb: implications for restoration genetics

Christopher G. Oakley*[1], Isabelle A. Turner[1,2], Juan Diego Rojas-Gutierrez[1], Bob Easter[3]

[1] Department of Botany and Plant Pathology, and the Center for Plant Biology, Purdue University, West Lafayette, IN 47907, USA

[2] Present address: Department of Biology, Miami University, Oxford, OH 45056, USA

[3] NICHES Land Trust, Lafayette, IN 47905, USA

Running title: Heterosis in remnant populations of a rare prairie forb

*Corresponding author: Christopher Oakley, oakleyc@purdue.edu

ABSTRACT

- *Background and Aims*

  A major goal in both basic and applied evolutionary genetics is understanding the role of drift in shaping genetic variation with fitness consequences. For organisms amenable to controlled crosses, the increased fitness of progeny from between- relative to within-population crosses (heterosis) provides a useful tool to infer historical fixation of partly recessive deleterious alleles by drift and/or bottlenecks. Additionally, heterosis may also provide a unique benefit in regional admixture provenancing for restorations of rare plant species, but empirical data to directly test this idea is lacking.

- *Methods*

  We quantified heterosis in crosses between small remnant populations of a rare prairie forb (*Silene regia* Sims). We measured early fitness components (seed number per fruit, germination, and juvenile survival) in a controlled environment. Adult fitness components (survival and reproduction) were quantified over two flowering seasons in two different environments, a field common garden simulating the initial stages of a restoration, and a greenhouse. To our knowledge, our study is the first to estimate cumulative fitness for a rare species in conditions of a newly established restoration in the field. Additionally, we provide a unique comparison of heterosis estimated under controlled vs. field conditions.

- *Key Results*

  The consequences of between-population crosses for cumulative fitness in the field were strongly positive in two of the populations (132-292% heterosis) and neutral in a third.

Heterosis was generally stronger when measured under field conditions and was more likely to be detected for adult fitness components.

- *Conclusions*

Regional admixture provenancing in restorations should be beneficial for this species. We advocate for more research on heterosis in rare species, particularly in restorations. Such restorations can be opportunities for experimental genetics in addition to reestablishing important ecosystem services.

Key words: Conservation, Deleterious mutations, Fragmented populations, Genetic drift, Genetic rescue, Heterosis, Prairie, Regional admixture provenancing, Restoration, Revegetation, *Silene regia*

INTRODUCTION

In nature, many plant populations are of limited size, spatially structured, and have a history of inbreeding, population bottlenecks, or both. Thus, genetic drift may play an important role in the adaptive potential and persistence of many populations, particularly in the face of anthropogenic effects such as habitat loss, fragmentation, and climate change. The loss of genetic variation and increased homozygosity brought about by drift will have two non-mutually exclusive consequences. Drift may result in limited quantitative genetic variation needed to adapt to fluctuating or changing conditions (Willi *et al.*, 2006), which could lead to greater extinction risk (Newman and Pilson, 1997), particularly if the environment changes rapidly (Gomulkiewicz and Houle, 2009). Additionally, drift is predicted to reduce the effectiveness of selection against deleterious mutations (Kimura *et al.*, 1963). Most new mutations are on average deleterious and partially recessive (Halligan and Keightley, 2009, Agrawal and Whitlock, 2011, Keightley *et al.*, 2014), but see (Rutter *et al.*, 2012, Weng *et al.*, 2021). Although many rare alleles are likely to be lost because of drift, those that do drift to fixation within populations are expected to be partly recessive and deleterious (Crow, 1948, Kimura *et al.*, 1963, Whitlock *et al.*, 2000, Glémin *et al.*, 2003, Charlesworth, 2018), which could decrease individual fitness and, under some circumstances, increase extinction risk (Lynch *et al.*, 1995).

Direct study of the effects of drift on genetic variation with fitness consequences is challenging. Heterosis, or the increased fitness in between- relative to within-population crosses is one way to probe the balance between selection and drift in nature (Fenster and Galloway, 2000, Paland and Schmid, 2003, Oakley and Winn, 2012, Lohr and Haag, 2015). Heterosis is common in crosses between natural populations (van Treuren *et al.*, 1993, Ouborg and Van Treuren, 1994, Luijten *et al.*, 2002, Willi and Fischer, 2005, Pickup *et al.*, 2012, Willi, 2013,

Weisenberger *et al.*, 2014, Koski *et al.*, 2019, Oakley *et al.*, 2019, Söderquist *et al.*, 2020). Observations of greater heterosis in small/less diverse compared to large/more diverse populations (Paland and Schmid, 2003, Escobar *et al.*, 2008, Oakley and Winn, 2012, Lohr and Haag, 2015, Spigler *et al.*, 2017) and in selfing compared to outcrossing populations (Busch, 2006, Oakley *et al.*, 2015a) is consistent with a role of drift in fixing partly recessive deleterious alleles (Charlesworth, 2018). Line cross approaches have also provided more direct evidence for a role of dominance complementation in heterosis (Oakley *et al.*, 2015b).

One open question is how much the magnitude of heterosis depends on the environmental assay conditions. The effect of the environment on inbreeding depression, due to partly recessive deleterious alleles segregating within a population, is well documented (Dudash, 1990, Fox and Reed, 2010, Cheptou and Donohue, 2011). However, relatively few estimates of heterosis for natural populations come from field plantings (but see for e.g., Maschinski *et al.*, 2013, Weisenberger *et al.*, 2014). Further, we are not aware of any comparisons of heterosis in field vs. controlled conditions. A final consideration is that if many mildly deleterious alleles make up heterosis (Whitlock *et al.*, 2000, Glémin *et al.*, 2003, Charlesworth, 2018), its expression may be weak for individual fitness components, particularly early in the life cycle (c.f., Husband and Schemske, 1996). Therefore, obtaining accurate estimates of heterosis requires quantifying cumulative fitness, encompassing as many fitness components as possible.

In addition to informing basic questions in evolutionary biology, heterosis is potentially useful in practical applications such as conservation and restoration of rare species. A major concern in conservation genetics is the negative fitness consequences of genetic load caused by deleterious recessive alleles (Frankham, 2005, Bell *et al.*, 2019). There have been repeated calls for the use of assisted gene flow/migration as a tool to provide both genetic rescue from

deleterious mutation load, as well as evolutionary rescue from limited quantitative genetic variation (Frankham, 2015, Ralls *et al.*, 2018, Bell *et al.*, 2019). Data thus far suggests that these efforts can be largely successful (Frankham, 2015, 2016) within certain guidelines to minimize reduced fitness due to either loss of local adaptation or genetic incompatibilities (Frankham *et al.*, 2011, Weeks *et al.*, 2011).

Similar issues are faced in the restoration/revegetation of rare plant species. Here, a major question revolves around where the source material should be collected. A traditional view is that seeds should be sourced as locally as possible to minimize the risk of disrupting adaptation to local environmental (extrinsic) conditions or outbreeding depression due to (intrinsic) genetic incompatibilities (McKay *et al.*, 2005, Gann *et al.*, 2019, Török *et al.*, 2024). On the other hand, potential seed sources for restorations are often small remnant populations which may harbor fixed deleterious mutations and have limited adaptive potential, particularly in the face of climate change. For these reasons many authors have called for some version of seed provenancing strategies that increase genetic variation while still minimizing the risk of disrupting local adaptation (Broadhurst *et al.*, 2008, Breed *et al.*, 2013, Bucharova *et al.*, 2019, St. Clair *et al.*, 2020, Hoffmann *et al.*, 2021, Stojanova *et al.*, 2021, Jordan *et al.*, 2023).

In "regional admixture provenancing" (Bucharova *et al.*, 2019) for example, seed would be sourced from multiple populations to increase genetic variation. These populations would however come from similar habitat in the same eco-region (possibly connected by gene flow prior to fragmentation) to minimize the risks of maladaptation and genetic incompatibilities. Additionally, heterosis following admixture in a restoration may provide at least a short term fitness boost that may aid establishment of rare species, analogous to the "catapult effect" proposed for the role of heterosis in the establishment of invasive species (Drake, 2006, Li *et al.*,

2018). Despite the potential benefits, there are few empirical tests of the regional admixture provenancing approach (but see St. Clair *et al.*, 2020), or the potential for heterosis during a newly established restoration.

Here we use crosses between small remnant populations of the rare prairie forb *Silene regia* Sims to address several questions. 1) What is the magnitude of heterosis, and does it vary among populations? 2) How does the magnitude of heterosis vary across different life history stages? And 3) Is heterosis in a field common garden greater than in controlled greenhouse conditions?

MATERIALS AND METHODS

Study system

Silene regia (royal catchfly) is a tap-rooted perennial in the Caryophyllaceae family (Dolan, 1994). It is endemic to a dozen states in eastern North America and is thought to have been broadly distributed across prairies prior to the widespread conversion of this habitat to agriculture (Menges, 1991). Contemporary populations across much of the native range are small and isolated and exist as remnants in prairie cemeteries and along railroad rights-of-way and roadsides (Dolan, 1994). This species does not vegetatively reproduce and achieves sexual reproduction via hummingbird pollination. It is self-compatible and some geitonogamous selfing may occur, but it is likely mostly outcrossing as the individual flowers are protandrous (Dolan, 1994, Menges, 1995).

