JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12612

On the potential strength and consequences for nonrandom gene flow caused by local adaptation in flowering time

A. E. WEIS

Department of Ecology and Evolutionary Biology, Koffler Scientific Reserve at Jokers Hill, University of Toronto, Toronto, ON, Canada

Keywords:

assortative mating;
Brassica rapa;
flowering schedule;
migrational load;
phenology.

Abstract

Gene flow is generally considered a random process, that is the loci under consideration have no effect on dispersal success. Edelaar and Bolnick (*Trends Ecol Evol*, 27, 2012, 659) recently argued that nonrandom gene flow could exert a significant evolutionary force. It can, for instance, ameliorate the maladaptive effects of immigration into locally adapted populations. I examined the potential strength for nonrandom gene flow for flowering time genes, a trait frequently found to be locally adapted. The idea is that plants that successfully export pollen into a locally adapted resident population will be a genetically biased subset of their natal population – they will have resident-like flowering times. Reciprocally, recipients will be more migrant-like than the resident population average. I quantified the potential for biased pollen exchange among three populations along a flowering time cline in *Brassica rapa* from southern California. A two-generation line cross experiment demonstrated genetic variance in flowering time, both within and among populations. Calculations based on the variation in individual flowering schedules showed that resident plants with the most migrant-like flowering times could expect to have up to 10 times more of the their flowers pollinated by immigrant pollen than the least migrant-like. Further, the mean flowering time of the pollen exporters that have access to resident mates differs by up to 4 weeks from the mean in the exporters' natal population. The data from these three populations suggest that the bias in gene flow for flowering time cuts the impact on the resident population by as much as half. This implies that when selection is divergent between populations, migrants with the highest mating success tend to be resident-like in their flowering times, and so, fewer maladaptive alleles will be introduced into the locally adapting gene pool.

Introduction

The adaptive divergence of populations along geographical gradients presented some of the earliest evidence for the power of natural selection (Clausen & Heisey, 1958; Endler, 1977). Divergent selection between sites leads to local adaptation (Linhart & Grant, 1996; Leimu & Fischer, 2008; Hereford, 2009). Gene flow, however, can counter selection when immigrants introduce maladaptive alleles into the local population. The net divergence between populations can rest upon a migration selection balance (Hendry *et al.*, 2001; Lenormand, 2002; Postma & van Noordwijk, 2005; Bolnick & Nosil, 2007; Paul *et al.*, 2011) in combination with genetic drift (Blanquart *et al.*, 2012).

Most considerations of migration's effect on local adaptation have treated gene flow as a random process, *that is* the dispersal characteristics of individuals are genetically uncorrelated with the locally adapted trait. But such correlations can occur. Edelaar & Bolnick (2012) have recently argued that when they do, the resulting nonrandom gene flow can be a significant force in evolution. Formally, nonrandom gene flow occurs when genetically variable traits (behavioural,

Correspondence: Arthur E. Weis, Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON, M5S 3B2 Canada.
Tel.: +1 416 476 4684; fax: +1 416 478 5878; e-mail: arthur.weis@utoronto.ca

physiological or morphological) influence migration success. As a result, a biased subset of alleles controlling the migration trait is passed to the resident population. A migrant's success entails not only movement from its natal population to a receiving population, but also, upon arrival, a nonzero mating success, *that is* its genes have to be incorporated into the recipient gene pool. In this view, realized gene flow will be biased by variation in fertility among arriving migrants because fecund individuals pass more genes into the resident population (Hendry, 2004; Lopez *et al.*, 2008; Guillaume, 2011). Biased mate choice can further affect realized gene flow. If migrants preferentially mate with low-fertility residents, fewer gene copies are passed into the recipient population. Mating preference for high-fertility individuals does the opposite. This study explores the potential for variation in flowering time within and between populations to cause nonrandom gene flow in an annual plant.

Nonrandom gene flow at flowering time loci

Plant populations frequently diverge in flowering phenology (Clausen & Heisey, 1958; Ågren & Schemske, 2012), often at short spatial scales (e.g. Antonovics & Bradshaw, 1970; Silvertown *et al.*, 2005; Lowry *et al.*, 2008b; Haggerty & Galloway, 2011; Kawai & Kudo 2011). Reciprocal transplant experiments show that phenological divergence is often, and perhaps usually, adaptive (Ellis *et al.*, 2006; Griffith & Watson, 2006; Colautti *et al.*, 2010). A literature review by Mazer & LeBuhn (1999) noted that flowering time differs among local plant populations more often than other life-history traits, and recent studies revealed that flowering time can evolve over tens of generations or less (Franks *et al.*, 2007; Colautti *et al.*, 2010).

Why is flowering phenology prone to local adaptation? Phenological assortative mating (Fox, 2003; Weis & Kossler, 2004) can be a contributing factor. First, assortative mating inflates the additive genetic variance in the assorting trait (Felsenstein, 1981), which facilitates response to local selection (Kirkpatrick, 2000). Second, asynchronous reproduction between residents and immigrants results in temporal reproductive isolation between populations (Lowry *et al.*, 2008a), thus restricting gene flow. This study argues that asynchrony also results in biased gene flow.

As with any trait, immigration can introduce maladaptive genetic variance for flowering time. Kirkpatrick (2000) presented a haploid, additive genetic model that evaluated the impact of immigration on traits that automatically mate assortatively, such as flowering time (as well as other traits such as body size in animals). The version of the model relevant to this study assumes free recombination, no genetic incompatibilities resulting in selection against hybrids and negligible stabilizing selection on the assorting trait. In Kirkpatrick's formulation, the change in the automatically assorting trait in the resident population after one generation can be written as

$$\Delta\bar{T}_R \approx \beta G + m(\bar{T}_M - \bar{T}_R)(1 - I),$$

where $\bar{T}_M$ is the trait mean of the resident population before selection and $\bar{T}_M$ is the mean of the immigrants. The product βG is the linear selection gradient acting on the trait multiplied by its genetic variance. Finally, m is the immigration rate and I is the intensity of selection acting against immigrants. The point is that natural selection moves $\bar{T}_R$ towards the local optimum, whereas immigration pulls it back towards $\bar{T}_M$ in proportion to the immigration rate and selection intensity. At equilibrium (i.e. at selection migration balance), the resident mean phenotype will be suboptimal, but less so than if migration was random.

Kirkpatrick's formulation is simple and elegant, but the interesting biology behind it may not be obvious. There is no explicit term for assortative mating: its impact is subsumed, in part, into the genetic variance term, which will be inflated relative to panmictic expectations. At first blush, one might suppose that m is the proportion of individuals in the resident site mating pool that are immigrants and that $\bar{T}_M$ is the mean phenotype of the immigrants' natal population (which is here denoted as $\bar{T}_N$). This would be correct for random gene flow. However, when the probability of successfully mating with a resident depends on the immigrant's phenotype, the effective proportion of immigrants in the mating pool and their phenotypic mean is changed from the random expectations. These effects are subsumed in the selection term, I.

This study dissects the factors that create the potential for nonrandom gene flow at flowering time loci. To set the stage, consider a verbal model that follows Kirkpatrick's logic, but in the context of diploid, plant populations exchanging genes through pollen. Imagine two adjacent populations that have adaptively diverged for flowering time; designate these as the E (early flowering) and L (late flowering) populations. Suppose that the mean flowering times sit away from their local optima, with E being later and L earlier than optimal (selection is directional within populations and divergent between). Assume also the following: the populations have equal genetic variances in flowering time, population divergence is small enough for some overlap in their flowering schedules, no additional traits influence mating success, and pollen is vectored freely and randomly within and between sites on every day of the flowering season. Under this scenario, realized gene flow at flowering time loci will be nonrandom. Pollen vectored away from the E site starts to arrive before the resident L-plants are in flower, and so, the earliest of the E plants sire no seed in the L population. As the season progresses, the flowering times of the earliest blooming resident

L-plants match those of the later-blooming *E*'s, and thus, the migrant pollen vectored by the latter can succeed. By the time the last *L* resident blooms, *E* is out of flower, and so, the latest resident plants mate only with other *L*'s. The bottom line is that even if pollen is vectored randomly in space on each day of the flowering season, the variation in reproductive timing between and within populations ensures that both the most successful migrant donors and resident recipients are genetically nonrandom subsets of their respective populations.

The nonrandomness of gene flow in this verbal model ameliorates the negative effects of migration on local adaptation in two ways. First, the *E* plants that are most successful at vectoring pollen to the *L*-site are themselves *L*-like and so closer to the *L* optimum than *E* plants generally. Second, the *L*-plants that are most likely to receive *E* pollen are themselves *E*-like and thus of low fitness under the local *L*-site conditions, and so, fewer resident × migrant offspring are produced than under random migration and mating. Alternative scenarios for selection and nonrandom gene flow can facilitate population differentiation (Edelaar & Bolnick, 2012; Bolnick & Otto, 2013). Other combinations of biased migration can lead to skewed distributions of genotypes across space (Haag *et al.*, 2005; Shine *et al.*, 2011). And in some cases, nonrandom gene flow may have asymmetric effects on local adaptation, as when directional selection on the migration trait acts in the same direction in the diverged populations; gene flow in one direction facilitates selection response but opposes it in the other.

