Heredity (2015) 115, 73–82

www.nature.com/hdy

ORIGINAL ARTICLE

Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*

CG Oakley[1], J Ågren[2] and DW Schemske[3]

Understanding the causes and architecture of genetic differentiation between natural populations is of central importance in evolutionary biology. Crosses between natural populations can result in heterosis if recessive or nearly recessive deleterious mutations have become fixed within populations because of genetic drift. Divergence between populations can also result in outbreeding depression because of genetic incompatibilities. The net fitness consequences of between-population crosses will be a balance between heterosis and outbreeding depression. We estimated the magnitude of heterosis and outbreeding depression in the highly selfing model plant *Arabidopsis thaliana*, by crossing replicate line pairs from two sets of natural populations (C ↔ R, B ↔ S) separated by similar geographic distances (Italy ↔ Sweden). We examined the contribution of different modes of gene action to overall differences in estimates of lifetime fitness and fitness components using joint scaling tests with parental, reciprocal F_1 and F_2, and backcross lines. One of these population pairs (C ↔ R) was previously demonstrated to be locally adapted, but locally maladaptive quantitative trait loci were also found, suggesting a role for genetic drift in shaping adaptive variation. We found markedly different genetic architectures for fitness and fitness components in the two sets of populations. In one (C ↔ R), there were consistently positive effects of dominance, indicating the masking of recessive or nearly recessive deleterious mutations that had become fixed by genetic drift. The other set (B ↔ S) exhibited outbreeding depression because of negative dominance effects. Additional studies are needed to explore the molecular genetic basis of heterosis and outbreeding depression, and how their magnitudes vary across environments.
Heredity (2015) **115,** 73–82; doi:10.1038/hdy.2015.18; published online 18 March 2015

INTRODUCTION

Understanding the genetic architecture of differentiation between natural populations is a major goal of evolutionary biology. The genetic basis of adaptive differentiation has been a central focus of evolutionary genetics since the modern synthesis (reviewed in Orr and Coyne, 1992). In finite populations, the number, mode of action and effect size of new mutations available for adaptive evolution will be influenced by the magnitude of genetic drift (reviewed in Frankham and Weber, 2000). In addition to reducing adaptive genetic variation, drift will also increase the chance fixation of deleterious alleles (Kimura *et al.*, 1963). One challenge to studying the effects of drift on genetic variation related to fitness is that it is often difficult to demonstrate drift as a causal mechanism. Indeed, few studies have demonstrated a role of drift in the evolution of conspicuous phenotypes (but see Husband and Barrett, 1992).

Perhaps the best available indicator of a historical role of genetic drift in shaping genetic variation underlying fitness is heterosis (Crow, 1948; Lynch, 1991; Whitlock *et al.*, 2000), the increased fitness of F_1 progeny from crosses between populations relative to crosses within populations. Alternatively for selfing taxa or inbred lines, heterosis can be defined as the increased fitness of F_1 progeny relative to the mean of the parental lines. Such an increase in fitness is attributed to the masking of deleterious recessive or nearly recessive alleles in the heterozygous state. Historical fixation of these alleles within populations or lines can only be attributed to genetic drift.

Many if not most new mutations are deleterious (Halligan and Keightley, 2009; but see Rutter *et al.*, 2012), and small effective population sizes should reduce the efficacy of selection in preventing the fixation of deleterious mutations (Kimura *et al.*, 1963). Many mildly deleterious recessive or nearly recessive mutations are predicted to become fixed when effective population sizes are modest and gene flow is limited (Whitlock *et al.*, 2000). Heterosis in F_1 crosses between natural populations is ubiquitous (for example, Fenster, 1991; Armbruster *et al.*, 1997; Edmands, 1999; Oakley and Winn, 2012).

The relative fitness of F_1 progeny will, however, be a balance between beneficial effects of dominance and other deleterious genetic effects. Many different intrinsic factors may affect the relative fitness of hybrids produced from crosses between widely diverged populations or lines. Genetic incompatibilities such as Dobzhansky–Muller incompatibilities are thought to be a common cause of outbreeding depression in hybrids between species or widely diverged populations (Lynch, 1991; Orr and Turelli, 2001). Such incompatibilities arise because new mutations are 'tested' against the genetic background of other loci present in the local population, and alleles without negative effects may become fixed. Introducing these alleles into a different genetic background by crossing can cause negative epistatic interactions. The frequency of epistatic incompatibilities is predicted to

[1]Department of Plant Biology, Michigan State University, East Lansing, MI, USA; [2]Department of Plant Ecology and Evolution, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden and [3]Department of Plant Biology and W K Kellogg Biological Station, Michigan State University, East Lansing, MI, USA
Correspondence: Dr CG Oakley, Department of Plant Biology, Michigan State University, 612 Wilson Road, East Lansing, MI 48824-1312, USA.
E-mail: coakley@msu.edu
Received 8 September 2014; revised 9 February 2015; accepted 18 February 2015; published online 18 March 2015

increase with increasing genetic distance between the parents (Lynch, 1991; Orr and Turelli, 2001; Matute *et al.*, 2010). Alternatively, decreased fitness in crosses between widely diverged populations could be due to chromosomal rearrangements (Lande, 1985; Charlesworth, 1992; Kirkpatrick and Barton, 2006), or to underdominance and/or negative epistatic interactions between closely linked loci (Schierup and Christiansen, 1996). These latter mechanisms could be particularly common in organisms that are capable of close inbreeding, such as self-fertilizing plants (Charlesworth, 1992; Schierup and Christiansen, 1996; Gimond *et al.*, 2013).

General insight into the genetic architecture (that is, net effects of different genetic modes of action) contributing to differentiation can be achieved using line cross techniques (Mather and Jinks, 1982; Lynch and Walsh, 1998). These approaches have been championed for their utility in exploring the genetic architecture underlying fitness differences in crosses between natural populations (Fenster *et al.*, 1997; Demuth and Wade, 2005, 2006, 2007; Demuth *et al.*, 2014). The line cross approach utilizes a set of crosses (minimally F_1 and F_2) between two parental lines to decompose the differences in a trait of interest into additive and non-additive genetic effects. Additional cross types can be performed, including backcrosses and reciprocal crosses, to partition different types of epistatic effects, and maternal and/or cytoplasmic effects, respectively.

Joint scaling analysis of line crosses is a powerful method for detecting genetic interactions (Mather and Jinks, 1982; Demuth and Wade, 2005, 2006). The line cross approach is therefore particularly well suited for investigating the modes of gene action underlying the balance between increased fitness (that is, heterosis) and decreased fitness (that is, outbreeding depression) of progeny derived from crosses between vs within populations. This approach could also be used to inform conservation genetic strategies, where predicting the success of genetic rescue efforts is of critical importance (Frankham *et al.*, 2011). Line cross techniques have successfully been applied to understanding the genetic architecture of fitness in crosses between natural populations in a variety of taxa, including crustaceans (Edmands, 1999), insects (Armbruster *et al.*, 1997; Demuth and Wade, 2007) and plants (Fenster and Galloway, 2000). Most of these examples uncovered both beneficial dominance effects and negative epistatic effects, at least in some environments.

It is commonly implied that hybrid incompatibilities arise because different substitutions are fixed by natural selection in different populations, but chance fixation because of genetic drift (*cf.*, Bomblies *et al.*, 2007), and a selfing mating system could also have a role. Selfing could hasten initial fixation because of increased homozygosity within populations, and decreased gene flow between populations (Gimond *et al.*, 2013). It has also been suggested that selfing will increase the likelihood of outbreeding depression because it suppresses recombination (Fenster *et al.*, 1997). To our knowledge, line cross analyses of crosses between natural populations of a selfing species have not been reported.

