THE ROYAL SOCIETY
PUBLISHING

The evolution of recombination rates in finite populations during ecological speciation

James Reeve, Daniel Ortiz-Barrientos and Jan Engelstädter

School of Biological Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia

DO-B, 0000-0002-7493-416X; JE, 0000-0003-3340-918X

Recombination can impede ecological speciation with gene flow by mixing locally adapted genotypes with maladapted migrant genotypes from a divergent population. In such a scenario, suppression of recombination can be selectively favoured. However, in finite populations evolving under the influence of random genetic drift, recombination can also facilitate adaptation by reducing Hill–Robertson interference between loci under selection. In this case, increased recombination rates can be favoured. Although these two major effects on recombination have been studied individually, their joint effect on ecological speciation with gene flow remains unexplored. Using a mathematical model, we investigated the evolution of recombination rates in two finite populations that exchange migrants while adapting to contrasting environments. Our results indicate a two-step dynamic where increased recombination is first favoured (in response to the Hill–Robertson effect), and then disfavoured, as the cost of recombining locally with maladapted migrant genotypes increases over time (the maladaptive gene flow effect). In larger populations, a stronger initial benefit for recombination was observed, whereas high migration rates intensify the long-term cost of recombination. These dynamics may have important implications for our understanding of the conditions that facilitate incipient speciation with gene flow and the evolution of recombination in finite populations.

1. Introduction

Ecological speciation is the evolution of reproductive isolation between populations adapting to contrasting environments [1–5]. Gene flow is a major impediment to this as it allows maladapted migrant alleles to infiltrate the local population [5–8]. A major factor determining the extent to which migrant alleles penetrate a population is recombination [9–11]. During ecological speciation with gene flow co-adapted gene complexes break down in recombined hybrids, resulting in chromosomes comprised of both adaptive and maladaptive alleles. Natural selection is not able to efficiently remove these maladaptive alleles because they are 'hidden' between adaptive alleles on the chromosome. As a result, alleles conferring local adaptation are expected to accumulate in regions of reduced recombination, such as population-specific chromosomal inversions [12–15].

The presence of maladaptive gene flow between populations is predicted to favour the suppression of recombination within populations (reviewed in [16]). Mathematical models aimed at understanding the evolution of sex demonstrate that a modifier allele reducing the recombination rate between selected loci (the modifier approach [17]) is generally favoured in the presence of maladaptive gene flow [18–20]. This can be interpreted as a mechanism of reinforcement [7,16] because reduced recombination rates will promote the genetic clustering of diverging populations. Similarly, models investigating the origin of species have shown that inversions that capture locally beneficial alleles should also be favoured by natural selection and thus spread within a population [14,21].

Empirical support for these theoretical predictions has only recently accumulated. For example, two host races of the North American apple maggot

(*Rhagoletis pomonella*) diverged following the spread of a chromosomal inversion due to natural selection in one host race [22,23]. Additional evidence comes from the *Drosophila* genus where closely related sympatric species are more often genetically separated by inversions than their allopatric relatives [24]. More broadly, chromosomal inversions have been implicated in many cases of divergence with gene flow, including sunflowers [25], monkeyflowers [26] and wild rabbits [27]. The capturing of adaptive loci in genomic regions of reduced recombination may be one mechanism that triggers the speciation process [12]. Overall, both theoretical and empirical data indicate that recombination antagonizes speciation, and that there is selection for mechanisms suppressing recombination in the presence of maladaptive gene flow that facilitate both the origin and maintenance of species.

Nevertheless, mechanisms increasing the rate of recombination can also be favoured in local populations adapting to new environments, as a consequence of Hill–Robertson interference [28,29]. This causes selection at one locus to reduce the efficacy of natural selection at another linked locus. This interference relies on stochastic processes (drift and mutation) operating in finite populations and hence is absent in idealized, infinitely large populations [28,29]. Linked modifier alleles that increase recombination between the selected loci will reduce Hill–Robertson interference and can spread through genetic hitchhiking [30–35]. A number of evolution experiments have demonstrated an advantage for sexual reproduction in adapting populations implicating variants of the Hill–Robertson effect as a cause [36–40], but definite evidence is still lacking.

Recombination thus would appear to have contrasting effects during ecological speciation with gene flow: on the one hand, recombination reduces Hill–Robertson interference and may therefore accelerate local adaptation, but on the other hand, recombination homogenizes diverging populations and thereby reduces local adaptation. As a consequence, modifier alleles that increase recombination rates can be favoured during local adaptation because of the Hill–Robertson effect [30–35], but disfavoured because of the maladaptive gene flow effect [18–20]. However, to date these two effects have only been investigated in isolation. A better understanding of the relative contribution and timing of these mechanisms when operating jointly may afford new insights into how recombination rate evolution can affect the early stages of ecological speciation with gene flow [16].

Here, we investigate a mathematical model of two finite populations exchanging migrants while adapting to two contrasting environments. We are interested in the spread of modifier alleles that affect the recombination rate between the loci under direct selection. Specifically, we ask whether the Hill–Robertson effect prevents the evolution of recombination suppression during the early stages of ecological speciation with gene flow. As expected, our model shows that migration selects for reduced recombination whereas finite population sizes can produce selection for increased recombination. However, these two effects operate at different times: selection for recombination through the Hill–Robertson effect is pronounced during the early phase of adaptation, whereas selection against recombination through the maladaptive gene flow effect predominates once the two populations have diverged. Our results suggest that depending on the migration rate and size of the populations, recombination rate evolution may affect both the early and late stages of ecological speciation.

2. Material and methods

(a) The genetic setting

We consider two populations of haploid organisms experiencing contrasting selection pressures while still connected by migration. Within a population, individuals have two biallelic loci (**A** and **B**), with a different allele favoured in each environment (*A* and *B* in population 1, *a* and *b* in population 2). In addition, a biallelic modifier locus **M** alters the recombination rate between **A** and **B**, without being directly affected by selection [17,20,41]. The loci are arranged in the order **MAB**. The resulting genotype frequencies (p_i and q_i in populations 1 and 2, where $i \in \{Mab, MAb, MaB, MAB, mab, mAb, maB, mAB\}$) were predicted each generation through a series of five recursion equations corresponding to migration, selection, mutation, recombination and random genetic drift. All parameters and variables of the model are summarized in table 1.