Study goals

How strongly does variation in flowering time between and within real populations bias gene flow at phenology loci? Here, I use population- and individual-level flowering schedules to dissect and quantify factors that would lead to nonrandom pollen exchange along a flowering time cline in *Brassica rapa*. A previous paper (Franke *et al.*, 2006) describes the cline, which consists of several populations across a soil moisture gradient along a 4 km stretch of the San Diego Creek drainage in Orange County, California. The median dates of first flowering for this hermaphroditic winter-annual are 10 to 20 days later at the wet-soil sites than at the dry-soil site. Flowering time differences are maintained in common garden experiments (Franke *et al.*, 2006; Franks, 2011). The extra soil moisture affords an extended growing season in this Mediterranean climate, which favours delayed flowering (Weis *et al.*, 2014). The differences in flowering phenology lead to partial temporal reproductive isolation among these populations (Franks & Weis, 2009), but flowering periods overlap enough to permit some gene flow. The discussion expands upon the relevance of these populations for this investigation.

I performed a two-generation line cross experiment that tested for (1) additive and dominance genetic components of population differences for day of first flower and (2) within-population additive genetic variance for day of first flower. In the nonrandom gene flow framework (Edelaar & Bolnick, 2012), the first experimental goal confirmed the genetic basis of clinal variation, whereas the second established within-population genetic variance in a trait that influences migration success. I then dissected the potential for nonrandom gene flow by devising calculations along the lines of those in Kirkpatrick's model. The first is $\bar{m}$, the mean exposure of residents to migrant pollen, which is the realized immigration rate. The second is $\bar{T}_M$, the mean flowering time of the producers of successfully migrating pollen. These indexes are based on the number of opportunities for pollen exchange, as derived from individual- and population-level flowering schedules.

Materials and methods

Crossing design and data collection

I performed line crosses (Lynch & Walsh, 1998) among the Back Bay, San Joaquin Marsh and Michelson populations, which occupy the earliest, middle and latest positions, respectively, of the cline (Franke *et al.*, 2006). The parental generation was planted in the University of California, Irvine, greenhouse with wild-collected seed. The 16 November planting date exposed the plants to a naturally encountered photoperiod, which progressed from ~9 h at emergence to ~11 h at senescence. Fifty-two plants per population were grown in 12-cm-wide and 35-cm-deep pots under ambient light. Pots were watered to saturation at least once daily and fertilized with 20 : 20 : 20 NPK soluble fertilizer applied as per label directions biweekly. Days from emergence to flowering were recorded for each individual. I will use the terms 'days to flowering' and 'flowering time' interchangeably.

Parents were crossed in a replicated factorial design that generated purebred and reciprocal hybrid progenies from the three parental populations (Fig. 1). Two randomly chosen plants from each of the three populations (six parents total) were assigned to each of 20 crossing blocks. Parents in each crossing block occupied adjacent locations on the greenhouse bench. Within a crossing block, one plant from each population was designated as sire and the other as dam. Each sire was mated to all three dams within its block by rubbing excised anthers over the stigmatic surface. Each crossing block thus produced three purebred sibships (diagonal, Fig. 1) and six reciprocal hybrid sibships (off-diagonals, Fig. 1). Time constraints prevented all possible crosses in several of the crossing blocks and a few sire–dam combinations appeared incompatible. In all, 18 of the crossing blocks produced sufficient seed for at

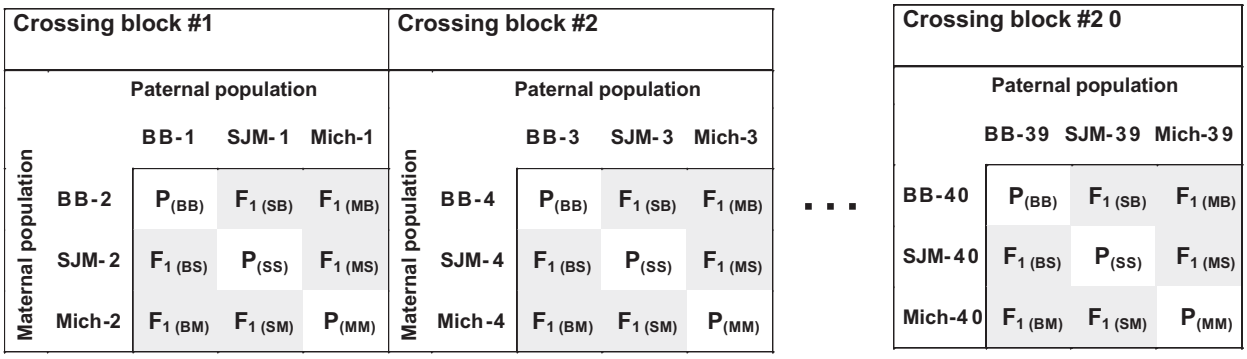

Fig. 1 Design of the line cross. Each crossing block included one sire and one dam from each population, and all possible sire–dam crossings were made. This yielded 3 pure-bred and six reciprocal hybrid sibship combinations per block.

least seven of the nine progeny types to evaluate their F_1 generation. The random assignment of mates avoided assortative mating for flowering time (correlation between mates for days to first flower, $r = 0.06$, *n.s.*), which would have otherwise distorted population-level differences. Early × late crosses were made possible by withholding pollination on the early parent, thereby extending its flowering period until its late partner bloomed. Previous work with these populations showed that maternal age at time of seed maturation has negligible effect on offspring flowering time (Weis & Kossler, 2004).

On the following 24 November, eight offspring per sibship were planted as seed in SuperCell Conetainers (Stewe & Sons, Corvalis, OR, USA). This gave a potential sample size of 1272 plants (8 × 9 × 18, minus 24 for 3 missing sibships). A second planting was made ten days later to replace germination failures. Subsequent mortality was very low and 960 plants were available for analysis. Conditions were otherwise the same as for the parental generation.

Bolting and flowering for each plant was noted in a daily census, and the times to each event are reported here as the number of days past emergence. The number of days to first flower for individuals is denoted as T_i, I also measured stem height, stem diameter, and the length and width of the longest leaf on the day of first flowering.

I recorded the flowering schedules for a subsample of 16 purebred offspring plants from each of the parental populations (14 plants from SJ Marsh due to accidental damage), all taken from the second planting. These were censused daily for first flowering. On every fifth day, starting at 45 days after sowing, the number of open flowers on each plant was counted. Note that the days to first flower, T_i, and the flower counts were measured independently. More than 6000 flowers had been counted by the last census on day 160. All plants except one were well past peak flowering by this day, although several still had a few unopened buds.

Statistical analysis

I applied a mixed model analysis of variance to test for genetic variation between and within the populations for the recorded traits (days to bolting and flowering, stem diameter and height, and the length and width of the longest leaf). Paternal population, maternal population and the 'paternal population × maternal population' interaction were treated as fixed effects because the intent was to draw inferences about these specific populations on this particular cline. Crossing block and the two interactions between crossing block and parental populations were treated as random effects. Planting date was used as a covariate. The analyses used PROC GLM of the SAS package, with type IV sums of squares.

The ANOVA terms are interpreted as follows: a significant 'paternal population' term indicates the populations differed due to genetic effects (Lynch & Walsh, 1998). A significant 'maternal population' effect would arise from either genetic differences among populations, population differences in maternal effects, or both. A genetic dominance component to population differences would lead to a significant 'paternal population × maternal population' term (Lynch & Walsh, 1998).

A significant 'paternal population × crossing block' term indicates genetic variance within populations, based on the following logic. One sire per parental population was randomly assigned to each crossing block. By chance, in some blocks, the three sires from the different populations could be phenotypically more similar to one another than predicted by their respective population means, whereas in other blocks, they could be more dissimilar. If the variation of sires about their population means is genetically based, the phenotypic means of the sibships they produce will likewise deviate from population means in proportion to the trait's heritability. Thus, genetic variation within the three populations should lead to greater variation among paternal sibships in some crossing blocks than others, evidenced

by a significant interaction effect. The corresponding 'maternal population × crossing block' term tests for combined genetic and maternal effect variation with populations. If significant, these two interaction terms, 'paternal population × crossing block' and 'maternal population × crossing block', serve as the error term for the paternal and maternal population effects, respectively.

These interaction terms cannot indicate whether genetic variance occurs within all three populations or only one. I tested for genetic variance within each population with parent–offspring regressions. Only days to flowering was analysed as it was the only trait measured in both parental and offspring generations.

Differences in the population-level flowering schedules were explored by comparing coefficients for nonlinear regressions of the number of open flowers over days since planting. I used the PROC NLIN of SAS to fit the flower census data to the Gaussian function

$$n_d = n_{\max}\left\{\exp\left[\frac{-(d-d_F)^2}{w}\right]\right\},$$

where n_d (dependent variable) is the number of open flowers on census day d (independent variable). Function parameters are as follows: $n_{\max}$ is the number of open flowers at peak flowering, d_F is the day at which the population reaches peak flowering, and w reflects the 'width' of the flowering schedule. I also attempted to apply this function to individual flowering schedules, but in many cases, the algorithm failed to converge onto reliable parameter estimates.

