Mathematical Biology

Adaptive gene introgression after secondary contact

Hildegard Uecker · Derek Setter · Joachim Hermisson

Abstract By hybridization and backcrossing, alleles can surmount species boundaries and be incorporated into the genome of a related species. This introgression of genes is of particular evolutionary relevance if it involves the transfer of adaptations between populations. However, any beneficial allele will typically be associated with other alien alleles that are often deleterious and hamper the introgression process. In order to describe the introgression of an adaptive allele, we set up a stochastic model with an explicit genetic makeup of linked and unlinked deleterious alleles. Based on the theory of reducible multitype branching processes, we derive a recursive expression for the establishment probability of the beneficial allele after a single hybridization event. We furthermore study the probability that slightly deleterious alleles hitchhike to fixation. The key to the analysis is a split of the process into a stochastic phase in which the advantageous alleles establishes and a deterministic phase in which it sweeps to fixation. We thereafter apply the theory to a set of biologically relevant scenarios such as introgression in the presence of many unlinked or few closely linked deleterious alleles. A comparison to computer simulations shows that the approximations work well over a large parameter range.

H. Uecker (✉)· J. Hermisson
Mathematics and Biosciences Group, Faculty of Mathematics and Max F. Perutz Laboratories,
University of Vienna, 1090 Vienna, Austria
e-mail: Hildegard.Uecker@univie.ac.at

J. Hermisson
e-mail: Joachim.Hermisson@univie.ac.at

H. Uecker
IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

D. Setter
Mathematics and Biosciences Group, Faculty of Mathematics and Vienna School of Population
Genetics, University of Vienna, 1090 Vienna, Austria
e-mail: Derek.Setter@univie.ac.at

1 Introduction

Hybridization between related species is a common phenomenon. Indeed, Mallet (2005) estimates that at least 25 % of plant species and 10 % of animal species still interbreed. The disappearance of natural habitat barriers following environmental change, the introduction of foreign species, the escape of domesticated animals into the wild, and the cultivation of crops all create new regions of species range overlap and consequently cause high rates of hybridization. Despite reproductive barriers, hybridization between related species is often not completely prohibited and leads to the production of viable and fertile offspring. In the course of backcrossing with a parental species, alien genetic material is lost, but some part of it may be permanently incorporated into the genome of the sister species. The introgression of genes from one species into another occurs over a wide range of taxa (Rhymer and Simberloff 1996; Lindner et al. 1998; Arnold et al. 1999; Arnold 2004; Miller et al. 2012; Ellstrand et al. 2013). The introgression of genes from feral to wild animals (Adams et al. 2003; Beaumont et al. 2001; Gottelli et al. 1994; Rhymer and Simberloff 1996) or from introduced to native species (Rhymer and Simberloff 1996; Fitzpatrick et al. 2010) poses ecological risks and if extensive may entail a loss of biodiversity.

In addition, evidence for the transfer of adaptations across species boundaries is growing (Arnold et al. 1999; Arnold 2004; Whitney et al. 2006; Schwenk et al. 2008; Arnold and Martin 2009; The *Heliconius* Genome Consortium 2012; Hedrick 2013). Introgression of genes can hence take direct influence on the evolutionary routes of a species and speed up adaptation. For example, the introduced sunflower species *Helianthus annuus* likely has acquired resistance genes from the native and locally adapted species *H. debilis* in Texas, allowing it to expand its species range southwards (Heiser 1951; Whitney et al. 2006). Similarly, Abi-Rached et al. (2011) suggest that positively selected immune system alleles from Neanderthals and Denisovans might have introgressed into modern humans. In agriculture, adaptive gene introgression can potentially constitute a major risk: adaptive herbivore, insecticide, or pathogen resistance genes from (possibly genetically modified) crops can spread to wild relatives, severely complicating weed control (Snow 2002; Snow et al. 2003; Stewart et al. 2003). Importantly, Snow et al. (2003) show that a transgene can reduce herbivory and increase fitness in a wild sunflower under natural conditions.