We selected three remnant focal populations in different counties in western Indiana (Tip = Tippecanoe; Fou = Fountain; Ver = Vermillion) for bulk seed collection in autumn 2019. Exact locations of these populations are not provided because this species is critically imperiled in

Indiana (NatureServe, 2024), but these populations are all separated by at least 14 Km. The smallest population was Tip (10 individuals), which has likely been small for an extended period. The largest population was Ver (50 individuals), and population size reductions are likely more recent at this site (B. Easter, pers. obs). The size of Fou was intermediate to the other populations (20 individuals).

Generation of experimental material

Bulk seed was collected from each population in September 2019. Our goal was to characterize population level heterosis, and the small number of individuals in these populations precluded sampling distinct maternal lineages. We sampled from most of the reproductive individuals in the two smaller populations and from a haphazard subset in the largest population.

In early October 2019, these seeds were surface sterilized in a 10% bleach solution for 10 minutes, followed by three rinses in deionized-autoclaved water. The seeds were then placed into moist paper towels in bags and cold stratified in the dark at 4°C for 10, 30, or 60 days. Different stratification periods were used because 60 days would be sufficient, but starting the plants sooner was preferrable due to logistical constraints. Because of poor germination we included individual maternal plants arising from all three seed stratification treatments. While there was some signature of maternal seed stratification environment affecting early fitness components of the next generation (not shown), this treatment did not qualitatively affect results related to heterosis (see statistical analysis below), so we omitted it from subsequent analyses.

After stratification, we sowed seeds (5 per cell) into a 72-cell flat insert (cell volume = 70 ml, Greenhouse Megastore, Danville, Illinois, USA), filled with a 3:1 mixture of MetroMix potting soil (Sun Gro Horticulture, Agawan, Massachusetts, USA) and autoclaved sand. A total

of 30-50 seeds were sown for each population-stratification treatment combination. The seeds were then germinated in an incubator (Percival scientific, Perry, Iowa, USA) set to a constant 27°C and a 12-hour photoperiod. After approximately 4 weeks, cells were thinned to one seedling per cell and the flat was moved to a heated greenhouse (minimum: 20°C) with supplemental lighting (12-hour photoperiod). Plants were grown there for approximately 6 weeks, watering as needed. Plants were then moved to a large refrigerator (4°C) with lights (10 hr photoperiod) for approximately one month to synchronize flowering. After the vernalization period, in mid-January 2020, plants were moved back to the greenhouse and transplanted into 17 cm circular pots (volume = 2.4 l, Greenhouse Megastore, Danville, Illinois, USA) filled with the same soil mixture. Plants were watered as needed and fertilized monthly with ¼ strength Hoagland's solution.

Once plants had begun to flower (early March 2020), we performed controlled hand pollinations within and between populations in all possible combinations and crossing directions. In mid-March, because of COVID related shutdowns, the plants (19, 13, and 10 individuals from Fou, Tip, and Ver respectively) were moved to a private residence and grown on light racks with a 14-hour photoperiod for the duration of the hand pollinations. Most of the pollinations were performed from mid-March through the end of June, 2020. Before each pollination we identified ovule donor flowers several days prior to anthesis and removed the upper part of the corolla tube, filaments, and anthers with scissors and forceps. Pollinations were completed when the stigma lobes of the ovule donor flowers appeared to be receptive (usually 3-4 days later) by removing stamens from a pollen donor flower with forceps and brushing the dehisced anthers against the stigma. Each pollination was tracked with a unique color combination of embroidery floss tied around the pedicel. In addition to hand pollinations, we also performed control

emasculations where flowers were emasculated but no pollen was applied to the stigma lobes. Of the 69 control emasculations, less than 3% (2 emasculations) resulted in a fruit, giving us confidence that the vast majority of the pollinations performed were successful crosses between the intended pollen and ovule donors.

The number of hand pollinations per maternal population ranged from 39-58 (mean = 46) for within-population crosses and 85-102 (mean = 91) for between-population crosses, for a grand total of 411. The approximately two-fold greater number of hand pollinations for between-population crosses is because each maternal population was crossed to two different paternal populations, and we wanted sufficient power to test for interactions between maternal and paternal populations.

Early fitness components

We collected fruits of all successful pollinations to be used in subsequent fitness assays. This species required an extended period between pollination and fruit maturity (mean = 56.8 days, SD =10.4). There were no patterns in fruit set or the rate of fruit development with respect to cross type, but fruits of Tip developed on average about 5 days faster than fruits from the other two populations. For each of 17-39 within-population cross, and 45-71 between-population cross fruits per maternal population (Table S1), we counted the number of seeds per fruit, not including small, shriveled seeds presumed to be inviable.

A subset of the viable seeds for each population and cross type was selected prioritizing fruits with enough viable seeds for germination assays, and the goal of proportionally balanced sample sizes. This subset included approximately 300 seeds each of the WIN and the two BET crosses for populations Tip and Ver, and approximately double those numbers for Fou (Table

S1). These seeds were cold stratified as previously described for a period of 7 weeks starting in mid-November of 2020. Seeds were then sown into 72 cell flat inserts filled with the same soil mix (5 flats in total) with 7-12 (mean = 9.96, SD = 0.34) seeds per cell in a randomized block design. The number of cells per flat varied by maternal population with an average of 6 and 12 for within-, and between-population crosses respectively from Tip and Ver, and approximately double the sample size for each cross type from Fou (Table S1). We recorded germination daily for a period of 24 days starting in mid-January 2021. Most seeds germinated within the first week, and only 5 new germinants were recorded during the second half of the scoring period.

Once seeds had germinated and seedlings had two true leaves, they were transplanted into individual cells of 72 cell flat inserts and moved to the greenhouse. The seedlings were transplanted into an incomplete randomized block design, with a total of 30 flats (blocks). The number of cells per flat varied by maternal population, with an average of 5-6 and 10 for within-, and between-population crosses respectively from Tip and Ver, and an average of 11 within- and 24 between-population crosses from Fou. Any seedlings that died within the first week were replaced with another seedling of the same population-cross type. This mortality was likely due to transplant shock and these individuals were omitted from subsequent analyses. Once roots started to poke through the bottom of the cells, individuals were transplanted first into 6 cm square (volume = 250 ml, Greenhouse Megastore, Danville, Illinois, USA) pots filled with the same soil mixture. Plants were watered as needed and fertilized twice with ¼ strength Hoagland's solution. Cumulative juvenile survival was recorded up to approximately seven weeks after stratification, excluding individuals that died within one week of the initial transplant. Surviving plants that were established in the 6 cm pots were then randomly divided

into two experiments, one in the greenhouse and one in a field common garden simulating a new restoration.

Greenhouse experiment

In late March 2021, seedlings were transplanted into 11 cm square (volume = 1 l, Greenhouse Megastore, Danville, Illinois, USA) pots filled with the same potting mix in an incomplete randomized block design consisting of 48 flats, each containing 7-8 pots. Each maternal population cross type combination was represented by an average of 1.3-2.6 plants in at least 19 different flats (blocks). The total size of the greenhouse experiment was 354 plants, with sample sizes for between population crosses being approximately double that of within population crosses, and sample sizes for Fou being approximately double that of the other populations (Table S1). Plants were grown for approximately 18 months, until mid-October of 2022, which included two flowering periods. During this time plants were watered as needed and fertilized once with Osmocote timed release 15-9-12 fertilizer at ½ the recommended dose after the first flowering period was over.

For each plant in the greenhouse experiment we scored adult fitness as the cumulative number of flowers produced over ~18 months (two flowering seasons), including zeros for plants that died or did not reproduce. We additionally quantified two components of this composite measure of adult fitness separately: combined cumulative probability of survival and reproduction, and fecundity (flower number) of reproducing plants.

Field experiment

Shortly after the start of the greenhouse experiment, we established an additional field common garden experiment to quantify adult fitness components under more realistic conditions. The site was located on the property of Arbor America LLC in West Point, Indiana (Tippecanoe County) USA. This location is within the extant range of *S. regia* and the field common garden experiment was planted on Coloma Sand (Mixed, mesic Lamellic Udipsamments) a soil type and series similar to the well-drained sites where *S. regia* naturally occurs (ISEE Network, 2024). The plot for this experiment had previously been treated to establish a prairie strip, with an application of Roundup in May 2020 to remove the turf, followed by controlled burning approximately 1 month later and another application of Roundup approximately 2 months after that. In early March 2021, the plot was hand seeded with a short stature high-diversity prairie mix from Spence Restoration Nursery (Muncie, IN, USA).

In early April 2021, seedlings were transplanted into a randomized complete block design consisting of six blocks containing 120 plants each. In total there were 720 plants, with approximately double the sample size of the greenhouse experiment for each maternal population-cross type combination (Table S1). Each block was a rectangular grid of 3 x 40 plants with 30 cm spacing between the plants. Plants were hand watered for the first week only, and plants that died during this first week were considered to have succumbed to transplant shock and were replaced. In July of both 2021 and 2022, other vegetation within 10 cm of the base of experimental plants was clipped at the base once to facilitate locating plants for the post-reproductive censuses.

