American Genetic Association

Journal of Heredity, 2016, 1–9
doi:10.1093/jhered/esx079
Symposium Article

OXFORD

Symposium Article

Genomic Admixture Between Locally Adapted Populations of *Arabidopsis thaliana* (mouse ear cress): Evidence of Optimal Genetic Outcrossing Distance

Kattia Palacio-Lopez, Stephen R. Keller and Jane Molofsky

From the Department of Plant Biology, University of Vermont, Burlington, VT 05405.

Address correspondence to Kattia Palacio-Lopez at the address above, or e-mail: kattia.palacio-lopez@uvm.edu.

Received March 10, 2017; First decision April 28, 2017; Accepted September 16, 2017.

Corresponding editor: Lynda Delph

Abstract

Admixture can break up divergent genetic architectures between populations, resulting in phenotypic novelty and generating raw material for environmental selection. The contribution of admixture to progeny trait variation and fitness varies based on the degree of genetic isolation between the parental populations, for which most studies have used geographic distance as a proxy. A novel approach is to estimate optimal crossing distance using the adaptive genetic distance between mates estimated from loci that contribute directly to local adaptation. Here, we aim to understand the effect of admixture on disrupting local adaptation of ecotypes of *Arabidopsis thaliana* separated along gradients of geographic, background, and locally adaptive genetic distances. We created experimental F_1 hybrids between ecotypes that vary in geographic distance and used SNP data to estimate background (putatively neutral) and adaptive genetic distance. Hybrids were grown under controlled conditions, and fitness, growth, and phenology traits were measured. The different traits measured showed a clear effect of adaptive genetic distance, but not geographic distance. The earliest bolting hybrids were intermediate in the adaptive genetic distance between their parents, and also had higher biomass and fitness in terms of fruit and seed production. Our results suggest that disruption of locally adaptive genomic loci decreases the performance of offspring between distantly related parents, but that crosses between very closely related parents also reduce performance, likely through the expression of deleterious recessive alleles. We conclude that during admixture, selection may have to balance the consequences of disrupting local adaption while also avoiding inbreeding depression.

Subject areas: Molecular adaptation and selection, Bioinformatics and computational genetics
Key words: adaptive genetic distance, *Arabidopsis thaliana*, background genetic distance, geographic distance, optimal outcrossing distance, SNPs

Natural populations across their distributional range exhibit genetic variation in locally adaptive traits as result of selection for different genotypes in different environments (Linhart and Grant 1996; Kawecki and Ebert 2004). Admixture, or intraspecific hybridization, recombines genomes between historically isolated lineages that have often diverged in their genetic architecture of fitness-related traits

(Verhoeven et al. 2011). Thus, natural admixture zones have long been recognized as important areas to study the evolutionary process because recombination can result in phenotypic novelty and reveal segregating genetic variance available for natural selection to act upon (Barton 2001; Rieseberg et al. 2003; Lavergne and Molofsky 2007; Keller and Taylor 2010; Friedman 2015; Goulet et al. 2017).

Admixture between genetically divergent and locally adapted populations can constrain or enhance the performance of hybrid offspring, depending on the degree of divergence separating populations and their history of inbreeding, drift, and selection (Lynch 1991; Verhoeven et al. 2011). If populations are highly inbred, then admixed individuals may benefit from an increase in heterozygosity which may bring heterosis (hybrid vigor), that is, the phenotypic superiority of hybrid genotypes compared with their parents (Lippman and Zamir 2007), as a result of sheltering the genetic load of recessive deleterious mutations or gene overdominance (Barton and Hewitt 1985; Lynch 1991; Prentis et al. 2008). Alternatively, genetically admixed individuals may have reduced fitness due to either inbreeding depression or outbreeding depression (Charlesworth and Charlesworth 1987; Lynch 1991; Oostermeijer et al. 1995; Byers 1998). Inbreeding depression is generally attributable to homozygosity of recessive deleterious alleles or loss of overdominance (Lynch 1991), whereas it is suggested that outbreeding depression can occur through 2 distinct mechanisms: 1) disrupting allelic coadaptation (underdominance or complementary epistasis) or 2) disrupting local adaptation to environmental conditions (Waser and Price 1989).

In sessile organisms such as plants, the geographic distance separating populations may be expected to lead to an optimal outcrossing distance that balances the fitness effects of inbreeding depression at short distances with disruption of coadapted alleles and loss of local adaptation at greater distances (Price and Waser 1979; Waser and Price 1989). Evidence for optimal outcrossing has been reported from natural plant populations across a range of geographic crossing distances (Fenster and Galloway 2000; Waser et al. 2000; Grindeland 2008). These studies typically use geographic distance as a proxy for relatedness and degree of shared local adaptation; yet surprisingly, few studies have tested for an optimal outcrossing distance that considers genetic distance (Edmands 1999; Mindaye et al. 2015), and to our knowledge, none of the studies explicitly consider genetic distances based on loci explicitly associated with local adaptation.

With the current availability of large population genomic datasets, it should be possible to parse out the contributions of geography, inbreeding/population history, and local adaptation on optimal outcrossing distance. For example, the overall genome-wide genetic distance calculated across many selectively neutral SNP loci (hereafter, "background genetic distance") should reflect demographic processes that may covary with geographic distance, such as inbreeding and genetic drift (Wright 1943). In addition, geographically separated populations often show elevated divergence (F_{ST}) or association with adaptive phenotypes for a subset of loci experiencing local selection (hereafter, "adaptive genetic distance") (Linhart and Grant 1996). Indeed, elevated divergence at selected loci relative to the background of the rest of the genome is a classic signature of local adaptation (Lewontin and Krakauer 1973; Whitlock 2015) and forms the basis of modern genome scans for local adaptation (Hoban et al. 2016). Comparing locally adapted genomes based on geographic distance with those based on genetic distance at different classes of genomic loci (background and adaptive) may reveal different contributions to the fitness of offspring produced by admixture between geographically separated populations, and provide a more mechanistic understanding of optimal outcrossing distances.

