PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Koski MH, Galloway LF, Busch JW. 2022 Hybrid breakdown is elevated near the historical cores of a species' range. Proc. R. Soc. B 289: 20220070. https://doi.org/10.1098/rspb.2022.0070

Received: 13 October 2021
Accepted: 11 February 2022

Subject Category:
Evolution

Subject Areas:
evolution, plant science, genetics

Keywords:
allopatry, ecological speciation, reproductive isolation, species range, standing genetic variation

Author for correspondence:
Matthew H. Koski
e-mail: mkoski@clemson.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5877637.

THE ROYAL SOCIETY PUBLISHING

Hybrid breakdown is elevated near the historical cores of a species' range

Matthew H. Koski[1], Laura F. Galloway[2] and Jeremiah W. Busch[3]

[1]Department of Biological Sciences, Clemson University, 134 Long Hall, Clemson, SC 29634, USA
[2]Department of Biology, University of Virginia, PO Box 400328, Charlottesville, VA 22904, USA
[3]School of Biological Sciences, Washington State University, PO Box 644236, Pullman, WA 99164, USA

MHK, 0000-0002-0274-3145

New species form when they become reproductively isolated. A classic model of speciation posits that derived mutations appear in isolated populations and reduce fitness when combined in hybrids. While these Bateson–Dobzhansky–Muller incompatibilities are known to accumulate as populations diverge over time, they may also reflect the amount of standing genetic variation within populations. We analysed the fitness of F_2 hybrids in crosses between 24 populations of a plant species (Campanula americana) with broad variation in standing genetic variation and genetic differentiation driven by post-glacial range expansions. Hybrid breakdown varied substantially and was strongest between populations near the historical cores of the species range where within-population genetic diversity was high. Nearly half of the variation in hybrid breakdown was predicted by the combined effects of standing genetic variation within populations, their pairwise genetic differentiation and differences in the climates they inhabit. Hybrid breakdown was enhanced between populations inhabiting distinct climates, likely reflecting local adaptation. Results support that the mutations causing hybrid breakdown, the raw material for speciation, are more common in long-inhabited areas of the species range. Genetic diversity harboured in refugial areas is thus an important source of incompatibilities critical to the speciation process.

1. Background

Speciation occurs when populations evolve reproductive isolation [1]. A diversity of intrinsic and extrinsic factors can cause populations to become reproductively isolated. Pre-zygotic reproductive barriers reduce the likelihood that hybrids are formed, while post-zygotic factors reduce the fitness of hybrids after they are formed, a phenomenon known as hybrid breakdown [2]. In the allopatric model of speciation, derived mutations are thought to accumulate over time in isolated populations and these interact to reduce the fitness of hybrids [3–5]. In this model, reproductive isolation evolves as a by-product when mutations at distinct loci, or Bateson–Dobzhansky–Muller incompatibilities (BDMIs), reach appreciable frequencies in a pair of populations due to neutral processes, adaptive differentiation of local populations, or histories of genomic conflict [6–9]. An impressive body of work in animals and plants shows that hybrid breakdown increases predictably as populations diverge over time [10–14]. However, all incompatibilities must first exist as polymorphisms, a fact implying that standing variation within populations may also contribute to hybrid breakdown [13–15].

BDMIs are expected to be polymorphic in theory and have been found to segregate within natural populations [14–16]. Because BDMIs reduce fitness through negative epistasis, they will be most common when selection against them is weak [14]. Given that any mutation has the potential to negatively interact with others, hybrid breakdown may reflect genetic differentiation between parental populations, but also their amounts of standing variation [10,16,17]. This perspective posits that much of hybrid breakdown reflects intrinsic factors

