E-Article

Parental Population Range Expansion before Secondary Contact Promotes Heterosis

Ailene MacPherson,[1,2,*,†] Silu Wang,[1,3,†] Ryo Yamaguchi,[4] Loren H. Rieseberg,[1] and Sarah P. Otto[1]

1. University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada; 2. University of Toronto, Toronto, Ontario M5S 1A5, Canada; and Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada; 3. University of California, Berkeley, California 94720; and University of California, Davis, California 95616; 4. Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan; and Hokkaido University, Sapporo, Hokkaido 060-8588, Japan

Submitted July 7, 2021; Accepted January 21, 2022; Electronically published May 26, 2022

Online enhancements: supplemental PDF. *Dryad data:* https://doi.org/10.5061/dryad.0vt4b8h07.

abstract: Population genomic analysis of hybrid zones is instrumental to our understanding of the evolution of reproductive isolation. Many temperate hybrid zones are formed by the secondary contact between two parental populations that have undergone postglacial range expansion. Here, we show that explicitly accounting for historical parental isolation followed by range expansion prior to secondary contact is fundamental to explaining genetic and fitness patterns in these hybrid zones. Specifically, ancestral population expansion can result in allele surfing, where neutral or slightly deleterious mutations drift to high frequency at the expansion front. If these surfed deleterious alleles are recessive, they can contribute to substantial heterosis in hybrids produced at secondary contact, counteracting negative effects of Bateson-Dobzhansky-Muller incompatibilities (BDMIs) and hence weakening reproductive isolation. When BDMIs are linked to such recessive deleterious alleles, the fitness benefit of introgression at these loci can facilitate introgression at the BDMIs. The extent to which this occurs depends on the strength of selection against the linked deleterious alleles and the distribution of recombination across the chromosome. Finally, surfing of neutral loci can alter the expected pattern of population ancestry; thus, accounting for historical population expansion is necessary to develop accurate null genomic models of secondary contact hybrid zones.

Keywords: hybrid zone, allele surfing, range expansion, genetic incompatibilities.

Introduction

Population expansions often lead to secondary contact and hybridization with closely related species (Szymura 1976; Weir and Schluter 2004; Arntzen et al. 2017; Duvernell et al. 2019). This is most commonly reported for population expansions out of separate glacial refugia, which has led to many of the hybrid zones we observe today across both aquatic and terrestrial habitats (Haffer 1969; Avise et al. 1998; Taberlet et al. 1998; Hewitt 1999; Weir and Schluter 2004; April et al. 2013; Gao et al. 2015; Licona-Vera et al. 2018; Duvernell et al. 2019; table S1). The well-documented association between invasive species and hybridization provides another example in which the rapid expansion of previously isolated species has led to widespread genetic mixing (Edmonds et al. 2004).

Range expansion can have significant evolutionary and ecological consequences (Edmonds et al. 2004; Excoffier et al. 2009). One important genetic consequence of range expansion is an increase in the fixation rate of alleles, including deleterious ones, at the range edge because of an increase in genetic drift. This phenomenon, termed "gene (allele) surfing" by Klopfstein et al. (2006), was first characterized by Edmonds (2004). While all alleles—beneficial, neutral, and deleterious—can surf, surfing of deleterious alleles can lead to a substantial reduction in population mean fitness at the range edge, termed "expansion load" (Peischl et al. 2013), and may even limit the extent of range expansion (Peischl and Excoffier 2015).

Here, we investigate the effect of postexpansion allele surfing on hybridization at secondary contact. Upon hybridization, deleterious alleles that have increased in frequency on the edge of one parental range can be masked by their beneficial counterparts present in the other parental population. When surfing has led to the rise in deleterious mutations in both parental populations, such masking may result in substantial heterosis (hybrid

* Corresponding author; email: ailenem@sfu.ca.
† These authors contributed equally to the article.

  https://doi.org/10.1086/719656

fitness advantage over the parental lineages) and alter the introgression dynamics that ultimately determine whether the diverged lineages could become independent evolutionary trajectories. Here, we explore the effect of parental population range expansion before secondary contact on the fitness regime in hybrids, both initially (among F_1 crosses) and on the long-term consequences for introgression.

There are a range of possible evolutionary outcomes following secondary contact of closely related lineages: (1) fusion of the hybridizing lineages, (2) emergence of a hybrid species, or (3) the maintenance and strengthening of the species boundary. Between these three distinct outcomes is the formation of a hybrid zone, a geographical region where divergent lineages hybridize (Slatkin 1973; May et al. 1975; Endler 1977; Barton and Hewitt 1985; Barton 2001). Among hybrid zones, there is extensive variation in the direction and strength of selection as well as the rate at which reproductive isolation evolves within them (Barton and Hewitt 1985; Servedio and Noor 2003; Roux et al. 2016). Models of hybrid zones typically assume that gene flow and selection are at equilibrium (Slatkin 1973; May et al. 1975; Endler 1977; Barton and Hewitt 1989; Barton and Gale 1993). A history of parental population expansion prior to secondary contact violates such assumptions. Hence, understanding the effect of allele surfing on hybridization will fill a gap in our understanding of speciation.

Theoretical Background

Models of Secondary Contact

The tension zone model (Barton 1979; Barton and Hewitt 1989) explicitly examines the impact of population structure on secondary contact. Ancestral populations are initially separated by a geographical barrier. Secondary contact of the parental populations following the removal of this barrier results in the formation of genetic clines due to the balance of migration and selection. The shape and dynamics of the resulting clines hence depend on the nature of selection and introgression at each genetic locus. Motivated by analysis of transects of hybrid zones and the observation of concordant genetic clines along these transects (Barton and Hewitt 1985), the classic tension zone model assumes that populations are arrayed along a linear habitat. Here, we consider a variant of this model where parental populations are initially isolated at opposite ends of this linear habitat separated by a wide geographic barrier and must undergo range expansion prior to coming into secondary contact.

The selection regimes that are often considered within hybrid populations include the null case of neutral loci and universal selection against deleterious alleles. In addition, there is selection specifically against hybrid genotypes as a result of underdominance or negative epistasis in the form of Bateson-Dobzhansky-Muller incompatibilities (BDMIs). In the model presented below, we will focus on the dynamics at neutral, universally deleterious, and pairs of epistatic BDMI loci. Our model will contrast two different types of BDMIs: a dominant BDMI for which all recombinant genotypes are less fit, affecting both F_1 and later hybrids, and a recessive BDMI for which breakdown is experienced only in F_2, backcross, and later crosses (Turelli and Orr 2000).

Range Expansion and Allele Surfing

Our expectations for the effect of range expansion, model design, and parameter choices are grounded in the results of Peischl and colleagues (Peischl et al. 2013; Peischl and Excoffier 2015; Peischl et al. 2015). Using a combination of individual-based simulations and analytical results, they modeled the decline in population mean fitness on the edge of an expanding range as a result of the accumulation of deleterious mutations by drift. While we do not analyze the dynamics of edge mean fitness directly, this quantity is directly related to heterosis observed among F_1 hybrids formed by crossing individuals at the edge of two parental ranges. Peischl and Excoffier (2015), as in our model below, consider deleterious (partially) recessive mutations initially at mutation-selection balance in the range core. As expansion progresses, population mean fitness initially declines rapidly as a result of the fixation of segregating variation. This rapid decline is followed by a period of gradual decline as a result of the continued fixation of new mutations (Peischl and Excoffier 2015). Analogously, we expect F_1 heterosis to initially increase rapidly as a function of the distance traveled by parental populations prior to secondary contact, followed by a continued increase in heterosis at a reduced rate.

Peischel and Excoffier (2015) found that the decline in mean fitness on the range edge is predominantly the result of alleles with small to intermediate effects for which $Ns \leq 2$. In addition to the strength of selection (s) and the population size (N), the accumulation of mutations is also impacted by selective interference between nonrecombining loci and hence by the pattern of recombination and segregation across the genome (Peischl et al. 2015). Specifically, they compare the expansion dynamics of a completely asexual population with 20 freely segregating but otherwise nonrecombining chromosomes. While segregation has only mild effects on mean fitness on the range edge itself, it reduces selective interference among chromosomes and allows beneficial core alleles to spread more readily toward the range edge, hence narrowing the region of reduced fitness on the range edge.

Introgression and Heterosis

Introgression in the tension zone model (Barton 1979; Barton and Hewitt 1989) is characterized by clines in allele frequency across the transect. Relatively insensitive to the nature of selection against hybrids, these clines are expected to be sigmoidal in shape at selection-migration balance with allele frequency at location x given by

$$p(x) = \frac{1}{1 + e^{-4d(x-c)}}.$$

Parameterized in this manner, the transition of the allele frequency from fixed in one parental population to absent in the other population is gradual and centered around the location c. The slope of the cline at c is given by d, from which the width of the cline is defined as $W = 1/d$. The transient dynamics of allele frequency at neutral and selected loci are also expected to be sigmoidal with a time-dependent center and width. In contrast to selected loci, the neutral alleles will continue to introgress, eventually equilibrating at an expected frequency of 0.5 across the range.