In our study, we predict different fitness effects for intraspecific hybrid offspring based on the geographic or genetic crossing distances separating the parental lines. Based on optimal outcrossing theory, we predict that all 3 distances (geographic, background genetic, adaptive genetic) are capable of generating highest fitness at intermediate distances, resulting in a quadratic fitness function (Figure 1). However, the set of genetic and evolutionary processes responsible for the fitness effects of each distance are distinct, and thus, different patterns observed in regression analyses of fitness traits on geographic or genetic distances yield different inferences on the evolutionary processes responsible (Table 1). In general, hybrids created from very close parental lines, either geographically or genetically, could experience a reduction in fitness due to inbreeding depression (Lande and Schemske 1985) (Figure 1A). The loss of fitness due to inbreeding depression may affect both background and adaptive genetic distances, if both types of distance are associated with genetic load of deleterious mutations. In background genetic distance, this may be attributable to the degree of relatedness between individuals, whereas in adaptive genetic distance, it may reflect slightly deleterious mutations that experienced hitchhiking selection during selective sweeps of adaptive loci (Hartfield and Otto 2011). In contrast, a reduction in fitness in crosses between very distant lines is indicative of outbreeding depression due to epistasis or loss of local adaptation (Figure 1E). It may be possible to tease apart the contributions of adaptive versus background genetic distance to identify the mechanisms responsible for outbreeding depression. We predict that outbreeding depression at large background genetic distances likely reflects the loss of beneficial epistasis and the breakup of coadapted alleles (Lynch 1991), whereas the contribution of large adaptive genetic distance likely reflects dilution of local adaptation

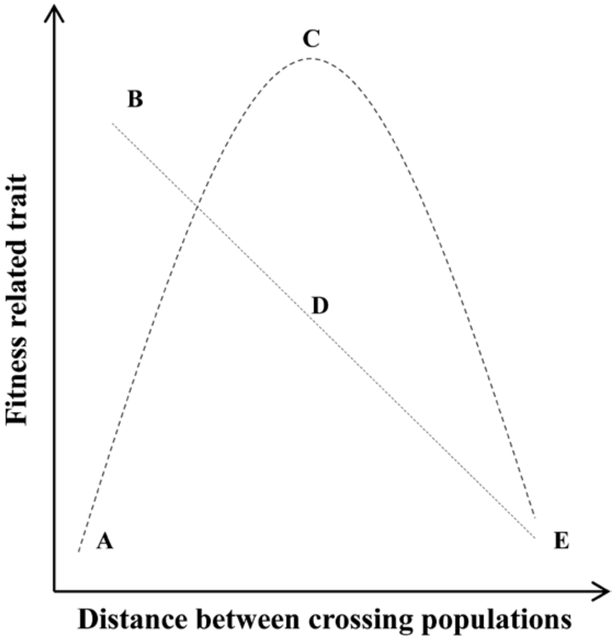

Figure 1. Predictions on the mechanisms that influence fitness hybrids due to geographic and genetic distances. (**A**) Decrease in fitness due to inbreeding depression. (**B**) Increase in fitness due to local adaptation. (**C**) Increase in fitness due to heterosis. (**D**) Intermediate fitness due to partial loss of local adaptation. (**E**) Decrease in fitness due to outbreeding depression (disruption of coadapted genes or loss of local adaptation).

Table 1. Theoretical predictions of fitness on hybrid offspring across a range of geographic and genetic crossing distances (low, mid, and high) due to different evolutionary processes

Crossing distance	Observed pattern	Inferred evolutionary processes	Regions in Figure 1
Geographic	Quadratic	Low: Inbreeding depression (↓) + local adaptation (↑)	A, B
		Mid: Heterosis (↑) + local adaptation (↑↓)	C, D
		High: Local adaptation (↓) + coadapted genes (↓)	E
Background genetic	Quadratic	Low: Inbreeding depression (↓)	A
		Mid: Heterosis (↑)	C
		High: Coadapted genes (↓)	E
Adaptive genetic	Linear or quadratic	Low: Local adaptation (↑) + inbreeding depression (↓)	B, A
		Mid: Local adaptation (↑↓) + heterosis (↑)	D, C
		High: Local adaptation (↓)	E

Arrows show the effect on fitness-related trait.

(Hufford and Mazer 2003; Verhoeven et al. 2011). As a result of these 2 extremes, we predict an optimal outcrossing distance should exist that reflects the benefits of heterosis and the maintenance of local adaptation at intermediate distances while avoiding loss of fitness due to inbreeding and outbreeding at larger distances (Lynch 1991; Oakley et al. 2015) (Figure 1C). In the absence of inbreeding depression, for example, if genetic load is purged in a highly selfing species, it is also possible that hybrid fitness shows an increment between very close lines (Figure 1B). A more linear decline with adaptive genetic distance as local adaptation becomes diluted in crosses between increasingly distant parental lines (Figure 1D).

The annual plant *Arabidopsis thaliana* L. (Brassicaceae) provides an interesting model system to study admixture and the contributions of outcrossing distance on fitness-related traits. *Arabidopsis thaliana* has a primarily selfing mating system in which natural outcrossing and admixture are rare but occur in nature frequently enough to influence population structure and generate clear signals of isolation by geographic distance (Platt et al. 2010). Furthermore, despite its mostly selfing mating system, *A. thaliana* is not immune to the accumulation of genetic load of deleterious mutations (Bustamante et al. 2002; Ågren et al. 2013), and viability loci exhibiting overdominance are known to contribute to heterosis during experimental outcrossing (Mitchell-Olds 1995). Lastly, multiple studies have shown that geographically diverse ecotypes of *A. thaliana* exhibit local adaptation (Rutter and Fenster 2007; Fournier-Level et al. 2011; Hancock et al. 2011; Ågren and Schemske 2012; Gaut 2012; Ågren et al. 2013; Huber et al. 2014) and population genomic studies have identified clear candidate genes contributing to local adaptation (Fournier-Level et al. 2011; Hancock et al. 2011).

To deepen our understanding of the fitness effects of admixture along geographic and genetic gradients, we created experimental F_1 hybrids between *A. thaliana* ecotypes across a range of geographic and genetic distances. Based on the optimal outcrossing hypothesis, we predicted either a negative linear or quadratic relationship between the adaptive genetic distance of the parents, calculated from loci under local adaptation, and the resulting performance of the offspring, where intermediate adaptive genetic distances would indicate an optimal crossing distance balancing local adaptation and inbreeding load. We also predicted that disruption of local adaptation would most strongly affect traits known to experience environmentally varying selection in natural populations, such as phenology and growth traits, whereas any intrinsic effects of disrupting coadapted alleles would more strongly affect traits closely associated with fitness.