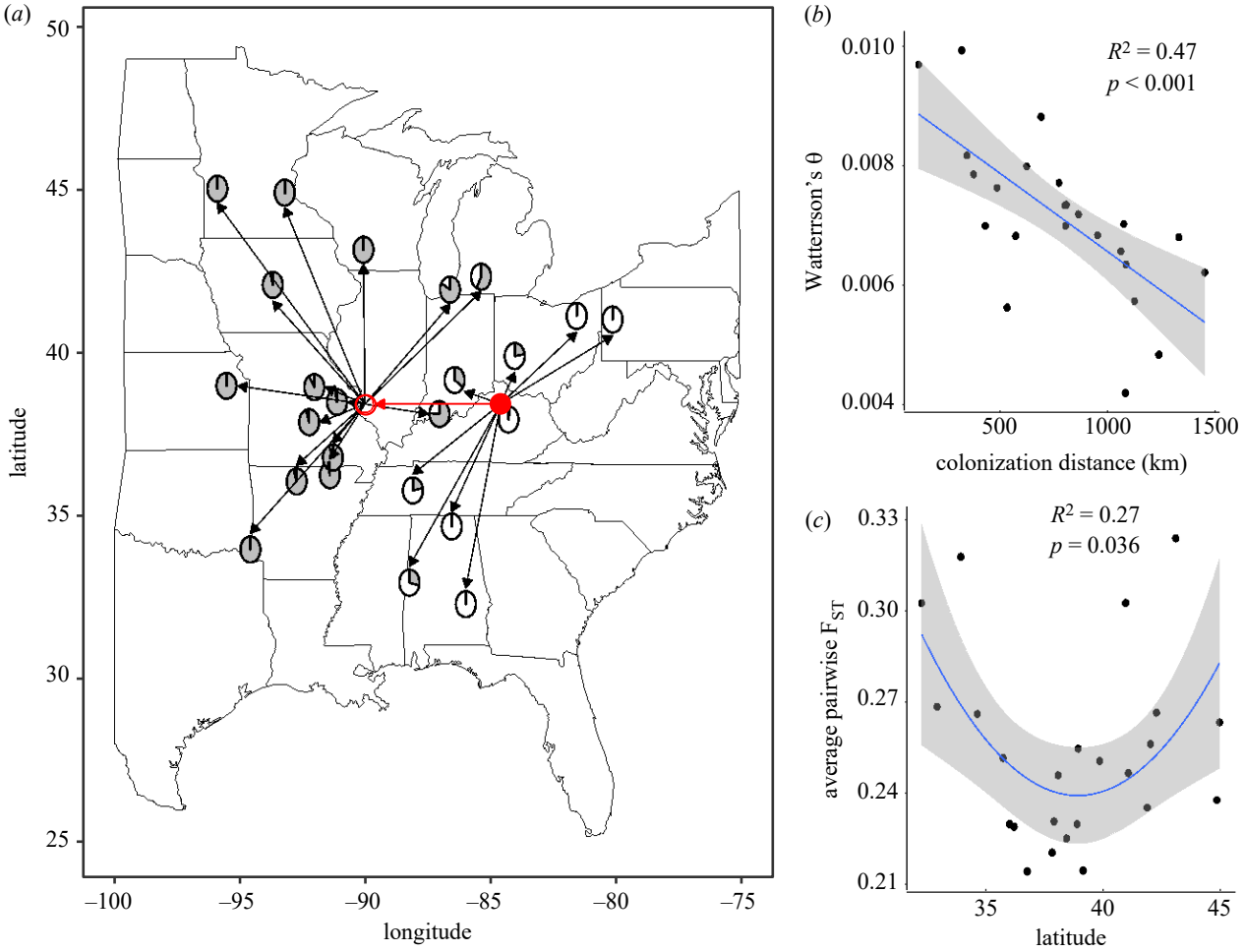

Figure 1. Spatial patterns of genetic variation in Campanula americana. (a) Two regions (red circles) were previously identified as sources for range expansion. Populations first expanded from a mid-latitude Pleistocene Appalachian refugium (closed red circle, right). A staging ground near the Mississippi River (open red circle, left) later served as a secondary origin of range expansion. Dashed black arrows depict the straight-line linear distance of each population from expansion origins. Population colours reflect ancestry from the most likely STRUCTURE model (K = 2). (b) Standing genetic variation in populations is predicted by their colonization distance (the total length of arrows) from the Appalachian refugium ($R^2 = 0.473$; $p < 0.001$). (c) Mid-latitude populations are the least genetically differentiated on average, given their proximity to mid-latitude refugia ($R^2 = 0.272$; $p = 0.036$). Shaded areas depict standard error.

(i.e. genetic aspects of a population's history [8,9]). Yet, the fitness of hybrids and their parents are also heavily dependent on the extrinsic environments in which mutations potentially influence fitness. Indeed, hybrid incompatibilities have been shown to accumulate as populations adapt to their unique ecological circumstances [18,19]. Therefore, to interpret the impacts of intrinsic genetic diversity and differentiation between populations on hybrid fitness, it is crucial to also consider the magnitude of environmental differences between parental populations as a proxy for local adaptation [20].

Historical range expansion frequently establishes geographic patterns in the distribution of genetic diversity within and between-populations. In the Northern Hemisphere in particular, lower latitude regions have repeatedly served as refugia during glaciation and thus the source of post-glacial range expansion [21]. These historical cores tend to harbour high levels of within-population genetic diversity, with serial founder events during range expansion causing declines in genetic diversity towards leading range edges [22,23]. Populations inhabiting refugial regions also typically exhibit substantial between-population genetic differentiation [24–26]. Large amounts of genetic diversity both within and between populations may set the stage for elevated hybrid breakdown near refugia, with less potential for breakdown far from these long-inhabited portions of the species range.

To study the factors that explain hybrid breakdown, ideal study systems should exhibit broad variation in their genetic diversity, differentiation and the selective regimes they experience [20,27,28]. Campanula americana is a flowering plant inhabiting a diverse spectrum of abiotic conditions in eastern North America. Across its range, there is a latitudinal cline in temperature and a longitudinal cline in aridity, generating a mosaic of environmental conditions across the landscape [29,30]. In C. americana, a recent range expansion has geographically structured patterns of genetic diversity and differentiation [31] (figure 1). Populations farther from a mid-latitude glacial refugium maintain fewer nucleotide polymorphisms (figure 1b; [32]), and contemporary populations at extreme latitudes are the most genetically differentiated (figure 1c). Phylogenomic studies and habitat suitability models also implicate the southern Gulf Coast as an important refugial region for C. americana [32,33]. We address the following questions: (i) does variability in F_2 hybrid breakdown covary with distance from regions of glacial refugia? (ii) What is the importance of standing genetic variation and genetic differentiation in determining the magnitude of hybrid breakdown? (iii) Does hybrid breakdown also reflect dissimilarity among the environments inhabited by populations?

2. Material and methods

(a) System and focal populations

American bellflower, Campanula americana L. (=Campanulastrum americanum Small, Campanulaceae), is a monocarpic herb found throughout the eastern United States. It is insect pollinated [34] and highly outcrossing, though fully self-compatible [35,36]. Plants are typically found in partially shaded, disturbed habitats including forest light gaps and margins as well as along riparian areas.

Campanula americana contains two main clades that are largely reproductively isolated. One occurs throughout the Appalachian Mountains (hereafter, 'Appalachian Clade'). The other occurs west of the Appalachians to the eastern portion of the Great Plains, with southern limits near the Gulf Coast and northern limits in the upper Midwest (hereafter, 'Western Clade'; [33]). The two clades are reproductively isolated due to cytonuclear incompatibilities that cause reductions in hybrid survival of up to 90% and arise in the F_1 generation [37–39]. There are no cytonuclear incompatibilities in hybrids between populations within the western clade [38].

Focal populations in the current study are in the Western clade. Genetic structure among populations in the Western clade and standing levels of within-population genetic diversity reflect a complex evolutionary history involving Pleistocene range expansion. Spatial signatures of genetic drift including declines in genetic diversity and the loss of rare alleles support westward colonization from a mid-latitude refugium near the southern Appalachian Mountains [31,32]. Within this westward expansion, a staging ground in the Mississippi River Valley was the source of subsequent colonization west of the Mississippi River, resulting in a genetic discontinuity among populations that largely coincides with the Mississippi River (figure 1; [32,33]). These historical cores of C. americana were identified using a time-difference of arrival analysis that used the frequency of derived alleles to identify origins of range expansion (see [22] for details). Another signature of genetic structure in the Western clade indicates ancient genetic subdivision between most populations and those at the southern range margin, where suitable Pleistocene habitats for C. americana were concentrated near the Gulf Coast [33]. Indeed, populations nearest the Gulf Coast are the earliest diverging lineages relative to all other populations in the Western clade [32,33] suggesting this genetic subdivision was preserved as the leading edge of the species' range expansion advanced in a primarily western direction.

(b) Population sampling

We collected seed of at least 25 maternal families from each of 24 C. americana populations in the Western clade during late Summer 2015. Populations were sampled across the latitudinal and longitudinal extent of the Western clade with an even division east and west of the Mississippi River to capture the relevant genetic structure (electronic supplementary material, table S1).

(c) Generation of experimental seed and fitness metrics

To evaluate hybrid breakdown, we created F_1 and F_2 hybrids between pairs of populations. Populations were roughly divided by latitude into the eight furthest south, eight at mid-latitudes and the eight furthest north (electronic supplementary material, figure S1). Each population was crossed with two other populations within its latitudinal cluster, once as a sire and once as a dam, for 24 independent hybrid crosses. Specific crosses are provided in electronic supplementary material, table S1 and figure S1. Eastern and western populations defined by geography also tended to form distinct genetic clusters based on STRUCTURE analysis, though admixture was common (figure 1; electronic supplementary material, figure S1). Our crossing design generated a range in pairwise genetic differentiation between the parental populations of hybrids (F_{ST}: 0.20–0.39). Crosses that involved populations with more ancestry in common had lower F_{ST} than those with more divergent ancestry (correlation between pairwise ancestry difference and pairwise F_{ST}, $r = 0.48$, $p = 0.01$). Our crossing design also generated hybrids between populations that displayed wide variation in standing genetic diversity, with populations closer to historical cores of the range having elevated within-population genetic diversity (figure 1c). Finally, our crossing design generated a wide degree of variation in pairwise geographic distance (electronic supplementary material, figure S1) and pairwise environmental distance between parental populations (electronic supplementary material, table S1).