The model plant species *Arabidopsis thaliana* (hereafter *Arabidopsis*) offers many advantages for studying the genetics of broad-scale population differentiation. *Arabidopsis* is a small, selfing annual with a broad native range that encompasses much of Europe and parts of Asia (Koornneef *et al.*, 2004). Differentiation for neutral genetic markers exhibits a pattern of isolation by distance over broad scales (Beck *et al.*, 2008), and reduced neutral genetic variation in northern populations is consistent with genetic drift because of population bottlenecks during range expansion following Pleistocene glaciations (Beck *et al.*, 2008). A pattern of reduced genetic variation in northern compared with southern populations has also been reported within Scandinavia (Lewandowska-Sabat *et al.*, 2010; Long *et al.*, 2013). Both the selfing mating system of *Arabidopsis* and the demographic history of populations from northern Scandinavia suggest a potential role for genetic drift in shaping genetic variation underlying fitness.

Previous investigations of the genetic basis (mode of gene action) and molecular genetic basis of heterosis in *Arabidopsis* (see, for example, Kusterer *et al.*, 2007a, b; Meyer *et al.*, 2010, 2012) have typically focused on a few laboratory accessions and/or traits not directly related to fitness in natural populations. Kusterer *et al.* (2007b) investigated the genetic basis of heterosis for biomass-related traits and found strong positive effects of dominance. Despite the expected positive correlation between biomass and fitness, another study using the same laboratory accessions found heterosis for biomass, but outbreeding depression for seed production (Barth *et al.*, 2003). This work thus contributes to our knowledge of the genetic basis of heterosis underlying agriculturally important traits, but cannot (and was not intended to) provide insight into the role of drift in shaping patterns of deleterious mutations within and among natural populations, or into the genetic basis of outbreeding depression between natural populations.

The genetic basis of outbreeding depression has also been investigated using *Arabidopsis*. Bomblies *et al.* (2007) generated >800 different F_1 progenies from 280 unique accessions and found a 2% incidence of hybrid necrosis in the rosette stage. Subsequent quantitative trait loci (QTL) mapping using *Col* as one of the parental lines indicated that the genetic basis of necrosis was due to negative epistatic interactions involving 2–4 loci (Bomblies *et al.*, 2007). Similar interactions, mapped to approximately the same regions, were found in a study of the genetic basis of necrosis in two different recombinant inbred line mapping populations with *Ler* as a parental line (Alcázar *et al.*, 2009).

The net effect on fitness of dominance (increasing fitness) and epistasis/underdominance (decreasing fitness) in crosses between natural populations of *Arabidopsis* is unknown. A recent QTL analysis of fitness in two natural populations of *Arabidopsis* from Italy and Sweden identified several cases where the local genotype was maladaptive (1 out of 13, and 4 out of 12 QTL in Italy and Sweden, respectively), consistent with the fixation of deleterious alleles because of drift (Ågren *et al.*, 2013).

Quantifying the effects of dominance on fitness in between-population crosses can provide insight into the magnitude of genetic drift because such deleterious mutations would not be driven to fixation by selection. The potential role of drift in outbreeding depression is less clear, but using multiple population pairs separated by similar geographic distances (and having likely undergone similar adaptive divergence to climatic factors) may provide additional insight. If selection is responsible for fixation of alleles resulting in outbreeding depression, we may expect crosses between these replicate population pairs to exhibit similar levels of outbreeding depression. If, for example, one population pair exhibits outbreeding depression, but another population pair exhibits heterosis, it suggests that some other mechanism may contribute to the fixation of alleles responsible for genetic incompatibilities.

Here we use crosses between two sets (pairs) of natural populations of *Arabidopsis* from Italy and Sweden to address the following questions: (1) What is the extent of heterosis and/or outbreeding depression in crosses between widely diverged populations? (2) Do populations separated by similar geographic distance display similar levels of heterosis and/or outbreeding depression? (3) Is heterosis and/or outbreeding depression consistent across different components of fitness?

MATERIALS AND METHODS

Study system

We selected four natural populations from close to the northern and southern (two each) margins of the native range as seed sources for this experiment. The southern populations were from central Italy (Castelnuovo di Porto—hereafter 'C'—42.1°N, 12.5°E; and Bolsena—hereafter 'B'—42.7°N, 12.0°E) and the northern populations were from northern Sweden (Rödåsen—hereafter 'R'—62.8°N, 18.2°E; and Skuleberget—hereafter 'S'—63.1°N, 18.4°E). All four populations occupy relatively intact habitats typical of the natural environments preferred by *Arabidopsis*, and all have a winter annual life history; seeds germinate in fall, plants overwinter as rosettes and flower in spring. Previous work with populations C (Italy) and R (Sweden) over multiple years has shown strong local adaptation between this population pair (Ågren and Schemske, 2012; Ågren *et al.*, 2013). These populations were also previously used for QTL mapping of fitness in the field (Ågren *et al.*, 2013). Swedish populations R and S are separated by about 30 km, Italian populations B and C are separated by about 70 km, and the Italian and Swedish populations are both separated by over 2400 km. These population sets are likely to replicate genetic distance and differences in adaptation to climatic factors.

Growing conditions and crossing design

We collected seeds from four lines per population and grew plants in the glasshouse for 1–2 generations to minimize environmental maternal effects. Seeds produced from these plants were surface sterilized, sown on nutrient agar in Petri dishes and cold stratified at 4 °C in the dark for 5 days to break dormancy. Dishes were then moved into a growth chamber (constant 22 °C, 16-h days). Twelve-day-old seedlings were transplanted into potting soil in 5.1 cm square pots, and plants were grown for 3 weeks at the same conditions. Plants were then vernalized for 8 weeks (4 °C, 10-h days) to promote synchronous flowering (Grillo *et al.*, 2013; CG Oakley, unpublished data). After vernalization, plants were moved to a growth chamber under the conditions described above until they flowered.

To characterize the genetic architecture underlying fitness and fitness components, we performed line crosses to generate plants for the joint scaling analyses. Two line pairs were chosen from each of two sets of populations. Having additional population sets and line pairs would be desirable, but because each line pair entails growing eight different cross types, and because large sample sizes are needed to have adequate power to detect epistasis (Demuth and Wade, 2006), this was logistically infeasible.

A first round of controlled hand pollinations was performed reciprocally between two line pairs from C ↔ R (C1 ↔ R1 and C2 ↔ R2) and between two line pairs from B ↔ S (B1 ↔ S1 and B2 ↔ S2) to produce the initial F_1 seed (Supplementary Figure S1). Autogamously selfed seed was also collected from parental lines (Supplementary Figure S1). For controlled pollinations, we first emasculated recipient flowers before anthesis to prevent accidental self-pollination. Emasculated controls that were not pollinated ($n = 40$) never produced successful fruits. In addition, previous paternity analysis using microsatellite markers on over 360 individuals from controlled crosses determined that our rate of accidental self-fertilization is $<1\%$ (CG Oakley, unpublished data).