Both the magnitude and the heterogeneity of recombination rate are known to shape the extent of introgression across the genome. Introgression is usually suppressed in regions of low recombination rate because of genetic linkage to loci that reduce hybrid fitness (Noor et al. 2001; Rieseberg 2001; Navarro and Barton 2003; Janoušek et al. 2015; Juric et al. 2016; Aeschbacher et al. 2017; Kirkpatrick 2017; Schumer et al. 2018; Martin et al. 2019).

Methods

Allele surfing and range expansion are complex eco-evolutionary processes (Peischl and Excoffier 2015; Peischl et al. 2015) involving feedbacks between ecological change in population size and evolutionary change in allele frequencies. We begin here by describing the genetics, life cycle, and geography of our eco-evolutionary model of secondary contact. Specifically, we consider secondary contact between two populations that were initially isolated in two distant refugia (of $n_{\text{core}} = 5$ demes each) followed by a subsequent period of range expansion before finally coming into secondary contact (fig. 1).

The life cycle of the species is characterized by discrete nonoverlapping generations with three stages: population census, reproduction, and migration. Selection occurs during reproduction such that each hermaphroditic individual with genotype i reproduces at a density-dependent rate, producing a Poisson-distributed number of gametes with mean $2(1 + \rho(1 - N_{\text{T}}/K))V_i$. These gametes, including those from the same hermaphrodite, are then combined randomly to produce diploid offspring (selfing is allowed). Here, $\rho = 0.01$ is the intrinsic growth rate, N_{T} is the total number of individuals in the local population, and $K = 200$ is a constant determining the strength of density dependence in each deme. Finally, V_i is a measure of the fertility of genotype i.

An individual's genotype consists of $n_{\text{N}} = 80$ neutral loci, $n_{\text{D}} = 100$ "selected" loci experiencing purifying selection, $n_{\text{B}} = 4$ two-locus BDMI pairs, and $n_{\text{FD}} = 12$ neutral loci that are artificially forced to exhibit fixed differences between parental populations, as in the classic tension zone model, for a total of $n_{\text{Loci}} = n_{\text{N}} + n_{\text{D}} + 2n_{\text{B}} + n_{\text{FD}} = 200$ loci. All loci are assumed to be biallelic and are arranged randomly along the genome. Selected loci are either fully recessive $h_{\text{D}} = 0$ or partially recessive $h_{\text{D}} = 0.25$. We examine three values of directional selection $s_{\text{D}} = 0.001$, $s_{\text{D}} = 0.003$, and $s_{\text{D}} = 0.005$, which correspond to values of Ks, a quantity equivalent to Ns when $V_i = 1$, of 0.2, 0.6, and 1, respectively. Hence, the strength of selection is weak to intermediate across the parameter range explored such that all selected loci are expected to experience extensive surfing during range expansion, as predicted by Peischl and Excoffier (2015). For reference purposes, we also examine the special case of neutral evolution ($s_{\text{D}} = 0.0$).

In addition to the selected loci, we consider epistatic selection between pairs of BDMI loci. Following the model and notation of Turelli and Orr (2000), for each pair of BDMI loci, B and C, we assume the parental populations are fixed for either the bbCC or BBcc genotype. The BDMI between the pair of loci B and C is characterized by three specific incompatibilities: H_0, the fitness of the double-heterozygous genotypes (BbCc); H_1, the fitness of the heterozygous-homozygous genotypes (bbCc, Bbcc, BBCc, BbCC); and H_2, the fitness of the double-homozygous genotype (bbcc). Breakdown in F_1 hybrids is determined by the value of H_0, whereas F_2 and backcross hybrids are affected by all three forms of incompatibilities. We consider two different types of BDMIs: "dominant BDMIs," for which $H_0 = H_1 = H_2 = 1 - s_{\text{B}}$, and "recessive BDMIs," where $H_0 = H_1 = 1$ and $H_2 = 1 - s_{\text{B}}$. We consider two strengths of epistatic selection at the BDMI loci: a weakly incompatible case ($s_{\text{B}} = 0.001$) and a strongly incompatible case ($s_{\text{B}} = 0.05$). See table 1 for a summary of model parameters.

Fertility across loci is multiplicative with a baseline value of V_0 such that the fertility of genotype i is given by the product

$$V_i = V_0 \times \underbrace{V_{\text{B}}}_{\text{BDMI}} \times \underbrace{\prod_{j=1}^{n_{\text{N}}+n_{\text{FD}}} 1}_{\text{neutral}} \times \underbrace{\prod_{j=1}^{n_{\text{D}}} \begin{cases} 1 & A_jA_j \\ 1 - h_{\text{D}}s_{\text{D}} & A_ja_j \\ 1 - s_{\text{D}} & a_ja_j \end{cases}}_{\text{selected}},$$

$$V_{\text{B}} = \prod_{j=1}^{n_{\text{B}}} \begin{cases} 1 & B_jB_j/c_jc_j \\ 1 & b_jb_j/C_jC_j \\ 1 - s_{\text{B}} & \text{otherwise} \end{cases} \quad \text{dominant BDMI,}$$

$$V_{\text{B}} = \prod_{j=1}^{n_{\text{B}}} \begin{cases} 1 - s_{\text{B}} & b_jb_j/c_jc_j \\ 1 & \text{otherwise} \end{cases} \quad \text{recessive BDMI.} \tag{1}$$

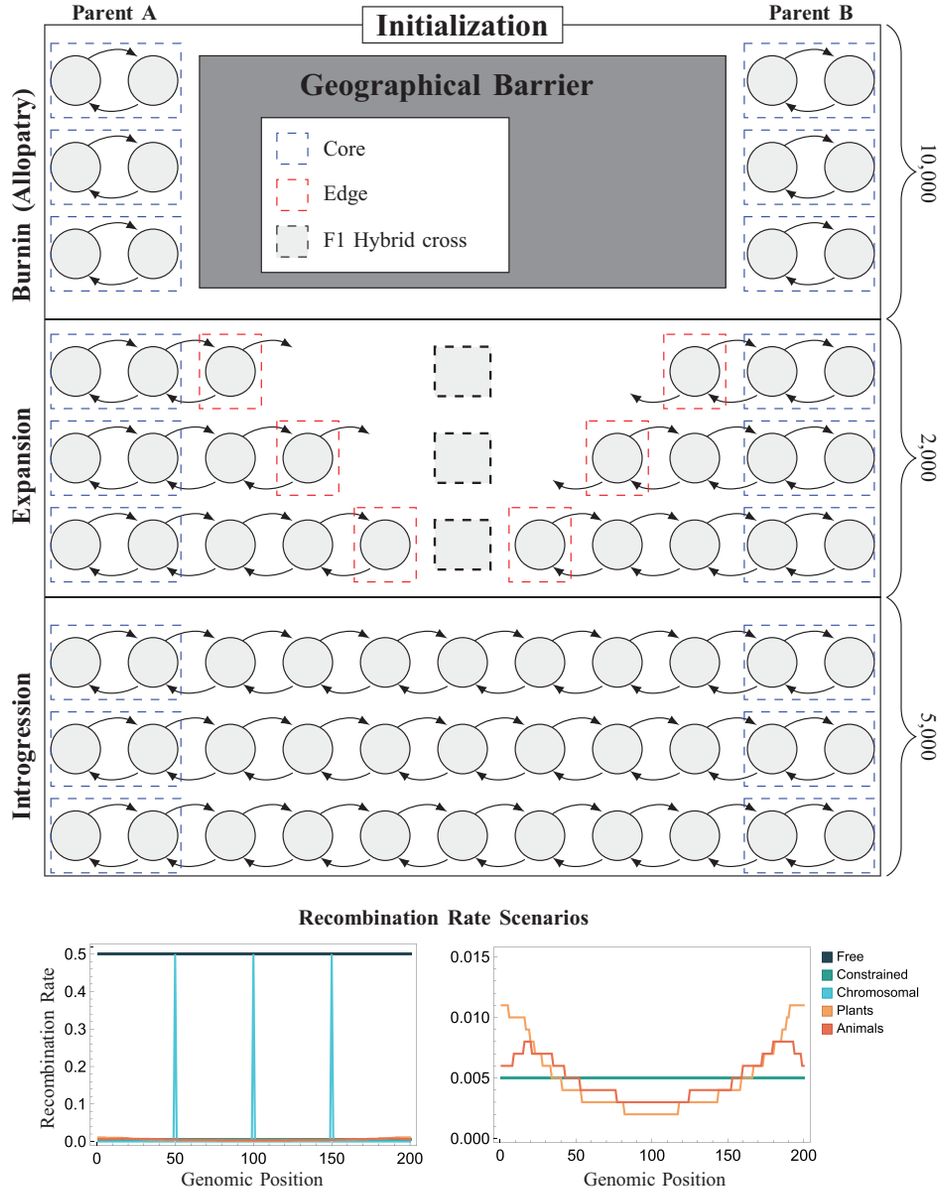


Figure 1: Schematic of model design. *Top*, individual-based simulation model proceeds in three phases: (1) evolution in allopatry, (2) parental range expansion, and (3) secondary contact and introgression. Geography is modeled as a 1D stepping-stone habitat. Occupied demes (gray circles, of which only a subset are shown) are connected via migration. Time is measured in units of generations of allopatry (10,000), generations of expansion (2,000), or generations of introgression (5,000), respectively. During range expansion, parental edge populations are crossed every 50 generations to create F_1 hybrids. *Bottom left-hand panel*, the five recombination rate scenarios shown on the same axis as a function of genomic position from locus 1 to locus 200. *Bottom right*, variability in recombination rate under the plants and animals scenario compared with uniform constrained recombination.