Seeds from the 24 populations were grown to flowering in the greenhouse. Two replicates of three field-collected seeds from each of approximately 25 maternal plants in each population were planted (1200 replicates = 24 pops × 25 mat fam/pop × 2 reps/fam) in plug trays filled with a 4 : 1 ratio of soilless potting medium and fritted clay. Trays were placed in growth chambers for 12 h (21°C day, 14°C) and kept moist. After 30 days, seedlings were thinned to one random plant per cell and moved to a 5°C cold room for 12 h. Following 45 days of vernalization, 25 seedlings from each population, distributed over as many maternal families as possible, were transplanted into tubular pots and placed in a greenhouse with 16 h day lengths. There they were watered as needed and fertilized alternate weeks until bolting and then weekly. Upon flowering, 20 plants per population, each from a different maternal line, were selected for crossing.

Within- and between-population crosses were conducted. Two flowers on each plant were tagged and emasculated. The following day, when in female phase, one flower was pollinated with pollen from another randomly selected individual from the same population to create parental-type seed. The second flower was pollinated with a randomly selected individual from the assigned paternal population (electronic supplementary material, table S1) to create between-population hybrid F_1 seed. In total, 960 crosses were done (24 pops × 20 plants/pop × 2 cross-types), with crosses representing a different field-collected maternal seed family. Fruits were collected when mature. F_1 seeds were then planted and grown to flowering using the procedures described above. One individual was grown for each parental and F_1 cross-type for each family and population for a total of 960 plants descended from different field-collected maternal seed families. Upon flowering, each plant was crossed to another plant of the same cross-type, creating a second generation of parental plants and an F_2 generation of hybrid crosses.

(d) Evaluation of parental, F_1 and F_2 performance

Fitness components were evaluated for parental populations and two generations of between-population crosses. Performance was evaluated over 2 years, with parental and F_1 plants grown the first year, and parental and F_2 plants the second year. Two replicates of five seeds from each parental and hybrid F_1 cross were planted, with their positions randomized across plug trays (1920 cells, two for each of the 960 crosses). The following year parental and F_2 seeds were planted, with two replicates of five seeds planted for 15–21 (mean 18.2) families per parental population or F_2 hybrid cross, for a total of 1728 cells (two for each of 864 crosses). Seeds were planted and germinated as described above.

Data were collected for three life-history traits. Germination was scored approximately four weeks after planting at which time there were few new germinants. Proportion germination was calculated for each cross as the number of seedlings (summed over the two cells planted) divided by the 10 seeds planted. A single seedling of each family was transplanted and grown in the greenhouse (954 Parental and F_1; 873 Parental and F_2). Survival to flowering was recorded. Day of first flower was assessed on alternate days, and the number of open flowers

was counted weekly for four weeks on each plant (flowering typically lasts a month [40]). The average flower number across the four weeks was calculated as an index of flower production. Cumulative fitness was estimated as the multiplicative combination of proportion germination, proportion survival and flower production for each cross-type and population.

(e) Population genomics

Nucleotide variation was studied in the 24 populations using a restriction-site associated DNA sequencing approach (RADseq). Details of DNA extraction, haplotyping and genotyping are described elsewhere [32]. Briefly, 6–7 plants were barcoded and 100 bp reads were generated after digestion with SbfI and size selection. Sequences were processed in iPyrad [41], with custom haplotyping of the sequencing reads to remove errors. After genotyping, loci missing data for more than 33% of populations were pruned. Only biallelic SNPs were considered, and SNPs with a frequency below 5% at the species level were also removed. The resulting dataset contained 2605 RAD loci with an average of 2.5 SNPs per RAD locus. Genotypic data were then used to count the number of nucleotide differences within populations using Watterson's and Tajima's estimators of θ [32]. Pairwise F_{ST} among populations was estimated using analysis of variance implemented in the StAMPP R package [42,43].

(f) Climate data and environmental distance

Climate data were acquired from the Prism Climate Group (Oregon State University, https://prism.oregonstate.edu). Annual averages for climate variables over the period of 1981–2010 were attained for the 800 m grid cell containing each of the 24 populations. We gathered elevation, rainfall (inches), mean temperature, minimum water vapour deficit (hPa) and maximum water vapour deficit (hPa). Other temperature variables were not considered because of their strong positive correlations with average temperature. Using these variables, we calculated two measures of the multivariate climate distance between populations. First, the Euclidean distance between points was calculated, though this could be biased by an underlying covariance between the variables. We then calculated the pairwise Mahalanobis distance between populations, which includes the underlying variance–covariance structure among populations when attaining distances. While the former distance can will be biased with a covariance between environmental variables, the latter distance factors out this covariance.

(g) Statistical analysis

Hybrid breakdown was inferred based on cumulative fitness for each between-population cross. Initially, means were calculated for cumulative fitness for each generation (Parental in 2 years, F_1, F_2) for each cross. Hybrid breakdown was estimated by Δ (equation (2.1); [44]), which is the difference between the F_2 and the average of the parental populations (P_1, P_2) and the F_1. Here the mid-parent value (MP) is the average of both parental population means for both years.

$$D = z(F_2) - \frac{z(MP) + z(F_1)}{2} \qquad (2.1)$$

Δ was standardized by dividing by the F_2 mean (hereafter, Δ/F_2).

We assessed whether latitude and colonization distance from the mid-latitude refugium predicted variation in Δ/F_2 using separate linear models for each predictor (R, 'lm'). Colonization distance for each population was calculated as the straight-line distance from the Appalachian Refugium, or the combined distance from the Southern Appalachian Refugium and the Mississippi River Staging Ground (figure 1a). We used mid-parent averages of each hybrid population (maternal value + paternal value/2) for colonization distance and latitude. To evaluate the joint influences of within-population genetic diversity, population genetic differentiation, and geographic and environmental distances between parental populations on Δ/F_2, we used a multiple linear regression (R, 'lm'). We used mid-parent average Watterson's θ as a metric of within-population genetic variation and pairwise F_{ST} as a metric of between-population genetic differentiation. The linear distance (km) between parental populations was used as geographic distance. For environmental distance, we used Mahalanobis D. Linear distance and environmental distance were uncorrelated ($r = 0.30$, $p = 0.143$) allowing for an evaluation of their independent effects on hybrid breakdown. Variance inflation factors for variables in the model were low (range: 1.24–3.01) indicating minimal impacts of multi-collinearity on the parameter estimates for each predictor variable. We then generated standardized effects of each parameter on Δ/F_2 by scaling all parameters in the model (Z-score).

A breakdown of fitness in the F_2 hybrid generation is posited to be caused by novel combinations of parental alleles formed only in the F_2. The sorting of such novel multi-locus genotypes should not be uniform across F_2 individuals. Thus, the expression of hybrid breakdown at the population level should be positively associated with variation in the magnitude of hybrid breakdown among individuals. We assessed the relationship between Δ/F_2 and its standard deviation by resampling individual fitness. Sample sizes varied among crosses, so the number of jackknifed individuals was modified to minimize the range of plants deleted (range: 7.1–13.0%). The standard deviation in Δ/F_2 was calculated over the 1000 replicate datasets and its Pearson product-moment correlation with mean population fitness was calculated (R, 'cor.test'). One standard deviation value was 117% higher than the average and was an outlier base on Grubb's Test ($G = 3.64$, $p < 0.0001$). This value was removed prior to the analysis.