Early-generation hybrids, even if not entirely infertile or inviable, frequently suffer from strongly reduced fitness. Often, hybrids display an intermediate phenotype that is maladapted to either parental niche. The low hybrid fitness can also result from genetic incompatibilities. By backcrossing with one of the parental species, alleles that prove to be deleterious in the foreign genetic background or cause maladaptation to the parental niches can be purged and fitness restored (Heiser 1951; Arnold et al. 1999).

The probability of successful gene introgression critically depends on the strength of this fitness bottleneck.

Theoretical models on adaptive gene introgression that take a reduction in hybrid fitness into account usually assume that a pre-defined number of backcrosses are required in order to lose the deleterious material and obtain a positively selected type (Demon et al. 2007; Gosh and Haccou 2010; Gosh et al. 2012a,b). This assumes that the deleterious effects are homogeneously spread over the genome of a diploid organism and that an appreciable amount of deleterious alleles is required to have a measurably impact on fitness. Focusing on other aspects of gene introgression, such as the impact of a temporally varying environment (Gosh et al. 2012b) or life history traits (Demon et al. 2007), these models greatly simplify the underlying genetics. A step towards more realistic population genetic models has been made by Gosh et al. (2012a). Their analysis remains, however, restricted to the most basic scenario in which a single deleterious allele is linked to the locus under positive selection (see also Barton 1979). For a neutral marker locus, in contrast, the impact of a genetic barrier consisting of an arbitrary number of deleterious alleles has been investigated (Bengtsson 1985; Barton and Bengtsson 1986).

In this paper, we focus on a single hybridization event and examine the impact of linked and unlinked deleterious alleles on the introgression process of an adaptive allele. The deleterious effect of alleles is caused by maladaptation to the new environment and is independent of the genetic background. We first set up a Moran-like model which describes the evolution of the population by genetic drift, selection, and recombination. In the first part of the model analysis, we apply the theory of reducible multitype branching processes to determine by how much deleterious alleles reduce the introgression probability of a favorable allele in dependence of the strength of selection and linkage. The second part considers the probability that closely linked deleterious alleles "hitchhike" to fixation. The analysis relies on a separation of the process into a strongly stochastic phase, in which a haplotype carrying the beneficial allele establishes, and a deterministic phase, in which it sweeps through the population, possibly losing deleterious material by recombination with wildtype individuals. These recombination events and the subsequent establishment or loss of haplotypes with fewer deleterious alleles are again subject to strong stochasticity. In this analysis, we again resort to the theory of branching processes. The derived approximations are applied to a variety of biological scenarios and complemented by computer simulations. We close the paper with a discussion.

2 Full model and simulations

We consider a large population of N haploid individuals. The theory also applies to diploids without dominance if we can assume Hardy-Weinberg equilibrium, but we use the haploid formalism throughout the paper. Through a single hybridization event, a hybrid individual is introduced to the population (for diploids, the alien alleles arrive in the foreign habitat at the haploid stage, i.e., for plants, by pollen dispersal). The hybrid carries an adaptive allele as well as a number of deleterious alleles. Deleterious alleles are either physically linked to the adaptive allele or unlinked. We assume that

this initial hybrid haplotype carries I and J linked deleterious alleles to the left and the right of the beneficial allele, respectively, and F unlinked alleles. By recombination with wildtype individuals, haplotypes with fewer introgressed alleles can be generated, leading to a hybrid swarm. Selection on the deleterious alleles relies on maladaptation to the environment and is independent of the genomic context. We assume that all wildtype individuals have the same fitness and introgressed alleles interact identically with all wildtype backgrounds. The fitness of an individual is thus fully determined by the introgressed alleles that it carries.

The evolution of the population is described by the following scheme, which represents a Moran model with recombination: At rate N, two individuals are chosen to reproduce and generate a single offspring. During reproduction, recombination can take place. In order to simplify our bookkeeping of genotypes in the analytical treatment, we restrict ourselves to single crossover among the linked alleles. Multiple crossover is unlikely to happen over recombination distances r with $r^2 \ll r$ so that the model captures scenarios of tight linkage. Considering larger recombination distances (where multiple crossover gets likely) or gene conversion requires a straightforward extension of the formalism, in which a larger number of genotypes can be generated by recombination with wildtype individuals. Unlinked alleles are inherited with probability one-half. The offspring replaces an individual that is chosen based on its fitness. For notational simplicity, we assign the numbers 1 to N to the individuals. Individual number k is then chosen with probability $\frac{1-\sigma^{(k)}}{\sum_{i=1}^{N}(1-\sigma^{(i)})}$, where $\sigma^{(k)} = 0$ for wildtype individuals. I.e., $\sigma^{(k)}$ is the Malthusian fitness of individual k in a wildtype population. All three individuals that are involved in a reproduction event are chosen with replacement, i.e., the same individual might be chosen twice.