A second round of crossing (Supplementary Figure S1) was conducted to generate all seeds for the fitness assay and joint scaling analyses (Mather and Jinks, 1982; Lynch and Walsh, 1998; Demuth and Wade, 2005) at the same time, in the same maternal environment. Seeds of parental lines and F_1 were germinated and raised to flowering as above. For each line pair, we generated parental (P1 and P2) and F_2 and rF_2 ($F_1 \times F_1$ and $rF_1 \times rF_1$, where F_1 is derived from P1 as the maternal plant and rF_1 is derived from P2 as the maternal plant) seed from emasculated controlled self-pollinations of the parental and the initial F_1 and rF_1 lines, respectively (Supplementary Figure S1). We also generated F_1 and rF_1 seed by controlled crosses between the parental lines (Supplementary Figure S1). We produced both directions of backcross (BC_1) seed from crosses with the parental lines as the maternal parents and pooled reciprocal F_1 as the paternal lines (Supplementary Table S1). Pooled reciprocal F_1 as the paternal parent for these crosses has no influence on the estimation of genetic effects we are interested in here because using either F1 or rF1 has the same expected genetic effects (Demuth and Wade, 2007).

The types of genetic effects that can be investigated in joint scaling analyses depend on the number and exact types of crosses used for a particular design, but in all cases what is estimated is the net effect(s) across the genome for particular genetic effects (for example, additive, dominance and so on). As such, it is not possible to distinguish true overdominance/underdominance from the effects of multiple closely linked loci.

Estimating fitness components

Seeds from all cross types for all four line pairs were germinated simultaneously following the protocol outlined above. For each of the line pairs, we sowed 40 seeds of each of the parental, F_1 and rF_1 lines. For both BC_1 and F_2, and rF_2 lines, we sowed 160 seeds each. We sowed proportionally more seeds for recombinant generations because we anticipated a need for increased power in estimating their fitness components (for example, Erickson and Fenster, 2006). At the time of transplant, the proportion of seeds germinating was recorded for all line pair × cross type combinations.

Seedlings were transplanted into 2.4 cm square, 4.5 cm deep cells of 200 cell flat inserts (P-200, Landmark Plastics, Akron, OH, USA) filled with a 1:1:1 Perlite:Vermiculite:Sure-Mix blend (MI Growers Products, Galesburg, MI, USA). Seedlings were planted in a checkerboard manner, such that plants were bordered on four sides by empty cells. In total, 1600 seedlings were transplanted in fully randomized manner across 16 flats. For each of the line pairs, we transplanted 20 seedlings of each of the parental, F_1 and rF_1 lines. For both BC_1 and F_2, and rF_2 lines, we transplanted 80 seedlings for each of the four line pairs. Seedlings that died within 5 days were replaced. We grew the plants in a growth chamber for 2 weeks and vernalized them for 8 weeks following the standard protocols outlined above. In mid-January 2013, plants were moved to a heated glasshouse with a maximum temperature of 26.7 °C during the day and a minimum of 15.6 °C at night with 16 h day^{-1} supplemental lighting. Plants were sub-irrigated as needed, and fertilized every other watering with ½ strength Hoagland's solution. The glasshouse environment is likely more uniform and less stressful than a natural environment, so the non-additive effects on fitness we measured under these conditions should be due to intrinsic genetic mechanisms expressed in a benign environment.

We measured several fitness components over the course of the experiment. Over all line pairs and cross types, mean germination was 97.45% (range of 92.29–100%), and mean survival was 97.42% (range of 90–100%). As germination and survival were uniformly high, we omit joint scaling results for these components. We counted the number of fruits produced by each plant and the number of seeds in one fruit per plant. A multiplicative estimate of fitness was obtained as the product of these two components. In total, 163 502 fruits and 66 438 seeds were counted for the 1558 surviving plants.

Statistical analyses

Multiplicative fitness and its two components were analyzed with analysis of variance (ANOVA) testing for the effects of population set (C ↔ R or B ↔ S), cross type (P1, P2, F_1, rF_1, F_2, rF_2, BC_1P1, and BC_1P2; Supplementary Table S1) and their interaction, as well as the effect of line pair (1 or 2) nested within-population set and the interaction of this term with cross type. All terms were treated as fixed effects because our main goal was to examine the genetic architecture of particular line pairs and the sample size needed for a given joint scaling analysis made large random samples of line pairs per population set infeasible.

For each line pair, we decomposed the total phenotypic variance for multiplicative fitness and fitness components into additive and non-additive components by joint scaling analyses (Mather and Jinks, 1982; Lynch and Walsh, 1998; Demuth and Wade, 2005). As we detected significant variation in the effect of cross type between line pairs within-population set (Table 1), we performed separate joint scaling analyses for each of the four line pairs. Our line crosses generated eight different cross types, each with different expected contributions of additive, dominant, epistatic, maternal genetic and cytoplasmic effects (Supplementary Table S1; see also Demuth *et al.*, 2014). Maternal additive effects (additive genetic effect of maternal genotype on progeny fitness, independent of progeny genotype) are expected to yield similar consequences as cytoplasmic effects, except that they would not be expressed in the F_2 (maternal parent was an F_1). Conversely, maternal dominance effects (effect of dominance

Table 1 Overall analysis of variance for fitness (=total number of seeds) and fitness components for line crosses between Italian and Swedish *Arabidopsis thaliana*

ANOVA effects	*Df*	*Fitness*	*Number of fruits*	*Seed number per fruit*
Population set	1	16.85***	28.53***	172.74***
Line pair (Pop. set)	2	0.40	3.33*	4.56*
Cross type	7	33.59***	60.24***	3.79***
Cross type * Pop. set	7	8.07***	6.62***	5.79***
Cross type * Line pair (Pop. set)	14	4.73***	4.21***	5.85***

Abbreviation: ANOVA, analysis of variance.
Table entries are F values, all denominator Df=1520. *$P<0.05$, ***$P<0.001$.

in maternal genotype on progeny fitness, independent of progeny genotype) are only expected in the F_2 generation (Supplementary Table S1).

Our methods closely follow Demuth *et al.* (2014), but we examined a wider range of possible models to be sure that the order of introduction of terms in a stepwise procedure did not influence the overall results. For each of the line pairs, we fitted least-square regressions of the composite genetic effects contributing to fitness differences among the eight cross types (Supplementary Tables S2–S4). All models contained two parameters, a mean and an additive effect. To determine the possible contribution of dominance, we constructed additional models adding this term. All additional (that is, more complex) models contained terms for mean, additive and dominance effects. To determine potential combinatorial effects of different types of epistatic, maternal genetic and cytoplasmic effects, we constructed models with each of these terms singly and in all possible two- and three-way combinations (Supplementary Tables S2–S4). Many more models are possible, but because more complex models were not necessary to explain the data, we used models with no >6 parameters (Supplementary Tables S2–S4). We assessed overall model fit by the goodness of fit statistic (X^2), hereafter GOF, for each model, with seven degrees of freedom (=number of cross types−number of parameters in the model).

The best model was identified as the simplest (fewest parameters) that sufficiently (nonsignificant GOF) explained the data. If more than one model of equal simplicity was sufficient, we compared the GOF statistics. In these cases, we considered a model to have a better fit if its GOF statistic was 3.84 less than another model. This is similar to a likelihood ratio test with 1 degree of freedom, except we compared two models with the same number of parameters, so this is a very conservative (only a model that was unambiguously a better fit is chosen) test. For each of the best models, significance of individual parameters was assessed using one-sample *t*-tests with 7 degrees of freedom (=number of cross types−1). In a few cases, we were unable to unequivocally choose the best model based on simplicity, sufficiency, and relative fit criteria, so we chose among the equivalent models based on which had the highest number of individually significant parameters (Supplementary Table S5; *cf.*, Demuth *et al.*, 2014).