Materials and Methods

Study System

We selected 17 distinct ecotypes sampled from across an array of geographic and genetic distances observed in natural populations of *A. thaliana* based on Atwell et al. (2010) (Table 2). Seeds were ordered from The Arabidopsis Information Resource (TAIR; available at www.arabidopsis.org) (TAIR, 2000). To reduce maternal effects, plants were grown for 1 generation under controlled long-day conditions (16 h light: 8 dark [16L: 8D], at 18 °C), in standard Metro-mix soil, with bottom watering provided every 2 days to reduce plant damage.

Crossing and Growing Conditions

We created F_1 hybrid progeny by emasculating pollen from flowers prior to anthesis and then outcrossing with pollen from a distinct ecotype. A total of 11 full-sib families were created this way among the 17 parental ecotypes, representing a range in geographic and genetic crossing distances (Table 3). After ripening and seed collection, first-generation hybrid seeds, along with seeds of the corresponding parental ecotypes, were vernalized in darkness for 5 days at 4 °C before being moved to growth chambers. Imbibed seeds were then transferred to soil and placed under 18 °C to stimulate germination, leaving 1 plant per 5-cm pot. In order to match growing conditions previously used to minimize maternal effects and to provide a benign growing environment, we grew 8 hybrid offspring per cross and 8 plants per each parent at 18 °C under long-day (16L: 8D) conditions in growth chambers. We measured offspring performance and tested for the effect of admixture along geographic and genetic gradients using 3 different types of traits: 1) phenology traits: bolting speed (1/bolting time) and number of leaves at bolting speed (1/number of leaves at bolting time), which covaries strongly with flowering time; 2) growth traits: aboveground dry mass (biomass) and stem length (height); and 3) fitness traits: total fruit production and seed weight.

To estimate the contribution of adaptive genetic distance between parents on the fitness of admixed offspring, we focused on 4 fitness QTL previously associated with local adaptation of *A. thaliana* populations under field conditions (Fournier-Level et al. 2011). Fournier-Level et al. (2011) used results from 4 geographically dispersed common garden trials in Europe and a corresponding genome-wide association study (GWAS) using 213 248 SNPs to identify 4 fitness QTL located near 8 candidate genes (*LAC1*, *CHR8*, *PHYB*, *Δ-TIP*, *NDF4*, *TRZ4*, *SAG21*, *PARP1*) that are associated with survival and silique number. SNPs marking these QTL also show higher frequencies of the fitness-associated allele in populations close to the common garden site where they were associated with fitness in the GWAS, strongly implicating their role in local adaptation (Fournier-Level et al. 2011). We obtained SNP genotypes for each of our parental ecotypes from a 1-kb window upstream and downstream of each QTL's position using publically available data from the Arabidopsis 1001 Genomes project (http://1001genomes.org/about.html), yielding a total of 748 SNPs. We chose to analyze all SNPs within a 1-kb window of each fitness QTL to capture the effects of linked selection acting on allele frequencies. Because linkage disequilibrium (LD) in *A. thaliana* extends to 10 kb (Kim et al. 2007), our use of a 1-kb window should yield SNPs in high LD with the selected locus, while also accommodating uncertainty in the exact location of the causal adaptive locus. In addition, we looked at the genotypes for the 4 individual SNP loci associated with local adaptation in Fournier-Level et al. (2011) and found congruence across all measured traits with the patterns that we reported for adaptive genetic distance measured from closely linked SNPs (results not shown). We used the 748 SNPs to estimate a locally adaptive genetic distance (calculated as 1 − identity by state) between each pair of ecotypes used as parents in the experimental crosses, using Plink v1.9 (Purcell et al. 2007).

Table 2. Information from TAIR on the 17 ecotypes of *Arabidopsis thaliana* selected from Atwell et al. (2010)

Ecotype	Origin	Latitude	Longitude
Wei-0	Switzerland (SUI)	47.25	8.26
Zu-1	Switzerland (SUI)	47.3667	8.55
Bu-0	Germany (GER)	50.5	9.5
Aa-0	Germany (GER)	50.9167	9.57073
Col-0[a]	Unknown	Unknown	Unknown
Ka-0	Austria (AUT)	47	14
Lp2-2	Czech Republic (CZE)	49.38	16.81
Tu-0[a]	Unknown	Unknown	Unknown
Ang-0	Belgium (BEL)	50.3	5.3
Ga-0	Germany (GER)	50.3	8
Lp2-6	Czech Republic (CZE)	49.38	16.81
Sq-1	United Kingdom (UK)	51.4083	−0.638
In-0	Austria (AUT)	47.5	11.5
Uod-7[a]	Unknown	Unknown	Unknown
Tsu-1[a]	Unknown	Unknown	Unknown
Gie-0[a]	Unknown	Unknown	Unknown
Oy-0	Norway (NOR)	60.39	6.19

[a] Ecotypes in which their origin is unknown (Anastasio et al. 2011).

In order to assess the neutral genetic contributions to admixed offspring due to variation in background relatedness and demographic history (i.e., caused by inbreeding or isolation by distance), we also calculated a genome-wide background genetic distance from SNP loci across the genome for the 17 parental ecotypes. SNP genotype data were converted to Variant Call Format (VCF) and filtered using VCFtools (Danecek et al. 2011) to keep only biallelic SNPs annotated as occurring within intergenic regions. We focused on intergenic regions in an attempt to avoid regions of the genome most likely to be under local selection, although we cannot rule out that these regions could contain regulatory functions. To obtain a genome-wide estimate of neutral relatedness, and to reduce effects of strong LD between closely spaced SNPs, we further thinned sites to a 1-kb minimum length between sites. This approach differed from our treatment of adaptive genetic distance, in which we included closely linked sites around fitness QTL to capture effects of linked selection, whereas here we are interested in sampling neutral genetic variation broadly across the genomic background. After filtering, we retained a total of 961 SNPs that were used to estimate background genetic distance (calculated as 1 − identity by state) using Plink v1.9 (Purcell et al. 2007).