The simulation program implements the successive events without consideration for the time spans between them. As we are only interested in probabilities, this does not influence the results. The number of replicates is chosen so that error bars vanish in the symbols in all plots. The simulation program is written in the C++ programming language, making use of the *Gnu Scientific Library* (Galassi et al. 2009).

The full model does not allow for an analytical treatment. In the following sections, we therefore consider approximations to the introgression process.

3 The early phase of spread

A single evolutionary step involves three individuals: two individuals that reproduce and one that dies. In a large population, as long as introgressed haplotypes are rare, it is unlikely that more than one hybrid individuals are involved in a single event (or that the same individual is involved twice). Formally, this means that terms of order (n_{intro}/N) in the transition rates, where n_{intro} denotes the number of individuals with introgressed alleles, are negligible. Consequently, hybrids suffer (nearly) independent fates in the early phase of spread, and the process is therefore well described by a multitype branching process. The branching process is strictly recovered in the limit $N \to \infty$.

A

B

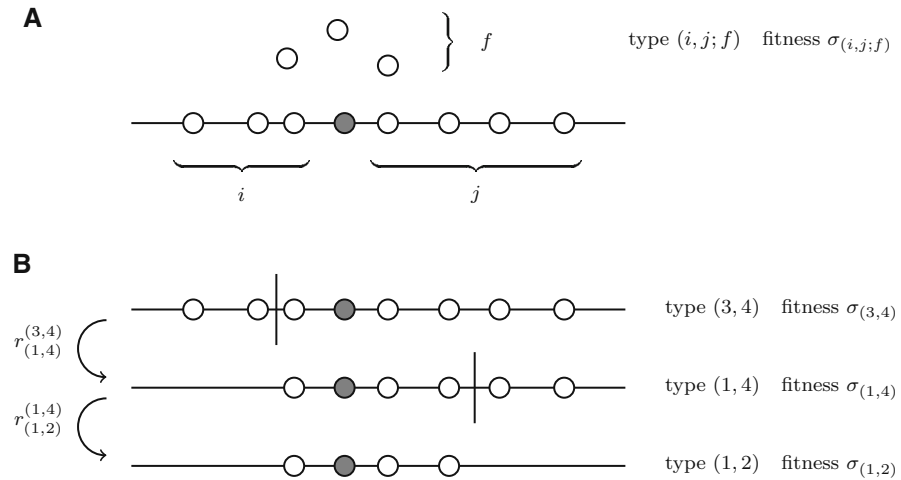**Fig. 1** An illustration of the branching model. The *dark dot* represents the advantageous allele, the *blank dots* represent deleterious alleles. i and j give the number of deleterious alleles to the *left* and *right* of the advantageous allele and f the number of unlinked deleterious alleles. Panel **b** illustrates how linked deleterious alleles are lost by recombination with wildtype individuals. "Fitness" refers to the Malthusian fitness

The lack of interaction among hybrids entails that types with introgressed material only recombine with wildtype individuals. This implies that by recombination, they can only lose, not gain deleterious alleles, and we encounter a special instance of a reducible multitype branching process (cf. also Barton and Bengtsson 1986; Demon et al. 2007; Gosh et al. 2012b). By recombination, types that carry only deleterious alleles but not the beneficial allele are generated. We do not consider these types in the following (within a branching process approach, they are doomed to extinction) but focus on carriers of the advantageous allele. In the main text of the paper, we assume that all unlinked deleterious alleles have the same effect size. A generalization of the main results to arbitrary effect sizes is given in Appendix D.