3. Results

(a) Geographic variation in hybrid breakdown

On average, crosses displayed modest hybrid breakdown for cumulative fitness ($\Delta/F_2 = -0.096$), though the magnitude of hybrid breakdown varied considerably across populations (figure 2a), ranging from a 38% reduction in F_2 fitness ($\Delta/F_2 = -0.38$) to a 17% increase in F_2 fitness over expectations ($\Delta/F_2 = 0.17$). The strength of hybrid breakdown tended to increase with each successive life stage (mean germination $\Delta/F_2 = -0.0058 \pm 0.0088$ s.e.; mean survival $\Delta/F_2 = -0.01 \pm 0.078$ s.e.; mean flower number $\Delta/F_2 = -0.085 \pm 0.027$ s.e.), with flower production contributing most strongly to cumulative fitness breakdown. Hybrid breakdown declined with the average colonization distance of parental populations from the mid-latitude glacial refugium ($R^2 = 0.30$; $p = 0.006$; figure 2b). Hybrid breakdown also declined with the average latitude of parental populations, a proxy for proximity to highly suitable environments throughout the Pleistocene and the Holocene ($R^2 = 0.41$; $p < 0.001$; figure 2c).

(b) Genetic and environmental predictors of hybrid breakdown

Hybrid breakdown was significantly associated with the standing genetic variation within populations (Watterson's θ; $p = 0.008$), the genetic differentiation between populations (F_{ST}; $p = 0.002$) and the geographic distance between them ($p = 0.005$). When considering the marginal effects of these variables, populations with the most standing variation (i.e.

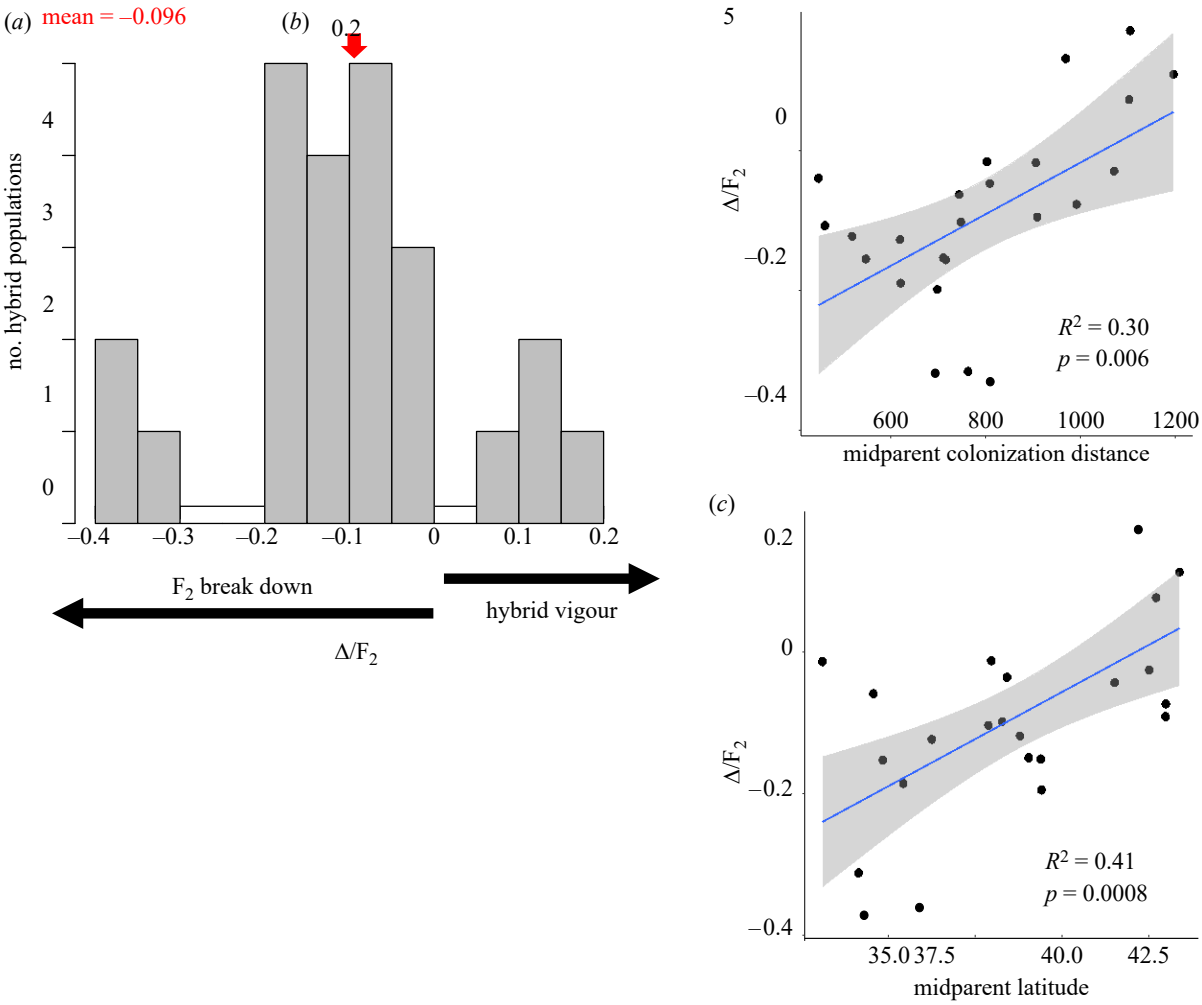

Figure 2. Variation and geographic patterns of F_2 hybrid breakdown for 24 hybrid crosses of Campanula americana. (a) Histogram depicting variation in hybrid breakdown (Δ/F_2) among population pairs. (b) Hybrid breakdown plotted against mid-parent colonization distance, which equals the distance between populations and a mid-latitude Pleistocene refugium in the Appalachians (figure 1). (c) Hybrid breakdown plotted against latitude, where southern latitudes were proximate to highly suitable habitats during the Pleistocene. Shaded areas represent standard error. (Online version in colour.)

higher Watterson's θ, a measure of nucleotide differences between sequences) produced hybrids with stronger hybrid breakdown (figure 3a). Moreover, populations that were more genetically differentiated, having higher pairwise F_{ST} produced hybrids with more severe reductions in fitness (figure 3b). After accounting for genetic variation within and between populations, populations that were geographically close to one another exhibited more hybrid breakdown (figure 3c). This pattern likely arises because paired populations nearest the historical core distribution had shorter inter-population distances (figure 3; electronic supplementary material, figure S1). Finally, hybrid breakdown was positively associated with the Mahalanobis climate distance between populations ($p = 0.028$; figure 3d). In total, these predictors account for nearly half of the total variation in hybrid breakdown measured in the F_2 generation ($R^2 = 0.498$; $p = 0.008$; figure 3). The standardized effect of F_{ST} on F_2 breakdown was 50% stronger than that of geographic distance and 137% stronger than that of environmental distance. Likewise, the standardized effect of Watterson's θ on F_2 breakdown was 47% stronger than that of geographic distance and 134% stronger than that of environmental distance.