Calculating the potential for nonrandom gene flow

Recalling the verbal model for nonrandom pollen exchange, the realized immigration rate will depend upon the differences in flowering schedules for plants in the resident and the migrant natal populations. Likewise, the phenotypic mean of the successful migrant pollen donors also depends on schedule overlap. I devised several calculations to quantify these effects. To understand what the terms m and $\bar{T}_M$ represent, imagine the ideal way for calculating them. Suppose that investigators were able to identify the source of every pollen grain reaching every resident stigma. With these data, m would be the proportion of grains of migrant origin. If they recorded the flowering time of each migrant grain's producer, then took the average across grains, they would have $\bar{T}_M$. If all arriving pollen grains have an equal chance of success, gene flow would be random. If mating success varies with T, and if T is heritable, gene flow at the contributing loci will be nonrandom. Such exact data are not readily collected, but m and $\bar{T}_M$ can be estimated prospectively from flowering schedules, which reveal the number of opportunities for successful mating by migrants to residents. (Note: in the present context, 'migrant' refers to a plant that exports pollen from its natal population to the resident population; migrants do not move, but their pollen does).

Because they vary in their flowering times, resident individuals will vary in their potential to receive migrant pollen. I will refer to this as a resident's 'exposure', quantified as

$$m_i = \sum_{d=1}^{D} r_{id} q_d$$

where r_{id} is the proportion of all flowers produced by resident i over the entire flowering season that opened on day d. The term q_d is the proportional contribution by immigrants to the pollen pool on day d. The flowering season ends on day D. The mean of this index, $\bar{m}$, reflects the proportion of resident mating opportunities that were with migrant sires and thus reflects the realized immigration rate (assuming no variance in flower number; see below). In the absence of an empirical estimate for the true arrival rate of immigrant pollen, these calculations assumed that integrated over the entire mating season, 10% of all pollen in circulation at the resident site was of migrant origin. I define this proportion as the seasonal contribution, k, to the resident pollen pool. (The calculations can be performed with any desired value for k, as shown below.)

Under complete temporal synchrony (panmixia), m_i for all individuals would be k (in this case, 0.1). When flowering is asynchronous within and between populations, however, the composition of the pollen pool changes over time. Early in the season, the earlier population contributes disproportionately to the pollen pool, since it flowers first. As the season progresses, the relative contribution shifts as flowering wanes in the earlier population and waxes in the later. If so, m_i will differ between early- and late-flowering plants within a resident population. Some migrant pollen arrives when there are no receptive residents, and so, $\bar{m}$ can be less than k.

To quantify the change in exposure with recipient flowering time, I performed a logistic regression of m_i over individual flowering date, using R (R Core Team, 2014). The predicted value at the resident mean flowering date, m_{mean}, and its standard error, was calculated for each of the six regressions with the msm package in R (Jackson, 2011). These were used to test the hypothesis that $m_{\text{mean}} = k$ (in this case, 0.1). Note that the interpretation of m_i can be reversed to denote a resident's potential to contribute pollen to the other population.

Average exposure, $\bar{m}$, accounts for the migration rate due to phenology alone. Fitness differences among *recipient* plants, reflected in the number of flowers produced, will affect their exposure, which in turn affects $\bar{m}$. Specifically, if flower production per plant covaries

with flowering date, residents with a high *proportion* of migrant mating opportunities will not necessarily engage in a large *number* of such matings. I tested for the flowering date–flower production covariance by linear regression. I then recalculated the mean of m_i, weighting each plant by its relative flower production ($R_i/\bar{R}$, where R is flower production by resident plants), giving it the notation $\bar{m}'$. The difference between weighted and unweighted mean exposure reflects the impact of fitness variance among residents on between-population mating.

Finally, I estimated the mean flowering time of migrants, $\bar{T}_M$. Two episodes of selection on migrants can cause $\bar{T}_M$ to differ from the mean of the natal population, $\bar{T}_N$. First, pollen production (flower number) can covary with flowering time, biasing the pollen available to be vectored to the resident site. I use the term $\bar{T}'_N$ to designate the mean flowering time of the migrants after this selection episode. Second, early and late migrants will have different numbers of mating opportunities because they will have different numbers of resident flowers available to them. The mean flowering time of the migrants, $\bar{T}_M$, accounts for both predispersal selection and selection through the variance in the number of mating opportunities.

I calculated $\bar{T}'_N$, as the mean of the migrants, weighted by their relative flower production ($M_i/\bar{M}$, where M is flower production by migrant plants). I then calculated $\bar{T}_M$ as another weighted mean, with the weighting for each migrant accounting for its pollen production (number of flowers) and the availability of recipient flowers. Specifically,

$$\omega_i = \left(\sum_{d=1}^{D} M_{id}\, r_d\right)\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{d=1}^{D} M_{id}\, r_d\right)^{-1}$$

where, M_{id} is the number of flowers on migrant i open on day d, and r_d is the proportion of all resident flowers across the season that were available for pollination on day d. Note that these two weightings are the relative fitness of migrant plants through the two selection episodes.

Directly testing the null hypotheses that $\bar{T}'_N = \bar{T}_N$ and $\bar{T}_M = \bar{T}_N$ is problematic because they both are derived from the observed values of individual days to flowering, T_i. Recall that T_i was measured independently from the flower counts, which were used to calculate relative fitness. This allowed me to test, through regression, the null hypothesis that pollen production and mating success are independent of T_i. T_i was z-transformed for testing. The selection differential, S, which measures the change in mean in days, was then calculated as the product of the regression slope and the population standard deviation in flowering date. In keeping with Kirkpatrick's model, I then calculated the difference between the mean flowering date of the resident population, $\bar{T}_R$, and the migrant means before and after the selection episodes, that is $\bar{T}_N$, $\bar{T}'_N$ and $\bar{T}_M$. This quantified the potential for migration to change the pollen pool at the resident site during successive stages of the flowering season.

Results

Genetic basis for variation among populations

The line cross experiment confirmed an additive genetic basis for the differences in bolting and flowering times among the San Diego Creek clinal populations. Significant 'paternal population' terms were found for both traits (Table 1), which are highly correlated. Tukey's test suggested that offspring sired by Back Bay plants flowered earlier than those sired by the other two populations, but the other two populations did not differ from one another. The 'maternal population' term was likewise significant, although weaker than the paternal term. The nonsignificant 'maternal × paternal' interaction indicates a lack of directional dominance (Table 1); Fig. 2 shows that the 95% confidence intervals for all hybrid combinations include the mid-point of their respective parents.

Stem height and diameter showed significant differences due to 'paternal population', with Back Bay being smaller that the others (Table S1 & S2). 'Maternal population' effects showed the same pattern for these traits. Leaf dimensions do not differ among populations.

Genetic variance for flowering time within populations

Parent–offspring regression demonstrated genetic variance for flowering time within all three populations (Table 2). Figure 3a–c shows the regressions of the

Table 1 Analysis of variance for phenological traits of offspring produced by a line cross among the Back Bay, San Joaquin Marsh and Michelson populations. Table entries are F-ratios based on type IV sums of squares. Significant 'paternal population' effects demonstrate additive genetic differences among populations, whereas significant 'maternal population' terms indicate additive genetic, maternal effect or both. Population differences due to the dominance component of genetic variance would be reflected in the 'paternal × maternal' term.

		F - ratio	
Source	d.f.	Days to bolting	Days to flowering
Crossing block	17	0.92	1.08
Paternal population	2	12.71***	21.80***
Paternal pop. × Cross. block	34	4.35***	5.09***
Maternal population	2	8.64***	9.31***
Maternal pop. × Cross. block	34	1.94**	2.71***
Maternal × Paternal	4	0.26	0.85

***$P < 0.001$, **$P < 0.01$.

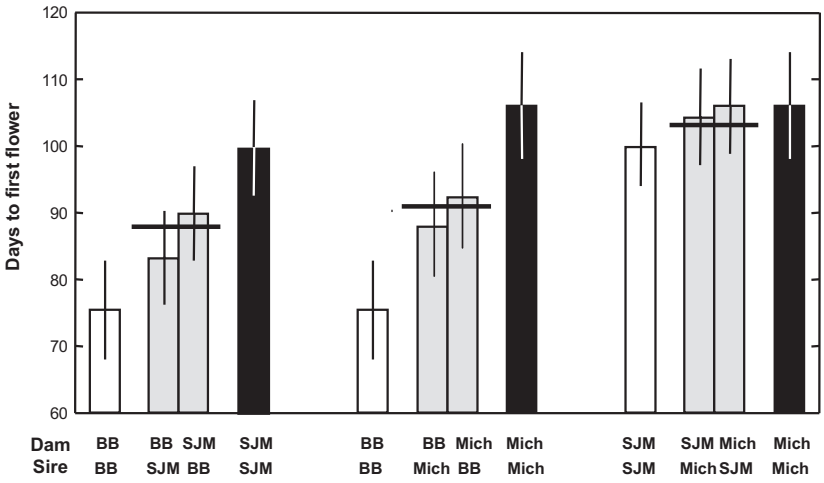

Fig. 2 Means flowering times (days from seedling emergence to opening of first flower). Thick horizontal lines thought hybrid bars indicated the expected hybrid mean in the absence of dominance. Error bars = 95% confidence intervals.

Table 2 Analysis of covariance *F*-ratios for parent–offspring regressions. Slopes are proportional to heritability. (A) Regression of offspring mean flowering time, by dam, over sire flowering time. Each sire was mated to one dam from each of the three populations. (B) Regression of offspring mean flowering time, by sire, over dam flowering time. Each dam was mated to one sire from each of the three populations. Degrees of freedom are 1 for paternal and maternal flowering times, 2 for paternal and maternal population. All interactions were nonsignificant and dropped from the final model.