We call an individual that carries the beneficial allele with i deleterious alleles to the left, j deleterious alleles to the right and f unlinked deleterious alleles an individual of type $(i, j; f)$ (cf. Fig. 1a). Its net selection coefficient is denoted by $\sigma_{(i,j;f)}$. We set $(i, j) \equiv (i, j; 0)$. The recombinant offspring of a type (i, j) individual are either of type (i, k) with $k < j$ or of type (k, j) with $k < i$. Such recombination events may happen with probabilities $r_{(i,k)}^{(i,j)}$ and $r_{(k,j)}^{(i,j)}$, respectively. An instance of repeated recombination events is depicted in Fig. 1b. The overall probability that a recombination event takes place is given by the sum $r_{(i,j)} = \sum_{k=0}^{j-1} r_{(i,k)}^{(i,j)} + \sum_{k=0}^{i-1} r_{(k,j)}^{(i,j)}$. The number of unlinked deleterious alleles that are inherited by an offspring individual is binomially distributed with parameter 0.5. We obtain for the per capita transition rates of the possible events in the branching process:

Death

$$P((i, j; f) \to \emptyset) = 1 - \sigma_{(i, j; f)}, \tag{1a}$$

Birth—with coupled recombination

$$P((i, j; f) \to \{(i, j; f), (i, j; g)\}) = \binom{f}{g} \left(\frac{1}{2}\right)^f (1 - r_{(i,j)}),$$

$$P((i, j; f) \to \{(i, j; f), (i, k; g)\}) = \binom{f}{g} \left(\frac{1}{2}\right)^f r_{(i,k)}^{(i,j)} \quad \text{for} \quad k < j,$$

$$P((i, j; f) \to \{(i, j; f), (k, j; g)\}) = \binom{f}{g} \left(\frac{1}{2}\right)^f r_{(k,j)}^{(i,j)} \quad \text{for} \quad k < i. \quad (1b)$$

4 The probability of adaptive gene introgression

First, we focus on the probability that the beneficial allele establishes in the population. Once the beneficial allele is sufficiently frequent, it is very unlikely to be lost again. The extinction probability of the branching process as described in the previous section is thus a good approximation for the extinction probability of the beneficial allele in the full model (and the introgression probability is the complementary probability). We denote by $Q_{(i,j;f)}$ the extinction probability of the process that is initiated by exactly one individual of type $(i, j; f)$.

Theorem 1 *The extinction probability $Q_{(I,J;F)}$ can be calculated by recursively solving the system of quadratic equations*

$$(1 - r_{(i,j)}) \left(\frac{1}{2}\right)^f Q_{(i,j;f)}^2$$

$$+ \sum_{g=0}^{f} \binom{f}{g} \left(\frac{1}{2}\right)^f \left\{ \sum_{k=0}^{i-1} r_{(k,j)}^{(i,j)} Q_{(k,j;g)} + \sum_{k=0}^{j-1} r_{(i,k)}^{(i,j)} Q_{(i,k;g)} \right\} Q_{(i,j;f)}$$

$$+ \sum_{g=0}^{f-1} \binom{f}{g} \left(\frac{1}{2}\right)^f (1 - r_{(i,j)}) Q_{(i,j;g)} Q_{(i,j;f)}$$

$$- (2 - \sigma_{(i,j;f)}) Q_{(i,j;f)} + 1 - \sigma_{(i,j;f)} = 0,$$

$$i \in \{0, \ldots, I\}, \quad j \in \{0, \ldots, J\}, \quad f \in \{0, \ldots, F\}, \quad (2)$$

where always the smaller root of the equation has to be used.

We only give an illustrative derivation of Eq. (2) here and move the full proof to Appendix A.

Consider a branching process initiated by an individual of type $(i, j; f)$. With probability $\frac{1 - \sigma_{(i,j;f)}}{2 - \sigma_{(i,j;f)}}$, the founding individual dies before it reproduces, in which case the lineage is immediately extinct. With probability $\binom{f}{g} \left(\frac{1}{2}\right)^f \frac{1 - r_{(i,j)}}{2 - \sigma_{(i,j;f)}}$, it reproduces and generates a non-recombinant offspring with g unlinked alleles, i.e., an offspring of type $(i, j; g)$. With probability $\binom{f}{g} \left(\frac{1}{2}\right)^f \frac{r_{(i,k)}^{(i,j)}}{2 - \sigma_{(i,j)}}$ (or $\binom{f}{g} \left(\frac{1}{2}\right)^f \frac{r_{(k,j)}^{(i,j)}}{2 - \sigma_{(i,j)}}$), it reproduces