Lastly, geographic distances between parental ecotypes were calculated using the latitude and longitude of their locations based on great circle distances (Table 3). Because there is evidence of misidentification in the geographic origin of some *A. thaliana* ecotypes (Anastasio et al. 2011), we first corroborated the origins of our 17 ecotypes with Anastasio et al. (2011). We found that 5 of our ecotypes (Col-0, Tu-0, Uod-7, Tsu-1, Gie-0) were considered by Anastasio et al. (2011) to have an erroneous geographic origin, identified as accessions that are genetically differentiated from their neighbors but genetically very similar to geographically distant individuals (Table 3). To attempt to correct for this misspecification of origin when calculating geographic distances in our crosses, we calculated an estimation of the geographic distance ("proxy geographic distance"). First, we identified the most closely related ecotype to our samples with unknown origin, based on the kinship matrix reported

Table 3. Crossing design based on parental geographic, locally adaptive, and background genetic distances

Crosses by ecotype ID	Crosses by their origin	Adaptive genetic distance	Background genetic distance	Geographic distance (km)	Closely related ecotype ID of the unknown ecotype
Lp2-6 × Lp2-2	CZE × CZE	0.0349	0.0231	0.00	
Wei-0 × Zu-1	SUI × SUI	0.0157	0.0300	25.43	
Aa-0 × Bu-0	GER × GER	0.0143	0.0294	46.60	
Aa-0 × Ga-0	GER × GER	0.0102	0.0244	130.30	
Ka-0 × In-0	AUT × AUT	0.0334	0.0406	196.70	
Ang-0 × Sq-1	BEL × UK	0.0035	0.0455	434.5	
Gie-0 × Oy-0	Unknown × NOR	0.0356	0.0392	0.00	Oy-0
Tsu-1 × Tu-0	Unknown × Unknown	0.0116	0.0457		N.F.
Col-0 × Sq-1	Unknown × UK	0.0202	0.0508	1077.00	H-55
Col-0 × Aa-0	Unknown × GER	0.0317	0.0368	638.60	H-55
Col-0 × Uod-7	Unknown × Unknown	0.0528	0.0294	868.70	H-55, Uod-1

Geographic distances are based on Euclidean distance using the latitude and longitude of their locations. Geographic origin was verified with the results of Anastasio et al. (2011). Geographic distance of ecotypes with unknown origin has been calculated after replacement with their closely related ecotype based on genome-wide estimates of pairwise kinship from Atwell et al. (2010). Countries where ecotypes come from GER (Germany), NOR (Norway), UK (United Kingdom), AUT (Austria), SUI (Switzerland), BEL (Belgium), and CZE (Czech Republic). Not found closely related ecotypes (N.F.).

by Atwell et al. (2010). We then used the geographic origin of the most closely related ecotype to calculate the proxy geographic distance between parents used in our crosses (Table 3). We recognize that this approach is an approximation of the true unknown geographic distances.

We assessed the nature of the response of fitness and locally adaptive traits to genetic and geographic distances by fitting multiple regression models in JMP, version 12.0. To adjust for non-normality, bolting speed and biomass were square root transformed, and number of leaves at bolting speed, number of fruits and seed weight were transformed using $\log_{10}$. We included both linear and quadratic terms in regression models to test our predictions. Using multiple regression allowed us to partition the relative importance of genetic and geographical distances, while controlling for partial correlations among these predictors, but could be biased if strong autocorrelation exists among our predictors. We assessed the degree of correlation between our 3 distances (adaptive, background, and geographic) using univariate regression in JMP 12.0, as well as through Mantel tests using PC-ORD 6.0. We found no significant evidence for correlation among predictors (all $P > 0.1$), justifying their inclusion in multiple regression. For simplicity, we plot results from univariate regressions to show the shape of the relationship between traits and distance but report statistical tests based on the multiple regression models in which we used Bonferroni correction to set the α significance level.

Results

We found support for a quadratic relationship between adaptive genetic distance and phenology, growth, and some of the fitness traits in F_1 hybrids, supporting the prediction of an optimal outcrossing distance (Table 4, Figure 2A). The earliest bolting hybrids were from crosses between ecotypes with intermediate adaptive genetic distance. The F_1 hybrids of *A. thaliana* that bolted early also showed a fitness advantage in terms of fruit production ($\beta \pm SE = 2.80279 \pm 0.840954$, $P = 0.0012$) (Figure 3). Hybrids from intermediate adaptive genetic distances also bolted with fewer numbers of leaves (Figure 2A). In addition, intermediate adaptive distance hybrids produced more seeds than hybrids from very close or far adaptive distances. A linear decrease in fruit number with the increment of adaptive genetic distance of the cross parents was evident for F_1 hybrids of *A. thaliana* (Table 4, Figure 2A). Height and biomass were not affected by adaptive genetic distance.

Effects of background genetic distance (inbreeding and disruption of epistatic interactions) were also observed for bolting speed and leaf number, showing a pattern similar to adaptive genetic distance (Table 4). Hybrids from crosses of intermediate background genetic distances bolted earlier with fewer numbers of leaves (Figure 2B). Background genetic distance showed a linear decrease in biomass and fruit number, whereas seed weight decreased up to an intermediate background genetic distance and then increased slightly among the most close and distant crosses (Figure 2B, Table 4).

To compare our analysis of genetic analysis of optimal outcrossing distances with previous studies, we also tested for an effect of geographical distance on outcrossing in the multivariate analysis. Geographical distance showed a different effect on phenology, growth, and fitness traits than the effect suggested by either genetic distance. The analysis using the proxy geographic distance indicated a significant positive quadratic relationship with phenology traits, with hybrids from intermediate geographic distances bolting later and having more leaves at bolting compared with hybrids from geographically close or distant parents (Figure 2C). This is opposite to the prediction for optimal outcrossing as well as opposite to the relationship observed for both adaptive and background genetic distances. For the remaining growth and fitness traits, the effect of geographic distance was generally weak and nonsignificant (after Bonferroni correction), with the exception of height and seed weight, which showed reduced trait values for crosses at close geographic distances (Figure 2C, Table 4).

Discussion

Independent assortment and recombination of locally adaptive genomic regions between distantly related parents can decrease the performance of offspring for environmentally selected traits (Waser et al. 2000; Grindeland 2008). However, crossing between closely related parents often results in inbreeding depression, also leading to reduced offspring performance (Lynch 1991). Thus, during outcrossing between parental ecotypes separated along a gradient of genetic distance, selection may have to balance the consequences of disrupting local adaption and coadapted genes at greater distances while also avoiding inbreeding depression and other dominance effects at closer distances, predicting an intermediate optimum crossing distance (Figure 1). Our study supports the existence of such an optimal outcrossing distance between *A. thaliana* ecotypes in which there is loss of local adaptation with adaptive genetic distance between the parents. Increased adaptive genetic distance resulted in reduced performance for most traits; however, we also observed lower performance at very close adaptive genetic distances, suggesting additional negative effects of inbreeding. Thus, our results suggest that selection history may favor recombinant offspring genotypes that come from an intermediate degree of adaptive genetic distance between the parents.