In late October 2021, we recorded whether each plant was still alive and had reproduced, and for plants that had reproduced we counted the number of fruits. Only about 11% of all plants had reproduced in the first year. We re-censused the same fitness components in October 2022.

We estimated adult fitness as the number of fruits (total across both years) produced per individual transplanted, including zeros for plants that died or did not reproduce. We additionally calculated the combined cumulative (across both years) probability of survival and reproduction for each plant, and cumulative fecundity (number of fruits) for each reproducing plant.

Statistical analyses and estimates of heterosis

We analyzed each fitness component using ANOVA. Preliminary analyses indicated no significant effects of a maternal population by paternal population interaction for any fitness component (results not shown), so the identity of the paternal population did not influence estimates of heterosis. We therefore used a simplified model of maternal population (Fou, Tip, and Ver), cross type (within- vs. between-population crosses) and their interaction. Our primary interests were the effects of cross type, which would indicate heterosis averaged over all populations, and the maternal population by cross type interaction, which would indicate that the presence and/or magnitude of heterosis varies among maternal populations. For all fitness components except seed number per fruit, we included a term for block in the analysis to account for micro-spatial variation within the experiment. All effects were treated as fixed effects. Inspection of residuals indicated no major issues with normality or heteroscedasticity for many fitness components, and for these we used normal error distributions. Juvenile survival and the combined probability of survival and reproduction in both the greenhouse and field experiments required a generalized linear model with a binomial error distribution. For adult fitness in the field, the large number of zeros required us to use a generalized regression with a zero-inflated Poisson error distribution, followed by comparisons of least squares means of between- and within- population crosses for each maternal population. An alternative, but less powerful,

analysis using block means and a normal error distribution yielded qualitatively similar results (not shown). All statistical analyses were performed in JMP Pro v. 16.1.0 (JMP, 1989-2021).

For each fitness component and composite estimate of fitness we calculated heterosis separately for each population. This was calculated as the percent increase in fitness of between population crosses relative to within population crosses ([(BET-WIN)/WIN]*100), where positive values indicate heterosis and negative values indicate outbreeding depression (Oakley *et al.*, 2015a). For each fitness component we used the means for each maternal population-cross type combination averaging over the block means (Table S1) for these calculations.

For cumulative fitness, calculated separately for the greenhouse and field experiments, we used the product of four estimates: 1) seed number per fruit, 2) proportion germination, 3) proportion juvenile survival, and 4) adult fitness. Because this multiplicative measure of cumulative fitness yields only one estimate per cross type per population, we could not directly test the effects of population cross type and their interaction as with the fitness components. Instead, we used bootstrap resampling to generate 20 replicate estimates of cumulative fitness for each population and cross type. For each replicate estimate we drew 10 values with replacement from the empirical values for each of the 4 fitness components indicated above, took the mean of those 10 values, and multiplied the four means. We then performed non-parametric ANOVA separately by greenhouse and field experiments, each using 120 total bootstrapped values for cumulative fitness and model terms as described above.

RESULTS

Early fitness components

In general, heterosis was weak and often non-significant for early fitness components (Figs. 1&2, Table 1). For seed number per fruit and germination, there were genetically based differences among maternal populations (Table 1), with Fou having greater overall seed number per fruit and germination than the other two populations (Fig. 1; Table S1). However, for both of these fitness components there were no significant effects of cross type or the interaction between maternal population and cross type (Table 1). On average, estimates of heterosis for seed number per fruit ranged from 2-19%, and estimates of heterosis (or outbreeding depression) for germination ranged from -14% to 17% (Fig. 2; Table S2). For juvenile survival there were significant effects of both cross type and the maternal population by cross type interaction (Table 1). Pairwise contrasts indicated significant heterosis (19%) only in Fou (Figs. 1&2, Tables 1&S2).

Greenhouse experiment

For adult fitness in the greenhouse experiment, there was no significant heterosis or variation in heterosis among populations (Fig. S1; Table S3). On average, estimates ranged from modest heterosis (36%) in Ver, to no effect in Fou, and weak outbreeding depression (-12%) in Tip (Fig. 2; Table S2). For the two fitness components, the only significant effect was that of cross type on fecundity (Fig. S1; Table S3), with about 19% heterosis averaged over all populations (Fig. 2; Table S2). For the combined probability of survival and reproduction on average, values of heterosis/outbreeding depression ranged from -15 to 0% (Fig. 2; Table S2).

Field experiment

For adult fitness in the field, there was no significant overall effect of cross type but there was a significant maternal population by cross type interaction (Fig. 3; Table 2). Within maternal populations, there was significant heterosis for both Fou (159%) and Tip (161%) and non-significant outbreeding depression (-27%) for Ver (Figs. 2&3; Tables 2&S2). There was also a significant main effect of maternal population (Table 2), with Tip having reduced fitness compared to the other two populations (Fig. 3; Table S1). Evidence of heterosis for individual fitness components also varied by maternal population. The combined probability of survival and reproduction was significantly affected by the interaction between maternal population and cross type (Fig. 3; Table 2). There was heterosis for both Fou (88%) and Tip (46%) on average (Fig. 2; Table S2), and this difference was significant for Fou (Fig. 3; Table 2). Significant outbreeding depression (-34%) was observed for Ver (Figs. 2&3; Tables 2&S2). There was also significant differentiation among maternal populations averaged over cross types (Table 2), with a greater chance of survival and reproduction for Ver compared to the other two populations. For fecundity, there was only a significant main effect of cross type (Table 2), with about 41% heterosis averaged across populations (Fig. 2; Table S2). Though the interaction between maternal population and cross type was not significant, there was nevertheless some evidence that the magnitude of heterosis varied among maternal populations (Figs. 2&3; Table 2&S2).

Cumulative fitness

Estimates of cumulative heterosis were substantial in both greenhouse and field experiments, but they were generally stronger in the field experiment depending on the maternal population. For the greenhouse experiment, heterosis was 52% for Fou and 78% for Ver, and there was weak outbreeding depression (-22%) for Tip (Fig. 2; Table S2). For the field common garden

experiment, there was 292% and 132% heterosis for Fou and Tip respectively, and very weak outbreeding depression (-4%) for Ver (Fig. 2; Table S2). Mean bootstrapped values for each fitness component for each population-cross type combination closely corresponded to the empirically estimated means (Fig. S2). Analyses of bootstrapped values of cumulative fitness indicated significant effects of both cross type and the maternal population by cross type interaction for both the greenhouse and field experiments (Fig 4, Table 3). Pairwise comparisons of cross type for each maternal population indicated that all instances of heterosis indicated above were significant, but the two estimates of outbreeding depression were not (Figs. 2&4; Table 3&S2). Averaged over all populations, estimates of heterosis were much greater in the field (140%) than in the greenhouse (36%) experiments (Table S2). Additionally, there was significant differentiation among maternal populations for both the greenhouse and field experiment (Table 3). For the greenhouse experiment this effect seems to be driven primarily by the greater mean values of early fitness components of Fou (Fig. 1), whereas for the field experiment, the lower average adult fitness of Tip is the likely driver (Fig. 3).

DISCUSSION

Heterosis is a useful tool to probe the balance between selection and drift on genetic variation with fitness consequences by indicating the relative degree of fixation of deleterious, partially recessive mutations. Such heterosis may also be an advantageous consequence of regional admixture provenancing in restorations with rare plant species, providing a short-term fitness boost in addition to the longer-term advantages of increased quantitative genetic variation. To our knowledge, our study is the first to quantify heterosis for cumulative fitness in a rare species in a newly established restoration. The consequences of between population crosses were

strongly positive in both Fou (292% heterosis) and Tip (132% heterosis) and neutral in Ver. This suggests a considerable amount of fixed deleterious recessive alleles within these small remnant populations. Regional admixture provenancing (Bucharova *et al.*, 2019) of remnants separated by short distances should therefore be a beneficial strategy for this species in restorations.

Heterosis for cumulative fitness under realistic field conditions is what is most applicable for restoration and conservation, but several aspects of variation in heterosis in our study also bear discussion. Heterosis was generally stronger when measured under field conditions than in the greenhouse. However, the opposite pattern was true for Ver, which had strong heterosis for cumulative fitness in the greenhouse compared to no effect in the field. Comparisons between field- and greenhouse-based estimates of heterosis are rare in the literature (but see Ouborg and Van Treuren, 1994), though some studies have indicated that heterosis may be stronger under more stressful conditions (Maschinski *et al.*, 2013, Hahn and Rieseberg, 2017, Stojanova *et al.*, 2021).

We note that our early fitness components were measured under controlled conditions, so our measures of heterosis for cumulative fitness in the greenhouse and field are not truly independent. This is unlikely to influence our conclusions because most of the differences in heterosis for cumulative fitness between the two experiments were driven by later fitness components. Starting seeds under controlled conditions was a practical constraint given the effort to produce hand pollinations combined with a very low (1-10%) expected germination rate in the field for this species (Menges, 1995). While starting field experiments from juvenile plants is common practice, it is possible that our experimental design bypassed selection (St. Clair *et al.*, 2020) on early life stages in the field. As a result, we might have underestimated heterosis for early fitness components, and therefore cumulative fitness. While heterosis has been reported for

germination in the field (Bowles *et al.*, 2015), another study found no significant effects (Luijten *et al.*, 2002). On the other hand, heterosis for germination has been reported for some species even in controlled environment conditions (Oakley and Winn, 2012), and can be quite large (Busch, 2006, Oakley *et al.*, 2015a).