Paternal population	Paternal days to flowering	Maternal population	Offspring–sire slope (SE)	Residual d.f.
(A) Offspring–sire regression				
Back Bay	21.43***	6.67**	0.5665 (0.1223)	38
San Joaquin Marsh	12.74**	6.34**	0.2817 (0.0831)	45
Michelson	9.09**	8.40**	0.2558 (0.0849)	43
Maternal population	**Maternal days to flowering**	**Paternal population**	**Offspring–dam slope (SE)**	**Residual d.f.**
(B) Offspring–dam regression				
Back Bay	7.45**	6.47**	0.4829 (0.1647)	44
San Joaquin Marsh	19.94***	11.19***	0.3633 (0.0816	47
Michelson	0.50 n.s.	12.92***	0.0285 (0.0847)	44

***$P < 0.001$, **$P < 0.01$.

means for the three half-sibships produced by each sire (columns in Fig. 1) over the sire's flowering time. ANCOVA results indicate that offspring flowering time is explained in part by the sire's flowering time. Additional variance in offspring means is explained by the 'maternal population' effect (Table 2). None of the 'maternal population × sire's flowering time' interactions were significant, so these terms were dropped from the final model. The slope for the offspring–sire regression was highest in the Back Bay population and lowest in the Michelson population. The corresponding dam–offspring regressions are illustrated in Fig. 3d–f. Differences in slope between the sire–offspring and dam–offspring regressions can indicate maternal effects on flowering time. Slopes did not differ between dam–offspring and sire–offspring regression for the Back Bay and SJ Marsh populations. For the Michelson population, there was no detectable influence of maternal flowering time on offspring flowering time (Fig. 3f).

Differences in population-level flowering schedules

Gaussian functions were fitted to the flowering schedules for the three study populations. As with the mean number of days to flowering, the Back Bay population was earliest, reaching peak flowering ~88 days after sowing, followed by SJ Marsh and Michelson at ~124 and ~136 days, respectively (Fig. 4 and Table S3). The Back Bay flowering schedule was also considerably narrower than the other two (Fig. 4 and Table S3). These flowering schedules show that the Back Bay had virtually ceased pollen production when the two later populations were at their peak of stigma availability. Reciprocally, a substantial number of flowers from SJ Marsh and Michelson were producing pollen during

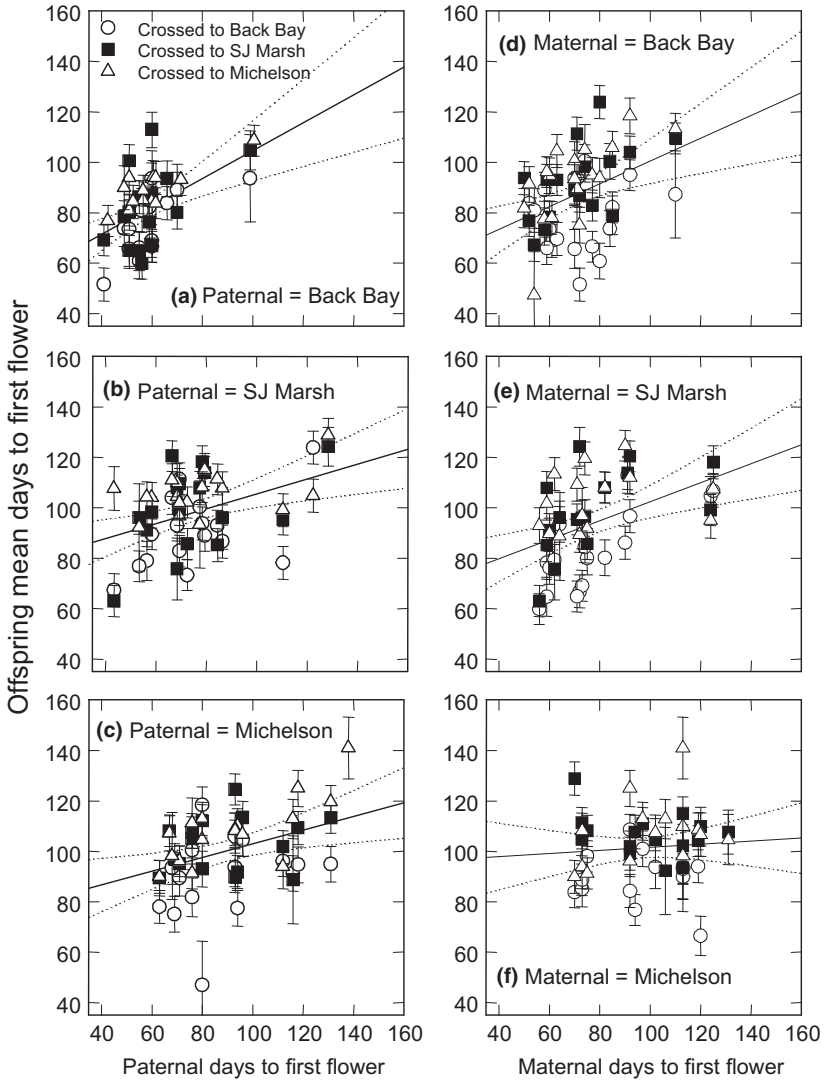

Fig. 3 Offspring–Parent regressions for line cross experiment. (a–c) Regression of full sibship mean (± standard error) over paternal days to first flowering. (d–f) Regression of full sibship mean (± standard error) over maternal days to first flowering. For statistics, see Table 2.

the time when Back Bay stigmas were at peak availability (Fig. 4).

Individual flowering time and the potential to receive migrant pollen

With the observed asynchronous flowering, only ~4% of the flowers on the earliest plant from the Back Bay would be exposed to SJ Marsh pollen, compared to ~40% of the flowers on the latest plant (Fig. 5a, Table S4). Reciprocally, the earliest SJ Marsh plant would have ~35% of its flowers exposed to Back Bay pollen, but this falls to ~2% for the latest plant. Figures obtained for the Back Bay–Michelson pairing were similar in value and symmetry (Fig. 5b, Table S4). Under synchronous flowering, plants would have 10% of their flowers exposed. In all cases involving the early-flowering Back Bay population, exposure at the mean flowering date, m_{mean}, was significantly less than k (i.e. less than 0.1; Table S4).

In contrast, individual variation in exposure to immigrant pollen was much lower for exchanges between the SJ Marsh and Michelson populations (Fig. 5c, Table S4). Predicted exposure at the mean flowering time, m_{mean}, did not differ significantly from 0.1 for either case. This makes sense in the light of their broadly overlapping flowering schedules (Fig. 4). As a reminder, Fig. 5 also reflects individual potential for pollen export to the specified population.

Flowering asynchrony not only induces individual variation in exposure to immigrant pollen, it can cause mean exposure, $\bar{m}$, to differ from panmictic expectations. Covariance between flowering time and total flower production (a fitness component) may further alter mean exposure. Total flower production was negatively associated with flowering time (Fig. S1). As a

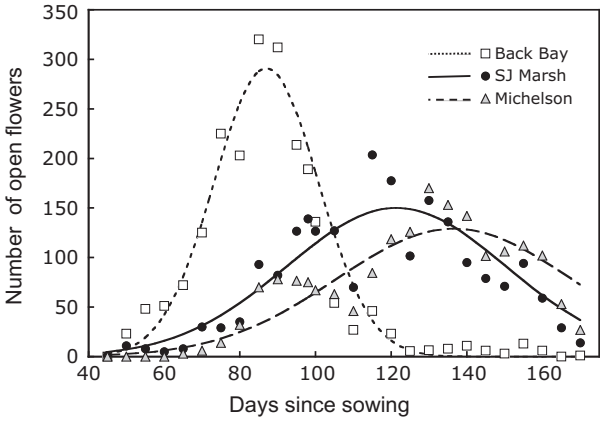

Fig. 4 Population-level flowering schedules. The *x*-axis depicts days after planting. Regression parameters and statistical tests are given in Table S3.

consequence, individuals with a high *proportion* of their flowers exposed to immigrant pollen will not necessarily have a large *number* of exposed flowers.

Figure 6 illustrates the mean exposure, $\bar{m}$, and fitness-weighted mean exposure, $\bar{m}'$, for the three populations. Again, these numbers assume that the seasonal proportional contribution of migrants to the resident pollen pool is $k = 0.1$, but that the proportion of migrant pollen in the pool on any given day depends on the flowering intensity of migrants and residents. The nonindependence of $\bar{m}$ and $\bar{m}'$ clouds their statistical comparison, but their numerical values (Fig. 6) suggest how selection on flowering time in the resident population (Fig. S1) affects exposure. Mean exposure is lowest between the Back Bay and Michelson populations (Fig. 6), as expected from the large divergence in flowering schedules (Figs 4 and 5b). The decline in flower number with later-flowering date is very weak for Back Bay (Fig. S1), and so, there is small difference between $\bar{m}$ and $\bar{m}'$ (Fig. 6, black and grey circles). Flower production falls with flowering date more strongly for Michelson (Fig. S1), and as a result, the resident plants with the greatest proportion of flowers exposed also have the greatest number exposed, inflating $\bar{m}'$ into (Fig. 6, white and grey squares). The greater similarity of flowering schedules for the Michelson and SJ Marsh populations led to mean exposure rates, $\bar{m}$, slightly above and below panmictic expectation for this pairing (Fig. 6, grey square, black triangle). After accounting for variation in recipient flower production, departures from random expectation are even less. The SJ Marsh plants most likely to receive pollen from Michelson population (Fig. 5c) were the ones with the fewest flowers (Fig, S1), reducing $\bar{m}'$ relative to $\bar{m}$. Reciprocally, the most likely Michelson recipients produced more flowers, increasing $\bar{m}'$ relative to $\bar{m}$.