To corroborate the finding that more genetically diverged populations exhibited stronger hybrid breakdown, we explored the relationship between d_{xy} [45] and hybrid breakdown. d_{xy} measures absolute genetic divergence between populations and is therefore less affected by the level of genetic diversity within populations being compared than F_{ST} [46]. Substituting d_{xy} for F_{ST} in the full model described above, hybrid breakdown was stronger in populations with higher d_{xy} ($b = -382.2$, $p = 0.053$), and the influence of within-population diversity (mid-parent Watterson's θ) remained strong ($b = -68.4$, $p = 0.026$) (electronic supplementary material, figure S2 and table S2).

The lack of reciprocal hybrids in the crossing design precluded a definitive test for asymmetry in hybrid breakdown, a hallmark of the contribution of cytonuclear interactions to incompatibilities [38]. However, breakdown was similar regardless of whether maternal populations were predominately from the eastern or western genetic cluster (figure 1a; majority eastern ancestry versus majority western ancestry, $F_{1,22} = 0.86$, $p = 0.36$).

(c) Within-population variation in hybrid breakdown

If hybrid breakdown is caused by mutations that segregate in the F_2 generation, then populations with the most fitness breakdown should also exhibit the most fitness variation

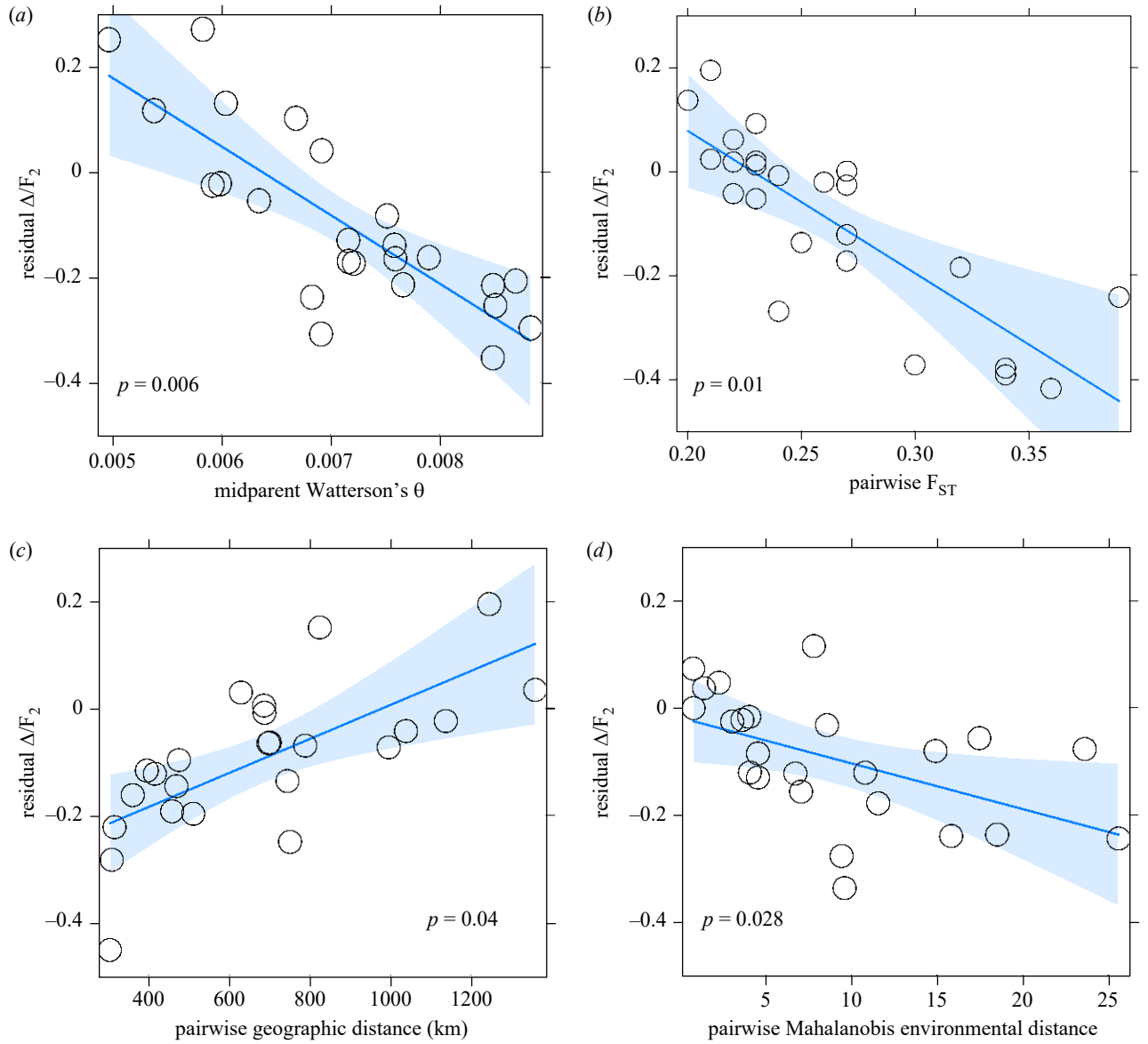

Figure 3. Effects of (a) within-population genetic diversity measured as mid-parent Watterson's θ, (b) between-population genetic diversity measured as pairwise F_{ST}, (c) pairwise geographic distance between parental populations and (d) Mahalanobis multivariate environmental distance on F_2 hybrid breakdown (Δ/F_2) for 24 hybrid crosses of Campanula americana. Plots depict marginal effects from a multiple linear regression including each factor. Full model: $R^2 = 0.498$, $p = 0.008$. (Online version in colour.)

among individuals. This prediction follows from the fact that hybrid breakdown is caused by some but not all hybrid genotypes having low fitness in the F_2 generation. There was a strong negative association between the observed hybrid breakdown (Δ/F_2) and the standard deviation of breakdown in each cross ($r = -0.66$; $p < 0.001$; figure 4). This relationship remained significant with the inclusion of one outlier ($r = -0.53$, $p = 0.007$).

4. Discussion

Geographic patterns of hybrid breakdown in Campanula americana were strongly structured by distance from glacial refugia. Colonization distance represents the distance travelled during a geographic range expansion, while latitude reflects a population's proximity to highly suitable environments throughout the Pleistocene and the Holocene [32,33]. Hybrid breakdown is therefore geographically structured and significantly enhanced between parental populations near the species' historical refugia. Conversely, hybrid incompatibilities were less often expressed at expanding range edges. While BDMIs are known to accumulate as populations diverge, they may also reflect the amount of standing variation within populations. In C. americana, post-glacial range expansion established clear geographic patterns in the amount of standing genetic variation and differentiation between populations, which in turn have structured spatial patterns in hybrid incompatibilities.