(b) Migration, selection and mutation steps

The migration step follows the classic two-island model where each generation a fraction m of individuals are replaced by migrants from the other equally sized population

$$p_i^{\text{mig}} = (1 - m)p_i + q_i m \tag{2.1a}$$

and

$$q_i^{\text{mig}} = (1 - m)q_i + p_i m. \tag{2.1b}$$

After migrating, the genotype frequencies are selected in the local environment according to their fitness values (w_i) in the respective population

$$p_i^{\text{sel}} = \frac{p_i^{\text{mig}} w_{ip}}{\bar{w}_p} \tag{2.2a}$$

and

$$q_i^{\text{sel}} = \frac{q_i^{\text{mig}} w_{iq}}{\bar{w}_q}. \tag{2.2b}$$

Here, the average population fitness is $\bar{w}_p = \sum_i p_i w_{ip}$ in population 1, and $\bar{w}_q = \sum_i q_i w_{iq}$ in population 2. For fitness values, we assume that beneficial alleles confer a fitness of $1 + s$ relative to a fitness of 1 in genotypes with the other allele, where s is the selection coefficient in both populations. The genotype fitness is calculated by multiplying the values of its composite alleles (i.e. we assume that there is no epistasis).

The mutation step is a two-locus extension of the single-locus reversible mutation model. We assume that the loci **A** and **B** mutate at the same rate μ. Furthermore, to eliminate any effects not attributable to recombination, the modifier locus is assumed not to mutate. Thus, the genotype frequencies after mutation are

$$p_i^{\text{mut}} = (1 - \mu)^2 p_{i_1}^{\text{sel}} + \mu(1 - \mu)p_{i_2}^{\text{sel}} + \mu(1 - \mu)p_{i_3}^{\text{sel}} + \mu^2 p_{i_4}^{\text{sel}} \tag{2.3a}$$

and

$$q_i^{\text{mut}} = (1 - \mu)^2 q_{i_1}^{\text{sel}} + \mu(1 - \mu)q_{i_2}^{\text{sel}} + \mu(1 - \mu)q_{i_3}^{\text{sel}} + \mu^2 q_{i_4}^{\text{sel}}. \tag{2.3b}$$

Here, the subscripts i_1, i_2, i_3 and i_4 refer to different genotypes; with i_1 being the focal genotype, i_4 the genotype that differs at both the **A** and **B** loci, and i_2 and i_3 being the genotypes one mutation step away at the **A** or **B** locus. For example, when $i = MAB$, we have $i_1 = MAB$; $i_2 = MAb$; $i_3 = MaB$ and $i_4 = Mab$.

Table 1.

Table 1. The notation of parameters and variables used in this model. Unless otherwise stated all parameters apply to the **A** and **B** locus.

parameter	description	standard values
m	migration rate	0.01
s	selection coefficient	0.05
μ	mutation rate	1×10^{-5}
r_M	recombination rate determined by M allele	0.0005
r_m	recombination rate determined by m allele	0.005
ρ	recombination rate between **M** and **A** locus	0
N	population size	100 000
variable	**description**	
p_i	genotype frequencies in population 1[a]	
q_i	genotype frequencies in population 2[a]	

[a]The subscript i takes values $i \in \{Mab, MAb, MaB, MAB, mab, mAb, maB, mAB\}$.

(c) Recombination step

The effects of recombination were calculated each generation according to a mating table of the outcomes of all possible crosses, assuming there is random mating between all genotypes. Three different recombination rates are required to cover every potential cross: r_M, r_m and ρ. The parameters r_M and r_m determine recombination rates between the **A** and **B** locus, depending on the allelic state at the **M** locus. Recombination takes place during a brief period of diploidy immediately following syngamy. We assume additive effects of the M and m alleles so that for example, the recombination rate in Mm heterozygotes is equal to $r_M + r_m$. ρ is the recombination rate between **M** and **A**.

Recombination is incorporated into the model through an $8 \times 8 \times 8$ tensor **T**, whose elements T_{jki} give the fraction of offspring i that are generated from a cross between parents j and k (see [17] for a full derivation of a similar tensor). The genotype frequencies after recombination were then calculated by the sum of all possible parental genotype frequencies multiplied by their corresponding element in the tensor

$$p_i^{rec} = \sum_j \sum_k p_j^{mut} \, p_k^{mut} \, T_{jki} \qquad (2.4a)$$

and

$$q_i^{rec} = \sum_j \sum_k q_j^{mut} \, q_k^{mut} \, T_{jki}. \qquad (2.4b)$$

(d) Random genetic drift

Random genetic drift was incorporated into the model by randomly sampling variants from a multinomial distribution. Specifically, we assume that the genotype abundances after drift follow a multinomial distribution with parameters p_i^{rec} and q_i^{rec} (genotype frequencies before drift, corresponding to probabilities of drawing these genotypes from an infinitely large pool of juveniles), and population size N (number of samples drawn). These genotype abundances are then

normalized by N to yield the genotype frequencies after drift. Removing this drift step from the model allowed us to compare our full model with a deterministic model that does not involve the Hill–Robertson effect (as studied in [18,20]).

(e) Computation and analysis

Recursion equations were solved numerically using the Mathematica 10 software package (Wolfram Research, Inc., Champaign, Illinois, USA). All simulations started with both modifier alleles in equal frequency and with the a and b alleles fixed in both populations (i.e. $p_{Mab} = p_{mab} = q_{Mab} = q_{mab} = 0.5$; all other frequencies are equal to zero). This corresponds to a situation where population 2 is already fully adapted and a new habitat is colonized by population 1. Moreover, equal modifier allele frequencies correspond to standing genetic variation in recombination rates, a situation frequently reported in natural populations [42–45]. Simulations were run for 10 000 generations with 2000 replicate runs. Throughout, only results for population 1 are presented, as population 2 was treated as an ancestral gene pool.