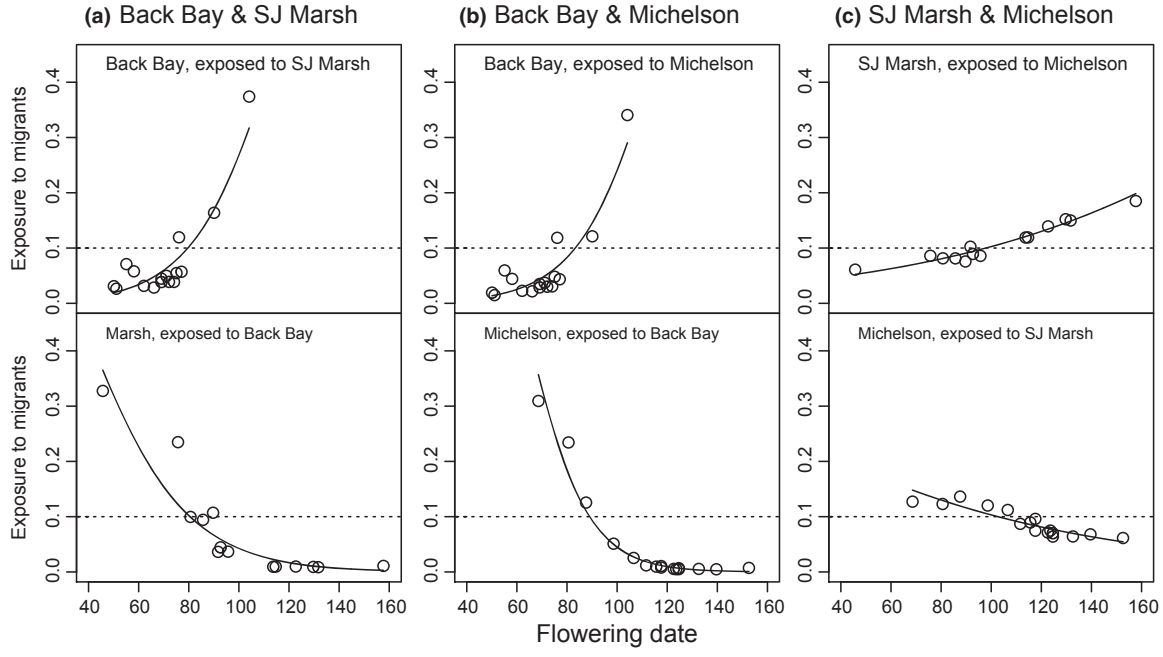

Fig. 5 Exposure to immigrant pollen, m_i, as a logistic function of individual days to flowering. (a) Exchange between Back Bay and SJ Marsh. (b) Exchange between Back Bay and Michelson. (c) Exchange between SJ Marsh and Michelson. The dashed horizontal line represents expected m_i if flowering was completely synchronous for all plants (i.e. panmixia). Regression parameters and statistical tests are given in Table S4.

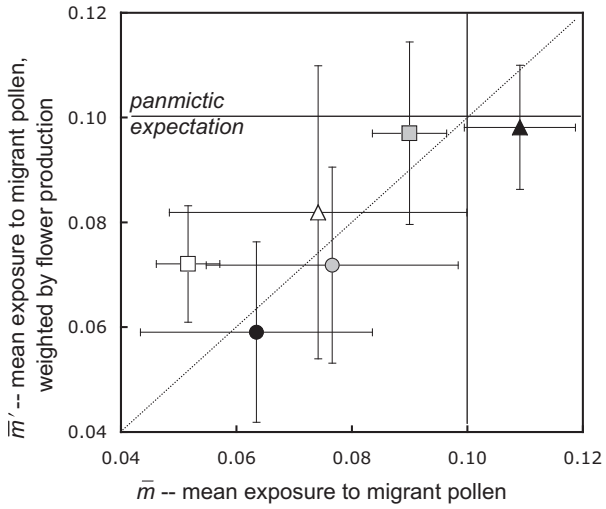

Fig. 6 Mean exposure of resident plants to migrant pollen, unweighted ($\bar{m}$) and weighted ($\bar{m}'$) by flower production. Symbol shape denotes which population acts as the resident, with circles for Back Bay, triangles of SJ Marsh and squares for Michelson. The colour of the symbol denotes the source of the immigrant pollen, with white, grey and black denoting Back Bay, SJ Marsh and Michelson, respectively. The horizontal and vertical lines mark the mean exposure expected under panmixia, that is when all plants flower synchronously. If $\bar{m}$ and $\bar{m}'$ are equal, they fall on the diagonal dashed line. Bars represent ±1 standard error.

Mean flowering time of migrants

Fitness differences due to variation in flower number, and in mating success, introduced a potential bias in migrant contributions to the resident mating pools. In the SJ Marsh and Michelson populations, there was a significant shift in mean flowering date through the pollen production selection episode (Table 3A), such that $\bar{T}'_N$ was 8.3 and 7.3 days earlier than $\bar{T}_N$, respectively. The Back Bay mean shifted to 2 days earlier, but the selection gradient analysis for this was not significant. When both the pollen production and mating success selection episodes are combined, the difference of $\bar{T}_M$ from $\bar{T}_N$ was significant in all cases, as tested by the selection gradient analysis (Table 3B). The mean migrant from Back Bay to the two later-flowering populations was a little over two days *later* than the natal population mean. For pollen moving in the opposite direction, the mean migrants from the SJ Marsh and Michelson populations were 23 and 33 days *earlier* than their natal means (Table 3A). In these cases, selection gradients for the combined selection episodes were significant. When considering migration from Michelson to SJ Marsh, a 7-day shift in the early direction was significant, whereas the 2-day shift for migration in the opposite direction was not (Table 3B). The direction of the shift in mean, in each case, indicated that successful migrants were more resident-like in flowering time than expected under panmixia.

The absolute values of $\bar{T}_M - \bar{T}_N$ for migration from Back Bay to the two later-flowering populations were an order of magnitude less than for the reverse migration into Back Bay. Two factors account for this. First, selection through pollen production and mating success act discordantly for Back Bay migrants, but concordantly on SJ Marsh and Michelson migrants migrating into Back Bay. The second factor emerges from the narrower range in flowering dates among the Back Bay plants. The spread in flowering dates for the earliest plant in the three populations was ~ 20 days (see *x*-axes in Fig. 5), whereas the spread in mean flowering dates was ~44 days (Table 3A). In other words, the compressed flowering schedule for Back Bay was partially nested within the lengthier schedules of the other two populations (Fig. 4). Given that individual plants' flowers for ~50 days, nearly all Back Bay plants had an opportunity to donate pollen to at least a few of the early SJ Marsh and Michelson recipients. In contrast, many SJ Marsh and Michelson plants started to flower after Back Bay had virtually finished (Fig. 4). Thus, successful migrants *into* Back Bay represent more strongly biased subsets of the plants in these two later-flowering populations.

Recalling Kirkpatrick's formulation (Kirkpatrick, 2000), the impact of gene flow on flowering time depends in part on the difference between the mean resident and the mean migrant, *that is* $\bar{T}_M - \bar{T}_R$. Figure S2 shows these differences. The natal means for SJ Marsh and Michelson were ~ 33 and 44 days later than the Back Bay resident population, but the mean migrants for these populations were only 11 and 13 days later. Thus, they were substantially more resident-like than random, thus reducing the potential consequences of gene flow in this direction. In the reverse direction, however, the 33 and 44 day differences were reduced only to 31 and 42 days for Back Bay migrants into the SJ Marsh and Michelson populations, respectively. This indicates that at a fixed value of *k*, gene flow from the Back Bay population to the later ones would change their mating pools more strongly than in the reverse; weaker selection through the two episodes makes gene flow in this direction closer to random.

How large is the bias in gene flow?

Having dissected out the elements of nonrandom gene flow at flowering time loci – realized migration rate and realize migrant mean phenotype – this section puts them back together to assess the potential magnitude of bias. Specifically, I estimate the change in composition of the resident mating pool caused by migration. Figure 7a illustrates the potential change in mean flowering time of sires from the SJ Marsh site when they export pollen to Back Bay and Michelson. Figure 7b illustrates the consequences for gene flow in the

Table 3 Mean flowering time of potential migrants. (A) Mean number of days from planting to flowering for the migrant's natal population before selection ($\bar{T}_N$), the natal population mean weighted for differences in flower production ($\bar{T}'_N$) and the mean of migrants ($\bar{T}_M$). (B) Univariate selection gradients, *b*, and selection differentials, *S*, for selection through flower number (premigration selection) and selection through the combined effects of flower number and mating opportunities.