In this study, we find support for positive effects of both genetic differentiation (pairwise F_{ST}) and standing variation (average Watterson's θ) on hybrid breakdown. While these measures are negatively correlated with each other ($r = -0.67$, $p = 0.002$), our general linear model implicates their independent influence on the magnitude of hybrid breakdown revealed in crosses. Moreover, the use of d_{xy} instead of F_{ST} as a metric of population differentiation provided similar support for the effects of both within- and between-population diversity on hybrid breakdown. While consistent with the allopatric model of speciation, it is

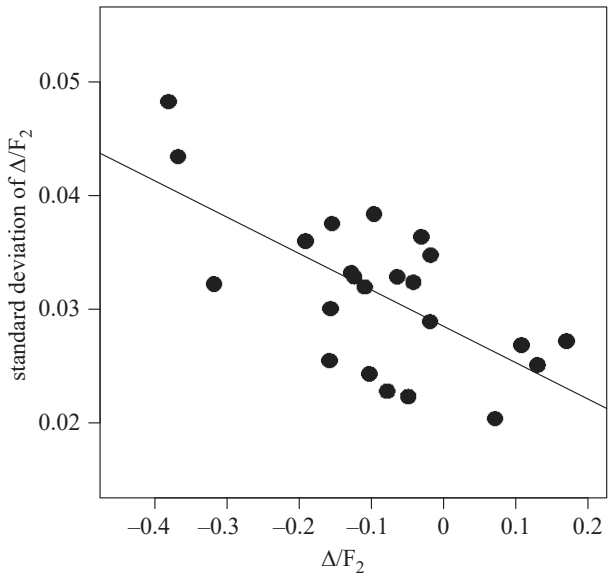

Figure 4. The relationship between variation in F_2 hybrid breakdown (SD) and the magnitude of F_2 hybrid breakdown (Δ/F_2). Crosses exhibiting greater breakdown (negative Δ/F_2) also exhibit greater fitness variation among individuals in the F_2 ($r = -0.66$; $p < 0.001$). The SD in breakdown was estimated with jackknife resampling.

important to note that this hypothesis concerns mutations that are uniquely derived in each population. Watterson's θ reflects nucleotide polymorphisms regardless of whether they are in fact unique to a population or have more ancient origins, yet those in the latter category may also contribute to hybrid breakdown [19].

In animals and plants, there is compelling support for the idea that hybrid breakdown accumulates over time as populations diverge [10,12–14]. This accumulation is typically thought to be caused by derived mutations that arise in isolated populations. This model of speciation is consistent with cases where large-effect mutations cause substantial declines in hybrid fitness, since polymorphic BDMIs would be eliminated by natural selection within a population. Selection would, however, be less effective in cases where BDMIs were subject to weak selection (e.g. hybrid breakdown caused by many small-effect mutations) [14]. In this study, the magnitude of hybrid breakdown was lowest during germination, moderate during vegetative growth and highest during the production of flowers near the end of life. Such a pattern is consistent with expectations based on evolutionary theories of ageing, since mutations expressed later in life have less influence on the reproductive contribution of individuals [47]. This argument has been previously applied to flowering plants, including C. americana, to explain the pattern of highest inbreeding depression for traits expressed late in life [48,49].

When BDMIs interact to reduce fitness, hybrid breakdown arises in the F_2 generation because some but not all genotypes express combinations of alleles that depress individual fitness [6,50]. This signature of BDMIs implies that crosses exhibiting more hybrid breakdown will also exhibit more variation in fitness among F_2 individuals. Our analyses support this prediction, demonstrating that the phenomenon of hybrid breakdown is consistent with the fitness effects of mutations segregating in F_2 hybrids. Such a pattern would not arise simply from a statistical scaling of the standard deviation in breakdown with its mean absolute value since crosses with the highest mean F_2 fitness express the least variation in fitness among individuals (figure 4). Hybrid breakdown is therefore strongly associated with individual fitness variation in the F_2 generation, in the specific direction predicted by when there is negative epistasis for fitness caused by BDMIs [6]. We note that our design was unlikely to detect the effects of all potential BDMIs in some of the hybrid populations. Because the number of replicates per hybrid population planted in the F_2 ranged from 15 to 21, the power to detect hybrid breakdown may have been limited. Given the expression of breakdown in C. americana, it is most likely that many loci of small effect underlie hybrid breakdown, though rare large-effect mutations would likely escape detection without larger samples.

Working models for the accumulation of hybrid incompatibilities have been greatly informed by dissecting the molecular basis of these interactions. Studies in both animal and plant model systems have revealed much about the nature of these interactions [18,50–54]. A complementary approach involves studying the spatial distribution of these incompatibilities to test hypotheses regarding their evolution [55–57]. While hybrid breakdown reported here is not directly connected to the underlying mutations, its variability among populations is valuable for testing hypotheses regarding the accumulation of reproductive isolation [58]. The work presented here provides a window into the early stages of the speciation process in natural populations of C. americana, where both genetic polymorphism and differentiation contribute to the expression of hybrid breakdown.

In tests of post-zygotic isolation, experiments frequently raise parents and hybrids in natural environments [2,59]. The motivation for this experimental design is to differentiate between intrinsic genetic incompatibilities and those that depend on the ecological context [60]. In cases of ecological speciation, there is an expectation that hybrid breakdown occurs when intermediate phenotypes reduce fitness in parental environments [61]. Such a pattern implies that natural selection contributes to divergence when hybrids are produced naturally [62,63]. In the current study, measurements of hybrid breakdown were made in a common greenhouse, so the degree to which hybrid fitness is contingent upon the environment is unknown. The fact that hybrid breakdown was exacerbated for populations inhabiting distinct climatic conditions (temperature, precipitation, aridity) implies, however, that there is at least intrinsic coadaptation that is associated with the climates experienced by populations across the species' geographic range [17]. We note that other factors not captured by our metric could also contribute to local adaptation. Regardless of whether post-zygotic reproductive isolation is caused by divergent selection among natural populations, its magnitude is associated with differences in the climatic conditions under which evolution has progressed in recent time.

5. Conclusion

Both polymorphism and differentiation contribute to the process of speciation, yet, their contributions are rarely studied together. Our study underscores the importance of standing genetic variation as a predictor of hybrid breakdown. That incompatibilities were most pronounced in refugial regions but were not expressed between populations at expanding range edges has important implications for speciation in the

face of climate change. Rear-edge populations frequently inhabiting lower latitudes in the northern hemisphere are particularly susceptible to extirpation as the climate warms [64,65]. These rear-edge populations generally harbour more genetic diversity than other parts of species' ranges and thus represent a hotspot for the accumulation of hybrid incompatibilities crucial for the process of species diversification. Leading-edge populations may be expected to persist in the wake of climate change, though their limited genetic diversity translates into lower speciation potential. Our study joins others in the call for strengthening conservation efforts of rear-edge populations [64–66].

Data accessibility. Data and code associated with this manuscript are provided here: https://zenodo.org/record/5566860#.YWbm7tnMIwQ.

The data are provided in the electronic supplementary material [67].