As there are strong eco-evolutionary feedbacks in the model, we wanted to limit the likelihood of random extinction of core populations during burn-in due to low initial absolute fitness followed by mutational meltdown. To do so, we set V_0 to ensure that the core populations would have an expected size of K at mutation-selection balance. Specifically, we set

$$V_0 = \frac{1}{(1+s_D)^{n_D}\int_{1/2K}^{1-(1/2K)}[x^2\phi(x,s_D,h_D,\mu,n_{core}\times K)]\,dx}, \quad (2)$$

Table 1: Parameter conditions used in individual-based simulations

Background selection, s_D	Background dominance, h_D	BDMI, s_B (h_B)	Recombination, $\bar{r}$
Neutral: .0	**Recessive: 0**	**No BDMI: 0 (NA)**	Free: .5
Weak: .001	Partially recessive: .25	Recessive weak: .01 (0)	**Constrained: .005**
Medium: .003		Dominant weak: .01 (1)	Chromosomal: .0075
Strong: .005		Recessive strong: .05 (0)	Plants: .005
		Dominant strong: .05 (1)	Animals: .005

Note: Individual-based simulations were run in a full factorial design with 10 replicate simulations per treatment. The genomic location of neutral, background, and Bateson-Dobzhansky-Muller incompatibility (BDMI) loci vary randomly across treatments but remained constant across the 10 replicates within each treatment. For all simulations, the following parameter values remained fixed: $\mu = 0.0001$, $m = 0.01$, $K = 200$, $\rho = 0.1$, $n_D = 100$, $n_N = 80$, $n_{FD} = 12$, and $n_B = 4$ (for a total of 200 loci). Unless otherwise specified, the parameters used are given in boldface type. Recombination rate scenarios are shown graphically in figure 1.

where $\phi(x, s, h, \mu, N_e)$ is Wright's distribution for mutation-selection-drift equilibrium. We initialize the simulations assuming $N_e = n_{core} \times K$ as an approximation for the effective population size in the range core. As shown in figure S1, burn-in simulations indicated that this definition of V_0 provides an equilibrium deme size near K.

The absolute fitness, W_i, and the relative fitness, w_i, of an individual with genotype i is determined by its fertility, V_i, in the following manner:

$$W_i = \left(1 + \rho\left(1 - \frac{N_T}{K}\right)\right) V_i,$$
$$w_i = \frac{V_i}{\bar{V}},$$

where $\bar{V}$ is the mean population fertility. Even though the absolute fitness, W_i, depends on the local population size, N_T, an individual's relative fitness, w_i, does not. We will use V_i as a proxy for fitness throughout, as it is proportional to absolute fitness and determines relative fitness.

Parents of genotype i produce a random Poisson-distributed number of gametes with mean $2W_i$. Given the previously identified importance of recombination to both the dynamics of expansion (Peischl et al. 2015) and genome-wide patterns of introgression (Noor and Bennett 2009), we consider five different recombination scenarios (fig. 1). In the first two cases, the recombination rate is uniform across the genome with either free ($r = 0.5$) or constrained ($r = 0.005$) recombination between any two consecutive loci. In the third case, the genome is divided into four segments of equal length with recombination occurring freely, $r = 0.5$, between them at three recombination "hot spots" and no recombination within each segment. This third scenario resembles the genomic structure used by Peischel et al. (2015) and can represent either highly heterogeneous recombination on a single chromosome or independent segregation among four otherwise nonrecombining chromosomes. The resulting genome-wide average recombination rate in this case is given by $\bar{r} = 0.0075$. In the fourth and fifth scenarios we model the evolution of a single chromosome where the recombination rate varies continuously across the genome as given by the average pattern observed across a wide range of plant and animal taxa, respectively (see Haenel et al. 2018, fig. 1, table 1). The genome-wide average recombination rate in these latter two scenarios was $\bar{r} = 0.005$ to facilitate comparison to the uniformly constrained and chromosomal segregation cases. Given this average rate, there is an expected ≈1 crossover events per genome per generation, on par with expectations for many diploid organisms (Stapley et al. 2017).

In addition to recombination, gametes carry novel mutations. The per-generation per-site mutation rate, with both forward and backward mutations allowed, was set to $\mu = 10^{-4}$ at both the neutral and the selected loci. To simplify the tracking of reproductive isolation, we do not allow mutation at the BDMI loci. To provide an analogous neutral comparison, mutation is also suppressed at the 12 fixed-difference loci. Following meiosis, mating is random with gametes combining randomly to produce diploid offspring. These offspring then migrate at a rate $m = 0.01$ to neighboring demes, as described in more detail below. Following migration, each population is censused, completing the generation.

We consider the demographic model depicted in figure 1, modeling first evolution in allopatry of populations A and B, followed by their expansion, and finally introgression between the two populations initially occupying opposite ends of a finite linear "stepping-stone" habitat. The model was coded in C++ and proceeds through the following stages (program and simulation results have been deposited in the Dryad Digital Repository; https://doi.org/10.5061/dryad.0vt4b8h07; MacPherson et al. 2022).

Initialization. The simulations begin by initializing the habitat with the leftmost n_{core} demes occupied by population A and the rightmost n_{core} populations occupied by population B. Individuals migrate among demes such

that they move to the left (right) neighboring deme with probability $m/2$ ($m = 0.01$ is used throughout). The leftmost and rightmost demes have reflective population boundaries such that migrants moving to the left (right) remain instead within the patch. These demes will be referred to as the core of populations A and B, respectively. We initialize each core deme with K individuals, drawing the initial allele frequencies at the selected loci using Wright's distribution (see eq. [2]; Crow and Kimura 1970). This initialization is an approximation to the steady state, ensuring that the burn-in phase efficiently reaches equilibrium.

Burn-in. For the core populations to reach the true eco-evolutionary equilibrium, we begin by simulating evolution for 10,000 burn-in generations. To test the quality of the burn-in, we evaluate the dynamics of the mean fertility $\bar{V}$ and the population size of each deme (see fig. S1).

Expansion. The third stage of the simulations is characterized by the expansion of parental populations A and B while remaining allopatric. Expansion proceeds for 2,000 generations. Each generation's individuals are allowed to migrate to the nearest neighboring demes in a stepping-stone manner. As described above, the geographic extent (the absolute x-axis distance covered) over the course of the 2,000 generations of expansion varies stochastically across simulations. During the expansion phase, the extent of the linear habitat was effectively infinite, such that no parental population reaches an opposite boundary during this time period. The total number of demes occupied by the two parental populations at the time of secondary contact was on average 80.5 ± 14.2. Although artificial, constraining expansion to a fixed time rather than distance allowed us to interchange parental populations, increasing the power of our analyses. To facilitate analysis of geographical patterns after expansion, we defined a normalized geographic distance, used for plotting and model fitting, for each simulation such that total range distance on secondary contact is 100 (e.g., see fig. 4*A*–4*C*).

To quantify the impact of allele surfing on population mean fitness, every 50 generations we measured expansion load, which is defined as the relative difference in mean fitness of range edge (i.e., the rightmost [leftmost] deme) and the mean fertility within the five demes of the range core. We defined the edge as a single deme because of the possibility of narrow reductions in fitness near the range edge (Peischl et al. 2015). By contrast, averaging fertility across the five core demes removes the impact of possible boundary effects. We then defined load as

$$\mathcal{L} = \frac{\bar{V}_{\text{core}} - \bar{V}_{\text{edge}}}{\bar{V}_{\text{core}}}. \tag{3}$$

Similarly, to quantify the effect of masking of deleterious recessive mutations in hybrids, every 50 generations we created an artificial population of 10 F_1 hybrids between the range edge of parental populations A and B. We then calculated heterosis by comparing mean F_1 hybrid fitness to the mean fitness of their parents:

$$\mathcal{H} = \bar{V}_{F_1} - \bar{V}_{\text{Par}}, \tag{4}$$

Introgression. Following expansion, we considered the dynamics of introgression between population A and population B allowing all neighboring populations to be connected by migration, as shown in figure 1. Following secondary contact, introgression proceeded for 5,000 generations. In this phase, we stopped mutation ($\mu = 0$). Although artificial, this ensured that all results observed during introgression are a result of expansion history and not de novo mutation during introgression itself.