3. Results

To put our work into context, we first briefly summarize previously derived analytical results for two infinitely large populations connected by migration and subject to divergent selection at two loci. Pylkov et al. [18] provided approximations for both the equilibrium allele frequencies and linkage disequilibrium ($\hat{D}$) under the resulting migration–selection balance. Using our notation and ignoring mutation, those approximations read

$$\left. \begin{array}{l} \hat{p}_A = \hat{p}_B = \hat{p} = \dfrac{1}{2} - \dfrac{(1+s)m}{s(1-2m)} + \sqrt{\dfrac{1}{4} + \dfrac{(1+s)m^2}{s^2(1-2m)^2}} \\[4mm] \text{and} \quad \hat{D} = \dfrac{s^2(1-m)\hat{p}^2(1-\hat{p})^2}{mr}. \end{array} \right\} \tag{3.1}$$

Note that $\hat{D}$ will always be positive, meaning there will be disproportionately many extreme genotypes (AB and ab) relative to intermediate ones (aB and Ab). Moreover, $\hat{D}$ is inversely proportional to r and is maximized at intermediate values of m. This is because with higher migration rates the allele frequencies in the two populations become increasingly similar. These approximations assume loose linkage between the loci and hence are in reasonable agreement with our simulation results, only for high recombination rates (not shown). We were able to derive an exact solution for the migration–selection balance accounting for arbitrary linkage, but the resulting formulae are long and uninformative and therefore not given here (Mathematica notebook available upon request).

Both Pylkov et al. [18] and Lenormand & Otto [20] also derived an approximation for the strength of selection acting on a rare modifier allele m that increases recombination, entering a population at migration–selection balance. Again using our notation, this approximation can be expressed as

$$s_m = -\hat{D} \frac{(r_m - r_M)s^2}{1 - (1-\rho)(1-r_M)} \left(\frac{1}{\rho} + \frac{1}{\rho + r_M - 2\rho r_M} - 1 \right). \tag{3.2}$$

As both $\hat{D}$ and the other two factors in this equation are positive there is always selection against a modifier increasing
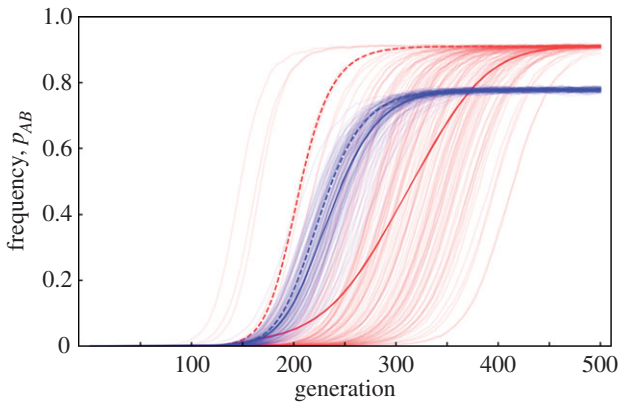
Figure 1. Change in the frequency of the fittest genotype *AB* in the presence and absence of recombination. Starting from the standard conditions (table 1), the genotype frequency of *AB* was tracked for 500 generations. The red lines represent simulations without recombination ($r = 0$), and the blue lines simulations with $r = 0.1$. The dashed lines are the deterministic model results ($N = \infty$). The faded lines are 2000 individual simulation runs with the solid line being the mean of all runs. See electronic supplementary material, figure S1 for similar dynamics with different parameters.

recombination rates, and the strength of this selection is directly proportional to $\hat{D}$. We will now explore how the 'maladaptive gene flow effect' quantified by these approximations interacts with the Hill–Robertson effect in our full, stochastic model.

We first performed simulations of our model for fixed recombination rates (no modifier). These simulations provided evidence for both the Hill–Robertson effect and the maladaptive gene flow effect (figure 1). In particular, using our standard set of parameters (table 1), we observed a faster spread of the optimal *AB* genotype with recombination; as predicted by the Hill–Robertson effect, this was only seen in finite populations. However, recombination also led to a lower equilibrium frequency of the *AB* genotype, as predicted by the maladaptive gene flow effect (figure 1). A stronger maladaptive gene flow effect occurred when either the migration rate was increased or the strength of selection was decreased (electronic supplementary material, figure S1*b,c*). By contrast, altering the mutation rate (electronic supplementary material, figure S1*d*) or the population size (electronic supplementary material, figure S1*e,f*) only affected the speed of adaptation and the magnitude of the Hill–Robertson effect.

We next investigated the full model including the modifier locus. Based on our initial simulations, we expected both selection for and against the *M* allele which reduces recombination. Indeed, simulations with our standard parameter set demonstrated that on average, the frequency of *M* declined initially, but then rose almost to fixation in the long term (electronic supplementary material, figure S2). To investigate the generality of this effect, we performed extensive simulations varying all six parameters of our model. The dynamics of each simulation were summarized by focusing on the range of the frequency of *M* for each set of parameters (figure 2). Increases to the population size *N* led to a more pronounced reduction in the minimum frequency of *M*, indicating that selection for recombination through the Hill–Robertson effect becomes stronger for *N* up to 10^6 (figure 2*a*). On the other hand, the maximum frequency of *M* also increased with increasing *N*, reflecting more efficient selection for *M* through the

maladaptive gene flow effect. We also observed that the Hill–Robertson effect was most notable when the migration rate was low and the selection pressure was strong, corresponding to the lowest frequencies of *M* (figure 2*b,c*).

The maladaptive gene flow effect was maximized at high migration rates, as recombination needs to be suppressed to prevent the populations rapidly homogenizing (figure 2*b*). Intriguingly, the maladaptive gene flow effect was maximized under intermediately strong selection, as weaker selection permits maladaptive alleles to persist in the population with little impact on the recombination modifier, while stronger selection would rapidly sweep maladaptive genes out of the population limiting the opportunities for recombining them onto resident backgrounds (i.e. pre-zygotic extrinsic reproductive isolation [3]).

The Hill–Robertson effect was maximized at intermediate mutation rates, as shown by the minimum frequency of *M* (figure 2*d*). This is because either mutations arise too rarely in the population to segregate simultaneously, or mutations arise too frequently rapidly producing double mutants and never creating linkage disequilibrium. By contrast, the maximum frequency of *M* increased with mutation rate (figure 2*d*), presumably because the migration–selection balance was reached more rapidly. The degree to which the *M* allele reduced recombination only had minor effects on the range of frequencies (figure 2*e*; values of $r_M < r_m$). When $r_M > r_m$, *M* enhanced recombination compared to *m*, leading to a reversal of selection through the Hill–Robertson and maladaptive gene flow effects. However, in line with Barton [46], selection for increased recombination became much weaker when the recombination rate was initially high (see right-hand side of figure 2*e*). Finally, as the genetic distance increased between the **M** and **A** locus (ρ), indirect selection on the modifier through both the Hill–Robertson and the maladaptive gene flow effect was weakened ([20,35] figure 2*f*; see also [46]).