Authors' contributions. M.H.K.: conceptualization, data curation, formal analysis, investigation, methodology, writing—original draft and writing—review and editing; L.F.G.: conceptualization, data curation, formal analysis, funding acquisition, project administration, supervision, writing—original draft and writing—review and editing; J.W.B.: data curation, formal analysis, funding acquisition, writing—original draft and writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by NSF DEB 1457037.

Acknowledgements. We thank D. Grossenbacher for collecting seed, K. Niedermaier, S. Adu and C. Waterhouse for assistance in the greenhouse. N. Layman and C. Prior contributed to analyses of genetic data. T. Buchloh, E. Hilpman, H. Landis and S. Tisinai provided helpful manuscript critiques.

References

1. Mayr E. 1942 Systematics and the origin of species. New York, NY: Columbia University Press.
2. Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH. 2008 The strength and genetic basis of reproductive isolating barriers in flowering plants. Phil. Trans. R. Soc. B 363, 3009–3021. (doi:10.1098/rstb.2008.0064)
3. Bateson W. 1909 Heredity and variation in modern lights. Cambridge, UK: Cambridge University Press.
4. Dobzhansky T. 1940 Speciation as a stage in evolutionary divergence. Am. Nat. 74, 312–321. (doi:10.1086/280899)
5. Muller HJ. 1942 Isolating mechanisms, evolution and temperature. Biol. Sympos. 6, 71–125.
6. Fishman L, Willis JH. 2001 Evidence for Dobzhansky–Muller incompatibilities contributing to the sterility of hybrids between Mimulus guttatus and M. nasutus. Evolution 55, 1932–1942. (doi:10.1111/j.0014-3820.2001.tb01311.x)
7. Edmands S. 2002 Does parental divergence predict reproductive compatibility? Trends Ecol. Evol. 17, 520–527. (doi:10.1016/S0169-5347(02)02585-5)
8. Johnson NA. 2010 Hybrid incompatibility genes: remnants of a genomic battlefield? Trends Genet. 26, 317–325. (doi:10.1016/j.tig.2010.04.005)
9. Coughlan JM, Matute DR. 2020 The importance of intrinsic postzygotic barriers throughout the speciation process. Phil. Trans. R. Soc. B 375, 20190533. (doi:10.1098/rstb.2019.0533)
10. Coyne JA, Orr HA. 1989 Patterns of speciation in Drosophila. Evolution 43, 362–381. (doi:10.1111/j.1558-5646.1989.tb04233.x)
11. Widmer A, Lexer C, Cozzolino S. 2009 Evolution of reproductive isolation in plants. Heredity 102, 31–38. (doi:10.1038/hdy.2008.69)
12. Matute DR, Butler IA, Turissini DA, Coyne JA. 2010 A test of the snowball theory for the rate of evolution of hybrid incompatibilities. Science 329, 1518–1521. (doi:10.1126/science.1193440)
13. Moyle LC, Nakazato T. 2010 Hybrid incompatibility 'snowballs' between Solanum species. Science 329, 1521–1523. (doi:10.1126/science.1193063)
14. Corbett-Detig RB, Zhou J, Clark AG, Hartl DL, Ayroles JF. 2013 Genetic incompatibilities are widespread within species. Nature 504, 135–137. (doi:10.1038/nature12678)
15. Cutter AD. 2012 The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. Trends Ecol. Evol. 27, 209–218. (doi:10.1016/j.tree.2011.11.004)
16. Orr HA. 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139, 1805–1813. (doi:10.1093/genetics/139.4.1805)
17. Edmands S, Timmerman CC. 2003 Modeling factors affecting the severity of outbreeding depression. Conserv. Biol. 17, 883–892. (doi:10.1046/j.1523-1739.2003.02026.x)
18. Presgraves DC, Balagopalan L, Abmayr SM, Orr HA. 2003 Adaptive evolution drives divergence of a hybrid inviability gene between two species of Drosophila. Nature 423, 715–719. (doi:10.1038/nature01679)
19. Schluter D, Conte GL. 2009 Genetics and ecological speciation. Proc. Natl Acad Sci. USA 106, 9955–9962. (doi:10.1073/pnas.0901264106)
20. Lowry DB, Willis JH. 2010 A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS Biol. 8, e1000500. (doi:10.1371/journal.pbio.1000500)
21. Hewitt G. 2000 The genetic legacy of the Quaternary ice ages. Nature 405, 907–913. (doi:10.1038/35016000)
22. Prugnolle F, Manica A, Balloux F. 2005 Geography predicts neutral genetic diversity of human populations. Curr. Biol. 15, R159–R160. (doi:10.1016/j.cub.2005.02.038)
23. González-Martínez SC, Ridout K, Pannell JR. 2017 Range expansion compromises adaptive evolution in an outcrossing plant. Curr. Biol. 27, 2544–2551.e4. (doi:10.1016/j.cub.2017.07.007)
24. Mitton JB, Kreiser BR, Latta RG. 2000 Glacial refugia of limber pine (Pinus flexilis James) inferred from the population structure of mitochondrial DNA. Mol. Ecol. 9, 91–97. (doi:10.1046/j.1365-294x.2000.00840.x)
25. Hampe A, Arroyo J, Jordano P, Petit RJ. 2003 Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. Mol. Ecol. 12, 3415–3426. (doi:10.1046/j.1365-294X.2003.02006.x)
26. Gómez A, Lunt D. 2007 Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In Phylogeogr. Southern Europ. Refugia (eds S Weiss, N Ferrand). Springer, Dordrecht. (doi:10.1007/1-4020-4904-8_5)
27. Fenster CB, Galloway LF. 2000 Inbreeding and outbreeding depression in natural populations of Chamaecrista fasciculata (Fabaceae). Conserv. Biol. 14, 1406–1412. (doi:10.1046/j.1523-1739.2000.99234.x)
28. Frankham R, Ballou JD, Eldridge MDB, Lacy RC, Ralls K, Dudash MR, Fenster CB. 2011 Predicting the probability of outbreeding depression. Conserv. Biol. 25, 465–475. (doi:10.1111/j.1523-1739.2011.01662.x)
29. Koski MH, Galloway LF. 2018 Geographic variation in pollen color is associated with temperature stress. New Phytol. 218, 370–379. (doi:10.1111/nph.14961)
30. Koski MH, Galloway LF. 2020 Geographic variation in floral color and reflectance correlates with temperature and colonization history. Front. Plant Sci. 11, 991. (doi:10.3389/fpls.2020.00991)
31. Koski MH, Layman NC, Prior CJ, Busch JW, Galloway LF. 2019 Selfing ability and drift load evolve with range expansion. Evol. Lett. 3, 500–512. (doi:10.1002/evl3.136)
32. Prior CJ, Layman NC, Koski MH, Galloway LF, Busch JW. 2020 Westward range expansion from middle latitudes explains the Mississippi River discontinuity in a forest herb of eastern North America. Mol. Ecol. 29, 4473–4486. (doi:10.1111/mec.15650)
33. Barnard-Kubow KB, Debban CL, Galloway LF. 2015 Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. Am. J. Bot. 102, 1842–1853. (doi:10.3732/ajb.1500267)