For seed number per fruit, one of the four line pairs had two sufficient models of equivalent simplicity with equal numbers of individually significant parameters. In this case, we chose the 'best' model as the one with the slightly lower GOF statistic (Supplementary Table S4). Both models share four of five terms suggesting that despite slight differences in the two models, there is good confidence in the importance of the four common terms. All calculations for the joint scaling analyses were performed using custom Python scripts, an example of which has been submitted with the data to Dryad.

RESULTS

We present results first for our estimate of multiplicative fitness, followed by fitness components. For each dependent variable, we first discuss the results of the overall ANOVA and general patterns of the means, followed by presentation of the results of the joint scaling analyses.

Fitness

The ANOVA for fitness indicated that there were significant differences in mean fitness between-population sets and among cross types, but not among line pairs within a population set (Table 1, Figure 1). Significant interactions with cross type indicate that the effect of cross type differed not only between-population sets, but also between line pairs from a given population set (Table 1, Figure 1).

For C↔R and B↔S, there were large differences in mean fitness between the two parental lines (Figure 1), with the Italian parental lines (C1, C2, B1 and B2) producing between about 2000 and 2800 (~60–100%) more seeds per plant on average than the Swedish parental lines (R1, R2, S1 and S2). Mean fitness of all cross types (except the parental lines) of C1↔R1 and C2↔R2 were greater than the additive expectation (Figures 1a and b), whereas all cross types of B1↔S1 and B2↔S2 had lower fitness compared with the additive expectation (Figures 1c and d). The significant interaction between cross type and line pair nested within-population set can be largely attributed to differences between B1↔S1 and B2↔S2 (Figures 1c and d). Notably, in B2↔S2, the F_1 but not the rF_1 had drastically reduced fitness (Figure 1d). All individuals of this F_1 exhibited a stunted phenotype observed only in some crosses from B↔S (Supplementary Table S6), which could have contributed to the reduced fitness of this cross type.

Joint scaling analyses revealed that the mode of gene action underlying fitness differs considerably between the two population pairs (Table 2). In both C1↔R1 and C2↔R2, mean fitness could be explained by a combination of additive and beneficial dominance effects (Table 2). For these pairs, there were positive effects of dominance on fitness, that is, heterosis (dominance coefficient=+11 and +23% of the means, respectively, for C1↔R1, and C2↔R2).

The mode of gene action underlying fitness in B1↔S1 and B2↔S2 was more complex than in C1↔R1 and C2↔R2, and there were also differences between B1↔S1 and B2↔S2. Like the previous population set, there were large additive differences between the mean fitness of the parental lines, with the Italian parental lines producing about 2500–2700 (~60–80%) more seeds per plant than the Swedish parental lines (Table 2, Figure 1). Of the two line pairs from B↔S, B2↔S2 had the simpler model in terms of the number of genetic effects detected, containing one additional term over the additive dominance model, a maternal additive effect (Table 2). The maternal additive effect estimate of −1061.4 indicates that P2, or rF_1 and BC_1 with P2 as the maternal parent produce on average 2122 more seeds per plant than do P1, and F_1 and BC_1 lines with P1 as the maternal parent, (Table 2, Figure 1d), all else being equal. Five parameters were required to adequately explain the pattern observed in B1↔S1. In addition to additive, dominance and maternal additive effects, both dominance × dominance (DD) epistasis and cytoplasmic effects were observed (Table 2). The direction of the maternal additive effect was reversed from that in B2↔S2. One striking difference between C↔R and B↔S is that in both B1↔S1 and B2↔S2, dominance had negative effects on fitness. In B1↔S1, this negative effect of dominance on fitness was in part recovered by DD epistasis (dominance coefficient=−43% of the mean for B2↔S2, and D+DD=−8% of the mean for B1↔S1).

Number of fruits

The analyses of number of fruits produced qualitatively similar results to those for multiplicative fitness (Table 1, Figure 2), except that the main effect of line pair nested within population was significant for number of fruits. There were large additive differences between the mean fitness of the parental lines; with the Italian parental lines

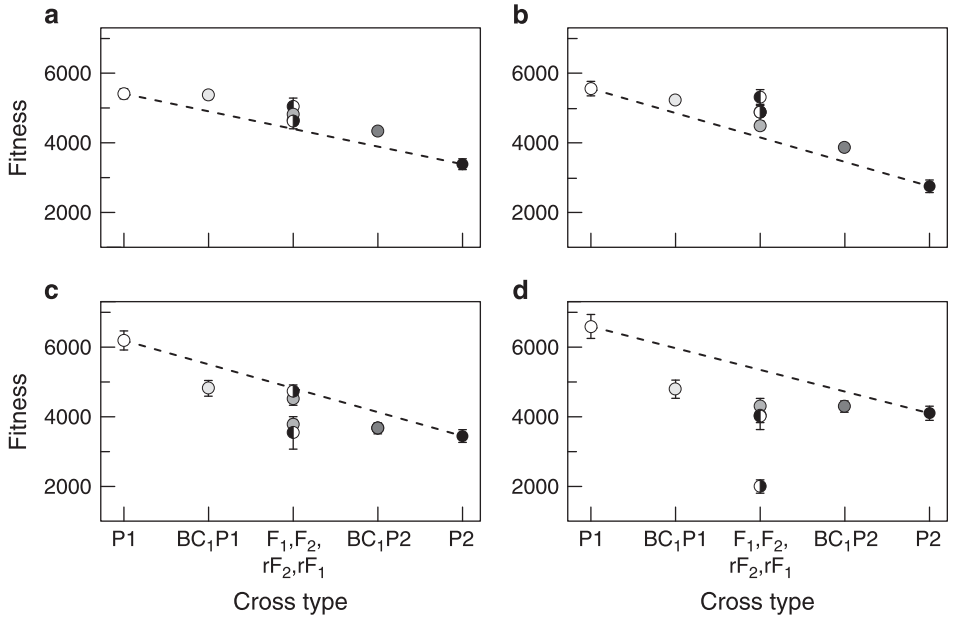

Figure 1 Line cross means for fitness (=total number of seeds) for each of four line pairs (**a**, C1↔R1; **b**, C2↔R2; **c**, B1↔S1; **d**, B2↔S2) of *Arabidopsis thaliana*. C↔R=Castelnuevo, Italy and Rödåsen, Sweden; B↔S=Bolsena, Italy and Skuleberget, Sweden. In all cases, the P1 parent is from Italy and the P2 parent is from Sweden. Four cross types (F_1, rF1, F_2, rF_2) have an additive expectation halfway between the two parents; in some cases the points are obscured (see Supplementary Table S7 for values). F_1 and F_2 have P1 as the maternal parent, rF_1 and rF2 have P2 as the maternal parent. Half-filled circles represent means of the different F_1's, the shading of the left half of the circle corresponds to the maternal parent. Error bars are 1 s.e., and in some cases are smaller than the size of the symbol. The dashed line is the expectation for purely additive gene action.