Although the range of the frequency of the *M* allele is informative about the relative strength of the Hill–Robertson and the maladaptive gene flow effect, it tells us nothing about the relative timing of these two effects. Therefore, we investigated the frequency of *M* at four explanatory time points over a range of values for the migration rate and population size (figure 3; electronic supplementary material, figure S3). Starting from equal frequencies, the first 100 generations were governed by random genetic drift (figure 3*a*), with reduced variance in the frequency of *M* in larger populations (electronic supplementary material, figure S3*a*). The variance in this frequency expanded by the 250th generation (roughly around the time when the *AB* genotype arose, which varies between simulation runs). This time also corresponded to the earliest evidence for a Hill–Robertson effect, except in the presence of frequent gene flow (figure 3*b*). The Hill–Robertson effect became strongest roughly by the time the *AB* genotype reached equilibrium (figure 3*c*). Once at equilibrium, selection for recombination suppression through the maladaptive gene flow effect became dominant, resulting in an increased frequency of *M* with stronger migration rates (figure 3*d*).

To explore the combined effects of two key parameters of our model, the migration rate and selection coefficient, we performed additional simulations for nine combinations of these parameters (figure 4). These simulations clearly show the limits of the maladaptive gene flow effect, as the rapid suppression of recombination only evolved when both *m*
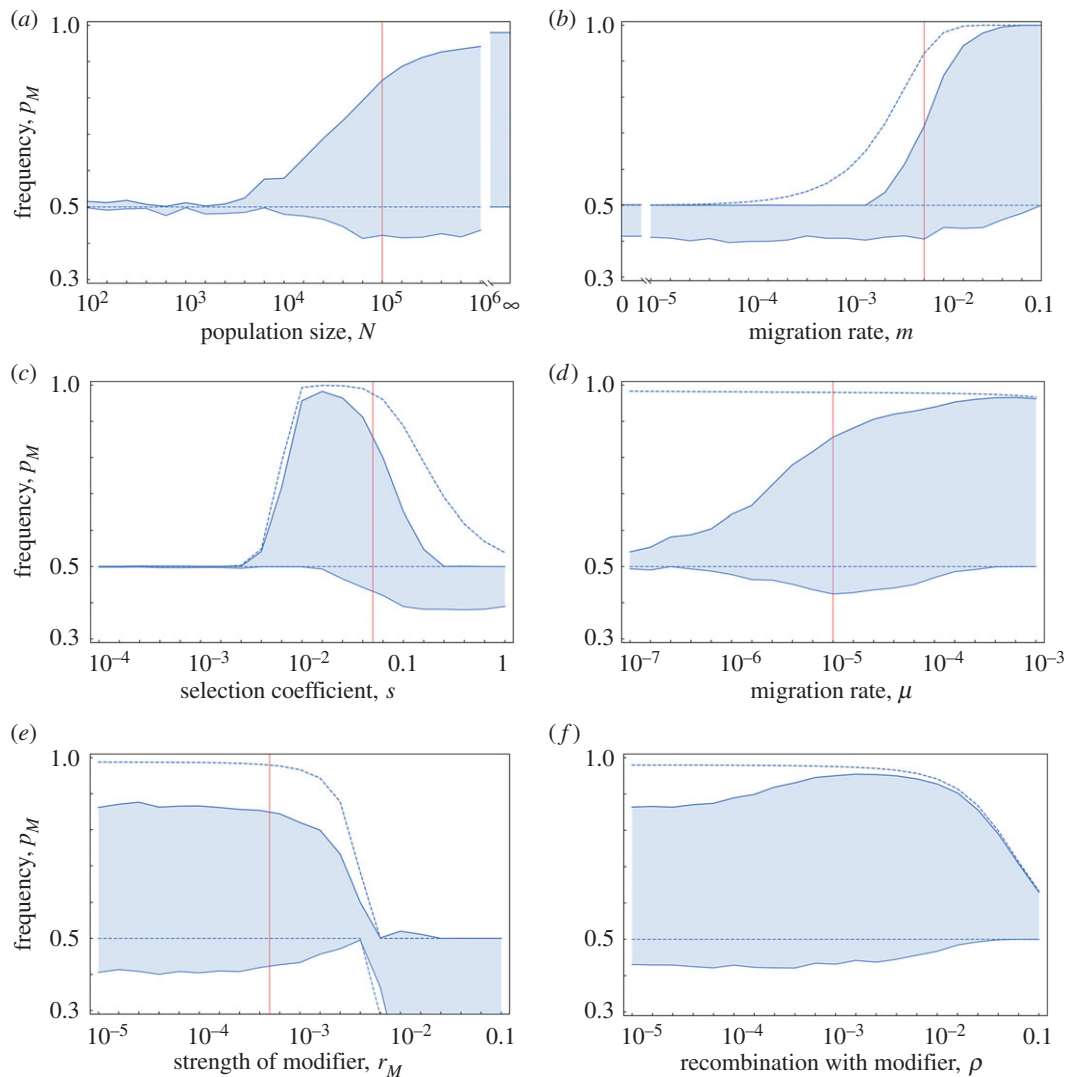
Figure 2. The range of the modifier frequency of the M allele for different parameters. Each panel shows the maximum and minimum frequency of the modifier through time, averaged across 2000 replicate simulations (solid lines; shading represents the range). In each panel (except panel (a)), the deterministic results from simulations that run without genetic drift ($N = \infty$) are shown as dashed lines. The vertical lines are the standard values for the focal parameter (table 1); note that because of the logarithmic scale, the standard value of $\rho = 0$ could not be indicated in panel (f). (Online version in colour.)

and s were high. This is because when m was high but s was not, allele frequencies in the two populations became very similar and the populations were close to linkage equilibrium, reducing selection for recombination suppression (see also the discussion of equations (3.1) and (3.2)). Conversely, when m was low but s was high, the few maladaptive migrant alleles that entered a population were very efficiently purged so that again there was only little linkage disequilibrium and hence selection for recombination suppression. In combination, the maladaptive gene flow effect can therefore be strongest at intermediate migration rates or selection coefficients when the other parameter is kept fixed (compare plots in left column or plots in middle row in figure 4). By contrast, early selection for increased recombination through the Hill–Robertson effect became stronger with increasing selection, but was diminished with increasing migration rates because this produces an earlier and stronger maladaptive gene flow effect.