The strongest evidence for optimal outcrossing we observed was for phenology traits (bolting time and leaves at bolting). The earliest bolting hybrids were produced by crosses that were intermediate in adaptive genetic distance between the parents, as predicted by the optimal outcrossing hypothesis. Abundant evidence exists from studies of *A. thaliana* for local adaptation at both the phenotypic and molecular levels (Mitchell-Olds and Schmitt 2006; Ågren and Schemske 2012; Ågren et al. 2013), and traits such as bolting and flowering time are clearly associated with locally adaptive ecological differentiation (McKay et al. 2003; Stinchcombe et al. 2004; Lasky et al. 2012, 2014). Bolting time shows evidence of adaptation to climatic conditions (Montesinos-Navarro et al. 2011) and is associated with fitness under field conditions (Korves et al. 2007). It has also been reported that selection on bolting time can constrain or enhance the ability of particular genotypes to colonize different areas (Griffith et al. 2004), for example, later flowering genotypes have more restricted range potentials and narrower niche breadths than earlier flowering genotypes (Banta et al. 2012).

Our observation of slower bolting speed at larger adaptive genetic distances suggests that recombination between locally adapted populations may generate a mismatch in the adaptive alleles in admixed genotypes, reducing their level of local adaptation (Verhoeven et al. 2011). Yet, phenology traits also experienced loss of performance between genetically close parents in our crosses, possibly the result of loss of overdominance (Mitchell-Olds 1995) or linked partially deleterious alleles in the regions of adaptive SNP loci (Bustamante et al. 2002). This suggests that some of the benefits of intermediate

Table 4. Multiple regression table including linear and quadratic effects of genetic and geographic distances

	Unstandardized coefficient		Standardized coefficient		
Response variable	*B*	SE	β	*t*	$P < 0.008$
Bolting speed (1/days)					
Intercept	0.2814	0.0151	0.0000	18.6700	**<0.0001**
Adaptive genetic distance	−1.7629	0.2012	−0.9134	−8.7600	**<0.0001**
Adaptive genetic distance^2	−38.1583	10.1493	−0.3097	−3.7600	**0.0003**
Background genetic distance	−2.2199	0.3172	−0.6992	−7.0000	**<0.0001**
Background genetic distance^2	−379.2405	40.5247	−0.8968	−9.3600	**<0.0001**
Proxy geographic distance	-8.37×10^{-7}	1.02×10^{-5}	−0.0104	−0.0800	0.9345
Proxy geographic distance^2	2.5×10^{-7}	2.77×10^{-8}	1.1707	9.0300	**<0.0001**
Number of leaves at bolting					
Intercept	−0.6891	0.0858	0.0000	−8.0300	**<0.0001**
Adaptive genetic distance	−10.2322	1.1460	−0.8407	−8.9300	**<0.0001**
Adaptive genetic distance^2	−213.5617	57.7979	−0.2749	−3.6900	**0.0004**
Background genetic distance	−12.3149	1.8062	−0.6152	−6.8200	**<0.0001**
Background genetic distance^2	−2960.5520	230.7798	−1.1103	−12.8300	**<0.0001**
Proxy geographic distance	-2.37×10^{-5}	5.78×10^{-5}	−0.0466	−0.4100	0.6831
Proxy geographic distance^2	1.67×10^{-6}	1.58×10^{-7}	1.2356	10.5600	**<0.0001**
Height (cm)					
Intercept	37.4910	5.9531	0.0000	6.3000	**<0.0001**
Adaptive genetic distance	−57.3551	79.5081	−0.1121	−0.7200	0.4724
Adaptive genetic distance^2	−7163.3520	4009.8910	−0.2194	−1.7900	0.0771
Background genetic distance	−201.2642	125.3067	−0.2392	−1.6100	0.1115
Background genetic distance^2	41 512.0100	16 010.9900	0.3703	2.5900	0.0110
Proxy geographic distance	0.0096	0.0040	0.4517	2.4000	0.0181
Proxy geographic distance^2	-3.13×10^{-5}	0.000011	−0.5528	−2.8600	**0.0051**
Biomass (mg)					
Intercept	20.8417	3.1422	0.0000	6.6300	**<0.0001**
Adaptive genetic distance	−68.0925	41.9665	−0.2448	−1.6200	0.1079
Adaptive genetic distance^2	−4603.8230	2116.5310	−0.2593	−2.1800	0.0320
Background genetic distance	−205.1340	66.1403	−0.4483	−3.1000	**0.0025**
Background genetic distance^2	16 041.9870	8451.0410	0.2632	1.9000	0.0606
Proxy geographic distance	0.0043	0.0021	0.3683	2.0200	0.0462
Proxy geographic distance^2	-1.27×10^{-5}	5.78×10^{-6}	−0.4108	−2.1900	0.0308
Number of fruits					
Intercept	2.5089	0.1959	0.0000	12.8100	**<0.0001**
Adaptive genetic distance	−8.2830	2.6164	−0.4695	−3.1700	**0.0021**
Adaptive genetic distance^2	−289.0316	131.9548	−0.2567	−2.1900	0.0309
Background genetic distance	−18.6301	4.1235	−0.6420	−4.5200	**<0.0001**
Background genetic distance^2	309.5684	526.8787	0.0801	0.5900	0.5582
Proxy geographic distance	0.0004	0.0001	0.4761	2.6500	0.0093
Proxy geographic distance^2	3.97×10^{-8}	3.60×10^{-7}	0.0203	0.1100	0.9125
Seed weight (mg)					
Intercept	1.9548	0.2381	0.0000	8.2100	**<0.0001**
Adaptive genetic distance	−1.9833	3.1780	−0.0909	−0.6200	0.5340
Adaptive genetic distance^2	−437.8378	156.4825	−0.3145	−2.8000	**0.0062**
Background genetic distance	−17.0705	4.9456	−0.4759	−3.4500	**0.0008**
Background genetic distance^2	2398.7780	640.6284	0.5006	3.7400	**0.0003**
Proxy geographic distance	0.0007	0.0002	0.7520	4.4200	**<0.0001**
Proxy geographic distance^2	-1.42×10^{-6}	4.35×10^{-7}	−0.5828	−3.2600	**0.0015**

P values in bold are significant after Bonferroni correction ($\alpha = 0.05/6 = 0.008$). SE, standard error.

outcrossing distances may accrue from heterosis, which may affect both adaptive and background genetic distances, based on the intermediate optima evident for both types of genetic distance (Figure 2).