Following secondary contact, we can use the allele frequencies at the n_N neutral loci that are not constrained to be fixed to define a measure of population ancestry based on the hybrid index. Adapting the likelihood-based measure developed by Buerkle (2005), let the allele frequency at neutral locus i in the core of population A (the leftmost deme) be r_i and the allele frequency in the core of population B (the rightmost deme) be s_i. Then the probability of observing a focal allele given a hybrid index of h is given by

$$\tilde{p}_i(h) = hr_i + (1 - h)s_i. \tag{5}$$

Given the observed allele frequency in deme d, p_d, we calculated the hybrid index of deme d by maximizing the likelihood

$$\mathcal{L}(h_d|\vec{p}_d) = \prod_{l=1}^{L} \binom{2N_d}{2N_d p_{d,l}} \tilde{p}(h_d)^{2N_d p_{d,l}} (1 - \tilde{p}(h_d))^{2N_d(1-p_{d,l})}, \tag{6}$$

where N_d is the population size of deme d. As with the individual hybrid index developed by Buerkle (2005), h_d accounts for both fixed and nonfixed differences at neutral loci between core demes. As is common in empirical studies, however, we restricted our inference to neutral loci exhibiting a difference in allele frequency between the core populations of $\text{Abs}[r_i - s_i] \geq 0.3$ (Corbett-Detig and Nielsen 2017; Matute et al. 2020; Wang et al. 2021).

Results

The core parental populations A and B reached eco-evolutionary equilibrium during the burn-in period (figure S1). As expected during the subsequent range expansion phase, allele surfing greatly increased the frequency

of deleterious mutations on the range edge and hence the number of fixed differences between the edges of populations A and B (fig. 2*A*). The fixation of deleterious mutations on the range edge results in substantial expansion load, as measured by equation (3) (fig. S2). The selection and dominance coefficient of the deleterious alleles at the selected loci are important determinants of the extent of expansion load. In contrast, the dynamics of expansion load are insensitive to the recombination rate scenario, with nearly identical accumulation of load even in the case of complete linkage with chromosomal segregation.

As the number of fixed differences between the range edges of populations A and B increases, so too does the observed heterosis among F_1 hybrids created in crosses of edge populations. The temporal dynamics of F_1 heterosis (fig. 2*B*) follow our expectation based on the previously characterized dynamics of fitness on the edge of expanding populations (Peischl and Excoffier 2015). The extent of heterosis that results from the masking of (partially) recessive deleterious mutations that have surfed to fixation or near fixation on the range edge of one parental population but not the other depends both on the strength of selection on the deleterious alleles, s_D, and the dominance of these alleles, h_D (see fig. 2*B*). As expected, dominant BDMIs decrease the fitness of F_1 hybrids, yet heterosis from masking deleterious recessive mutations can partially or even completely compensate for the deleterious fitness effects of these incompatibilities (fig. 2*C*). Finally, heterosis from the masking of deleterious mutations in the F_1 generation is relatively insensitive to the recombination scenario (fig. 2*D*). That said, we would expect the rate of recombination to be important in determining the amount of heterosis in advanced-generation hybrids. Specifically, in regions of low recombination, chromosomes bearing deleterious mutations from one parent or the other will be maintained in linkage disequilibrium for longer. In such regions, there

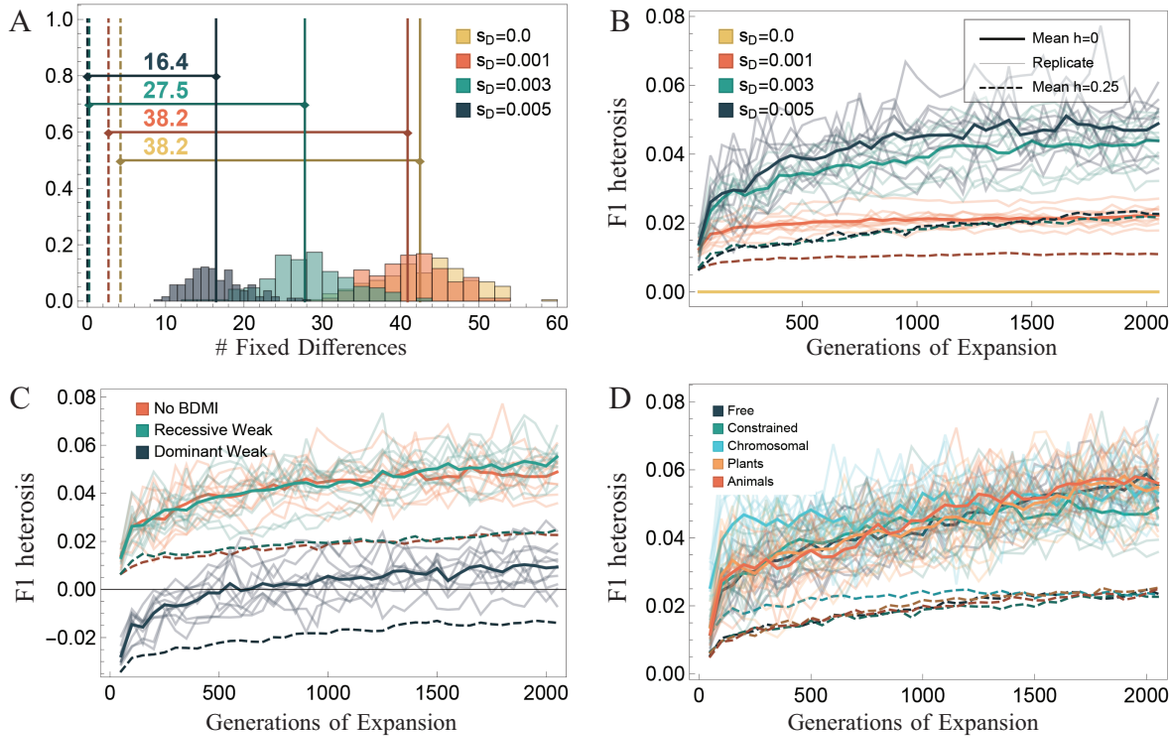


Figure 2: Impact of allele surfing on F_1 heterosis during range expansion. *A*, Number of selected loci fixed for different alleles between parental core populations at the end of evolution in allopatry (mean = dashed vertical line) verses parental edge populations at the end of expansion (histogram; mean = solid vertical line) for different strengths of selections. *B–D*, F_1 heterosis as measured by equation (4) in crosses between parental edge populations during range expansion. Dark solid curves give mean dynamics averaged across individual replicates (solid light curves) when deleterious loci are recessive ($h_D = 0$). Dark dashed lines give mean dynamics for a partially recessive case ($h_D = 0.25$). *B*, Effect of the strength of selection (case of $s_D = 0$ not shown) exhibits no heterosis, as expected. *C*, Effect of Bateson-Dobzhansky-Muller incompatibility (BDMI) presence/absence and dominance, shown here for the case of a weak BDMI ($s_B = 0.01$). *D*, Effect of recombination scenario. Unless specified otherwise, parameters are given by boldface values in table 1.

would be longer-lasting associative overdominance (Frydenberg 1963), contributing to the fitness of advanced generation hybrids and the likelihood of introgression.

The dynamics of population mean fitness during range expansion and on secondary contact are shown in figure 3. As with F_1 heterosis, on secondary contact there is a substantial increase in fitness in the hybrid zone, which expands from the former range edge toward the core as wild-type alleles masking accumulated deleterious alleles introgress. While the presence of BDMIs does not prevent the introgression of wild-type alleles, these incompatibilities result in a persistent reduction in population mean fitness at the point of secondary contact. Relative to dominant BDMIs, recessive incompatibilities result in a shallower and more diffuse reduction in fitness (fig. 3*B*, 3*C*).

Deleterious mutations fixed on the edge of one parental population and not the other are selected against following secondary contact (recall that the simulation design assumes that no mutations occur following secondary contact), favoring the ultimate fixation of the corresponding wild-type allele across the range. While the deleterious alleles persist, however, these loci will exhibit a cline in allele frequency across the hybrid zone. Unlike the sigmoidal clines observed in a tension zone (Barton 1979), however, the shape of the allele frequency cline is affected both by introgression and by historical allele surfing. If we consider all loci for which the deleterious allele has a frequency $p^{\mathrm{A}}_{\mathrm{edge}} < 0.5$ at the range edge of parental population A and $p^{\mathrm{B}}_{\mathrm{edge}} > 0.5$ at the edge of parental population B, then the cline in average allele frequency across these loci peaks just to the left of the contact zone (see fig. 4*A*).