4. Discussion

It has long been recognized that recombination antagonizes adaptive divergence and speciation with gene flow [9–11].

This is because recombination homogenizes migrant and resident genotypes that have diverged by natural selection. Theoretical studies of speciation with gene flow have often found that natural selection favours the spread of suppressors of recombination (e.g. chromosomal inversions and modifier alleles; reviewed in [16]) due to maladaptive gene flow [14,15,18–21]. However, these studies have ignored the stochastic processes that operate in finite populations and that are known to produce selection for increased recombination via the Hill–Robertson effect [29–32,35,47]. Here, we have investigated the interplay between these two effects and demonstrated that the Hill–Robertson effect favours increased recombination between adaptive loci during the early stages of adaptive divergence, whereas the maladaptive gene flow effect leads to suppression of recombination after populations have diverged.

Our results caution against the view that suppressors of recombination will always spread in the presence of maladaptive gene flow. Although in our two-locus model selection for increased recombination is only transitory, the magnitude of the Hill–Robertson effect is expected to increase with a larger number of loci under selection [30,31,33,47]. We would also expect the duration of the
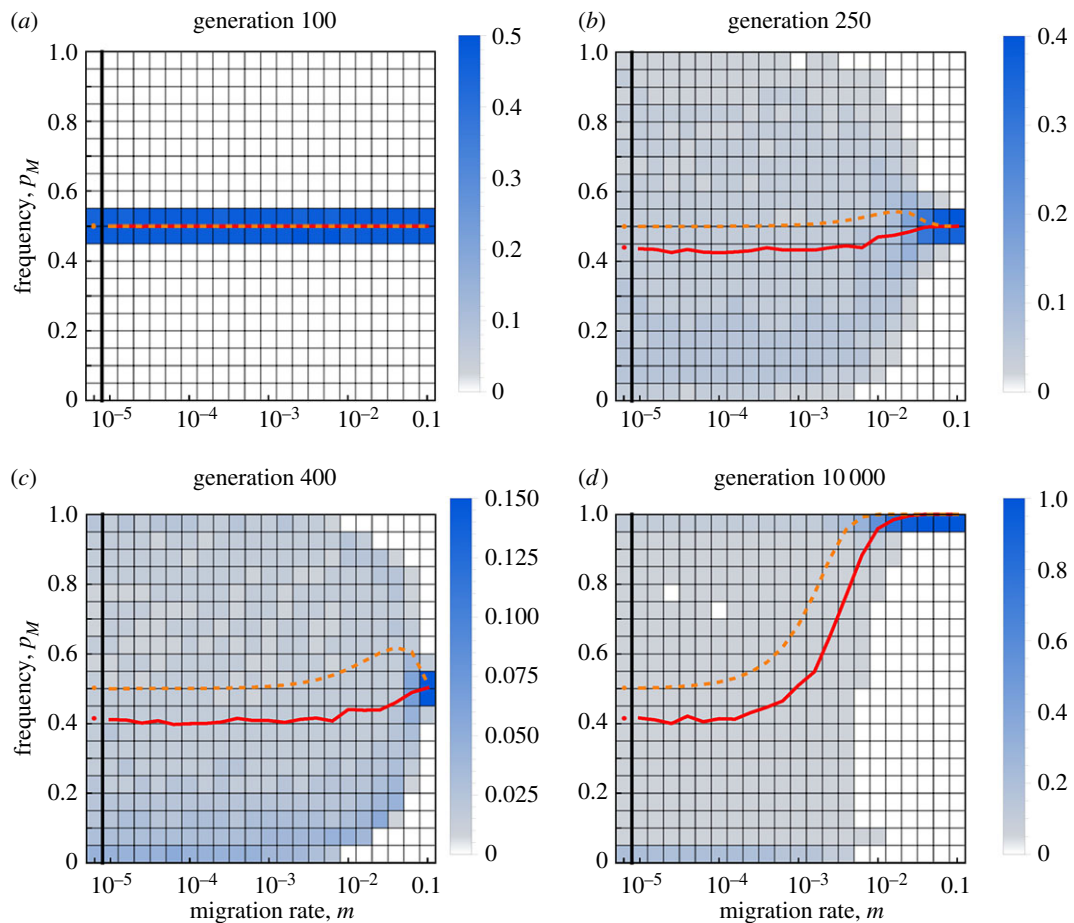
Figure 3. Distribution of the frequency of a modifier allele M reducing recombination over a range of migration values at four time points. The charts show the frequencies of the modifier p_M at four different time points that describe the evolution of recombination. The y-axis of each chart is split up into 20 equally sized bins (horizontal lines) covering all frequencies from 0 to 1, in increments of 0.05. The shading of each cell indicates the fraction of runs, out of 2000 replicates, that the modifier frequency was in that bin at any given migration rate (see legend). The left-most value for migration rate (left of the thick vertical line) represents $m = 0$. The solid red line shows the mean of all replicates for a given migration rate, and the dashed orange line indicates the result of the deterministic simulation ($N = \infty$).

Hill–Robertson effect to be extended relative to the maladaptive gene flow effect if there is ongoing selection acting on many loci. In addition, even after the population has reached a fitness peak, selection for increased recombination may continue as a consequence of the mutation–selection balance [33]. Finally, selection for the suppression of recombination can be opposed or even reversed if there is negative epistasis between the selected loci [20]. More generally, the evolution of supressed recombination during the early stages of ecological speciation with gene flow depends on a number of factors including not only genetic architecture (e.g. number of loci and epistatic interactions) and migration rate, but also mutation rate and population size. Whether or not the initial conditions that favour the evolution of reproductive isolation between locally adapted populations in the face of gene flow are common in nature remains an empirical problem.