Table 2 Joint scaling results for fitness (=total number of seeds) for two line pairs from each population set of *Arabidopsis thaliana* (C↔R=Castelnuovo, Italy and Rödåsen, Sweden; B↔S=Bolsena, Italy and Skuleberget, Sweden)

Line pair	*Param.*	*Est.*	*t*	P-*value*	*Full model goodness of fit*
C1↔R1					$X^2=6.83$, $P=0.234$, df=5
	Mean	4489.8	47.95	<0.0001	
	A	1006.0	11.46	<0.0001	
	D	501.8	2.77	0.028	
C2↔R2					$X^2=9.19$, $P=0.102$, df=5
	Mean	4137.1	36.63	<0.0001	
	A	1375.4	14.03	<0.0001	
	D	934.5	4.39	0.003	
B1↔S1					$X^2=4.72$, $P=0.094$, df=2
	Mean	4764.7	29.32	<0.0001	
	A	973.2	3.39	0.012	
	D	−1774.3	−2.69	0.031	
	DD	1375.4	1.99	0.087	
	Ma	646.3	2.80	0.027	
	Cyt	−390.5	−3.34	0.012	
B2↔S2					$X^2=4.70$, $P=0.320$, df=4
	Mean	5461.5	33.61	<0.0001	
	A	2372.1	9.10	<0.0001	
	D	−2347.0	−8.78	<0.0001	
	Ma	−1061.4	−6.59	<0.0001	

Individual parameters are as follows: additive (A), dominance (D), Epistatic (AA, AD and/or DD), cytoplasmic (Cyt), maternal additive (Ma) and maternal dominance (Md).

producing 56–125% more fruits per plant than the Swedish parental lines over all line pairs (Table 3, Figure 2). As with fitness, all crosses from C1↔R1 and C2↔R2 had greater mean number of fruits than the additive expectation, whereas most cross types from B1↔S1 and B2↔S2 produced fewer fruits than the additive expectation (Figure 2).

Results for joint scaling analyses of number of fruits were also similar to those for multiplicative fitness in that C1↔R1 and C2↔R2 were similar to each other, B1↔S1 and B2↔S2 were different from C1↔R1 and C2↔R2, and B1↔S1 and B2↔S2 were also different from each other (Tables 3 and 4, Figure 2). Line pairs C1↔R1 and C2↔R2 both had relatively simple genetic architectures, with beneficial effects of dominance (+16% and +13% increase over the mean for C1↔R1 and C2↔R2, respectively). For line pairs B1↔S1 and B2↔S2, there were negative effects of dominance (−8 and −32% for B1↔S1 and B2↔S2, respectively). Qualitative differences between the results for number of fruits and the results for fitness include a weak cytoplasmic effect for C2↔R2, the absence of any DD epistasis, a maternal dominance effect for B1↔S1 and a strongly negative additive×additive epistatic effect for B2↔S2 (Tables 3 and 4).

Seed number per fruit

As with number of fruits, all ANOVA effects were significant for seed number per fruit (Table 1). Unlike fitness and number of fruits, there were only slight differences between the parental lines in all four of the line pairs (Figure 3). All cross types for C1↔R1 and C2↔R2 were close to the additive expectation, but many cross types from B1↔S1 and B2↔S2 (especially B2↔S2) had substantially reduced mean seed number per fruit compared with the additive expectation (Figure 3).

For the joint scaling analyses, only very weak additive, dominance and cytoplasmic effects were detected in C1↔R1 and C2↔R2 (Table 5, Figure 3). Very different models were found for the two line pairs from B↔S. For B1↔S1, results were similar to those for fitness except for the absence of the cytoplasmic effect, and the addition of a strongly negative additive × additive epistatic term for seed number per fruit (Tables 4 and 5). For B2↔S2 by contrast, no epistasis was

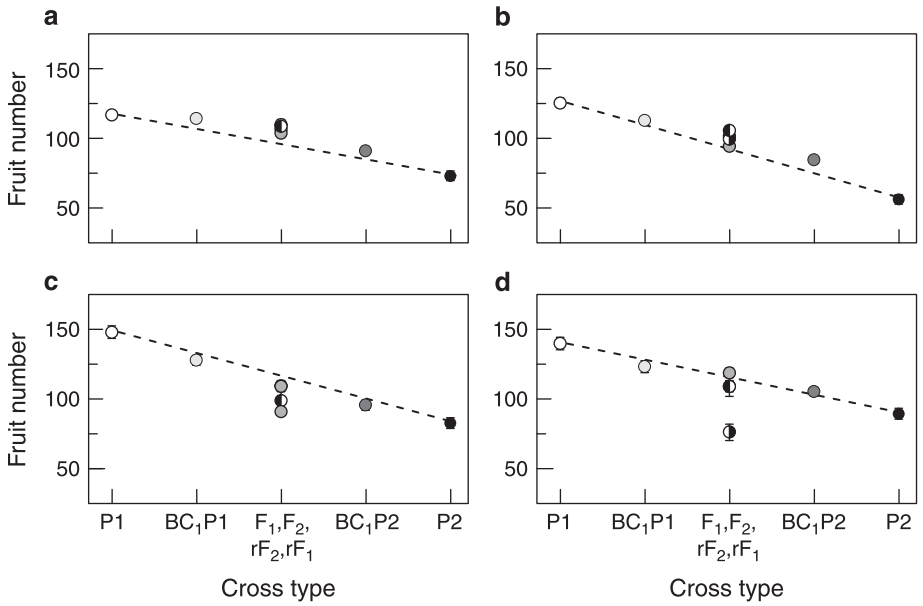

Figure 2 Line cross means for number of fruits for each of the four line pairs. (**a**, C1↔R1; **b**, C2↔R2; **c**, B1↔S1; **d**, B2↔S2). Letters, symbols and error bars as in Figure 1. The dashed line is the expectation with purely additive gene action.

Table 3 Joint scaling results for number of fruits, abbreviations as in Table 2

Line pair	*Param.*	*Est.*	*t*	P-*value*	*Full model goodness of fit*
C1↔R1					$X^2=3.68$, $P=0.596$, df=5
	Mean	95.7	53.85	<0.0001	
	A	22.3	13.77	<0.0001	
	D	15.0	4.41	0.003	
C2↔R2					$X^2=1.49$, $P=0.828$, df=4
	Mean	92.1	48.85	<0.0001	
	A	37.4	15.49	<0.0001	
	D	11.6	3.38	0.012	
	Cyt	−3.9	−3.41	0.011	
B1↔S1					$X^2=0.08$, $P=0.961$, df=2
	Mean	115.7	43.22	<0.0001	
	A	28.9	5.98	<0.0001	
	D	−9.8	−2.03	0.082	
	Ma	12.0	3.00	0.020	
	Md	−10.8	−3.74	0.007	
	Cyt	−9.0	−4.00	0.005	
B2↔S2					$X^2=6.67$, $P=0.083$, df=3
	Mean	138.8	20.52	<0.0001	
	A	40.5	7.87	<0.0001	
	D	−44.7	−4.38	0.003	
	AA	−23.1	−3.12	0.017	
	Ma	−14.0	−3.97	0.005	

detected, and results were similar to those for fitness with the addition of a maternal dominance term for seed number per fruit (Tables 4 and 5). As with fitness and number of fruits, the effects of dominance in B1↔S1 and B2↔S2 were negative (dominance coefficient = −34% of the mean for B2↔S2, and D+DD = −28% of the mean for B1↔S1).

DISCUSSION

The extent to which genetic drift has a role in shaping genetic variation related to fitness is difficult to address empirically, but estimates of heterosis in crosses between natural populations provide one such avenue of investigation. As the relative fitness of crosses between widely diverged populations is expected to be a balance between the beneficial effects of dominance complementation and the negative effects of genetic incompatibilities, joint scaling analyses of line cross data are well suited to estimate the relative magnitudes of these different genetic effects.

We found markedly different genetic architectures (that is, modes of gene action) for fitness and fitness components for crosses between two sets of populations separated by similar geographic distances. In one population set (C↔R), there were consistently positive effects of dominance, consistent with the masking of recessive or nearly recessive mildly deleterious mutations fixed by genetic drift. The other population set (B↔S) exhibited outbreeding depression due largely to underdominance (or pseudo-underdominance).