			($\bar{T}_M$) (SE), by resident population		
Migrant population	($\bar{T}_N$) (SE)	($\bar{T}'_N$) (SE)	BB	SJM	Mich
(A) Migrant mean flowering date					
Back Bay	66.00 (3.71)	64.49 (3.56)	–	68.43 (3.91)	68.22 (3.85)
SJ Marsh	98.57 (7.45)	90.87 (24.18)	77.20 (4.92)	–	96.00 (5.18)
Michelson	110.06 (5.33)	103.10 (5.22)	79.09 (3.47)	103.08 (4.84)	–

	Selection episode			
	Flower number	Flower number plus mating opportunities, by resident population		
Migrant population		BB	SJM	Mich
(B) Univariate selection gradients and selection differentials on migrant flowering date				
Back Bay				
b (SE)	−0.116 (0.081)	–	0.0187 (0.068)*	0.171 (0.066)*
S	−1.61	–	2.59	2.37
SJ Marsh				
b (SE)	−0.298 (0.112)*	−0.826 (0.182)***	–	−0.099 (0.142)
S	−8.27	−23.01	–	−2.77
Michelson				
b (SE)	−0.349 (0.113)**	−1.155 (0.239)***	−0.349 (0.129)*	–
S	−7.33	−33.02	−7.43	–

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

opposite direction, that is sires from the Back Bay and Michelson sites when they export to SJ Marsh.

Calculations are made for three scenarios: (1) the fully random case, where $m = k$, and $\bar{T}_M = \bar{T}_N$; (2) the case where selection on flowering time alone, such that $m = k$ and $\bar{T}_M = \bar{T}'_N$; and (3) the fully nonrandom case, where $m = \bar{m}$, and T_M is the migrant mean weighted for fitness differences and asynchronous flowering. I assume that both migrant and resident populations produce the same number of flowers over the season. For the general calculation given above, I set k to 0.1. Here, I vary it from 0.05 to 0.5. Importantly, k will vary with the distance between populations – the nearer in space, the higher will be the seasonal contribution of migrants to the resident pollen pool. As a final note, variation in mating phenology causes assortative mating by necessity, which in turn causes some degree of nonrandom gene flow (Bolnick and Kirkpatrick 2012). Thus, the 'fully random' case presented here in reality pertains to a hypothetical trait with the same mean and variance as flowering time, but which has no effect on the propensity to migrate, or on mating success at the resident site. I nonetheless call the focal trait 'flowering time' in all cases to ease the flow of argument.

Migration out of SJ Marsh would shift the mean sire at the Back Bay site to a later-flowering date (Fig. 7a). If there were no spatial isolation between sites ($k = 0.5$), random migration shifts the mean sire by 16.2 days, compared to no migration. Selection through differential flower production before migration reduces this to 12.4 days, making the mean producer of migrant pollen more resident-like. When differential mating success, which acts concordantly with flower production, is factored in, the shift falls to 3.6 days. The nonrandomness of migration thus reduces its impact on the resident pollen pool by 80% in this case. With greater spatial isolation (lower k), impacts are reduced more or less proportionately, but the nonrandomness of migration always reduces its impact. Interestingly, the impact of SJ Marsh migrants on the pollen pool at Michelson is not different from the random case. Here, selection due to flower production would increase the impact, as the mean pollen producer is later than the mean plant (Table 3A). However, the discordant selection through mating opportunity changes both the realized values for the mean sire and migration rate, cancelling this difference.

Consider now pollen migrating in the reverse directions, that is from Back Bay and Michelson into SJ Marsh. Impacts are of opposite sign but similar, yet different, magnitudes. When migration goes out of Back Bay into SJ Marsh, selection during the flower production episode changes the mean migrant from 16.2 to 17 days. Discordant selection through mating opportunity brings the impact down to 8 days, a 50% reduction. Finally, migrants from Michelson into SJ Marsh

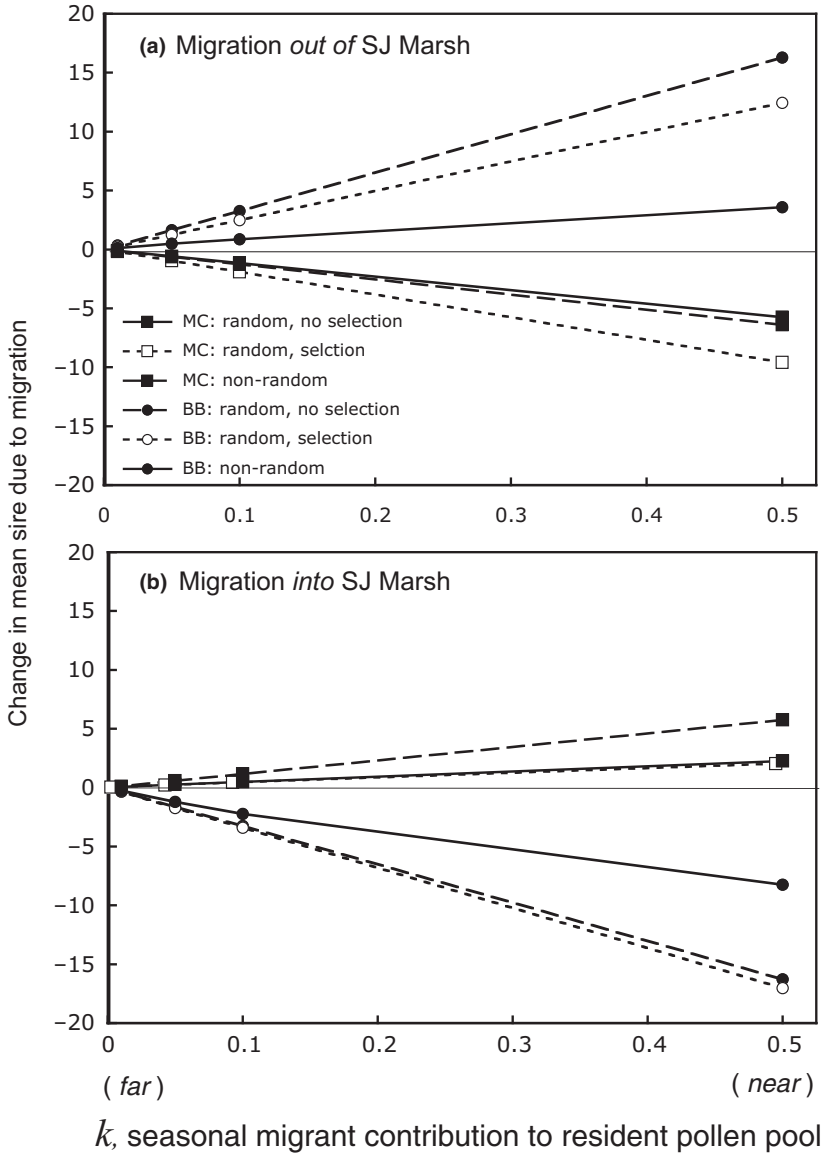

Fig. 7 Potential change in the composition of the pollen pool at resident sites due to gene flow. (a) The change at the Back Bay and Michelson sites due to migration out of SJ Marsh. (b) The change at SJ Marsh due to immigrant pollen arriving from Back Bay and Michelson. The parameter k is the proportion of migrant pollen in the pool at the resident site, integrated across the entire flowering season; k is expected to increase with decreasing distance between migrant and resident sites.

shift the mean sire to an earlier flowering date, but this shift is explained by selection through flower production alone because realized migration rate, $\bar{m}$, is virtually equal to k.

Discussion

Gene flow is an important force in plant evolution, keeping species cohesive evolutionary units (Ellstrand, 2014) and spreading globally adaptive genes from their population of origin across the species range. With respect to local adaptation, however, theoretical considerations of gene flow (migration) have largely focused on its negative effect (e.g. Slatkin, 1975; Antonovics, 1976; Lenormand, 2002; Lopez *et al.*, 2008): the rate at which maladaptive alleles flow into a population can exceed the rate at which selection eliminates them, thereby depressing mean fitness. Recently, an appreciation of the more diverse effects that gene flow has on population differentiation has emerged (Edelaar & Bolnick, 2012). When migration success depends on genotype, 'migrational load' can be relaxed or even reversed (Bolnick & Otto, 2013). Other traits that show a propensity for assortative mating, such as body size in animals, may offer additional examples of biased gene flow. Variation in dispersal ability *per se* leads to a form of biased gene flow resulting in the spatial sorting of genotypes (Haag *et al.*, 2005; Shine *et al.*, 2011).

This study establishes that the conditions for nonrandom gene flow, with respect to flowering time loci,

exist among a set of locally adapted populations. Flowering time is genetically variable between and within populations. The resulting asynchrony of mating opportunities generates variation in mating success, selecting for migrants that are more resident-like in flowering date. Thus, genetic contributions from the migrant's natal population to the resident are less than for comparable traits that do not influence migration success.

As a caution, some late-blooming plants still had a few unopened buds when the experiment was terminated; the potential for late × late mating was probably underestimated. However, Weis & Kossler (2004) found that the last flowers produced by a *B. rapa* plant seldom set seed, which would counterbalance this bias. Truncation would also inflate the negative covariance between flowering time and flower production. Regardless, the calculations of m_i and $\bar{T}_M$ presented here are an essential first step in dissecting the sources of nonrandom gene flow for flowering time – a trait that frequently diverges among populations (Mazer and LeBuhn, 1999).