The strength of the Hill–Robertson and the maladaptive gene flow effect as well as the timing of the transition between these effects was determined by the parameters we used. We selected our standard parameters so that both effects were observed, including complete linkage between the modifier locus and the loci under selection (as figure 2f indicates, this does not seem to be crucial though as long as linkage is reasonably tight). Relatively large populations and strong selection were required for increased

recombination rates to evolve under the Hill–Robertson effect, whereas strong migration and selection were necessary for recombination suppression to evolve under the maladaptive gene flow effect. We expect that both of these conditions should be met in many systems. In particular, in hybrid zones between populations with large effective population size recombination suppression should evolve readily. For example, *Helianthus* sunflowers and *Drosophila* species that occur in sympatry are often separated by chromosomal inversions [24,25] and strong selection [48,49]. These taxa also frequently have N_e values above one million [48,50,51] and are experiencing, or have experienced, extensive gene flow [50,52,53]. Across other systems, regions of suppressed recombination also separate hybridizing species or locally adapted populations [54–56], but whether these regions evolved during speciation with gene flow or are ancestral genomic landscapes (e.g. [57]) remains to be determined (see [22] for a more detailed discussion).

If populations exposed to maladaptive gene flow can establish suppressed recombination, then it is possible that they might also continue to diverge, becoming new species. However, the specific conditions for continued divergence can be restrictive. For example, Flaxman *et al*. [58] found that recombination suppression aids speciation when selection is weak but migration is high, presumably because selection effects acting upon each locus are extended to
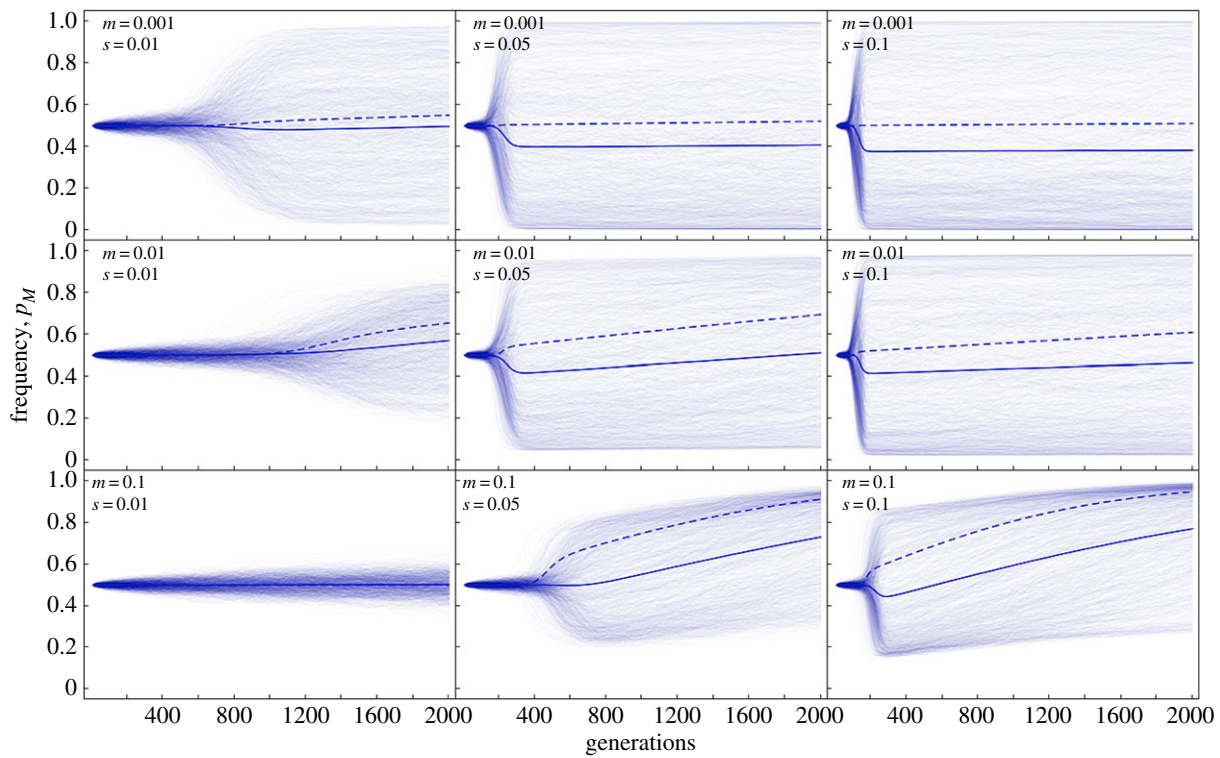
Figure 4. Changes in the frequency of a modifier allele M reducing recombination depending on the strength of migration and selection. Simulations were run using three different migration rates (rows) and three different selection coefficients (columns); other parameters were maintained at standard conditions (table 1). The faded lines show 1000 replicate runs for each simulation, with the solid line showing the average and the dashed line the deterministic result ($N = \infty$). (Online version in colour.)

larger regions of the genome [59]. In our model, the recombination rate itself evolves and the levels of migration and selection that lead to its suppression appear to partially differ from those favouring divergence in these speciation models. For instance, under moderate selection and strong migration, the modifier of recombination did not spread. However, when s and m were both high and of similar magnitude, the modifiers of recombination spread easily (figure 4). This disjunction between models might suggest that cases of speciation with gene flow where migration is stronger than selection might be unlikely because the required recombination landscape is difficult to evolve in the first place. Conversely, cases of speciation with gene flow where selection is strong are likely to proceed without the suppression of recombination, although recombination suppression is nonetheless expected to evolve. We therefore speculate that the evolution of suppressed recombination might not be required in some cases, but nonetheless will be a signature of speciation with gene flow. This is because the most favourable conditions for speciation with gene flow are also the most favourable conditions for the spread of modifier alleles that suppress recombination (i.e. when selection is strong). Overall, cases of sympatric and parapatric speciation might require strong selection, and they will generally be characterized by coldspots of recombination.