Despite the early fitness components only being measured in the greenhouse, we observed that significant effects on the relative fitness of interpopulation crosses were more likely to be detected for adult fitness components. This was expected based on similar patterns for inbreeding depression in selfing species, which is presumably due to many smaller effect deleterious alleles (Husband and Schemske, 1996). For fecundity, there was significant overall heterosis in both the field and greenhouse experiments, but the magnitude varied among populations. For the combined probability of survival and reproduction in the field, we found strong and significant heterosis in Fou, but significant modest outbreeding depression in Ver. We note that this was the only significant estimate of outbreeding depression for any fitness component for any population. Furthermore, heterosis for other fitness components like fecundity neutralized this effect for cumulative fitness for Ver in the field.

Despite the tremendous effort to estimate cumulative fitness across multiple fitness components, one caveat is that our experiment was still not long enough to obtain a comprehensive picture of adult fitness in this long-lived perennial. Although there are some notable exceptions (Luijten *et al.*, 2002, Wagenius *et al.*, 2010, Maschinski *et al.*, 2013, Bowles *et al.*, 2015), most studies that quantify heterosis for rare or fragmented species (van Treuren *et al.*, 1993, Ouborg and Van Treuren, 1994, Sheridan and Karowe, 2000, Paland and Schmid, 2003, Busch, 2006, Willi *et al.*, 2007, Oakley and Winn, 2012, Pickup *et al.*, 2012) are of similar duration to the present study. These studies were either conducted under controlled conditions in

the greenhouse, or over 1-2 years in the field. Logistical constraints based on the average length of grant funding cycles or student degree programs likely contribute to this pattern.

Despite the limited duration of the present study, our results suggest that we likely underestimate heterosis for cumulative fitness in the field for both Fou and Tip. The consistent pattern of heterosis for both adult fitness components for these populations indicate that cumulative heterosis would continue to increase with each additional year. The potential long-term results for Ver are less certain, but some evidence suggests that heterosis for this population could increase over time in the field as well. Cumulative survival to the end of the field experiment for Ver was very similar for plants from both within- and between-population crosses (76.7% compared to 72.5%), so most of the outbreeding depression observed for this population was because of plants from between-population crosses that had yet to reproduce. In the greenhouse experiment, plants from between-population crosses in Ver had both greater survival and greater probability of reproduction by the end of the experiment. Therefore, if plants from between-population crosses from Ver continue to survive at similar rates to plants from within-population crosses and are just slower to reach reproductive maturity in the field, then heterosis would increase over time for this population. Heterosis for fecundity, at least on average, was consistent across all populations in both experiments, also supporting this possibility.

Quantitatively comparing our results overall to those of other studies is challenging because of differences in the biology of the study systems, and methodological differences including study duration, environment, and the fitness components measured and how cumulative fitness was calculated. In some studies, heterosis was not explicitly calculated and the means for populations and cross types are not reported. In general, however, it seems that in many studies most similar to ours, positive effects of between-population crosses are the most

common outcome. This echoes the conclusions of Frankham (2015) reviewing a much broader dataset.

Another possible generalization from our specific subset is that the strongest examples of heterosis come from self-compatible populations or species (Busch, 2006, Oakley and Winn, 2012, Weisenberger *et al.*, 2014, Oakley *et al.*, 2015a). In fact, heterosis was by far the most common outcome in self-compatible species (van Treuren *et al.*, 1993, Sheridan and Karowe, 2000, Paland and Schmid, 2003, Maschinski *et al.*, 2013, Spigler *et al.*, 2017). Only one study on a self-compatible species presented evidence for modest outbreeding depression (or heterosis) depending on the population (Ouborg and Van Treuren, 1994). There were fewer studies on self-incompatible species, but the fitness outcome of between-population crosses ranged from positive (Luijten *et al.*, 2002, Willi *et al.*, 2007, Pickup *et al.*, 2012) to neutral (Wagenius *et al.*, 2010) to negative (Bowles *et al.*, 2015). Although a more balanced sampling of mating system types is needed to make general conclusions, it may be that strong selection for outcrossing in self-incompatible species slows the accumulation of deleterious, partially recessive alleles. On the other hand, for very small populations of self-incompatible species where fixation of a single self-incompatibility allele prevents sexual reproduction (DeMauro, 1993, Weekley *et al.*, 2002), crossing between populations seems the only likely option to prevent certain extinction.

In conclusion, the strong heterosis reported here for crosses between these small remnant populations in conditions realistic for a newly established restoration strongly suggest that regional admixture provenancing would be a beneficial strategy. The short-term fitness boost provided by heterosis is in addition to any potential long-term benefits of increased genetic variation necessary to adapt to environmental heterogeneity and climate change. We agree with previous authors that outbreeding depression is unlikely to have serious negative consequences

for restoration and conservation efforts (Frankham, 2015, Ralls *et al.*, 2018, Bell *et al.*, 2019, Hoffmann *et al.*, 2021, Jordan *et al.*, 2023) if appropriate guidelines are followed (Frankham *et al.*, 2011, Weeks *et al.*, 2011). Even if there is some outbreeding depression in recombinant generations in this system, selection should be able to quickly remove deleterious combinations (Luijten *et al.*, 2002, Erickson and Fenster, 2006, Aitkin and Whitlock, 2013, Hamilton and Miller, 2016). Regardless, many open questions about heterosis remain, and we agree with Menges (2008) that restorations should be seen as experimental opportunities.

ACKNOWLEDGEMENTS

We thank P. Chidambaram, C. Gallick, G. Lee, M. Lopp, N. Ryan, T. Soto, and J. Will for assistance with the experiments. We also thank G. Pardillo and Arbor America for permission to conduct the common garden experiment on their property, and K. Jacobsen for agreeing to house the plants for crossing for nearly four months during the COVID shutdown. This work was funded in part by USDA Hatch grant 7000281 to CGO via Purdue College of Agriculture, and a research grant from the Indiana Native Plant Society to IAT.

LITERATURE CITED

Agrawal AF, Whitlock MC. 2011. Inferences about the distribution of dominance drawn from yeast knockout data. *Genetics,* **187**: 553-566.

Aitkin S, Whitlock MC. 2013. Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology Evolution and Systematics,* **44**: 367-388.

Bell DA, Robinson ZL, Funk WC, Fitzpatrick SW, Allendorf FW, Tallmon DA, Whiteley AR. 2019. The exciting potential and remaining uncertainties of genetic rescue. *Trends in Ecology & Evolution,* **34**: 1070-1079.

Bowles ML, McBride JL, Bell TJ. 2015. Long-term processes affecting restoration and viability of the federally threatened Mead's milkweed (*Asclepias meadii*). *Ecosphere,* **6**: 1-21.

Breed MF, Stead MG, Ottewell KM, Gardner MG, Lowe AJ. 2013. Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. *Conservation Genetics,* **14**: 1-10.

Broadhurst LM, Lowe A, Coates DJ, Cunningham SA, McDonald M, Vesk PA, Yates C. 2008. Seed supply for broadscale restoration: maximizing evolutionary potential. *Evolutionary Applications,* **1**: 587-597.

Bucharova A, Bossdorf O, Hölzel N, Kollmann J, Prasse R, Durka W. 2019. Mix and match: regional admixture provenancing strikes a balance among different seed-sourcing strategies for ecological restoration. *Conservation Genetics,* **20**: 7-17.

Busch JW. 2006. Heterosis in an isolated, effectively small, and self-fertilizing population of the flowering plant *Leavenworthia alabamica*. *Evolution,* **60**: 184-191.

Charlesworth B. 2018. Mutational load, inbreeding depression and heterosis in subdivided populations. *Molecular Ecology,* **27**: 4991-5003.

Cheptou P-O, Donohue K. 2011. Environment-dependent inbreeding depression: its ecological and evolutionary significance. *New Phytologist,* **189**: 395-407.

Crow JF. 1948. Alternative hypotheses of hybrid vigor. *Genetics,* **33**: 477-487.

DeMauro MM. 1993. Relationship of breeding system to rarity in the lakeside daisy (*Hymenoxyx acaulis var. glabra*). *Conservation Biology,* **7**: 542-550.

Dolan RW. 1994. Patterns of isozyme variation in relation to population size, isolation, and phytogeographic history in royal catchfly. *American Journal of Botany,* **81**: 965-972.

Drake JM. 2006. Heterosis, the catapult effect and establishment success of a colonizing bird. *Biology Letters,* **2**: 304-307.

Dudash MR. 1990. Relative fitness of selfed and outcrossed progeny in a self-compatible protandrous species, *Sabatia angularis* L. (Gentianaceae): a comparison in three environments. *Evolution,* **44**: 1129-1139.

Erickson DL, Fenster CB. 2006. Intraspecific hybridization and the recovery of fitness in the native legume *Chamaecrista fasciculata*. *Evolution,* **60**: 225-233.

Escobar JS, Nicot A, P. D. 2008. The different sources of variation in inbreeding depression, heterosis and outbreeding depression in a metapopulations of *Physa acuta*. *Genetics,* **180**: 1593-1608.