Potential impact of nonrandom gene flow along the San Diego Creek cline

Direct pollen exchange among the San Diego Creek populations is probably rare. The SJ Marsh site is located ~2.6 km inland from the Back Bay, and Michelson a further 1.6 km inland from that. However, small *B. rapa* patches are scattered between these sites, such that gene flow via pollen could occur in stepping stone fashion. Regardless of how much spatial isolation restricts pollen movement, these populations provide a concrete scenario for thinking through the potential strength and consequences of nonrandom gene flow.

These populations appear to have diverged some time during the mid-20th century, after several of the creek's tributaries were impounded to form an artificial freshwater marsh for waterfowl conservation (the SJ March and Michelson areas). The Back Bay area is unaffected by this disturbance and so occupies a site with a lower water table. The annual southern California drought starts here in mid-spring. Further inland, the impoundments capture run-off, raise the water table and keep soil moist much later into the spring (Franke *et al.*, 2006). This lengthens the growing season, thus selecting for later flowering at the moister sites (Franks *et al.*, 2007; Weis *et al.*, 2014). This is as predicted from life-history theory (King & Roughgarden, 1982; Amir & Cohen, 1990; Kozlowski, 1992), which basically posits that optimal flowering time reflects a trade-off between time allotted to accumulating resources for reproduction and time allotted to converting those resources into offspring. The longer the growing season, the later the optimal date for switching from accumulation to conversion. The soil moisture gradient thus generated divergent selection along this cline and may over the long run generate stabilizing selection within local sites (Weis *et al.*, 2014). If so, these populations would conform to the verbal model presented earlier. Populations at the extreme ends, Back Bay and Michelson, would receive genes only from populations with later- and earlier-than-optimal flowering times, respectively. However, mating asynchrony reduces both the proportion of migrant pollen grains that are successful and shifts their mean genetic contribution to be more resident-like, compared to panmictic expectations. This would relax migration load.

A slightly different scenario would have played out in these populations between 1997 and 2004, when several years of reduced winter precipitation shortened the growing seasons at all sites. In a resurrection experiment, plants grown from seed derived from the post-drought generation (descendants) flowered earlier than those from predrought seed (ancestors), demonstrating an evolutionary shift (Franks *et al.*, 2007). Fitness functions constructed from field data during this prolonged drought (Franke *et al.*, 2006) showed selection favouring earlier flowering at both Back Bay and the Arboretum sites (adjacent to the SJ Marsh site). However, the evolutionary change towards early flowering was greater at the moist end of the gradient than the dry (8.5 vs. 1.9 days, Franks *et al.*, 2007). Note that the selection regime imposed by greenhouse conditions is similar to that imposed by reduced precipitation. This affords an opportunity to speculate on how the experimental results could apply in a real-world situation.

Consider SJ Marsh pollen vectored into the Back Bay. The arriving pollen would come from later-flowering plants (Table 3) and so carry *later*-flowering genes into a population where *early* flowering is favoured. This would contribute to migrational load. But, that pollen disproportionately fertilizes later-flowering resident plants, which have low fecundity (Figs S1 and 5a). This reduces load, but does not eliminate it (Fig. 7a). In contrast, pollen vectored from Back Bay to SJ Marsh would carry early-flowering genes into a population where early flowering is favoured. Rather than putting a drag on selection response, gene flow would enhance it. Yet, the positive impact is still less than under panmixia (Fig. 7b). Migrants disproportionately mate with the more fecund residents (Figs S1 and 5a), but much of the pollen arrives too early fertilize many of the resident plants. The positive effect of gene flow is reduced.

Future challenges

The estimates I have presented for immigration rate, $\bar{m}$, and the phenotypic difference between residents and migrants, $\bar{T}_R - \bar{T}_M$, are *prospective* in nature, *that is* they are based on the opportunities for pollen exchange. This approach is essential for dissecting the factors that create the conditions for nonrandom gene flow. However, seasonal shifts in the abundance and diversity of pollinators, or their behavioural responses to the wax

and wane of floral abundance could cause realized gene flow to deviate from the prospective estimate. *Retrospective* analyses, constructed from a genetic analysis on two successive generations, are required to see how much of the potential is realized. Substantial practical challenges face this course of research. This final section sketches a way forward.

Retrospective estimates of $\bar{m}$ and $\bar{T}_M - \bar{T}_R$ are theoretically possible through genetic paternity analysis of dams, a sample of their offspring, and *all* potential sires. But of course, most plant populations are too large to make this approach practical, although the size and spatial distribution of some tropical tree species (e.g. Loveless *et al.*, 1998) may approach feasibility. Realistically, more could be learned about the effects of nonrandom gene flow on local adaptation in flowering time using small experimental populations wherein the position, phenotype and the multilocus marker genotype of each plant are known. Migration rates could be manipulated from near zero to very high levels by varying the spatial distances among paired populations. A Bayesian full probability model that includes marker genotypes, parental phenotypes and distance (Austen, 2014; Hadfield *et al.*, 2006; see also Morgan & Conner, 2001) could estimate population parameters $\bar{T}_M$ and $\bar{m}$ and perhaps m_i for each maternal plant.

A complementary approach for estimating paternal phenotype uses the following relationship. If flowering time phenotype is expressed as deviation from the raw mean of the *resident* population, the expected flowering time for seed offspring of a given resident plant is

$$\mathrm{E}\left[T^*_{\mathrm{R}i}\right] = (1 - m_i)\,\frac{h^2(T_{\mathrm{R}i} + \rho T_{\mathrm{R}i})}{2} + m_i\,\frac{C(T_{\mathrm{R}i} + \tilde{T}_{\mathrm{M}i})}{2}$$

where $T_{\mathrm{R}i}$ is the i-th resident's phenotype, and $\tilde{T}_{\mathrm{M}i}$ is the mean phenotype of the migrants donating pollen to the i-th resident. To understand this relationship, consider how it simplifies under zero migration and random mating. Without migration, the seed offspring of a given resident plant will have an expected flowering time that depends only upon the dam's flowering time, the expected sire's flowering time and the flowering time heritability. With random mating, the expected paternal flowering time is that same as the transformed population mean (i.e. 0). Under these conditions, the relationship reduces to the regression of offspring phenotype onto dam phenotype ($\mathrm{E}[T^*_{\mathrm{R}i}] = {}^1/_2\, h^2 T_{\mathrm{R}i}$). If we bring in assortative mating, the sire's expected flowering time now deviates from the mean by the amount $\rho T_{\mathrm{R}i}$, where ρ is the phenotypic correlation between mates (Lynch & Walsh, 1998). If h^2 is known, ρ can be calculated from the inflation of dam–offspring regression relative to panmictic expectations (i.e. ${}^1/_2 h^2$; Weis & Kossler, 2004). The mean phenotype of resident sires contributing pollen to the plant is $\rho T_{\mathrm{R}i}$.

The second term on the right hand side of the equation is equivalent to the first, but concerns resident–migrant matings. Heritability is replaced by the term C, the regression of offspring phenotype over the mid-parent value when the parents come from different breeding populations. When populations differ exclusively due to genes of additive effect, the average of the two heritabilities may suffice as an estimate of C. If population differences involve dominance and epistasis, C could differ substantially from the average heritability. When C is known, $\tilde{T}_{\mathrm{M}i}$ can be estimated from offspring and maternal phenotypes.

Although h^2 and C can be estimated in an open pollination experiment, assuming no paternity assignment error (see Hoffman & Amos, 2005), they can be more reliably estimated from controlled crosses. In the field experiment described above, a set of resident plants would be reserved for a controlled pollination scheme similar to the line crosses used in the greenhouse study. These parameter estimates could then be incorporated into an analytical framework that yields a credible measure of the genetically based bias in migration success.

As Edelaar & Bolnick (2012) note, 'The study of nonrandom gene flow and dispersal is still in its infancy' and much empirical and theoretical work is needed to evaluate its general importance. This includes understanding the variety of mechanisms that lead to genotype-dependent dispersal success. I have presented a study taking the first steps in dissecting the factors that potentially contribute to nonrandom gene flow at loci controlling flowering time, a key plant life-history trait that frequently differs genetically among local populations. Difficult work lies ahead in determining how factors such as spatial barriers, pollinator behaviour and correlated plant traits can cause realized gene flow to vary from its potential.

Acknowledgments

I thank M Cuevas, A Calusen, D Franke, M Freshwater and M Vega for assistance in the greenhouse; E Austen, J Ison and S Wadgymar provided valuable comments on the manuscript; JM Weis helped with proofreading, although any lingering mistakes are my own. Writing space was graciously allotted by Covernotes. Grants from the National Science Foundation (DEB98-15873 and DEB03-45030) and the National Science and Engineering Research Council provided financial support.

References

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

Amir, S. & Cohen, D. 1990. Optimal reproductive efforts and the timing of reproduction of annual plants in randomly varying environments. *J. Theor. Biol.* **147**: 17–42.

Antonovics, J. 1976. The nature of limits to natural selection. *Ann. Mo. Bot. Gard.* **62**: 224–247.

Antonovics, J. & Bradshaw, A.D. 1970. Evolution in closely adjacent plant populations part 8: clinal patterns at a mine boundary. *Heredity* **25**: 349–362.