Selection favors different alleles over the geographic range of a species in which local adaptation has played an important role maintaining adaptive natural variation (Feulner et al. 2015). Although this has led to the common assumption that geographic distance should be a good predictor for disruption of local adaptation (Price and Waser 1979; Schmidt and Levin 1985; Waser and Price 1989), we found geographic distance to be a relatively poor fit to the optimal outcrossing hypothesis compared with adaptive and background genetic distances (Figure 2, Table 4). Genomic tools have provided extensive evidence for the genetic basis of local adaptation (e.g., Lasky et al. 2014; Yoder et al. 2014; Rellstab et al. 2017). With the growing availability of methods to uncover locally adapted regions of the genome, we expect that future studies of admixture across a gradient of adaptive genetic distances will uncover similar findings, especially when allelic interactions within or between loci

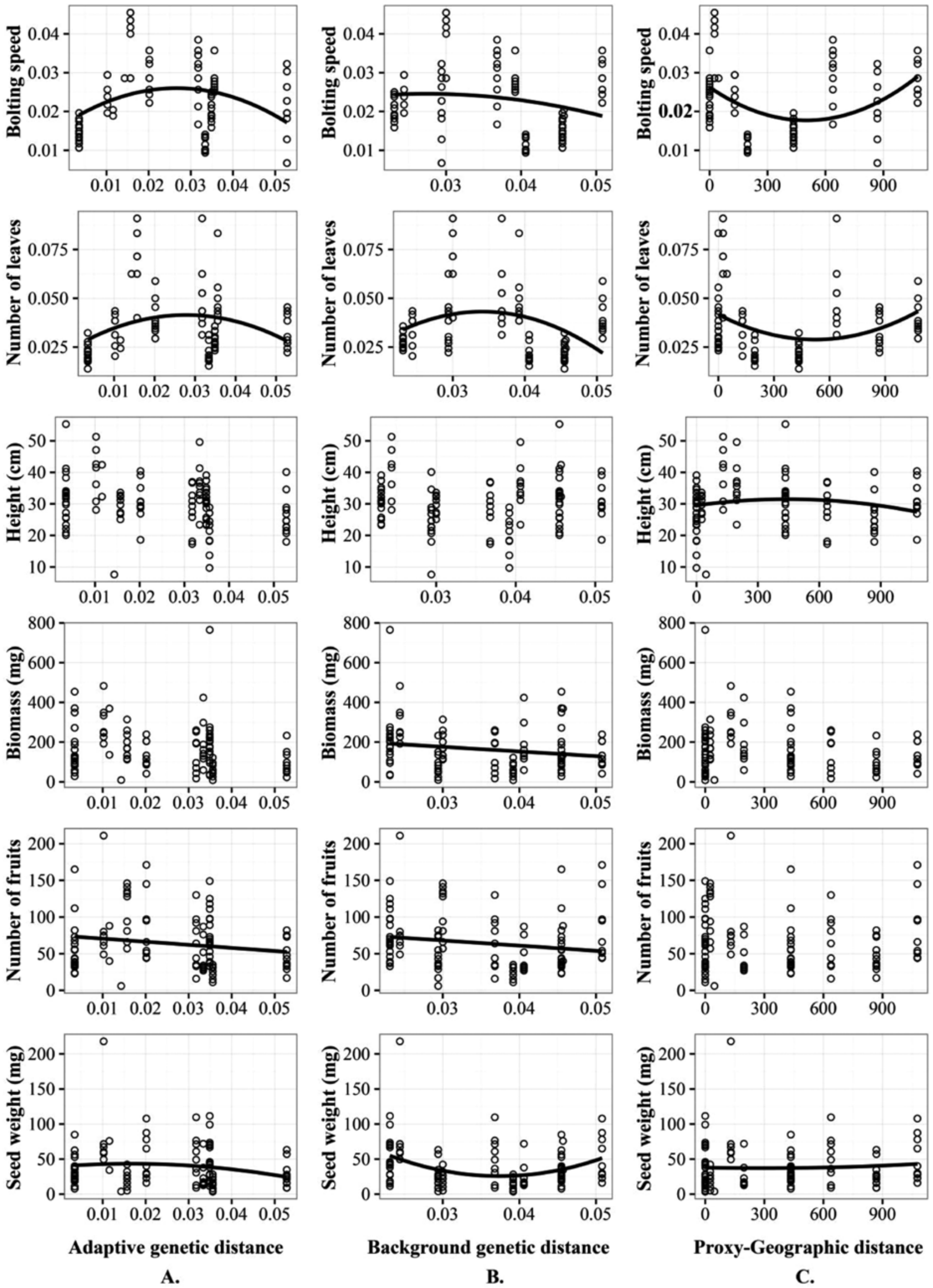

Figure 2. Relationship between phenology, growth, and fitness traits and different parental distance metrics (**A**, adaptive genetic distance; **B**, background genetic distance; **C**, proxy geographic distance) in F_1 hybrids of *Arabidopsis thaliana*.

are implicated in the genetic architecture of adaptation. For example, a recent study with different populations of *A. thaliana* across its native range reported evidence of strong selection on a defense trait against herbivores (glucosinolate profiles) (Brachi et al. 2015). This study suggests 2 genes (MAM1 and GS-OH) to be the targets of divergent selection between Eastern and Western Europe driven by the local herbivore community (Brachi et al. 2015). Because these 2 genes are part of the same biosynthetic pathway and have epistatic effects on fitness, these results suggest that selection played a role in locking the genome into locally favorable combinations of alleles (Brachi et al. 2015). Thus, hybridization could break out the locally adaptive genome bringing negative fitness consequences.