Fenster CB, Galloway LF. 2000. Population differentiation in an annual legume: genetic architecture. *Evolution,* **54**: 1157-1172.

Fox CW, Reed DH. 2010. Inbreeding depression increases with envionmental stress: and experimental study and meta-analysis. *Evolution,* **65**: 246-258.

Frankham R. 2005. Genetics and extinction. *Biological conservation,* **126**: 131-140.

Frankham R. 2015. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology,* **24**: 2610-2618.

Frankham R. 2016. Genetic rescue benefits persist to at least the F3 generation, based on a meta-analysis. *Biological Conservation,* **195**: 33-36.

Frankham R, Ballou JD, Eldridge MDB, Lacy RC, Ralls K, Dudash MR, Fenster CB. 2011. Predicting the probability of outbreeding depression. *Conservation Biology,* **25**: 465-475.

Gann GD, McDonald T, Walder B, Aronson J, Nelson CR, Jonson J, Hallett JG, Eisenberg C, Guariguata MR, Liu J, Hua F, Echeverría C, Gonzales E, Shaw N, Decleer K, Dixon KW. 2019. International principles and standards for the practice of ecological restoration. Second edition. *Restoration Ecology,* **27**: S1-S46.

Glémin S, Ronfort J, Bataillon T. 2003. Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics,* **165**: 2193-2212.

Gomulkiewicz R, Houle D. 2009. Demographic and genetic constraints on evolution. *The American Naturalist,* **174**: E218-E229.

Hahn MA, Rieseberg LH. 2017. Genetic admixture and heterosis may enhance the invasiveness of common ragweed. *Evolutionary Applications,* **10**: 241-250.

Halligan DL, Keightley PD. 2009. Spontaneous mutation accumulation studies in evolutionary genetics. *Annual Review of Ecology, Evolution, and Systematics,* **40**: 151-172.

Hamilton JA, Miller JM. 2016. Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology,* **30**: 33-41.

Hoffmann AA, Miller AD, Weeks AR. 2021. Genetic mixing for population management: from genetic rescue to provenancing. *Evolutionary Applications*: 634-652.

Husband BC, Schemske DW. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution,* **50**: 54-70.

ISEE Network. 2024. Soil Explorer http://SoilExplorer.net.

JMP. 1989-2021. *JMP Version 16*. Cary, NC: SAS Institute Inc.

Jordan R, Harrison PA, Breed MF. 2023. The eco-evolutionary risks of not changing seed provenancing practices in changing environments. *Ecology Letters,* **24**: e14348.

Keightley PD, Ness RW, Halligan DL, Haddrill PR. 2014. Estimation of spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics,* **196**: 313-320.

Kimura M, Maruyama T, Crow JF. 1963. Mutation load in small populations. *Genetics,* **48**: 1303-1312.

Koski MH, Layman NC, Prior CJ, Busch JW, Galloway LF. 2019. Selfing ability and drift load evolve with range expansion. *Evolution Letters,* **3**: 500-512.

Li Y, Stift M, van Kleunen M. 2018. Admixture increases performance of an invasive plant beyond first-generation heterosis. *Journal of Ecology,* **106**: 1595-1606.

Lohr JN, Haag CR. 2015. Genetic load, inbreeding depression, and hybrid vigor covary with population size: An empricial evaluation of theoretical predictions. *Evolution,* **69**: 3109-3122.

Luijten SH, Kéry M, Oostermeijer JGB, Den Nijs JCM. 2002. Demographic consequences of inbreeding and outbreeding in *Arnica montana*: a field experiment. *Journal of Ecology,* **90**: 593-603.

Lynch M, Conery J, Bürger R. 1995. Mutation accumulation and the extinction of small populations. *American Naturalist,* **146**: 489-518.

Maschinski J, Wright SJ, Koptur S, Pinto-Torres EC. 2013. When is local the best paradigm? Breeding history influences conservation reintroduction surivval and population trajectories in times of extreme climate events. *Biological Conservation,* **159**: 277-284.

McKay JK, Christian CE, Harrison S, Rice KJ. 2005. "How local is local?" - A review of practical and conceptual issues in the genetics of restoration. *Restoration Ecology,* **13**: 432-440.

Menges ES. 1991. Seed germination percentage increases with population size in a fragmented prairie species. *Conservation Biology,* **5**: 158-164.

Menges ES. 1995. Factors limiting fecundity and germination in small populations of *Silene regia* (*Caryophyllaceae*), a rare hummingbird-pollinated prairie forb. *American Midland Naturalist,* **133**: 242-255.

Menges ES. 2008. Restoration demography and genetics of plants: when is a translocation successful? *Australian Journal of Botany,* **56**: 187-196.

NatureServe. 2024. https://explorer.natureserve.org/Taxon/ELEMENT_GLOBAL.2.128777/Silene_regia.

Newman D, Pilson D. 1997. Increased probability of extinction due to decreased genetic effective populaiton size: experimental populaitons of *Clarkia pulchella*. *Evolution,* **51**: 354-362.

Oakley CG, Lundemo S, Ågren J, Schemske DW. 2019. Heterosis is common and inbreeding depression absent in natural populations of *Arabidopsis thaliana*. *Journal of Evolutionary Biology,* **32**: 592-603.

Oakley CG, Spoelhof JP, Schemske DW. 2015a. Increased heterosis in selfing populations of a perennial forb. *AoB Plants,* **7**: plv122.

Oakley CG, Winn AA. 2012. Effects of population size and isolation on heterosis, mean fitness, and inbreeding depression in a perennial plant. *New Phytologist,* **196**: 261-270.

Oakley CG, Ågren J, Schemske DW. 2015b. Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*. *Heredity,* **115**: 73-82.

Ouborg NJ, Van Treuren R. 1994. The significance of genetic erosion in the process of extinction. 4. Inbreeding load and heterosis in relation to population size in the mint *Salvia pratensis*. *Evolution,* **48**: 996-1008.

Paland S, Schmid B. 2003. Population size and the nature of genetic load in *Gentianella germanica*. *Evolution,* **57**: 2242-2251.

Pickup M, Field DL, Rowell DM, Young AG. 2012. Source population characteristics affect heterosis following genetic rescue of fragmented plant populations. *Proceedings of the Royal Society B,* **280**: 20122058.

Ralls K, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, Lacy RC, Sunnucks P, Frankham R. 2018. Call for a paradigm shift in the genetic management of fragemented populations. *Conservation Letters,* **11**: 1-6.

Rutter MT, Roles A, Conner JK, Shaw RG, Shaw FH, Schneeberger K, Ossowski S, Weigel D, Fenster CB. 2012. Fitness of *Arabidopsis thaliana* mutation accumulation lines whose spontaneous mutations are known. *Evolution,* **66**: 2335-2339.

Sheridan PM, Karowe DN. 2000. Inbreeding, outbreeding, and heterosis in the yellow pitcher plant, *Sarracenia flava* (Sarraceniaceae), in Virginia. *American Journal of Botany,* **87**: 1628-1633.

Spigler RB, Theodorou K, Chang SM. 2017. Inbreeding depression and drift load in small populations at demographic disequilibrium. *Evolution,* **71**: 81-94.

St. Clair AB, Dunwiddie PW, Fant JB, Kaye TN, Kramer AT. 2020. Mixed source populations increasses genetic diversity of restored rare plant populations. *Restoration Ecology,* **28**: 583-593.

Stojanova B, Münzbergová Z, Pánková H. 2021. Inbreeding depression and heterosis vary in space and time in the serpentinophyte perennial *Minuartia smejkalii*. *Preslia,* **93**: 149-168.

Söderquist L, Broberg A, Rosenberg V, Sletvold NS. 2020. Predicting heterosis and inbreeding depression from population size and density to inform management efforts. *Journal of Applied Ecology,* **57**: 1459-1468.

Török K, Valkó O, Deák B. 2024. Ecosystem restoration with local or broad seed provenancing: debates and perceptions in science and practice. *Biological Conservation,* **293**: 110535.

van Treuren R, Bijlsma R, Ouborg NJ, van Delden W. 1993. The significance of genetic erosion in the process of extinction. 4. Inbreeding depression and heterosis effects caused by selfing and outcrossing in *Scabiosa-columbaria*. *Evolution,* **47**: 1669-1680.

Wagenius S, Hangelbrook HH, Ridley CE, Shaw RG. 2010. Biparental inbreeding and interremnant mating in a perennial prairire plant: Fitness consequences in their first eight years. *Evolution,* **64**: 761-771.

Weekley CW, Kubisiak TL, Race TM. 2002. Genetic impoverishment and cross-incompatibility in remnant genotypes of *Ziziphus celeta* (Rhamnaceae), a rare shrub endemic to the Lake Wales Ridge, Florida. *Biodiversity and Conservation,* **11**: 2027-2046.

Weeks AR, Sgro CM, Young AG, Frankham R, Mitchell NJ, Miller KA, Byrne M, Coates DJ, Eldridge MD, Sunnucks P, Breed MF, James EA, Hoffmann AA. 2011. Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications,* **4**: 709-725.