Genetic architecture of fitness and components in C↔R

The genetic architecture of fitness in both line pairs from C↔R was relatively simple. The difference between the parental lines had a largely additive genetic basis, although some dominance complementation (that is, heterosis) was also observed (Table 2). The estimates of heterosis from these crosses (about 10–20%) are modest, and may be considered trivial in an agricultural context, but do suggest that genetic drift can have an influence on the distribution of genetic variation influencing fitness within and among natural populations of this species.

Results for fitness components for C↔R suggest that differences in fitness were largely a result of differences in fruit production, and not differences in seed production per fruit (Table 4). If heterosis was due to many mildly deleterious alleles, it should be expressed relatively evenly across different fitness components, because such alleles would be unlikely to accumulate in one region of the genome. Observing heterosis for number of fruits, but not seed number per fruit, could therefore indicate the fixation of one or a few strongly deleterious mutations underlying this fitness component. Alternatively, there may simply be a greater number of loci underlying fruit number (through effects on growth and branching, for example) than underlying seed

number per fruit. Correlations between number of fruits and seed number per fruit for all individual plants, done separately by line pair, are weak, ranging from −0.14 in C1↔R1 to 0.25 in B2↔S2, suggesting that they have mostly independent genetic bases. Overdominance of one or few loci is an alternative hypothesis for the genetic basis of heterosis (Crow, 1948), and although overdominance is generally thought to be uncommon (for example, Charlesworth and Willis, 2009), we acknowledge that we cannot with the present data rule it out as a mechanism for the heterosis observed here.

Table 4 Qualitative comparison of genetic architecture between fitness, and fitness components for the four different line pairs

Trait and line pair	*A*	*D*	*AA*	*AD*	*DD*	*Ma*	*Md*	*Cyt*
Fitness								
C1↔R1	++	++						
C2↔R2	++	++						
B1↔S1	++	−−			++	++		−
B2↔S2	++	−−				−−		
Number of fruits								
C1↔R1	++	++						
C2↔R2	++	++						−
B1↔S1	++	−				++	−	−
B2↔S2	++	−−	−−			−−		
Seed number per fruit								
C1↔R1	$+^{ns}$	$-^{ns}$						$-^{ns}$
C2↔R2	$-^{ns}$	$+^{ns}$						
B1↔S1	$-^{ns}$	−−	−−		$++^{ns}$	+		
B2↔S2	+	−−				−−	−	

Abbreviations as in Table 2. Sign indicates the sign of the parameter estimate; doubled signs indicate that the parameter estimate was large relative to the mean (>10%). Nonsignificant parameter estimates are indicated with ^{ns}.

Previous studies have reported heterosis in *Arabidopsis*, but in all cases, the crosses involved laboratory accessions. In crosses between laboratory accessions (*Col-0* and *C24*), Barth *et al.* (2003) reported considerably stronger heterosis (60–69%) for biomass and rosette diameter than what we found here for fitness. Similarly, in a cross between the same pair of lines, Kusterer *et al.* (2007b) found 49% heterosis for biomass yield. Direct comparison of these results with ours is difficult because heterosis between laboratory accessions may reflect fixation of alleles during cultivation rather than the historical effects of genetic drift in natural populations. Moreover, despite strong heterosis for biomass traits in the *Col-0* and *C24* cross, F_1 individuals in one of these same studies (Barth *et al.*, 2003) actually produced 19% fewer seeds than the mean of the parents (that is, 19% outbreeding depression for fitness).

Heterosis is commonly observed in crosses between natural populations separated by wide genetic and/or geographic distances. In wide crosses between populations of the pitcher plant mosquito *Wyeomyia smithii* grown at low density, heterosis for fitness was of similar magnitude to what we report here, although in a high-density treatment much stronger heterosis was observed (Armbruster *et al.*, 1997). In wide crosses between natural populations of the annual plant *Chamaecrista fasciculata* grown at different field sites in multiple years (Fenster and Galloway, 2000), heterosis for fitness was of equal or greater magnitude to what we report here, although statistical significance varied among planting sites and between years. Common to these examples is that reduced fitness (that is, outbreeding depression) was often observed in recombinant generations despite increased fitness in the F_1. In our crosses from C1↔R1 and C2↔R2, no such outbreeding depression was observed (Figure 1). This suggests a role for genetic drift in fixing deleterious mutations, but despite wide divergence, no genetic incompatibilities have arisen.

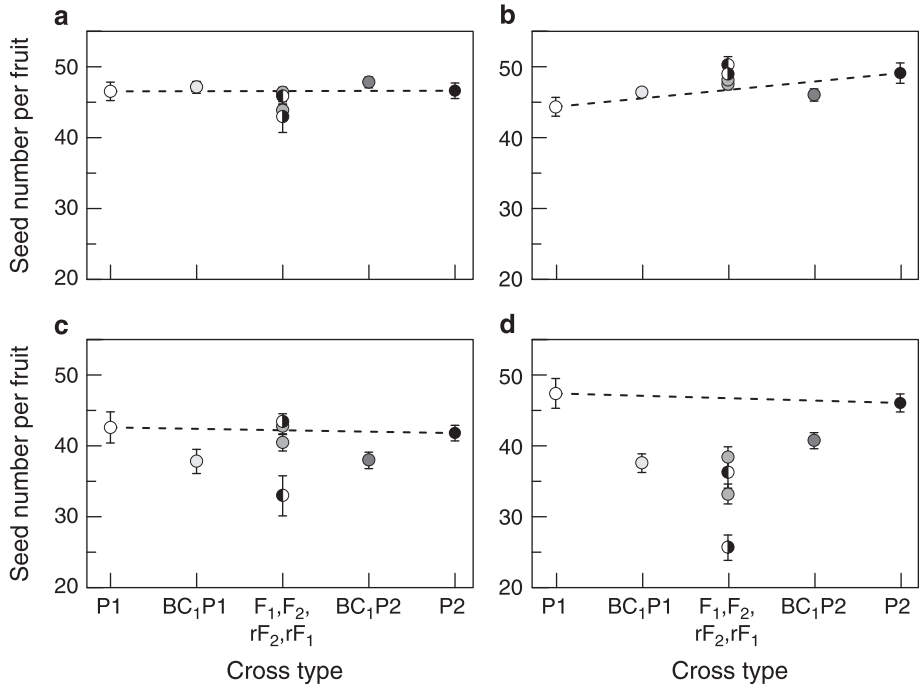

Figure 3 Line cross means for seed number per fruit for each of the four line pairs. (**a**, C1↔R1; **b**, C2↔R2; **c**, B1↔S1; **d**, B2↔S2). Letters, symbols and error bars as in Figure 1. The dashed line is the expectation with purely additive gene action.