To quantify the transient dynamics of these allele frequency clines at each type of locus (e.g., neutral, selected, fixed difference, BDMI) and their variation across parameter space, we fitted them phenomenologically with one of the following functions as determined by Akaike information criterion–based model selection:

$$\bar{p}_1(x) = \begin{cases} \dfrac{d}{b_{\mathrm{A}}}\mathrm{Sech}(a_{\mathrm{A}}(x-c))\mathrm{Tanh}(b_{\mathrm{A}}(x-c)) + e & \text{if } x < c, \\ \dfrac{d}{b_{\mathrm{B}}}\mathrm{Sech}(a_{\mathrm{B}}(x-c))\mathrm{Tanh}(b_{\mathrm{B}}(x-c)) + e & \text{if } x > c, \end{cases}$$
$$\bar{p}_2(x) = \frac{a}{1+\exp[(-4d/a)(x-c)]} + e - \frac{a}{2}. \tag{7}$$

These model fits are defined such that the geographic center of the cline is given by c and height e, with the slope

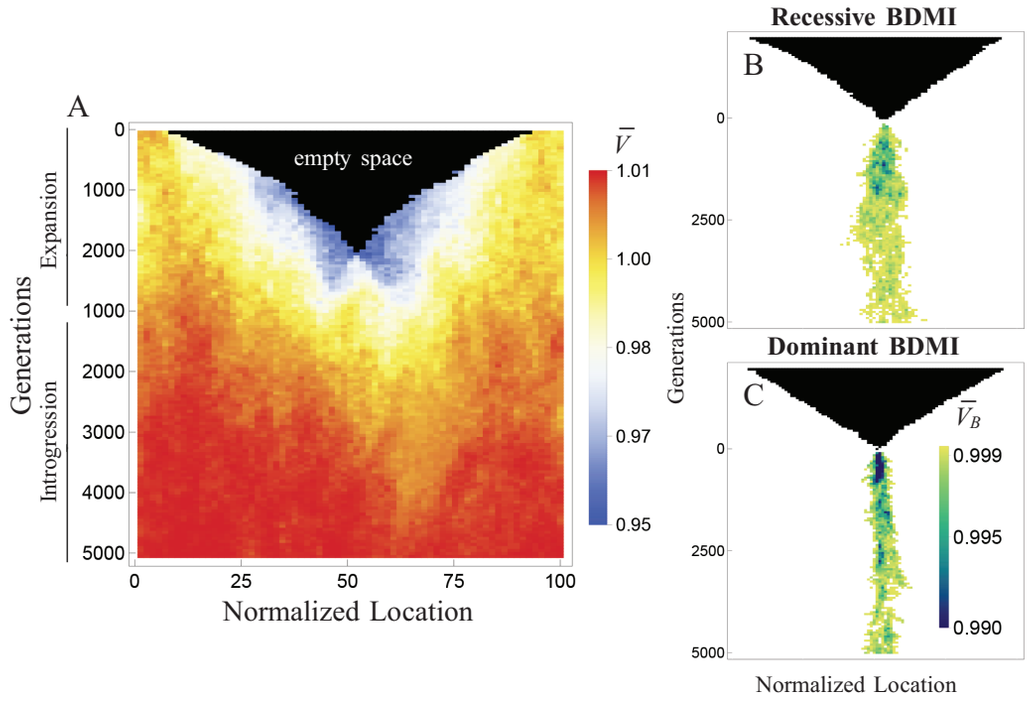


Figure 3: Dynamics of mean fitness and the effect of Bateson-Dobzhansky-Muller incompatibilities (BDMIs) on secondary contact. *A*, Exemplary dynamics of mean population fertility $\bar{V}$ during range expansion and introgression in the absence of a BDMI (see boldface parameters in table 1). *B*, *C*, Fitness effect of BDMI as measured by $\bar{V}_{\mathrm{B}}$ in equation (1) for a weak recessive (*B*) and weak dominant (*C*) BDMI.

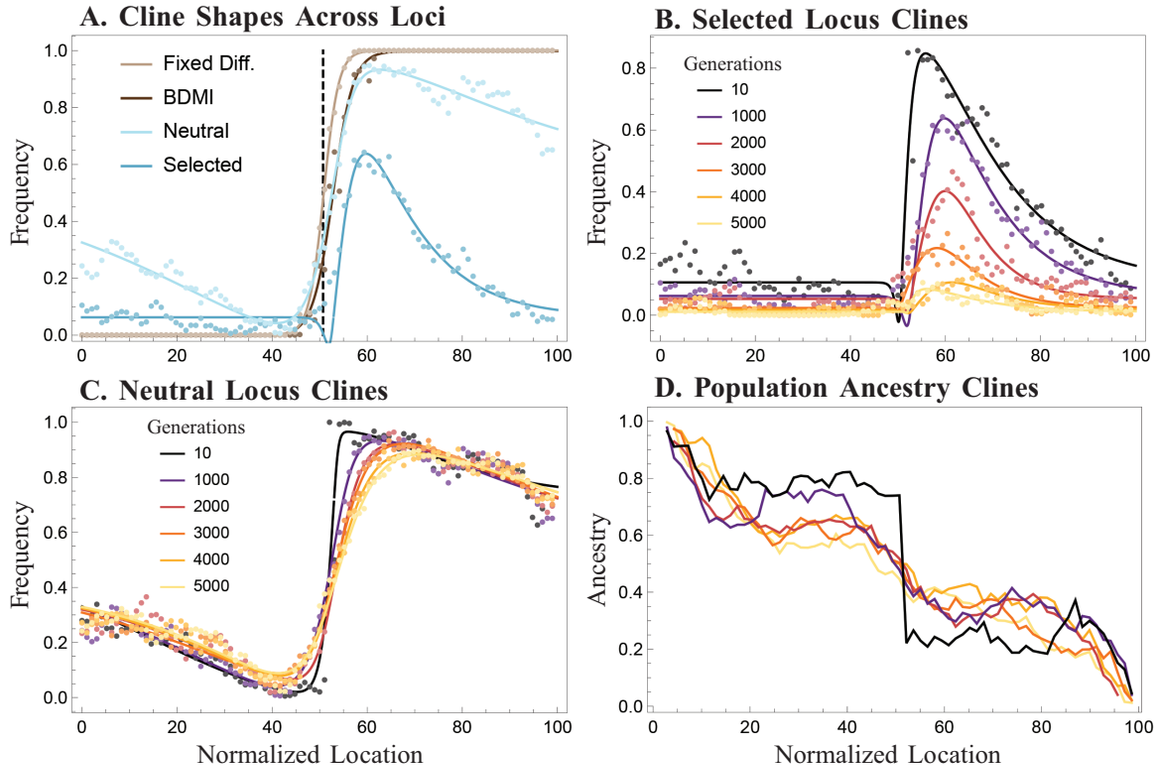


Figure 4: Clinal dynamics and population ancestry. Clines in average allele frequency are calculated by averaging across loci of a specific type (e.g., neutral, selected, fixed difference, Bateson-Dobzhansky-Muller incompatibility [BDMI]) for which $p^{A}_{edge} < 0.5$ and $p^{B}_{edge} > 0.5$. *A*, Shape of the allele frequency cline for each locus type 1,000 generations after secondary contact. Brown versus blue clines exemplify the difference in fit between Sech ($\bar{p}_1$) and sigmoidal ($\bar{p}_2$) curves given by equation (7). *B*, Temporal dynamics of clines at the selected loci throughout introgression. *C*, Temporal dynamics of allele frequency clines at neutral loci following secondary contact. *D*, Temporal dynamics of the cline in population ancestry, as defined by hybrid index in equation (6). See *C* for color legend. Parameters are given by boldface values in table 1.

(derivative) of the cline at this point by d. Following the definition of Barton and Hewitt (1989), the width of the cline is given by $W = |1/d|$. For $\bar{p}_1(x)$, the value of a_A (a_B) determines the slope in allele frequency from the range core to its maximum (minimum) near the point of secondary contact, as determined by the extent of allele surfing. The remaining parameters (b_A, b_B) determine the height and symmetry of the cline. By contrast, $\bar{p}_2(x)$ describes a sigmoidal function, as observed when parental population have fixed differences. The height of this symmetric sigmoid is given by a. The functions $\bar{p}_1$ versus $\bar{p}_2$ were developed to match the shape of the allele frequency clines in the case of the neutral and selected versus the fixed difference and BDMI clines, respectively (see fig. 4*A*), although neither model is preferred exclusively in either of these cases across all parameter conditions.