Weisenberger LA, Weller SG, Sakai AK. 2014. Remnants of populations provide effective source material for reintroduction of an endangered Hawaiian plant, *Schiedea kaalae* (Caryophyllaceae). *American Journal of Botany,* **101**: 1954-1962.

Weng M-L, Ågren J, Imbert E, Nottebrock H, Rutter MT, Fenster CB. 2021. Fitness effects of mutation in natural populations of *Arabidopsis thaliana* reveal a complex influence of local adaptation. *Evolution,* **75**: 330-348.

Whitlock MC, Ingvarsson PK, Hatfield T. 2000. Local drift load and the heterosis of interconnected populations. *Heredity,* **84**: 452-457.

Willi Y. 2013. Mutational meltdown in selfing *Arabidopsis lyrata*. *Evolution,* **67**: 806-815.

Willi Y, Fischer M. 2005. Genetic rescue in interconnected populations of small and large size of the self-incompatible *Ranunculus reptans*. *Heredity,* **95**: 437-443.

Willi Y, Van Buskirk J, Hoffmann AA. 2006. Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution and Systematics,* **37**: 433-458.

Willi Y, van Kleunen M, Dietrich S, Fischer M. 2007. Genetic rescue persists beyond first-generation outbreeding in small populations of a rare plant. *Proceedings of the Royal Society B-Biological Sciences,* **274**: 2357-2364.

Table 1. ANOVA results for early fitness components. Models for both seed number per fruit and percent germination (of approximately 10 seeds per cell) employed normal error distributions, whereas the model for juvenile survival employed a binomial error distribution. Contrasts between cross types within a maternal population are only reported for models with a significant cross type or maternal population by cross type interaction (WIN = within-population cross, BET = between-population cross).

	Seed number per fruit		Germination		Juvenile Survival	
	F	*P*	*F*	*P*	χ^2	*P*
Maternal population	15.2	**<0.001**	6.6	**0.002**	1.1	0.574
Cross type	1.4	0.235	0.1	0.747	5.8	**0.016**
Mat. Pop * type	0.2	0.808	2.2	0.118	15.2	**0.001**
Block	-	-	4.6	**0.001**	65.4	**<0.001**
Contrast	*F*	*P*	*F*	*P*	χ^2	*P*
Fou: BET vs. WIN	-	-	-	-	30.6	**<0.001**
Tip: BET vs. WIN	-	-	-	-	0	-
Ver: BET vs. WIN	-	-	-	-	0	-

Table 2. ANOVA results for the field experiment. The model for the combined probability of survival and reproduction employed a binomial error distribution, whereas the model for fecundity (fruits per reproducing plant) employed a normal error distribution. The model for fruits per seedling planted (including zeros) was analyzed with a zero-inflated Poisson regression.

	Survival and reproduction		Fecundity		Fruits per seedling planted	
	χ^2	*P*	*F*	*P*	χ^2	*P*
Maternal population	19.2	**<0.001**	1.2	0.313	20.7	**<0.001**
Cross type	1.3	0.252	6.0	**0.015**	2.4	0.118
Mat. Pop * type	20.9	**<0.001**	1.0	0.358	20.1	**<0.001**
Block	23.2	**<0.001**	1.0	0.418	47.5	**<0.001**
Contrast	χ^2	*P*	*t*	*P*	*t*	*P*
Fou: BET vs. WIN	15.6	**<0.001**	1.9	**0.065**	5.4	**<0.001**
Tip: BET vs. WIN	2.3	0.126	1.9	**0.058**	6.0	**<0.001**
Ver: BET vs. WIN	7.4	**0.007**	0.4	0.715	1.6	0.119

Table 3. Nonparametric ANOVA results for bootstrapped values of cumulative fitness separately for the greenhouse and field experiments.

	Greenhouse expt. cumulative fitness		Field expt. cumulative fitness	
	F	*P*	*F*	*P*
Maternal population	22.7	**<0.001**	48.3	**<0.001**
Cross type	38.6	**<0.001**	48.9	**<0.001**
Mat. Pop * type	8.4	**<0.001**	22.3	**<0.001**
Contrast	*t*	*P*	*t*	*P*
Fou: BET vs. WIN	5.3	**<0.001**	8.8	**<0.001**
Tip: BET vs. WIN	0.2	0.244	3.9	**<0.001**
Ver: BET vs. WIN	5.2	**<0.001**	-0.7	0.492

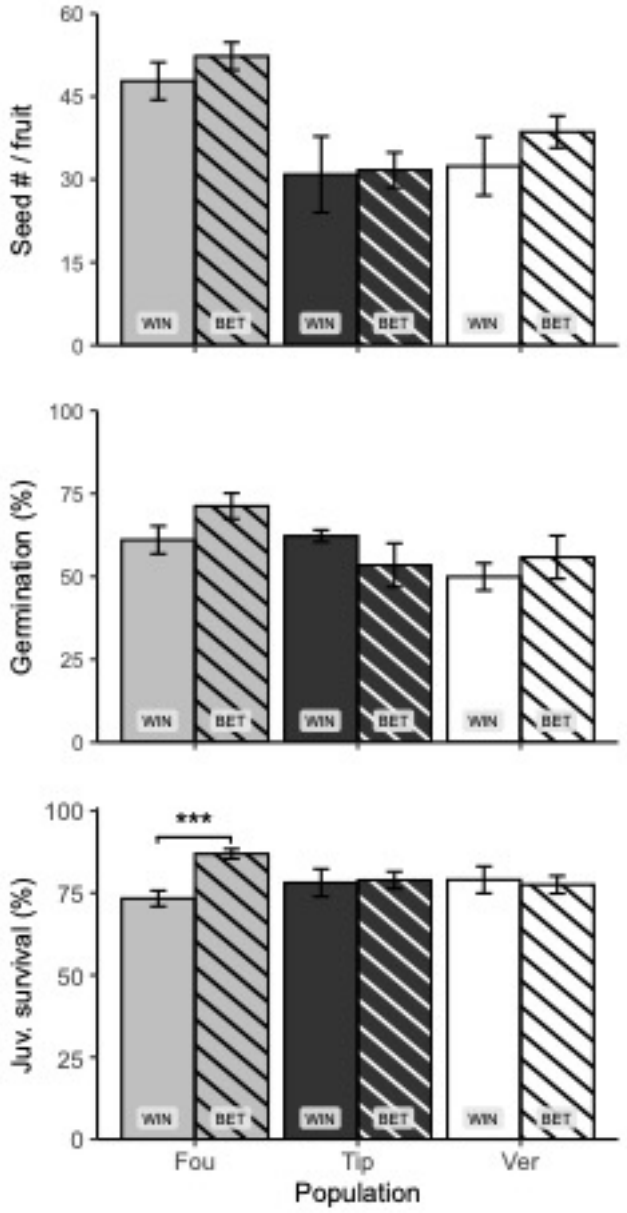

Figure 1. Mean fitness components for each maternal population (Fou, Tip, and Ver) and cross type (WIN = within-population cross, BET = between-population cross) for seed number per fruit of the hand pollinations, germination, and juvenile survival. Means and standard errors were calculated as the mean of block means (see Table S1), except for seed number per fruit which are the raw means and SE. Within a population, a significant contrast of BET vs. WIN is indicated by a bar and asterisks. *** $P < 0.001$ (see Table 1).