34. Koski MH, Grossenbacher DL, Busch JW, Galloway LF. 2017 A geographic cline in the ability to self-fertilize is unrelated to the pollination environment. Ecology 98, 2930–2939. (doi:10.1002/ecy.2001)
35. Galloway LF, Etterson JR, Hamrick JL. 2003 Outcrossing rate and inbreeding depression in the herbaceous autotetraploid, Campanula americana. Heredity 90, 308–315. (doi:10.1038/sj.hdy.6800242)
36. Koski MH, Galloway LF, Busch JW. 2019 Pollen limitation and autonomous selfing ability interact to shape variation in outcrossing rate across a species range. Am. J. Bot. 106, 1240–1247. (doi:10.1002/ajb2.1342)
37. Etterson JR, Keller SR, Galloway LF. 2007 Epistatic and cytonuclear interactions govern outbreeding depression in the autotetraploid Campanulastrum americanum. Evolution 61, 2671–2683. (doi:10.1111/j.1558-5646.2007.00234.x)
38. Barnard-Kubow KB, So N, Galloway LF. 2016 Cytonuclear incompatibility contributes to the early stages of speciation. Evolution 70, 2752–2766. (doi:10.1111/evo.13075)
39. Barnard-Kubow KB, McCoy MA, Galloway LF. 2017 Biparental chloroplast inheritance leads to rescue from cytonuclear incompatibility. New Phytol. 213, 1466–1476. (doi:10.1111/nph.14222)
40. Galloway LF, Burgess KS. 2009 Manipulation of flowering time: phenological integration and maternal effects. Ecology 90, 2139–2148. (doi:10.1890/08-0948.1)
41. Eaton DAR. 2014 PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30, 1844–1849. (doi:10.1093/bioinformatics/btu121)
42. Weir BS, Cockerham CC. 1984 Estimating F-Statistics for the analysis of population structure. Evolution 38, 1358–1370. (doi:10.2307/2408641)
43. Pembleton LW, Cogan NOI, Forster JW. 2013 StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol. Ecol. Res. 13, 946–952. (doi:10.1111/1755-0998.12129)
44. Lynch M, Walsh B. 1998 Genetics and analysis of quantitative traits. Sunderland, MA: Sinauer Associates Inc.
45. Smith J, Kronforst MR. 2013 Do Heliconius butterfly species exchange mimicry alleles? Biol. Lett. 9, 20130503. (doi:10.1098/rsbl.2013.0503)
46. Cruickshank TE, Hahn MW. 2014 Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. Mol. Ecol. 23, 3133–3157. (doi:10.1111/mec.12796)
47. Williams GC. 1957 Pleiotropy, natural selection, and the evolution of senescence. Evolution 11, 398–411.
48. Husband BC, Schemske DW. 1996 Evolution of the magnitude and timing of inbreeding depression in plants. Evolution 50, 54–70. (doi:10.1111/j.1558-5646.1996.tb04472.x)
49. Galloway LF, Etterson JR. 2007 Inbreeding depression in an autotetraploid herb: a three cohort field study. New Phytol. 173, 383–392. (doi:10.1111/j.1469-8137.2006.01909.x)
50. Sweigart AL, Fishman L, Willis JH. 2006 A simple genetic incompatibility causes hybrid male sterility in Mimulus. Genetics 172, 2465–2479. (doi:10.1534/genetics.105.053686)
51. Ting CT, Tsaur SC, Wu ML, Wu CI. 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. Science 282, 1501–1504. (doi:10.1126/science.282.5393.1501)
52. Sicard A, Kappel C, Josephs EB, Lee YW, Marona C, Stinchcombe JR, Wright SI, Lenhard M. 2015 Divergent sorting of a balanced ancestral polymorphism underlies the establishment of gene-flow barriers in Capsella. Nat. Commun. 6, 7960. (doi:10.1038/ncomms8960)
53. Mack KL, Nachman MW. 2017 Gene regulation and speciation. Trends Genet. 33, 68–80. (doi:10.1016/j.tig.2016.11.003)
54. Forejt J. 1996 Hybrid sterility in the mouse. Trends Genet. 12, 412–417. (doi:10.1016/0168-9525(96)10040-8)
55. Zuellig MP, Sweigart AL. 2018 A two-locus hybrid incompatibility is widespread, polymorphic, and active in natural populations of Mimulus. Evolution 72, 2394–2405. (doi:10.1111/evo.13596)
56. Teeter KC et al. 2008 Genome-wide patterns of gene flow across a house mouse hybrid zone. Genome Res. 18, 67–76. (doi:10.1101/gr.6757907)
57. Maheshwari S, Barbash DA. 2011 The genetics of hybrid incompatibilities. Annu. Rev. Genet. 45, 331–355. (doi:10.1146/annurev-genet-110410-132514)
58. Scopece G, Lexer C, Widmer A, Cozzolino S. 2010 Polymorphism of postmating reproductive isolation within plant species. Taxon 59, 1367–1374. (doi:10.1002/tax.595004)
59. Campbell DR, Waser NM, Aldridge G, Wu CA. 2008 Lifetime fitness in two generations of Ipomopsis hybrids. Evolution 62, 2616–2627. (doi:10.1111/j.1558-5646.2008.00460.x)
60. McKinnon JS, Rundle HD. 2002 Speciation in nature: the threespine stickleback model systems. Trends Ecol. Evol. 17, 480–488. (doi:10.1016/S0169-5347(02)02579-X)
61. Gow JL, Peichel CL, Taylor EB. 2007 Ecological selection against hybrids in natural populations of sympatric threespine sticklebacks. J. Evol. Biol. 20, 2173–2180. (doi:10.1111/j.1420-9101.2007.01427.x)
62. Egan SP, Funk DJ. 2009 Ecologically dependent postmating isolation between sympatric host forms of Neochlamisus bebbianae leaf beetles. Proc. Natl Acad. Sci. USA 106, 19 426–19 431. (doi:10.1073/pnas.0909424106)
63. Melo MC, Grealy A, Brittain B, Walter GM, Ortiz-Barrientos D. 2014 Strong extrinsic reproductive isolation between parapatric populations of an Australian groundsel. New Phytol. 203, 323–334. (doi:10.1111/nph.12779)
64. Hampe A, Petit RJ. 2005 Conserving biodiversity under climate change: the rear edge matters. Ecol. Lett. 8, 461–467. (doi:10.1111/j.1461-0248.2005.00739.x)
65. Provan J, Maggs CA. 2012 Unique genetic variation at a species' rear edge is under threat from global climate change. Proc. R. Soc. B 279, 39–47. (doi:10.1098/rspb.2011.0536)
66. Bennett S, Wernberg T, Arackal Joy B, de Bettignies T, Campbell AH. 2015 Central and rear-edge populations can be equally vulnerable to warming. Nat. Commun. 6, 10280. (doi:10.1038/ncomms10280)
67. Koski MH, Galloway LF, Busch JW. 2022 Hybrid breakdown is elevated near the historical cores of a species' range. Figshare.