We next focus on the dynamics of cline width at the selected and BDMI loci. As introgression proceeds, the width of the cline increases at both the selected (fig. 5*A*) and BDMI loci (fig. 5*B*), although less so for the BDMI loci, as expected. Because of selection favoring the introgression of wild-type mutations, the cline width at the selected loci increases with the strength of selection s_D for the parameters explored. Interestingly, the cline width at the selected loci is less than that of the corresponding neutral loci. This is a result of the low frequency of the deleterious alleles in the parental range core and the subsequent spread of alleles from the range core toward the hybrid zone following secondary contact. Essentially, selective recovery from deleterious allele surfing occurs from two directions: from secondary contact with the other range-edge population and from the core of the same population. For further details, see supplement A.

The cline width at the BDMI loci decreases with increasing strength of selection against the BDMI incompatibilities s_B, with the width of BDMI clines less than

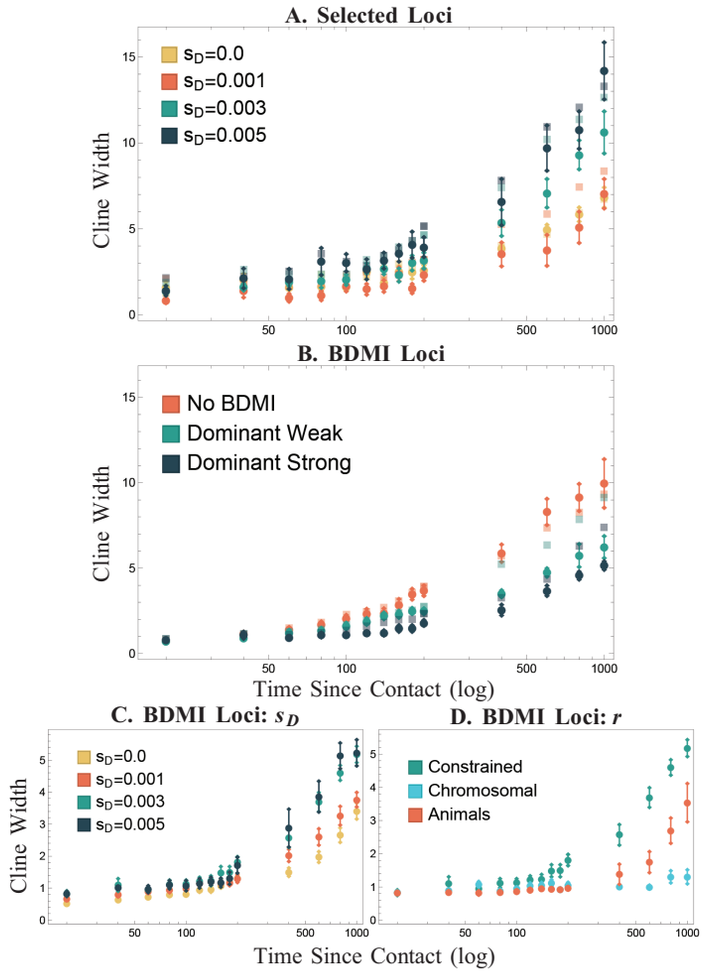


Figure 5: Early temporal dynamics and determinants of cline width at selected and Bateson-Dobzhansky-Muller incompatibility (BDMI) loci. *A*, Temporal dynamics of cline width at selected loci. Shown is the effect of the strength of selection s_D on cline width of selected loci. Dark points give mean width across replicates with 1 SE whiskers. Opaque squares show mean cline width at neutral loci in the same genome. *B*, Cline width at BDMI loci (dark points show mean width with 1 SE whiskers). Opaque squares show mean for fixed difference loci in the same genome. *C*, Effect of the strength of selection on linked deleterious loci on cline width of BDMI loci. *D*, Effect of recombination scenario on cline width at BDMI loci. The plants case is very similar to that of animals and is therefore not shown for simplicity. Unless specified otherwise, parameters are given by boldface values in table 1.

the clines at the neutral fixed-difference loci. Cline width at the BDMI loci is also impacted by the strength of linked selection against deleterious alleles (s_D; see fig. 5*C*), with strong linked selection facilitating introgression, at least over short timescales (0–1,000 generations of introgression). Finally, the cline width at the BDMI loci is impacted by the distribution of recombination across the chromosome(s). When associative overdominance is absent (fig. 5*D*; $s_D = 0$), the more uniform recombination is, the more the BDMI loci introgress. This is consistent with the more efficient breakdown of linkage between multiple BDMIs. In contrast, when selection is strong ($s_D = 0.005$), long-lasting associative overdominance maintained in the case of chromosomal segregation can lead to extensive introgression at the BDMI loci (see fig. S4).

Finally, we characterize the population ancestry across the geographic range following secondary contact using the population-wide hybrid index at the n_N neutral loci that were not constrained to be fixed (eq. [6]). The cline in population ancestry is not sigmoidal, as would be expected in the tension zone model. This is a result of the nonsigmoidal shape of the underlying neutral loci with which ancestry is measured. Affected by allele surfing during range expansion, the cline in population

ancestry is gradual, showing a less marked shift in ancestry at the time of secondary contact than in a traditional tension zone (black line in fig. 4*D*). This gradual change in ancestry across the range persists over the course of introgression, obscuring the location of initial contact and making the hybrid zone appear older than it is. These results suggest that incorporating historical range expansion is important for the accurate reconstruction of the biogeography of hybrid zones, even at neutral loci.

Discussion

Here, we investigated the effect of parental population range expansion and allele surfing prior to secondary contact on hybridization and introgression. As found by previous models of range expansion (e.g., Peischl et al. 2013; Peischl et al. 2015; Peischl and Excoffier 2015), we find that the edge of two expanding parental populations can accumulate a significant number of deleterious mutations and hence expansion load. If these deleterious alleles are at least partially recessive, hybridization of range-edge populations results in extensive heterosis, as wild-type alleles present in one population mask the effects of deleterious mutations accumulated in the other. This masking effect can compensate for the reduction in hybrid fitness due to genetic incompatibilities (e.g., BDMI) between parental populations facilitating introgression over generations after the initial contact. We further uncovered selective interference between the loci with surfed alleles and BDMI alleles in shaping introgression across species boundaries. The strength of the interference is dependent on the selection against surfed alleles and the distribution of the recombination rate across the genome. This observation sheds light on many unexplained empirical patterns of hybridization and introgression (table S1).

In this model, we have considered only a small subset of the possible selection regimes acting in postexpansion secondary contact hybrid zones. Namely, we have not explored the evolution of loci experiencing environment-dependent or underdominant selection. In addition, we have assumed that each parental population is fixed for a particular combination of compatible BDMI genotypes. In reality, the genotypes at these loci are also subject to allele surfing during expansion, which could alter the prevalence of incompatible genotypes among hybrids and the permeability of these loci during introgression.

Hybridization and Hybrid Zones

Our results may help explain previously reported patterns seen in hybrids and hybrid zones. One commonly observed pattern is variation among cross combinations (both within and between species) in the level of heterosis among hybrid offspring (Lowry et al. 2008). This variation has been exploited by plant breeders to develop heterotic groups for hybrid crop production (Crow 1998). Despite its importance, the extent of heterosis remains difficult to measure and predict in natural populations (Pickup et al. 2013). Results from the present study imply that the history of the parental populations may be key. If divergence occurs in large, adjacent populations, little masking of recessive deleterious alleles is expected on contact, whereas divergence in distant locations followed by range expansion with extreme drift, such as explored here, creates genomic conditions that facilitate heterosis and introgression via the masking of deleterious recessive alleles that counteract genetic incompatibilities.

Another partially unexplained phenomenon is the relatively high frequency of introgressed regions that appear to be positively selected in hybrid populations (Rieseberg et al. 1999; Hamilton et al. 2013; Corbi et al. 2018). This could be a by-product of adaptation in finite populations, such that some favorable alleles were previously in only one of the parental populations (Barton 2001) or the environment may be different in the hybrid zone center, favoring a different combination of alleles. Alternatively, it seems likely that a substantial fraction of favorable introgressed regions are due to the masking of deleterious mutations, as shown here. For example, ancestral range expansion of *Anopheles gambiae* mosquitoes and concomitant accumulation of deleterious load might have facilitated the introgression of an inversion from *A. arabeinsis* (Powell et al. 1999; Besansky et al. 2003; Sharakhov et al. 2006; White et al. 2007).

The spatial distribution of population ancestry is usually modeled as a sigmoidal cline, the expected shape assuming selection-dispersal balance in the tension zone model (Barton 1979). However, here we found distorted shapes of genomic clines (fig. 4*D*) as a result of allele surfing. This pattern might be difficult to detect in many empirical systems if the spatial genetic sampling within parental ranges was limited. For future investigations in hybrid zones, our results emphasize the importance of broad spatial sampling centered around the core of admixture in detecting signatures and effects of allele surfing in hybrid zone ancestry dynamics.