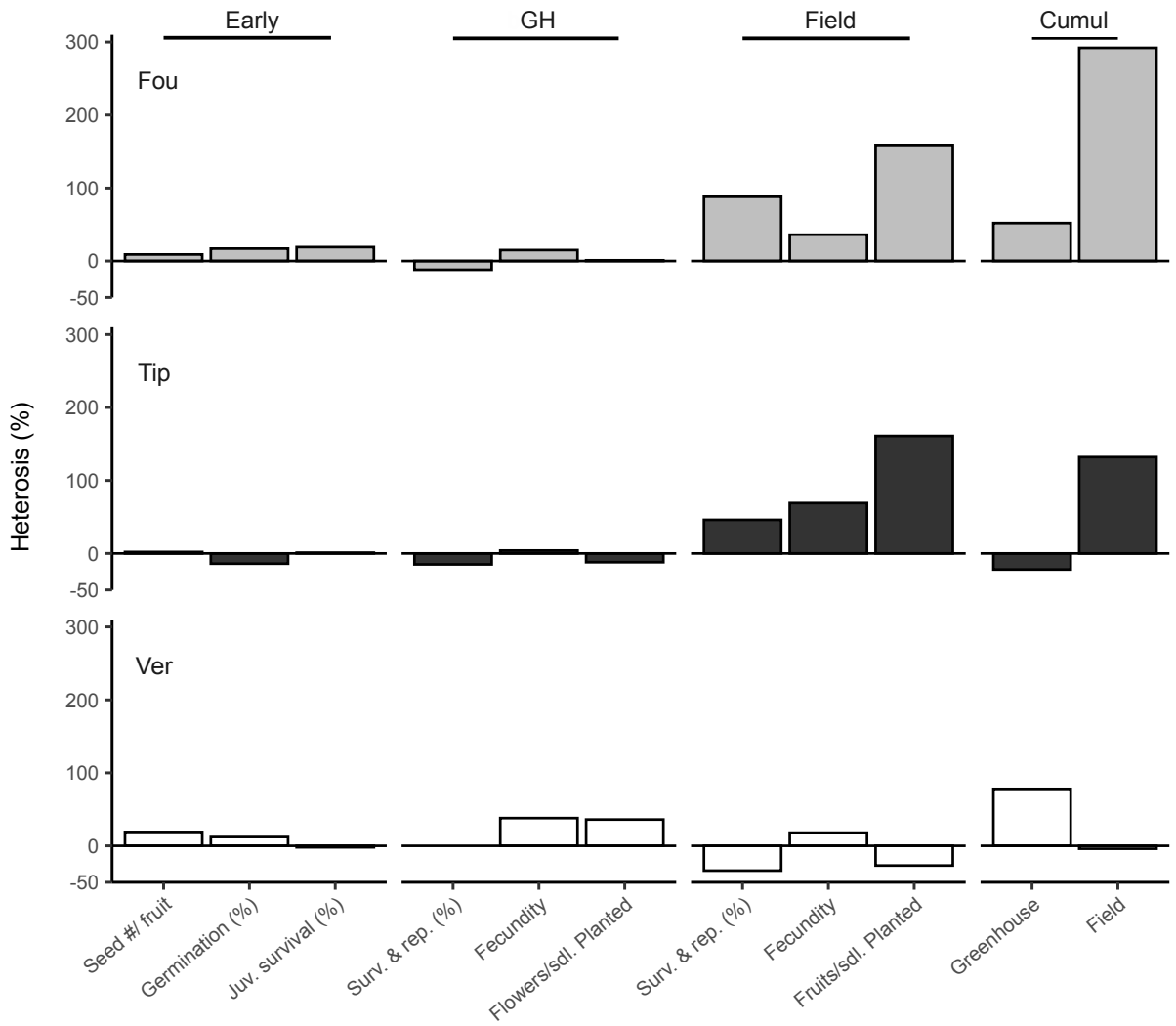

Figure 2. Estimates of heterosis (positive values) or outbreeding depression (negative values) for each of the three populations and for each fitness component and composite estimate of fitness (see also Table S3). For both the greenhouse and field experiment, estimates of adult fitness per seedling planted incorporated probability of survival and reproduction, and fecundity of reproducing plants. Cumulative fitness is the product of all individual fitness components (see Table S1). Note that cumulative fitness for the greenhouse and field common garden experiment share the estimates for early fitness components in common.

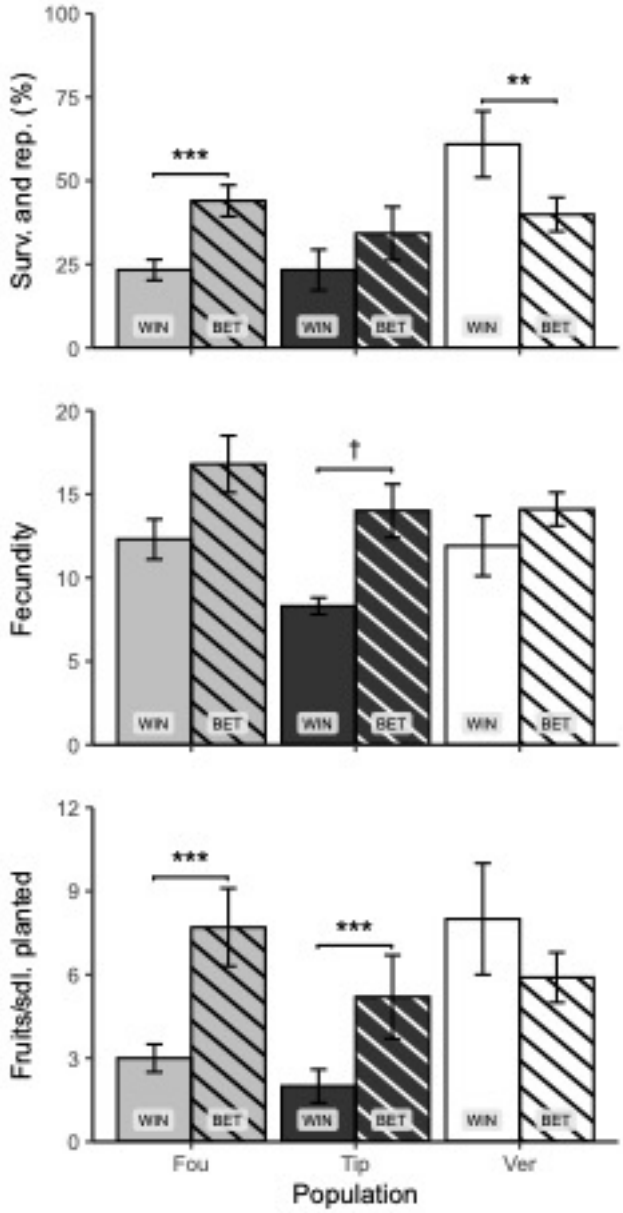

Figure 3. Mean fitness components for each maternal population (Fou, Tip, and Ver) and cross type (WIN = within-population cross, BET = between-population cross) for the field common garden experiment. Means and standard errors were calculated as the mean of block means (see Table S1). Within a population, a significant contrast of BET vs. WIN is indicated by a bar and asterisks. † $0.10 < P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (see Table 3).

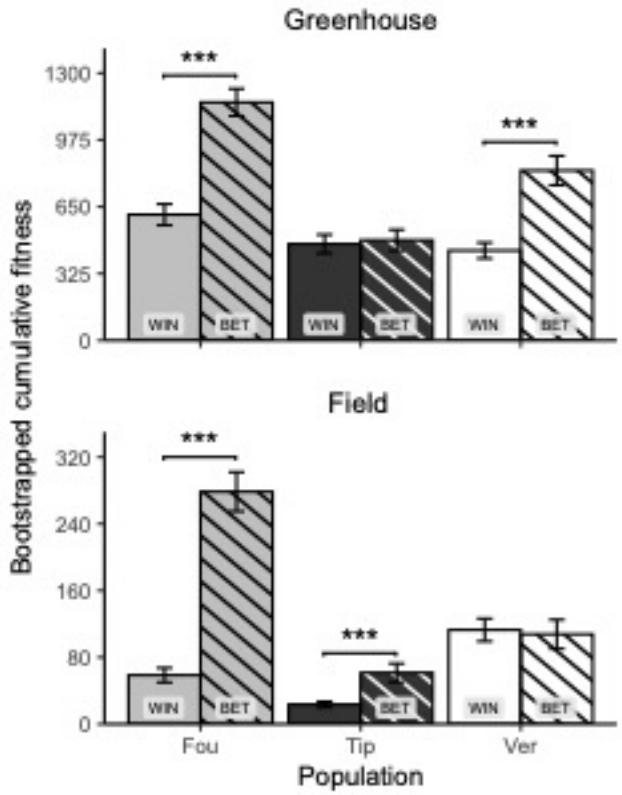

Figure 4. Mean bootstrapped cumulative fitness for each maternal population (Fou, Tip, and Ver) and cross type (WIN = within-population cross, BET = between-population cross) for the greenhouse and field common garden experiments. Note the difference in scale of the y axis between the two experiments. Error bars are 1 SE. Within a population, a significant contrast of BET vs. WIN is indicated by a bar and asterisks. *** $P < 0.001$ (see Table 3).

SUPPLEMENTAL MATERIAL

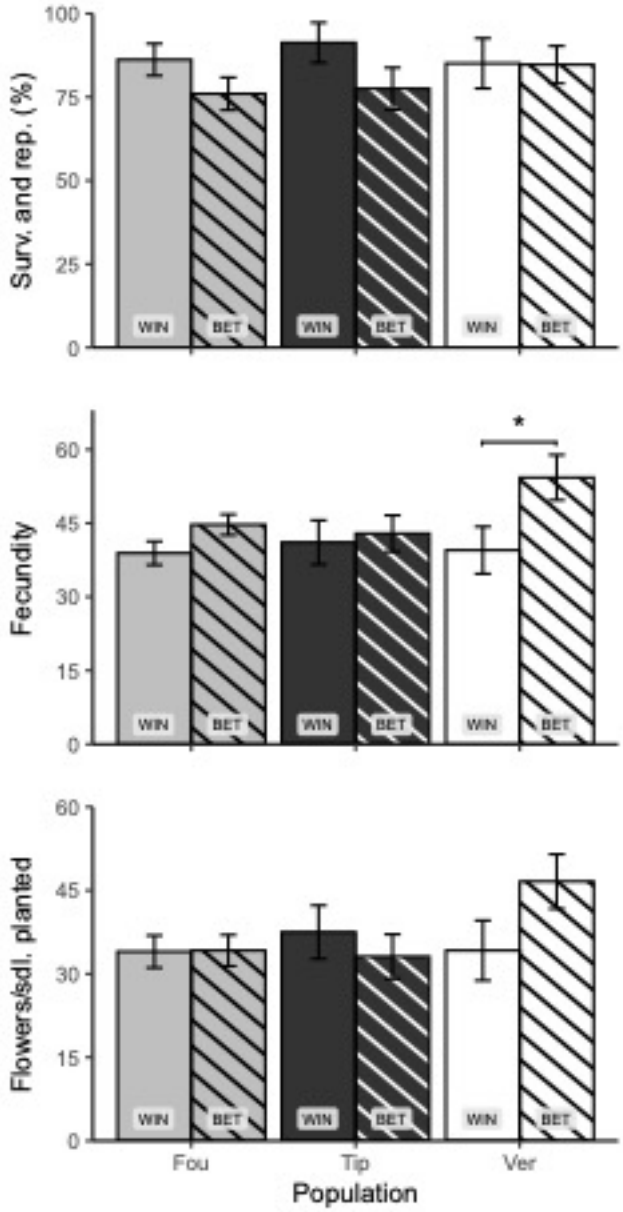

Figure S1. Mean fitness components for each maternal population (Fou, Tip, and Ver) and cross type (WIN = within-population cross, BET = between-population cross) for the greenhouse experiment. Means and standard errors were calculated as the mean of block means (see Table S1). Except for fecundity, there were no significant cross type or maternal population by cross type interactions for fitness components, so no within-population contrasts of BET vs. WIN were performed (see Table S3). Within a population a significant contrast of BET vs. WIN for fecundity is indicated by a bar and asterisk. * $P < 0.05$ (see Table S3).