Table 5 Joint scaling results for seed number per fruit, abbreviations as in Table 2

Line pair	*Param.*	*Est.*	*t*	P-*value*	*Full model goodness of fit*
C1 ↔ R1					$X^2 = 9.36$, $P = 0.053$, df = 4
	Mean	46.99	66.27	<0.0001	
	A	1.13	1.22	0.262	
	D	−1.82	−1.37	0.213	
	Cyt	−1.07	−2.29	0.056	
C2 ↔ R2					$X^2 = 9.71$, $P = 0.084$, df = 5
	Mean	45.79	59.92	<0.0001	
	A	−1.36	−1.89	0.101	
	D	3.08	2.27	0.058	
B1 ↔ S1					$X^2 = 5.93$, $P = 0.052$, df = 2
	Mean	54.46	10.26	<0.0001	
	A	−4.22	−2.18	0.066	
	D	−36.11	−2.54	0.039	
	AA	−12.90	−2.46	0.044	
	DD	20.99	2.19	0.065	
	Ma	3.55	2.76	0.028	
B2 ↔ S2					$X^2 = 6.86$, $P = 0.076$, df = 3
	Mean	46.99	44.53	<0.0001	
	A	5.94	3.29	0.013	
	D	−16.01	−8.88	<0.0001	
	Ma	−4.89	−4.26	0.004	
	Md	−3.30	−2.75	0.029	

Genetic architecture of fitness and components in B ↔ S

The genetic architecture of fitness in B1 ↔ S1 and B2 ↔ S2 was relatively complex, and there were considerable differences between the two line pairs from the same population set. In contrast to the line pairs C1 ↔ R1 and C2 ↔ R2, both B1 ↔ S1 and B2 ↔ S2 displayed large negative effects of dominance on fitness. For B2 ↔ S2, fitness was reduced by an average of 44% in the F_1 and rF_1. We calculate outbreeding depression as the relative fitness of the F_1 compared with the midparent for the sake of comparison with C1 ↔ R1 and C2 ↔ R2, but note that the presence of maternal additive genetic effects, cytoplasmic effects in both B1 ↔ S1 and B2 ↔ S2, and dominance-by-dominance epistasis in B1 ↔ S1, complicate such a comparison. For B1 ↔ S1, fitness of the F_1 and rF_1 were reduced by an average of about 15% relative to the midparent. The opposite sign of the dominance and the dominance-by-dominance epistatic term in this line pair indicates that underdominant/pseudo-underdominant effects were somewhat reduced by interacting underdominant/pseudo-underdominant loci.

Results for number of fruits and seed number per fruit suggest that both of these components contribute to the overall pattern for fitness, but in somewhat different ways (Figures 1–3). Joint scaling results indicate some differences in the contribution of different types of epistasis and maternal dominance effects among fitness components and fitness, but the effect of dominance is consistently negative in all cases. Thus, the negative effects on fitness result from underdominant/pseudo-underdominant effects on both number of fruits and seed number per fruit. Weak correlations between number of fruits and seed number per fruit suggest that these underdominant/pseudo-underdominant effects have different causes. We are not aware of any examples of generalized negative effects of dominance across many loci (that is, the opposite of heterosis in the strictest sense), so we interpret these results as underdominance/pseudo-underdominance. As with any line cross design, however, we cannot distinguish between true underdominance and the effects of multiple, tightly linked loci.

Although several other studies of intraspecific crosses have reported outbreeding depression (Armbruster *et al.*, 1997; Edmands, 1999; Fenster and Galloway, 2000; Demuth and Wade, 2007), it is usually attributable to epistasis and often manifest in the F_2 or other recombinant generations following some heterotic effects in the F_1. Strong outbreeding depression in the F_1 of crosses between different populations of selfing species of *Caenorhaditis* has recently been reported (Gimond *et al.*, 2013), but without a line cross design such as used here, it is not possible to distinguish between underdominance/pseudo-underdominance and negative epistatic interactions between unlinked loci.

In *Arabidopsis*, hybrid necrosis in the F_1 generation has been reported in about 2% of crosses between different accessions (Bomblies *et al.*, 2007). Genetic mapping of stunted growth and necrotic phenotypes for a small subset of these crosses using laboratory strains such as *Col* and *Ler* as one of the parents implicated epistasis between a few loci involved in pathogen defense (Bomblies *et al.*, 2007; Alcázar *et al.*, 2009). These stunted and necrotic phenotypes are likely to have negative effects on fitness, such as we observed here, but relative fitness (fruit and seed production) of the F_1 has not been quantified for most of the crosses reported in Bomblies *et al.* (2007). Using a cross (*Bla* x *Sha*) previously identified as producing a necrotic phenotype, Smith *et al.* (2011) generated transgenic lines containing one or both of the *Bla* and *Sha* alleles for a single protein kinase (*OAK*) gene in a *Col* background. They found that transgenic lines carrying one or the other parental allele had similar fitness to *Col*, but transgenic lines that were heterozygous for this gene had an 80% reduction in fitness compared with *Col*. Fitness estimates for the parental lines and their F_1 were not presented, but these results suggest that strongly underdominant effects on fitness could underlie F_1 outbreeding depression in this cross.

We are not aware of other studies of intraspecific crosses between natural populations separated by similar geographic distances that display genetic architectures as different as what we found. Although it is fairly common for the results of wide crosses to be somewhat idiosyncratic based on the population pair chosen (Armbruster *et al.*, 1997; Fenster and Galloway, 2000; Demuth and Wade, 2007), the magnitude of heterosis in the F_1 is usually a matter of degree rather than a difference in sign as we see between C1 ↔ R1 and C2 ↔ R2 and B1 ↔ S1 and B2 ↔ S2. We found idiosyncratic differences in heterosis vs outbreeding depression between two parallel population sets that are separated by the same geographic distance, and that are likely to be locally adapted to the same climatic conditions (Ågren and Schemske, 2012). This suggests a role for stochastic processes such as mutation and drift in shaping not only patterns of fixation of deleterious mutations, but potentially also outbreeding depression. *Arabidopsis* is highly selfing, which increases the likelihood of genetic drift, and ancestral populations of present day northern populations of *Arabidopsis* likely experienced genetic bottlenecks during northern range expansion following Pleistocene glaciation (Beck *et al.*, 2008). Heterosis in C1 ↔ R1 and C2 ↔ R2 is consistent with the action of genetic drift in shaping genetic variation of selective importance. In previous work on this set of populations, the presence of QTL for fitness where the local allele is maladaptive (in C1 ↔ R1), and the greater incidence of maladaptive QTL in Sweden (Ågren *et al.*, 2013), also suggest a role for genetic drift in partially constraining adaptation. There is much less direct evidence for the role of genetic drift in generating outbreeding depression in B1 ↔ S1 and B2 ↔ S2. However, chromosomal inversions and/or tight linkage resulting in underdominant or

pseudo-underdominant effects on fitness are predicted to be more likely to become fixed in selfing populations, or those experiencing bottlenecks (Lande, 1985; Charlesworth, 1992; Schierup and Christiansen, 1996). Moreover, we cannot rule out very local selective pressures as a mechanism for outbreeding depression.

It is important to note that the magnitudes of both heterosis and outbreeding depression are likely to depend on the environment in which fitness is measured. For example, Fenster and Galloway (2000) found that the magnitude of heterosis varied with planting site. In addition, hybrid necrosis found in *Arabidopsis* was partially alleviated at warm (23 °C) temperatures (Bomblies *et al.*, 2007; Alcázar *et al.*, 2009; Smith *et al.*, 2011). In wide crosses, it is difficult to identify a single relevant environment in which to measure fitness, and it was not feasible for us to replicate the experiment across four or even two parental environments. Our approach was to grow all plants in a single, relatively benign, environment to determine the magnitudes of heterosis (and outbreeding depression) without the complications of varied environments. Water and nutrients are not limiting in the greenhouse, and the potential for competition is thus greatly reduced compared with field conditions, as was any potential interactions with herbivores or pathogens. Our estimates of heterosis and outbreeding depression may therefore underestimate what may be expressed in the native field environments. Future work will address the role of the environment on the magnitude of heterosis and outbreeding depression in this system, and pursue a more detailed understanding of the molecular genetic basis of these quantities.