Genomic Architecture of Speciation

The effects of recombination rate heterogeneity in this expansion hybridization context alters the conditions under which speciation can be impeded by introgression. Here, we find elevated introgression at the BDMI loci due to hitchhiking with flanking loci, where surfed deleterious alleles were masked by heterospecific alleles.

Interestingly, such interference-facilitated BDMI introgression is modulated by the distribution of recombination rates, holding the mean magnitude of recombination rate constant. Observed variation in cline width among simulation replicates with the same recombination map, particularly for the case of chromosomal segregation, may be due in part to the variation of BDMI loci along the chromosome.

Our observation extends the existing understanding of the effect of recombination on introgression by accounting for the distribution (in addition to the magnitude) of recombination and the interference among neutral loci, BDMI loci, and loci with surfed deleterious alleles during parental range expansion. Both allele surfing and introgression are expected to be shaped by the magnitude and distribution of recombination events across the genome. In the expansion context, constrained recombination can reduce absolute fitness at the range edge, suppressing population sizes and even resulting in substantial reductions in expansion rate (Peischl et al. 2015). In the introgression context, local recombination suppression due to chromosomal rearrangements can prevent introgression and facilitate differentiation (Noor et al. 2001; Rieseberg 2001; Navarro and Barton 2003).

In contrast to previous results (Peischl and Excoffier 2015), we find that recombination has relatively little impact on the accumulation of expansion load (fig. S2) and thus F_1 heterosis (fig. 2*D*), even in the case of chromosomal segregation. The contrast between these results may indicate that only small amounts of recombination are necessary to dramatically reduce selective interference during range expansion. However, for introgression we find that genetic linkage and the distribution of recombination have large impacts on introgression, effects that can persist for hundreds of generations after the onset of hybridization (fig. 5). Specifically, introgression at neutral loci is facilitated by selection at linked selected loci. Even the introgression at BDMI loci is elevated under stronger selection at those deleterious loci (fig. 5). On the other hand, heterogeneity in the recombination rate suppresses introgression at BDMI loci, helping maintain barriers to hybridization. Therefore, the expansion-introgression dynamics could be variable in organismal systems with different levels of recombination heterogeneity in the genomes. These results highlight how the recombination landscape and range expansion interact to shape genomic ancestry at species boundaries.

Hybrid Speciation

Our results also have implications for long-term outcomes of natural hybridization. Most obviously, our simulations indicate that levels and heterogeneity of introgression may be greater than that predicted by standard hybrid zone models (Barton and Hewitt 1985). Extensive masking of deleterious alleles could contribute to the weakening of reproductive barriers, sustained hybrid swarms/species, or even the fusion of previously isolated populations and reduce the strength of reinforcement. This leads to the interesting prediction that reinforcement might be more likely when populations have diverged in adjacent areas rather than following extensive range expansion. This differs from the classic scenario that ignores allele surfing, in which reinforcement "completes speciation" following range expansion and contact between previously allopatric populations (Dobzhansky 1937).

On the other hand, expansion contact–induced heterosis might sustain hybrid populations for extended periods of time, allowing hybrid populations to diverge from parental lineages under divergent selection (Rieseberg 1997; Buerkle et al. 2000). One compelling example resides in a North American wood warblers species complex, where hybrid populations existed long after postexpansion hybridization between the parental species (Krosby and Rohwer 2009; Wang et al. 2021). Heterosis could sustain an ancient hybrid population despite genetic incompatibilities between the parental populations. For example, in the *Setophega* warbler complex (Wang et al. 2021), hybrid populations have persisted and diverged from parental lineages at genomic regions related to climatic adaptation. The expansion hybridization–resulted fusion, hybrid swarms, or hybrid species could be prevalent yet hidden in many extant populations with histories of historical range expansion prior to admixture.

Genomic Invasion

The masking effect of postexpansion hybridization can also facilitate introgression of native species' genetic material between native and invasive species. In this case, if invasive species have undergone the most recent range expansion, deleterious alleles may have surfed to higher frequencies, facilitating introgression from stable native populations. Here, we did not model invasion explicitly (asymmetry of expansion, in which the invasive species expanded its range while the native species did not). However, we might expect masking of deleterious alleles in the expanding invasive species following hybridization with its native congener. Such hybridization-induced masking could be the genomic mechanism underlying hybridization-accelerated anthropogenic invasions, including the process in which modern humans replaced Neanderthals (Mallet 2005; Mesgaran et al. 2016). Ironically, native species could potentially rescue invaders that are ecologically detrimental to them. Our result raises questions about the conservation of native populations

in the face of anthropogenic invasion of closely related and rapidly expanding invasive populations.

Acknowledgments

We thank Steven Wright, Darren Irwin, and Brett Payseur for their helpful feedback in the development of this project and Matt Pennell for his helpful comments on the manuscript. A.M. was supported by a fellowship from the University of British Columbia and an Ecology and Evolutionary Biology Department Postdoctoral Fellowship from the University of Toronto. The authors received additional support for this work from the Natural Sciences and Engineering Research Council of Canada (PGS D 331015731 to S.W. and NSERC RGPIN-2016-03711 to S.P.O.).

Statement of Authorship

This work was conceptualized by S.W., A.M., and R.Y. A.M. coded the simulations and performed the data analysis. A.M. and S.W. lead the writing, editing, and reviewing. L.H.R. and S.P.O. provided conceptual feedback and supervision and assisted in writing and revising the manuscript.

Data and Code Availability

Program and simulation results have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.0vt4b8h07; MacPherson et al. 2022).

Literature Cited

Aeschbacher, S., J. P. Selby, J. H. Willis, and G. Coop. 2017. Population-genomic inference of the strength and timing of selection against gene flow. Proceedings of the National Academy of Sciences of the USA 114:7061–7066.

April, J., R. H. Hanner, A.-M. Dion-Côté, and L. Bernatchez. 2013. Glacial cycles as an allopatric speciation pump in northeastern American freshwater fishes. Molecular Ecology 22:409–422.

Arntzen, J. W., W. de Vries, D. Canestrelli, and I. Martínez-Solano. 2017. Hybrid zone formation and contrasting outcomes of secondary contact over transects in common toads. Molecular Ecology 26:5663–5675.

Avise, J. C., D. Walker, and G. C. Johns. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. Proceedings of the Royal Society B 265:1707–1712.

Barton, N. H. 1979. The dynamics of hybrid zones. Heredity 43:341–359.

———. 2001. The role of hybridization in evolution. Molecular Ecology 10:551–568.

Barton, N. H., and K. Gale. 1993. Genetic analysis of hybrid zones. Pages 13–45 *in* R. Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, Oxford.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology, Evolution, and Systematics 16:113–148.

———. 1989. Adaptation, speciation and hybrid zones. Nature 341:497–503.

Besansky, N. J., J. Krzywinski, T. Lehmann, F. Simard, M. Kern, O. Mukabayire, D. Fontenille, Y. Touré, and N. Sagnon. 2003. Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. Proceedings of the National Academy of Sciences of the USA 100:10818–10823.

Buerkle, C. A. 2005. Maximum-likelihood estimation of a hybrid index based on molecular markers. Molecular Ecology Notes 5:684–687.

Buerkle, C. A., R. J. Morris, M. A. Asmussen, and L. H. Rieseberg. 2000. The likelihood of homoploid hybrid speciation. Heredity 84:441–451.

Corbett-Detig, R., and R. Nielsen. 2017. A hidden Markov model approach for simultaneously estimating local ancestry and admixture time using next generation sequence data in samples of arbitrary ploidy. PLoS Genetics 13:e1006529.

Corbi, J., E. J. Baack, J. M. Dechaine, G. Seiler, and J. M. Burke. 2018. Genome-wide analysis of allele frequency change in sunflower crop–wild hybrid populations evolving under natural conditions. Molecular Ecology 27:233–247.

Crow, J. F. 1998. 90 years ago: the beginning of hybrid maize. Genetics 148:923–928.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Blackburn, Caldwell, NJ.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Duvernell, D. D., E. Westhafer, and J. F. Schaefer. 2019. Late Pleistocene range expansion of North American topminnows accompanied by admixture and introgression. Journal of Biogeography 46:2126–2140.

Edmonds, C. A., A. S. Lillie, and L. L. Cavalli-Sforza. 2004. Mutations arising in the wave front of an expanding population. Proceedings of the National Academy of Sciences of the USA 101:975–979.

Endler, J. A. 1977. Geographic variation, speciation, and clines. Princeton University Press, Princeton, NJ.

Excoffier, L., M. Foll, and R. J. Petit. 2009. Genetic consequences of range expansions. Annual Review of Ecology, Evolution, and Systematics 40:481–501.

Frydenberg, O. 1963. Population studies of a lethal mutant in *Drosophila melanogaster*. Hereditas 50:89–116.