Because fitness traits are sensitive to environmental conditions, it is important to be cautious when interpreting the performance of our experimental hybrids measured in a single test environment (18 °C under 16L: 8D). In particular, we recognize that our experimental

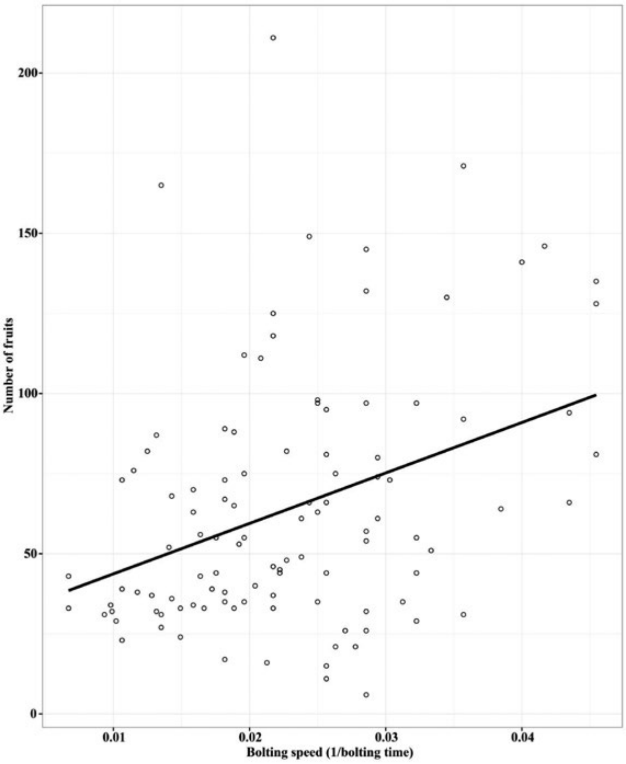

Figure 3. Relationship between phenology and fitness traits in F_1 hybrids of *Arabidopsis thaliana*.

growing conditions do not reflect the full range of natural environments of all the parental ecotypes used in our study, which can constrain the response of the different traits measured. However, growth temperatures around 18 °C with no water limitations have been used in other studies as a benign growth environment for *A. thaliana* (Lee et al. 2013; Posé et al. 2013). Favorable growing conditions allowed us to assess the effects of genetic and geographic crossing distances under conditions conducive to plant growth, but probably missed differential responses to stress that may be important aspects of local adaptation. Future studies should incorporate limiting abiotic factors or other forms of stress in order to have a more realistic scenario, and ideally, hybrids should be measured in both native parental environments in a reciprocal transplant design in order to compare how hybrid fitness compares with parental fitness under the conditions to which the parental genotypes adapted.

This contribution is relevant to studies that wish to make predictions of plant performance in hybrid crosses between different genetically divergent lines. Based on our results, integrating information on adaptive genetic distances based on genome scans for local adaptation provides additional information on plant performance beyond what is attributable to genome-wide measures of overall kinship or geographic distance. As population genomic studies identify additional candidates for local adaptation, future work testing the relevance of adaptive genetic distance on crosses should evaluate the context dependency of optimal outcrossing across a range of growth environments and for a larger fraction of the locally adaptive portion of the genome. Studies of optimal outcrossing based on accumulated knowledge of the genomic basis of local adaptation have great potential to reveal the fitness effects of recombination between locally adapted populations, and the implications this has for species experiencing admixture when expanding their ranges.

Funding

Funding for this study was provided by a Fulbright Scholarship to K.P.L., NIFA-HATCH to J.M., and NSF-1461868 to S.K.

Acknowledgments

We gratefully acknowledge the organizers of the 2016 AGA Symposium on Local Adaptation, especially L. Delph and A. Baker. We also thank the 2 anonymous reviewers for helpful comments that improved the manuscript.

Conflict of interest

None declared.

References

Ågren J, Oakley CG, McKay JK, Lovell JT, Schemske DW. 2013. Genetic mapping of adaptation reveals fitness tradeoffs in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA*. 110:21077–21082.

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

Anastasio AE, Platt A, Horton M, Grotewold E, Scholl R, Borevitz JO, Nordborg M, Bergelson J. 2011. Source verification of mis-identified *Arabidopsis thaliana* accessions. *Plant J*. 67:554–566.

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, *et al.* 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*. 465:627–631.

Banta JA, Ehrenreich IM, Gerard S, Chou L, Wilczek A, Schmitt J, Kover PX, Purugganan MD. 2012. Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in *Arabidopsis thaliana*. *Ecol Lett*. 15:769–777.

Barton NH. 2001. The role of hybridization in evolution. *Mol Ecol*. 10:551–568.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annu Rev Ecol Syst*. 16:113–148.

Brachi B, Meyer CG, Villoutreix R, Platt A, Morton TC, Roux F, Bergelson J. 2015. Coselected genes determine adaptive variation in herbivore resistance throughout the native range of *Arabidopsis thaliana*. *Proc Natl Acad Sci USA*. 112:4032–4037.

Bustamante CD, Nielsen R, Sawyer SA, Olsen KM, Purugganan MD, Hartl DL. 2002. The cost of inbreeding in *Arabidopsis*. *Nature*. 416:531–534.

Byers D. 1998. Effect of cross proximity on progeny fitness in a rare and a common species of *Eupatorium* (Asteraceae). *Am J Bot*. 85:644.

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

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* 2011. The variant call format and VCFtools. *Bioinformatics*. 27:2156–2158.

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

Fenster CB, Galloway LF. 2000. Inbreeding and outbreeding depression in natural populations of *Chamaecrista fasciculate* (Fabaceae). *Conserv Biol*. 14:1406–1412.

Feulner PG, Chain FJ, Panchal M, Huang Y, Eizaguirre C, Kalbe M, Lenz TL, Samonte IE, Stoll M, Bornberg-Bauer E, *et al.* 2015. Genomics of divergence along a continuum of parapatric population differentiation. *PLoS Genet*. 11:e1004966.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science*. 334:86–89.

Friedman E. 2015. Consequences of hybridization and heterozygosity on plant vigor and phenotypic stability. *Plant Sci*. 232:35–40.

Gaut B. 2012. *Arabidopsis thaliana* as a model for the genetics of local adaptation. *Nat Genet*. 44:115–116.

Griffith C, Kim E, Donohue K. 2004. Life-history variation and adaptation in the historically mobile plant *Arabidopsis thaliana* (Brassicaceae) in North America. *Am J Bot*. 91:837–849.

Grindeland JM. 2008. Inbreeding depression and outbreeding depression in *Digitalis purpurea*: optimal outcrossing distance in a tetraploid. *J Evol Biol*. 21:716–726.