DATA ARCHIVING

Means and variances needed to perform the joint scaling analyses are given in Supplementary Table S7. Raw data and an example analysis script are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.v6b8c.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank a team of dedicated undergraduates for assistance with this work, particularly J Schlang who performed the crosses, and A Harris and E Sandroc who led the effort to count fruits and seeds. We thank Don Griffin and three anonymous reviewers whose comments improved the manuscript. This work was funded by an NSF grant (DEB 1022202) to DW Schemske, and a grant from the Swedish Research Council to J Ågren.

Ågren J, Oakley CG, McKay JK, Lovell JT, Schemske DW (2013). Genetic mapping of adaptation reveals fitness trade-offs in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **110**: 21077–21088.

Ågren J, Schemske DW (2012). Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. *New Phytologist* **194**: 1112–1122.

Alcázar R, García AV, Parker JE, Reymond M (2009). Incremental steps toward incompatibility revealed by *Arabidopsis* epistatic interactions modulating salicylic acid pathway activation. *Proc Natl Acad Sci USA* **106**: 334–339.

Armbruster P, Bradshaw AD, Holzapfel C (1997). Evolution of the genetic architecture underlying fitness in the pitcher-plant mosquito, *Wyeomyia smithii*. *Evolution* **51**: 451–458.

Barth S, Busimi AK, Utz HF, Melchinger AE (2003). Heterosis for biomass yeild and related traits in five hybrids of *Arabidopsis thaliana* L. Heynh. *Heredity* **91**: 36–42.

Beck JB, Schmuths H, Schaal BA (2008). Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics. *Mol Ecol* **17**: 902–915.

Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller type incompatibility syndrome in plants. *PLoS Biol* **5**: 1962–1972.

Charlesworth B (1992). Evolutionary rates in partially self-fertilizing species. *Am Naturalist* **140**: 126–148.

Charlesworth D, Willis JH (2009). The genetics of inbreeding depression. *Nat Rev Genet* **10**: 783–796.

Crow JF (1948). Alternative hypotheses of hybrid vigor. *Genetics* **33**: 477–487.

Demuth JP, Flanagan RJ, Delph LF (2014). Genetic architecture of isolation between two species of *Silene* with sex chromosomes and Haldane's rule. *Evolution* **68**: 332–342.

Demuth JP, Wade MJ (2005). On the theoretical and emperical framework for studying genetic interactions within and among species. *Am Naturalist* **165**: 524–536.

Demuth JP, Wade MJ (2006). Experimental methods for measuring gene interactions. *Annu Rev Ecol Evol System* **37**: 289–316.

Demuth JP, Wade MJ (2007). Population differentiation in the beetle *Tribolium castaneum*. I. Genetic architecture. *Evolution* **61**: 494–509.

Edmands S (1999). Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution* **53**: 1757–1768.

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

Fenster CB (1991). Gene flow in *Chamaecrista fasciculata* (Leguminosae) II. Gene establishment. *Evolution* **45**: 410–422.

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

Fenster CB, Galloway LF, Chao L (1997). Epistasis and its consequences for the evolution of natural populations. *Trends in Ecology & Evolution* **12**: 282–286.

Frankham R, Ballou JD, Eldridge MDB, Lacy RC, Ralls K, Dudash MR *et al.* (2011). Predicting the probability of outbreeding depression. *Conservation Biology* **25**: 465–475.

Frankham R, Weber K (2000). Nature of quantitative genetic variation. In: Singh R, Krimbas CB (eds) *Evolutionary Genetics: From Molecules to Morphology*. Cambridge University Press: Cambridge.

Gimond C, Jovelin R, Han S, Ferrari C, Cutter AD, Braendle C (2013). Outbreeding depression with low genetic variation in selfing *Caenorhabditis* nematodes. *Evolution* **67**: 3087–3101.

Grillo MA, Li C, Hammond M, Wang L, Schemske DW (2013). Genetic architecture of flowering time differentiation between locally adapted populations of *Arabidopsis thaliana*. *New Phytologist* **197**: 1321–1331.

Halligan DL, Keightley PD (2009). Spontaneous mutation accumulation studies in evolutionary genetics. *Annu Rev Ecol Evol System* **40**: 151–172.

Husband BC, Barrett SCH (1992). Genetic drift and the maintenance of the style length polymorphism in tristylous populations of *Eichornia paniculata* (Pontederiaceae). *Heredity* **69**: 440–449.

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

Kirkpatrick M, Barton NH (2006). Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419–434.

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu Rev Plant Biol* **55**: 141–172.

Kusterer B, Muminovi M, Utz HF, Piepho HP, Barth S, Heckenberger M *et al.* (2007a). Analysis of a triple testcross design with recombinant inbred lines reveals a significant role of epistasis in heterosis for biomass-related traits in *Arabidopsis*. *Genetics* **175**: 2009–2017.

Kusterer B, Piepho HP, Utz HF, Schön CC, Muminovic J, Meyer RC *et al.* (2007b). Heterosis for biomass related traits in *Arabidopsis* investigated by quantitative trait loci analysis of the triple testcross design with recombinant inbred lines. *Genetics* **177**: 1839–1850.

Lande R (1985). The fixation of chromosomal rearrangements in a subdivided population with local extinction and colonization. *Heredity* **54**: 323–332.

Lewandowska-Sabat AM, Fjellheim S, Rognli OA (2010). Extremely low genetic variability and highly structured local populations of *Arabidopsis thaliana* at higher latitudes. *Mol Ecol* **19**: 4753–4764.

Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A *et al.* (2013). Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. *Nat Genet* **45**: 884–891.

Lynch M (1991). The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* **45**: 622–629.

Lynch M, Walsh B (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc.: Sunderland, MA.

Mather K, Jinks JL (1982). *Biometrical Genetics. The Study of Continuous Variation* 3rd edn. University Press: Cambridge.

Matute DR, Butler IA, Turissini DA, Coyne JA (2010). A test of the snowball theory for the rate of evolution of hybrid incomaptibilities. *Science* **329**: 1518–1521.

Meyer RC, Kusterer B, Lisec J, Steinfath M, Becher M, Scharr H *et al.* (2010). QTL analysis of early stage heterosis for biomass in *Arabidopsis*. *Theor Appl Genet* **120**: 227–237.

Meyer RC, Witucka-Wall H, Becher M, Blacha A, Boudichevskaia A, Dörmann P *et al.* (2012). Heterosis manifestation during early *Arabidopsis* seedling development is characterized by intermediate gene expression and enhanced metabolic activity in the hybrids. *Plant J* **71**: 669–683.

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.

Orr HA, Coyne JA (1992). The genetics of adaptation: a reassessment. *Am Naturalist* **140**: 725–742.

Orr HA, Turelli M (2001). The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**: 1085–1094.

Rutter MT, Roles A, Conner JK, Shaw RG, Shaw FH, Schneeberger K *et al.* (2012). Fitness of *Arabidopsis thaliana* mutation accumulation lines whose spontaneous mutations are known. *Evolution* **66**: 2335–2339.

Schierup MH, Christiansen FB (1996). Inbreeding depression and outbreeding depression in plants. *Heredity* **77**: 461–468.

Smith LM, Bomblies K, Weigel D (2011). Complex evolutionary events at a tandem cluster of Arabidopsis thaliana genes resulting in a single-locus genetic incompatibility. *PLoS Genet* **7**: 1–13.

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

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)