Gao, Y.-D., Y. Zhang, X.-F. Gao, and Z.-M. Zhu. 2015. Pleistocene glaciations, demographic expansion and subsequent isolation promoted morphological heterogeneity: a phylogeographic study of the alpine *Rosa sericea* complex (Rosaceae). Scientific Reports 5:11698.

Haenel, Q., T. G. Laurentino, M. Roesti, and D. Berner. 2018. Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. Molecular Ecology 27:2477–2497.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science 165:131–137.

Hamilton, J. A., C. Lexer, and S. N. Aitken. 2013. Differential introgression reveals candidate genes for selection across a spruce (*Picea sitchensis* × *P. glauca*) hybrid zone. New Phytologist 197:927–938.

Hewitt, G. M. 1999. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68:87–112.

Janoušek, V., P. Munclinger, L. Wang, K. C. Teeter, and P. K. Tucker. 2015. Functional organization of the genome may shape the species boundary in the house mouse. Molecular Biology and Evolution 32:1208–1220.

Juric, I., S. Aeschbacher, and G. Coop. 2016. The strength of selection against Neanderthal introgression. PLoS Genetics 12: e1006340.

Kirkpatrick, M. 2017. The evolution of genome structure by natural and sexual selection. Journal of Heredity 108:3–11.

Klopfstein, S., M. Currat, and L. Excoffier. 2006. The fate of mutations surfing on the wave of a range expansion. Molecular Biology and Evolution 23:482–490.

Krosby, M., and S. Rohwer. 2009. A 2000km genetic wake yields evidence for northern glacial refugia and hybrid zone movement in a pair of songbirds. Proceedings of the Royal Society B 276:615–621.

Licona-Vera, Y., J. F. Ornelas, S. Wethington, and K. B. Bryan. 2018. Pleistocene range expansions promote divergence with gene flow between migratory and sedentary populations of *Calothorax* hummingbirds. Biological Journal of the Linnean Society 124:645–667.

Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. Philosophical Transactions of the Royal Society B 363:3009–3021.

MacPherson, A., S. Wang, R. Yamaguchi, L. H. Rieseberg, and S. P. Otto. 2022. Data from: Parental population range expansion before secondary contact promotes heterosis. American Naturalist, Dryad Digital Repository, https://doi.org/10.5061/dryad.0vt4b8h07.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in Ecology and Evolution 20:229–237.

Martin, S. H., J. W. Davey, C. Salazar, and C. D. Jiggins. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biology 17:e2006288.

Matute, D. R., A. A. Comeault, E. Earley, A. Serrato-Capuchina, D. Peede, A. Monroy-Eklund, W. Huang, C. D. Jones, T. F. C. Mackay, and J. A. Coyne. 2020. Rapid and predictable evolution of admixed populations between two *Drosophila* species pairs. Genetics 214:211–230.

May, R. M., J. A. Endler, and R. E. McMurtrie. 1975. Gene frequency clines in the presence of selection opposed by gene flow. American Naturalist 109:659–676.

Mesgaran, M. B., M. A. Lewis, P. K. Ades, K. Donohue, S. Ohadi, C. Li, and R. D. Cousens. 2016. Hybridization can facilitate species invasions, even without enhancing local adaptation. Proceedings of the National Academy of Sciences of the USA 113:10210–10214.

Navarro, A., and N. H. Barton. 2003. Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evolution 57:447–459.

Noor, M. A. F., and S. M. Bennett. 2009. Islands of speciation or mirages in the desert? examining the role of restricted recombination in maintaining species. Heredity 103:439–444.

Noor, M. A. F., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Proceedings of the National Academy of Sciences of the USA 98:12084–12088.

Peischl, S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier. 2013. On the accumulation of deleterious mutations during range expansions. Molecular Ecology 22:5972–5982.

Peischl, S., and L. Excoffier. 2015. Expansion load: recessive mutations and the role of standing genetic variation. Molecular Ecology 24:2084–2094.

Peischl, S., M. Kirkpatrick, and L. Excoffier. 2015. Expansion load and the evolutionary dynamics of a species range. American Naturalist 185:E81–E93.

Pickup, M., D. L. Field, D. M. Rowell, and A. G. Young. 2013. Source population characteristics affect heterosis following genetic rescue of fragmented plant populations. Proceedings of the Royal Society B 280:20122058.

Powell, J. R., V. Petrarca, A. della Torre, A. Caccone, and M. Coluzzi. 1999. Population structure, speciation, and introgression in the *Anopheles gambiae* complex. Parassitologia 41:101–113.

Rieseberg, L. H. 1997. Hybrid origins of plant species. Annual Review of Ecology and Systematics 28:359–389.

———. 2001. Chromosomal rearrangements and speciation. Trends in Ecology and Evolution 16:351–358.

Rieseberg, L. H., J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics 152:713–727.

Roux, C., C. Fraïsse, J. Romiguier, Y. Anciaux, N. Galtier, and N. Bierne. 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. PLoS Biology 14: e2000234.

Schumer, M., C. Xu, D. L. Powell, A. Durvasula, L. Skov, C. Holland, S. Sankararaman, P. Andolfatto, G. G. Rosenthal, and M. Przeworski. 2018. Natural selection interacts with the local recombination rate to shape the evolution of hybrid genomes. Science 360:656–660.

Servedio, M. R., and M. A. Noor. 2003. The role of reinforcement in speciation: theory and data. Annual Review of Ecology, Evolution, and Systematics 34:339–364.

Sharakhov, I. V., B. J. White, M. V. Sharakhova, J. Kayondo, N. F. Lobo, F. Santolamazza, A. della Torre, F. Simard, F. H. Collins, and N. J. Besansky. 2006. Breakpoint structure reveals the unique origin of an interspecific chromosomal inversion (2La) in the *Anopheles gambiae* complex. Proceedings of the National Academy of Sciences of the USA 103:6258–6262.

Slatkin, M. 1973. Gene flow and selection in a cline. Genetics 75:733–756.

Stapley, J., P. G. D. Feulner, S. E. Johnston, A. W. Santure, and C. M. Smadja. 2017. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. Philosophical Transactions of the Royal Society B: Biological Sciences 372:20160455.

Szymura, J. M. 1976. Hybridization between discoglossid toads *Bombina bombina* and *Bombina variegata* in southern Poland as revealed by the electrophoretic technique. Journal of Zoological Systematics and Evolutionary Research 14:227–236.

Taberlet, P., L. Fumagalli, A.-G. Wust-Saucy, and J.-F. Cosson. 1998. Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology 7:453–464.

Turelli, M., and H. A. Orr. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663–1679.

Wang, S., M. J. Ore, E. K. Mikkelsen, J. Lee-Yaw, D. P. L. Toews, S. Rohwer, and D. Irwin. 2021. Signatures of mitonuclear

coevolution in a warbler species complex. Nature Communications 12:4279.

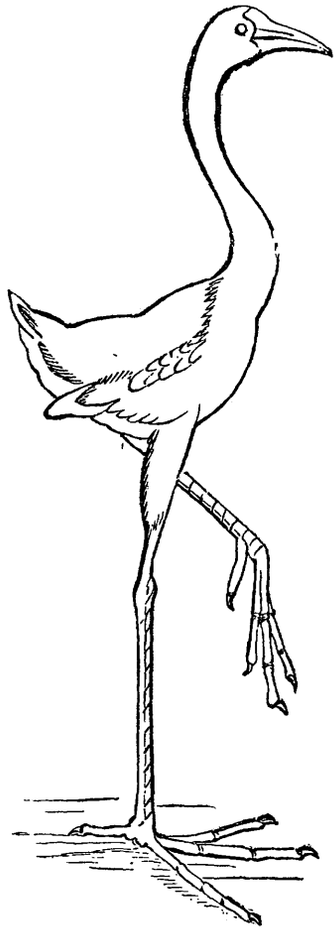
Weir, J. T., and D. Schluter. 2004. Ice sheets promote speciation in boreal birds. Proceedings of the Royal Society B 271:1881–1887.

White, B. J., M. W. Hahn, M. Pombi, B. J. Cassone, N. F. Lobo, F. Simard, and N. J. Besansky. 2007. Localization of candidate regions maintaining a common polymorphic inversion (2La) in *Anopheles gambiae*. PLoS Genetics 3:e217.

References Cited Only in the Online Enhancements

Weir, J. T., O. Haddrath, H. A. Robertson, R. M. Colbourne, and A. J. Baker. 2016. Explosive ice age diversification of kiwi. Proceedings of the National Academy of Sciences of the USA 113:E5580–E5587.

Associate Editor: Daniel R. Matute
Editor: Erol Akçay



"These singular birds characterizing the land fauna of these islands, of which Mauritius is the largest, seem like the gigantic birds of New Zealand, as Schlegel remarks, to have replaced the mammals, of which these two groups of islands are destitute." From "The Gigantic Birds of the Mascarene Islands" (*The American Naturalist*, 1868, 1:615).