A number of studies have also investigated the reduction of recombination by chromosomal inversions during speciation with gene flow [14,21,22,24,25,27]. For instance, Kirkpatrick & Barton [14] showed that a chromosomal inversion would spread when it captures loci under divergent selection, as an inversion will prevent hybridization with maladapted migrants. Such an inversion and the modifier model we used in this study are similar, except for one key difference: inversions only prevent recombination in heterokaryotypes. As a consequence, shuffling between resident and migrant alleles is prevented but shuffling of different resident alleles can still occur. Therefore, we would expect the extent to which the Hill–Robertson effect opposes the maladaptive gene flow effect to be reduced, making an inversion a powerful mechanism that simultaneously allows efficient local adaptation and limits maladaptive gene flow between populations during speciation with gene flow. In addition, populations separated by fixed chromosomal inversions might be able to diverge further during secondary contact even if gene flow was very high and selection was weak. As noted by Flaxman et al. [58], these are the conditions where recombination suppression has the greatest contribution to speciation driven by divergent natural selection.

In order to isolate the specific effects on recombination modifiers caused by gene flow as opposed to population size, we intentionally kept our model as simple as possible. Future models should consider a larger number of loci and integrate additional factors that are not part of our model. These could include different types of recombination modifiers (local versus global allelic modifiers and also chromosomal rearrangements), epistasis, diploidy and multiple populations with variable spatial selection and migration rates. It would also be interesting to construct models that explore the interplay between selection on recombination rates and other mechanisms of reinforcement such as selection for assortative mating or migration rate in response to maladaptive gene flow [7,16].

In conclusion, our model indicates that the dynamics of recombination rate evolution during adaptive divergence are governed not only by deterministic evolutionary forces

such as selection and migration, but also by stochastic forces such as random genetic drift and mutation. Much of the theoretical literature on adaptive speciation with gene flow has overlooked the interplay between these forces, and consequently often saw recombination as a hindrance to adaptive divergence and speciation. The fact that the benefits of recombination in finite populations are well established within the research community studying the evolution of sex highlights the need of increased dialogue between these fields.

References

1. Schluter D. 2009 Evidence for ecological speciation and its alternative. Science 323, 737–741. (doi:10.1126/Science.1160006)

2. Nosil P, Harmon LJ, Seehausen O. 2009 Ecological explanations for (incomplete) speciation. Trends Ecol. Evol. 24, 145–156. (doi:10.1016/j.tree.2008.10.011)

3. Rundle HD, Nosil P. 2005 Ecological speciation. Ecol. Lett. 8, 336–352. (doi:10.1111/J.1461-0248.2004.00715.X)

4. Schluter D. 2001 Ecology and the origin of species. Trends Ecol. Evol. 16, 372–380. (doi:10.1016/S0169-5347(01)02198-X)

5. Nosil P. 2012 Ecological speciation, 280. Oxford, NY: Oxford University Press.

6. Nosil P, Vines TH, Funk DJ. 2005 Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution 59, 705–719. (doi:10.1111/j.0014-3820.2005.tb01747.x)

7. Lenormand T. 2012 From local adaptation to speciation: specialization and reinforcement. Int. J. Ecol. 2012, 508458. (doi:10.1155/2012/508458)

8. Bolnick DI, Nosil P. 2007 Natural selection in populations subject to a migration load. Evolution 61, 2229–2243. (doi:10.1111/j.1558-5646.2007.00179.x)

9. Felsenstein J. 1981 Skepticism towards Santa Rosalia, or why are there so few kinds of animals. Evolution 35, 124–138. (doi:10.2307/2407946)

10. Butlin RK. 2005 Recombination and speciation. Mol. Ecol. 14, 2621–2635. (doi:10.1111/J.1365-294x.2005.02617.X)

11. Bürger R, Akerman A. 2011 The effects of linkage and gene flow on local adaptation: a two-locus continent-island model. Theor. Popul. Biol. 80, 272–288. (doi:10.1016/j.tpb.2011.07.002)

12. Faria R, Navarro A. 2010 Chromosomal speciation revisited: rearranging theory with pieces of evidence. Trends Ecol. Evol. 25, 660–669. (doi:10.1016/J.Tree.2010.07.008)

13. Navarro A, Barton NH. 2003 Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evolution 57, 447–459. (doi:10.1554/0014-3820(2003)057[0447:Apigip]2.0.Co;2)

14. Kirkpatrick M, Barton N. 2006 Chromosome inversions, local adaptation and speciation. Genetics 173, 419–434. (doi:10.1534/Genetics.105.047985)

15. Feder JL, Nosil P. 2009 Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? Evolution 63, 3061–3075. (doi:10.1111/j.1558-5646.2009.00786.x)

16. Ortiz-Barrientos D, Engelstädter J, Rieseberg LH. 2016 Recombination rate evolution and the origin of species. Trends Ecol. Evol. 31, 226–236. (doi:10.1016/j.tree.2015.12.016)

17. Nei M. 1967 Modification of linkage intensity by natural selection. Genetics 57, 625–641.

18. Pylkov KV, Zhivotovsky LA, Feldman MW. 1998 Migration versus mutation in the evolution of recombination under multilocus selection. Genet. Res. 71, 247–256. (doi:10.1017/S0016672398003243)

19. Charlesworth D, Charlesworth B. 1979 Selection on recombination in clines. Genetics 91, 581–589.

20. Lenormand T, Otto SP. 2000 The evolution of recombination in a heterogeneous environment. Genetics 156, 423–438.

21. Trickett AJ, Butlin RK. 1994 Recombination suppressors and the evolution of new species. Heredity 73, 339–345. (doi:10.1038/hdy.1994.180)

22. Feder JL, Xie XF, Rull J, Velez S, Forbes A, Leung B, Dambroski H, Filchak KE, Aluja M. 2005 Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in Rhagoletis. Proc. Natl Acad. Sci. USA 102, 6573–6580. (doi:10.1073/pnas.0502099102)

23. Michel AP, Sim S, Powell THQ, Taylor MS, Nosil P, Feder JL. 2010 Widespread genomic divergence during sympatric speciation. Proc. Natl Acad. Sci. USA 107, 9724–9729. (doi:10.1073/pnas.1000939107)

24. Noor MAF, Grams KL, Bertucci LA, Reiland J. 2001 Chromosomal inversions and the reproductive isolation of species. Proc. Natl Acad. Sci. USA 98, 12 084–12 088. (doi:10.1073/pnas.221274498)

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

26. 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:e1000500 10.1371/journal.pbio.1000500)

27. Carneiro M, Ferrand N, Nachman MW. 2009 Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (Oryctolagus cuniculus). Genetics 181, 593–606. (doi:10.1534/genetics.108.096826)

28. Hill WG, Robertson A. 1966 Effect of linkage on limits to artificial selection. Genet. Res. 8, 269–294. (doi:10.1017/S0016672300010156)

29. Felsenstein J. 1974 Evolutionary advantage of recombination. Genetics 78, 737–756.

30. Otto SP, Barton NH. 2001 Selection for recombination in small populations. Evolution 55, 1921–1931. (doi:10.1017/S0016672300010156)

31. Iles MM, Walters K, Cannings C. 2003 Recombination can evolve in large finite populations given selection on sufficient loci. Genetics 165, 2249–2258.