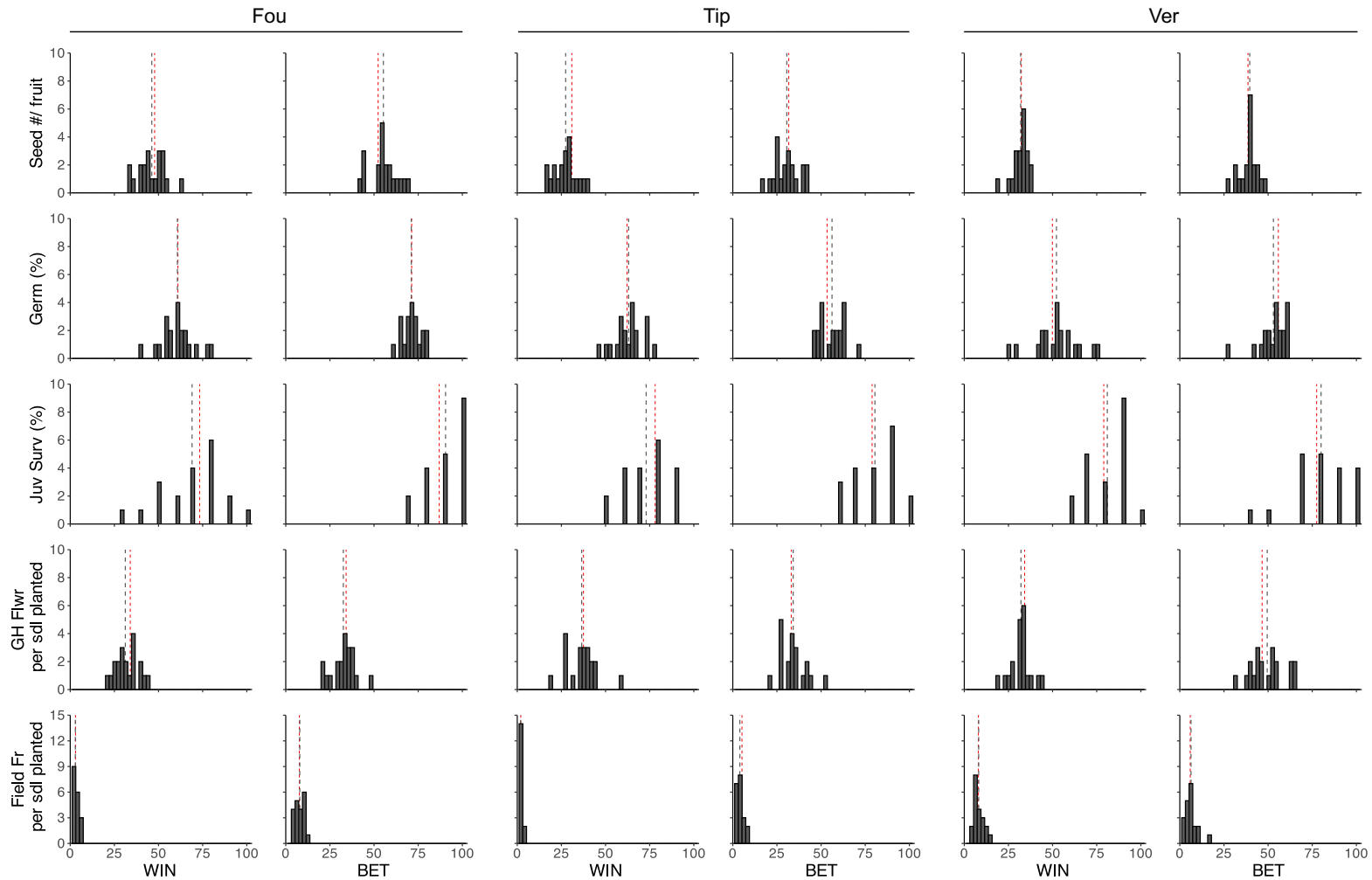

Figure S2. Histograms of 20 bootstrap resampled estimates of fitness components for each population (Fou, Tip, and Ver) and cross type (WIN = outcrossed within populations, BET = outcrossed between populations). The dashed black line indicates the mean of the bootstrapped estimates. The dashed red line indicates the empirical mean. Seed#/fruit (seed number per fruit of the hand pollinations), Germ (%) (germination), Juv Surv (%) (Juvenile survival), GH Flwr per sdl planted (Greenhouse – number of flowers per seedling planted) & Field Fr per sdl planted (Field – number of fruits per seedling planted).

Table S1. Sample sizes and mean values (averaged over block where appropriate) for fitness components for each maternal population and cross type. For seed number per fruit, sample size indicates the number of fruits. For germination, mean is the percent germination per cell (usually starting with 10 seeds) averaged over five flats, and sample size represents the total number of cells over all flats. For juvenile survival, the mean percent survival over the means of the 30 flats is given, and sample size represents the total number of seedlings over all flats. For the greenhouse experiment, sample sizes are the total number of plants at the start of the experiment over all 48 flats, and means are the mean over flat means. For the field experiment, sample sizes are the total number of plants at the start of the experiment over all 6 blocks, and means are the mean over block means. For fecundity in both experiments, sample sizes are the number of plants that reproduced.

	Fou				Tip				Ver			
	BET		WIN		BET		WIN		BET		WIN	
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
Early												
Seed # / fruit	71	52.3	39	47.8	45	31.7	17	30.9	47	38.6	19	32.4
Germination (%)	120	71.2	60	61.0	60	53.4	30	62.2	59	55.8	30	49.9
Juv. survival (%)	723	86.9	337	73.3	307	78.9	170	78.1	291	77.5	149	79.0
Greenhouse												
Surv. and rep. (%)	118	76.0	59	86.2	59	77.5	30	91.3	59	84.7	29	85.1
Fecundity	112	44.7	55	38.9	53	42.9	28	41.1	55	54.3	26	39.5
Flowers/sdl. planted	118	34.2	59	34.0	59	33.1	30	37.5	59	46.6	29	34.2
Field												
Surv. and rep. (%)	239	44.0	120	23.3	120	34.2	60	23.3	118	39.9	59	60.9
Fecundity	105	16.8	28	12.3	41	14.0	14	8.3	47	14.1	36	11.9
Fruits/sdl. planted	239	7.7	120	3.0	120	5.2	60	2.0	118	5.9	59	8.0

Table S2. Estimates of heterosis (outbreeding depression for negative values) for each maternal population for fitness components and composite fitness estimates.

	Fou	Tip	Ver
Early			
Seed # / fruit	9	2	19
Germination (%)	17	-14	12
Juv. survival (%)	19	1	-2
Greenhouse			
Surv. and rep. (%)	-12	-15	0
Fecundity	15	4	38
Flowers/sdl. planted	1	-12	36
Field			
Surv. and rep. (%)	88	46	-34
Fecundity	36	69	18
Fruits/sdl. planted	159	161	-27
Cumulative			
Greenhouse	52	-22	78
Field	292	132	-4

Table S3. ANOVA results for the greenhouse experiment. The model for the combined probability of survival and reproduction employed a binomial error distribution, whereas the models for fecundity (flowers per reproducing plant) and flowers per seedling planted (including zeros) employed normal error distributions. Contrasts between cross types within a maternal population are only reported for models with a significant cross type or maternal population by cross type interaction (WIN = within-population cross, BET = between-population cross).

	Survival and reproduction		Fecundity		Flowers per seedling planted	
	χ^2	*P*	*F*	*P*	*F*	*P*
Maternal population	1.6	0.455	0.5	0.595	0.9	0.415
Cross type	2.2	0.138	6.2	**0.013**	1.0	0.312
Mat. Pop * type	2.9	0.235	1.2	0.301	2.2	0.112
Block	45.5	0.534	1.3	0.141	1.2	0.202
Contrast	χ^2	*P*	*t*	*P*	*t*	*P*
Fou: BET vs. WIN	-	-	1.2	0.246	-	-
Tip: BET vs. WIN	-	-	0.5	0.621	-	-
Ver: BET vs. WIN	-	-	2.5	**0.014**	-	-