Goulet BE, Roda F, Hopkins R. 2017. Hybridization in plants: old ideas, new techniques. *Plant Physiol*. 173:65–78.

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, Toomajian C, Roux F, Bergelson J. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*. 334:83–86.

Hartfield M, Otto SP. 2011. Recombination and hitchhiking of deleterious alleles. *Evolution*. 65:2421–2434.

Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC. 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am Nat*. 188:379–397.

Huber CD, Nordborg M, Hermisson J, Hellmann I. 2014. Keeping it local: evidence for positive selection in Swedish *Arabidopsis thaliana*. *Mol Biol Evol*. 31:3026–3039.

Hufford KM, Mazer SJ. 2003. Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends Ecol Evol*. 18:147–155.

JMP, Version 12.0. 1989–2007. Cary (NC): SAS Institute Inc.

Kawecki TJ, Ebert D. 2004. Conceptual issues in local adaptation. *Ecol Lett*. 7:1225–1241.

Keller SR, Taylor DR. 2010. Genomic admixture increases fitness during a biological invasion. *J Evol Biol*. 23:1720–1731.

Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M. 2007. Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nat Genet*. 39:1151–1155.

Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, Purugganan MD, Schmitt J. 2007. Fitness effects associated with the major flowering time gene FRIGIDA in *Arabidopsis thaliana* in the field. *Am Nat*. 5:141–157.

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

Lasky JR, Des Marais DL, Lowry DB, Povolotskaya I, McKay JK, Richards JH, Keitt TH, Juenger TE. 2014. Natural variation in abiotic stress responsive gene expression and local adaptation to climate in *Arabidopsis thaliana*. *Mol Biol Evol*. 31:2283–2296.

Lasky JR, Des Marais DL, McKay JK, Richards JH, Juenger TE, Keitt TH. 2012. Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Mol Ecol*. 21:5512–5529.

Lavergne S, Molofsky J. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc Natl Acad Sci USA*. 104:3883–3888.

Lee JH, Ryu HS, Chung KS, Posé D, Kim S, Schmid M, Ahn JH. 2013. Regulation of temperature-responsive flowering by MADS-box transcription factor repressors. *Science*. 342:628–632.

Lewontin RC, Krakauer J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*. 74:175–195.

Lippman ZB, Zamir D. 2007. Heterosis: revisiting the magic. *Trends Genet*. 23:60–66.

Linhart YB, Grant MC. 1996. Evolutionary significance of local genetic differentiation in plants. *Annu Rev Ecol Syst*. 27:237–277.

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

McCune B, Mefford MJ. 2011. *PC-ORD: multivariate analysis of ecological data, Version 6*. Gleneden Beach (OR): MjM Software.

McKay JK, Richards JH, Mitchell-Olds T. 2003. Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. *Mol Ecol*. 12:1137–1151.

Mindaye TT, Mace ES, Godwin ID, Jordan DR. 2015. Genetic differentiation analysis for the identification of complementary parental pools for sorghum hybrid breeding in Ethiopia. *Theor Appl Genet*. 128:1765–1775.

Mitchell-Olds T. 1995. Interval mapping of viability loci causing heterosis in *Arabidopsis*. *Genetics*. 140:1105–1109.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature*. 441:947–952.

Montesinos-Navarro A, Wig J, Pico FX, Tonsor SJ. 2011. *Arabidopsis thaliana* populations show clinal variation in a climatic gradient associated with altitude. *New Phytol*. 189:282–294.

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

Oostermeijer JGB, Altenburg RGM, den Nijs JCM. 1995. Effects of outcrossing distance and selfing on fitness components in the rare *Gentiana pneumonanthe* (Gentianaceae). *Acta Bot Neerl*. 44:257–268.

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, *et al*. 2010. The scale of population structure in *Arabidopsis thaliana*. *PLoS Genet*. 6:e1000843.

Posé D, Verhage L, Ott F, Yant L, Mathieu J, Angenent GC, Immink RG, Schmid M. 2013. Temperature-dependent regulation of flowering by antagonistic FLM variants. *Nature*. 125:1–16.

Prentis PJ, Wilson JR, Dormontt EE, Richardson DM, Lowe AJ. 2008. Adaptive evolution in invasive species. *Trends Plant Sci*. 13:288–294.

Price MV, Waser NM. 1979. Pollen dispersal and optimal outcrossing in *Delphinium nelsoni*. *Nature*. 277:294–297.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, *et al*. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 81:559–575.

Rellstab C, Fischer MC, Zoller S, Graf R, Tedder A, Shimizu KK, Widmer A, Holderegger R, Gugerli F. 2017. Local adaptation (mostly) remains local: reassessing environmental associations of climate-related candidate SNPs in *Arabidopsis halleri*. *Heredity (Edinb)*. 118:193–201.

Rieseberg LH, Widmer A, Arntz AM, Burke JM. 2003. The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philos Trans R Soc Lond B Biol Sci*. 358:1141–1147.

Rutter MT, Fenster CB. 2007. Testing for adaptation to climate in *Arabidopsis thaliana*: a calibrated common garden approach. *Ann Bot*. 99:529–536.

Schmidt KP, Levin DA. 1985. The comparative demography of reciprocally sown populations of *Phlox drummondii* Hook. I. Survivorships, fecundities, and finite rates of increase. *Evolution*. 39:396–404.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD, Schmitt J. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proc Natl Acad Sci USA*. 101:4712–4717.

Verhoeven KJ, Macel M, Wolfe LM, Biere A. 2011. Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proc Biol Sci*. 278:2–8.

Waser NM, Price MV. 1989. Optimal outcrossing in *Ipomopsis aggregata*: seed set and offspring fitness. *Evolution*. 43:1097–1109.

Waser NM, Price MV, Shaw RG. 2000. Outbreeding depression varies among cohorts of *Ipomopsis aggregata* planted in nature. *Evolution*. 54:485–491.

Whitlock MC. 2015. Modern approaches to local adaptation. *Am Nat*. 186:1–4.

Wright S. 1943. Isolation by distance. *Genetics*. 28:114–138.

Yoder JB, Stanton-Geddes J, Zhou P, Briskine R, Young ND, Tiffin P. 2014. Genomic signature of adaptation to climate in *Medicago truncatula*. *Genetics*. 196:1263–1275.