Austen, E.J. 2014. *The Nature of Selection on Flowering Time: Integrating Fitness Contributions through Male and Female Function*. PhD diss. University of Toronto, Toronto, Canada.

Blanquart, F., Gandon, S. & Nuismer, S.L. 2012. The effects of migration and drift on local adaptation to a heterogeneous environment. *J. Evol. Biol.* **25**: 1351–1363.

Bolnick, D.I. & Kirkpatrick, M. 2012. The relationship between intraspecific assortative mating and reproductive isolation between divergent populations. *Curr. Zool.* **58**: 484–492.

Bolnick, D.I. & Nosil, P. 2007. Natural selection in populations subject to a migration load. *Evolution* **61**: 2229–2243.

Bolnick, D.I. & Otto, S.P. 2013. The magnitude of local adaptation under genotype-dependent dispersal. *Ecol. Evol.* **3**: 4722–4735.

Clausen, J. & Heisey, W.M. 1958. *Experimental Studies on the Nature of Species. IV. Genetic Structure of Ecological Races*. Carnegie Institution of Washington, Washington, DC.

Colautti, R.I., Eckert, G.G. & Barrett, S.C.H. 2010. Evolutionary constraints on adaptive evolution during range expansion in an invasive plant. *Proc. R. Soc. B Biol. Sci.* **277**: 1799–1806.

Edelaar, P. & Bolnick, D.I. 2012. Non-random gene flow: an underappreciated force in evolution and ecology. *Trends Ecol. Evol.* **27**: 659–665.

Ellis, A.G., Weis, A.E. & Gaut, B.S. 2006. Evolutionary radiation of "stone plants" in the genus *Argyroderma* (Aizoaceae): unraveling the effects of landscape, habitat, and flowering. *Evolution* **60**: 39–55.

Ellstrand, N.C. 2014. Is gene flow the most important evolutionary force in plants? *Am. J. Bot.* **101**: 737–753.

Endler, J.A. 1977. *Geographic Variation, Speciation, and Clines Princeton Monographs in Population Biology*, Vol. 10. Princeton University Press, Princeton, NJ.

Felsenstein, J. 1981. Continuous-genotype models and assortative mating. *Theor. Popul. Biol.* **19**: 341–357.

Fox, G.A. 2003. Assortative mating and plant phenology: evolutionary and practical consequences. *Evol. Ecol. Res.* **5**: 1–18.

Franke, D.M., Ellis, A.G., Dharjwa, M., Freshwater, M., Fujikawa, M., Padron, A. *et al.* 2006. A steep cline in flowering time for *Brassica rapa* in southern California: population-level variation in the field and the greenhouse. *Int. J. Plant Sci.* **167**: 83–92.

Franks, S.J. 2011. Plasticity and evolution in drought avoidance and escape in the annual plant *Brassica rapa*. *New Phytol.* **190**: 249–257.

Franks, S.J. & Weis, A.E. 2009. Climate change alters reproductive isolation and potential gene flow in an annual plant. *Evol. Appl.* **2**: 481–488.

Franks, S.J., Sim, S. & Weis, A.E. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proc. Natl. Acad. Sci. USA* **104**: 1278–1282.

Guillaume, F. 2011. Migration–induced phenotypic divergence: the migration–selection balance of correlated traits. *Evolution* **65**: 1723–1738.

Griffith, T.M. & Watson, M.A. 2006. Is evolution necessary for range expansion? Manipulating reproductive timing of a weedy annual transplanted beyond its range. *Am. Nat.* **167**: 153–164.

Haag, C.R., Saastamoinen, M., Marden, J.H. & Hanski, I. 2005. A candidate locus for variation in dispersal rate in a butterfly metapopulation. *Proc. R. Soc. B Biol. Sci.* **272**: 2449–2456.

Hadfield, J.D., Richardson, D.S. & Burke, T. 2006. Towards unbiased parentage assignment: combining genetic, behavioural and spatial data in a Bayesian framework. *Mol. Ecol.* **15**: 3715–3730.

Haggerty, B.P. & Galloway, L.F. 2011. Response of individual components of reproductive phenology to growing season length in a monocarpic herb. *J. Ecol.* **99**: 242–253.

Hendry, A.P. 2004. Selection against migrants contributes to the rapid evolution of ecologically dependent reproductive isolation. *Evol. Ecol. Res.* **6**: 1219–1236.

Hendry, A.P., Day, T. & Taylor, E.B. 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. *Evolution* **55**: 459–466.

Hereford, J. 2009. A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* **173**: 579–588.

Hoffman, J.I. & Amos, W. 2005. Microsatellite genotypic errors: detection approaches, common sources and consequences for paternal exclusion. *Mol. Ecol.* **14**: 599–612.

Jackson, C.H. 2011. Multi-state models for panel data: the msm package for R. *J. Stat. Softw.* **38**: 1–29.

Kawai, Y. & Kudo, G. 2011. Local differentiation of flowering phenology in an alpine-snowbed herb Gentiana nipponica. *Botany* **89**: 361–367.

King, D. & Roughgarden, J. 1982. Graded allocation between vegetative and reproductive growth for annual plants growing in seasons of random length. *Theor. Popul. Biol.* **22**: 1–16.

Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. *Proc. R. Soc. B Biol. Sci.* **267**: 1649–1655.

Kozlowski, J. 1992. Optimal allocation of resources to growth and reproduction: implications for age and size at maturity. *Trends Ecol. Evol.* **7**: 15–19.

Leimu, R. & Fischer, M. 2008. A meta-analysis of local adaptation in plants. *PLoS One* **3**: e4010.

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17**: 183–189.

Linhart, Y.B. & Grant, M.C. 1996. Evolutionary significance of local genetic differentiation in plants. *Ann. Rev. Ecol. Syst.* **27**: 237–277.

Lopez, S., Rousset, F., Shaw, F.H., Shaw, R.G. & Ronce, O. 2008. Migration load in plants: role of pollen and seed dispersal in heterogeneous landscapes. *J. Evol. Biol.* **21**: 294–309.

Loveless, M.D., Hamrick, J.L. & Foster, R.B. 1998. Population structure and mating system in *Tachigali versicolor*, a monocarpic neotropical tree. *Heredity* **81**: 134–143.

Lowry, D.B., Rockwood, R.C. & Willis, J.H. 2008a. Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution* **62**: 2196–2214.

Lowry, D.B., Modliszewski, J.L., Wright, K.M., Wu, C.A. & Willis, J.H. 2008b. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philos. Trans. R. Soc. B Biol. Sci.* **363**: 3009–3021.

Lynch, M. & Walsh, J.B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Press, Sunderland, MA.

Mazer, S.J. & LeBuhn, G. 1999. Genetic variation in life history traits: evidence within and between populations, and lessons learned. In: *Life History Evolution in Plants* (T. Vuorisalo & P. Mutikainen, eds), Kluwer Academin, Dordrecht.

Morgan, M.T. & Conner, J.K. 2001. Using genetic markers to directly estimate male selection gradients. *Evolution* **55**: 272–281.

Paul, J.R., Sheth, S.N. & Angert, A.L. 2011. Quantifying the impact of gene flow on phenotype-environment mismatch: a demonstration with the scarlet monkeyflower *Mimulus cardinalis*. *Am. Nat.* **178**: S62–S79.

Postma, E. & van Noordwijk, A.J. 2005. Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature* **433**: 65–68.

R Core Team. 2014. R: A Language and Environment for Statistical Computing, *R Foundation for Statistical Computing*, Vienna, Austria. http://www.R-project.org.

SAS Institute Inc. 1990. *SAS/STAT User's Guide. Version 6*, Fourth edn. SAS Institute, Inc, Cary, NC.

Shine, R., Brown, G.P. & Phillips, B.L. 2011. An evolutionary process that assembles phenotypes through space rather than through time. *Proc. Natl. Acad. Sci. USA* **108**: 5708–5711.

Silvertown, J., Servaes, C., Biss, P. & Macleod, D. 2005. Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. *Heredity* **95**: 198–205.

Slatkin, M. 1975. Gene flow and selection in a two-locus system. *Genetics* **81**: 787–802.

Weis, A.E. & Kossler, T.M. 2004. Genetic variation in flowering time induces phenological assortative mating: quantitative genetic methods applied to *Brassica rapa*. *Am. J. Bot.* **91**: 825–836.

Weis, A.E., Wadgymar, S.M., Sekor, M. & Franks, S.J. 2014. The shape of selection: using alternative fitness functions to test predictions for selection on flowering time. *Evol. Ecol.* **28**: 885–904.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1 Analysis of Variance for phenotypic traits of offspring produced by a line cross 1 among the Back Bay.

Table S2 Means and Confidence Limits for morphological traits of offspring produced by a line cross among the Back Bay.

Table S3 Parameter estimate for fit of the gausian function to population-level flowering schedules in the California clinal *B. rapa* populations.

Table S4 Logistic regression coefficients of exposure, *mi*, over flowering time.

Table S5 Mean flowering time of potential migrants.

Figure S1 Number of flowers counted over the flowering season as a function of flowering time (days from seedling emergence to first flower).

Figure S2 Difference between mean flowering time of resident populations and migrants.

Received 19 October 2014; revised 20 January 2015; accepted 5 February 2015