32. Barton NH, Otto SP. 2005 Evolution of recombination due to random drift. Genetics 169, 2353–2370. (doi:10.1534/genetics.104.032821)

33. Keightley PD, Otto SP. 2006 Interference among deleterious mutations favours sex and recombination in finite populations. Nature 443, 89–92. (doi:10.1038/nature05049)

34. Martin G, Otto SP, Lenormand T. 2006 Selection for recombination in structured populations. Genetics 172, 593–609. (doi:10.1534/genetics.104.039982)

35. Roze D, Barton NH. 2006 The Hill–Robertson effect and the evolution of recombination. Genetics 173, 1793–1811. (doi:10.1534/Genetics.106.058586)

36. Becks L, Agrawal AF. 2012 The evolution of sex is favoured during adaptation to new environments. PLoS Biol. 10, e1001317. (doi:10.1371/Journal.Pbio.1001317)

37. Cooper TF. 2007 Recombination speeds adaptation by reducing competition between beneficial mutations in populations of Escherichia coli. PLoS Biol. 5, 1899–1905. (doi:10.1371/journal.pbio.0050225)

38. Goddard MR, Godfray HCJ, Burt A. 2005 Sex increases the efficacy of natural selection in experimental yeast populations. Nature 434, 636–640. (doi:10.1038/nature03405)

39. Colegrave N. 2002 Sex releases the speed limit on evolution. Nature 420, 664–666. (doi:10.1038/nature01191)

40. McDonald MJ, Rice DP, Desai MM. 2016 Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature* **531**, 233 – 236. (doi:10.1038/nature17143)

41. Karlin S, McGregor J. 1974 Towards a theory of evolution of modifier genes. *Theor. Popul. Biol.* **5**, 59 – 103. (doi:10.1016/0040-5809(74)90052-5)

42. Smukowski CS, Noor MA. 2011 Recombination rate variation in closely related species. *Heredity* **107**, 496 – 508. (doi:10.1038/hdy.2011.44)

43. Kong A *et al.* 2010 Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **467**, 1099 – 1103. (doi:10.1038/nature09525)

44. Bauer E *et al.* 2013 Intraspecific variation of recombination rate in maize. *Genome Biol.* **14**, R103. (doi:10.1186/gb-2013-14-9-r103)

45. Gion JM, Hudson CJ, Lesur I, Vaillancourt RE, Potts BM, Freeman JS. 2016 Genome-wide variation in recombination rate in Eucalyptus. *BMC Genomics* **17**, 590. (doi:10.1186/s12864-016-2884-y)

46. Barton NH. 1995 A general-model for the evolution of recombination. *Genet. Res.* **65**, 123 – 144. (doi:10.1017/S0016672300033140)

47. Hartfield M, Otto SP, Keightley PD. 2010 The role of advantageous mutations in enhancing the evolution of a recombination modifier. *Genetics* **184**, 1153 – 1164. (doi:10.1534/genetics.109.112920)

48. Sambatti JB, Strasburg JL, Ortiz-Barrientos D, Baack EJ, Rieseberg LH. 2012 Reconciling extremely strong barriers with high levels of gene exchange in annual sunflowers. *Evolution* **66**, 1459 – 1473. (doi:10.1111/j.1558-5646.2011.01537.x)

49. Ortiz-Barrientos D, Counterman BA, Noor MA. 2004 The genetics of speciation by reinforcement. *PLoS Biol.* **2**, e416. (doi:10.1371/journal.pbio.0020416)

50. Strasburg JL, Rieseberg LH. 2008 Molecular demographic history of the annual sunflowers Helianthus annuus and *H. petiolaris*—large effective population sizes and rates of long-term gene flow. *Evolution* **62**, 1936 – 1950. (doi:10.1111/j.1558-5646.2008.00415.x)

51. Ortiz-Barrientos D, Chang AS, Noor MA. 2006 A recombinational portrait of the *Drosophila pseudoobscura* genome. *Genet. Res.* **87**, 23 – 31. (doi:10.1017/S0016672306007932)

52. Schaeffer SW, Miller EL. 1992 Estimates of gene flow in *Drosophila pseudoobscura* determined from nucleotide sequence analysis of the alcohol dehydrogenase region. *Genetics* **132**, 471 – 480.

53. Machado CA, Kliman RM, Markert JA, Hey J. 2002 Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Mol. Biol. Evol.* **19**, 472 – 488. (doi:10.1093/oxfordjournals.molbev.a004103)

54. Saether SA *et al.* 2007 Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* **318**, 95 – 97. (doi:10.1126/science.1141506)

55. White BJ, Cheng C, Simard F, Costantini C, Besansky NJ. 2010 Genetic association of physically unlinked islands of genomic divergence in incipient species of Anopheles gambiae. *Mol. Ecol.* **19**, 925 – 939. (doi:10.1111/j.1365-294X.2010.04531.x)

56. Berg PR, Star B, Pampoulie C, Sodeland M, Barth JM, Knutsen H, Jakobsen KS, Jentoft S. 2016 Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Sci. Rep.* **6**, 23246. (doi:10.1038/srep23246)

57. Renaut S, Grassa CJ, Yeaman S, Moyers BT, Lai Z, Kane NC, Bowers JE, Burke JM, Rieseberg LH. 2013 Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nat. Commun.* **4**, 1827. (doi:10.1038/ncomms2833)

58. Flaxman SM, Wacholder AC, Feder JL, Nosil P. 2014 Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Mol. Ecol.* **23**, 4074 – 4088. (doi:10.1111/mec.12750)

59. Rieseberg LH. 2001 Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **16**, 351 – 358.