and gives birth to a type $(i, k; g)$ (or $(k, j; g)$) individual with $0 \leq kj$ (or $0 \leq k < i$). After reproduction, there are now two individuals, each of which is again the founding individual of a lineage. In order for the original lineage to go extinct, both of these lineages have to die out. It therefore holds for the extinction probability $Q_{(i,j;f)}$:

$$
Q_{(i,j;f)} = \frac{1 - \sigma_{(i,j;f)}}{2 - \sigma_{(i,j;f)}} + \sum_{g=0}^{f} \binom{f}{g} \left(\frac{1}{2}\right)^f \frac{1 - r_{(i,j)}}{2 - \sigma_{(i,j;f)}} Q_{(i,j;g)} Q_{(i,j;f)}
$$

$$
+ \sum_{g=0}^{f} \binom{f}{g} \left(\frac{1}{2}\right)^f \left\{ \sum_{k=0}^{i-1} \frac{r_{(k,j)}^{(i,j)}}{2 - \sigma_{(i,j;f)}} Q_{(k,j;g)} + \sum_{k=0}^{j-1} \frac{r_{(i,k)}^{(i,j)}}{2 - \sigma_{(i,j;f)}} Q_{(i,k;g)} \right\} Q_{(i,j;f)}.
$$

$$(3)$$

By rearranging terms, we obtain Eq. (2).

For the special case $F = 0$, Eq. (2) simplifies to

$$
0 = (1 - r_{(i,j)})Q_{(i,j)}^2 + \sum_{k=0}^{j-1} r_{(i,k)}^{(i,j)} Q_{(i,j)} Q_{(i,k)} + \sum_{k=0}^{i-1} r_{(k,j)}^{(i,j)} Q_{(i,j)} Q_{(k,j)}
$$

$$
+ 1 - \sigma_{(i,j)} - (2 - \sigma_{(i,j)})Q_{(i,j)}, \quad i \in \{0, \ldots, I\}, \quad j \in \{0, \ldots, J\}. \quad (4)
$$

If all deleterious alleles are unlinked (i.e., $I = J = 0$; abbreviate $(0, 0; g) \equiv g$), Eq. (2) yields:

$$
0 = \left(\frac{1}{2}\right)^f Q_f^2 + \left(\frac{1}{2}\right)^f \sum_{g=0}^{f-1} \binom{f}{g} Q_g Q_f + 1 - \sigma_f - (2 - \sigma_f)Q_f, \quad f \in \{0, \ldots, F\}.
$$

$$(5)$$

Implications of these results are discussed in Sect. 6.

In Appendix B, we relate our results to results by Bengtsson (1985) and Barton and Bengtsson (1986) on the impact of a cline on the spread of a neutral marker allele. Our results are consistent with Bengtsson (1985) and Barton and Bengtsson (1986) when all loci are only loosely linked or unlinked but deviate for tight linkage of the deleterious allele.

5 The hitchhiking probability

5.1 General idea

If the effects of closely linked deleterious alleles are not too harmful, namely if $\sigma_{(i,j)} - r_{(i,j)} > 0$, the beneficial allele can drag (some of) these deleterious alleles along to fixation. In this section, we develop a framework for determining the hitchhiking probabilities conditioned on fixation of the beneficial allele. The approach is based on a split of the process into two phases. After the original hybridization event, the beneficial allele must establish in the population. We call this the "stochastic phase". In

the previous section, we were concerned with the establishment probability itself. Here, we further derive which haplotype (i, j) will escape stochastic loss in this initial phase conditioned on survival of the process. We assume that only one haplotype escapes. This is a very likely outcome of the stochastic phase under many circumstances because the establishment probability of each type is low. Since this establishment happens while the introgressed types are rare, we base the derivation on the multitype branching process as before. Once established, type (i, j) increases in frequency approximately as predicted by deterministic growth. If no further recombination events were to happen, it would rise to fixation following the logistic equation

$$\dot{x}_{(i,j)}(t) = \sigma_{(i,j)} x_{(i,j)}(t)(1 - x_{(i,j)}(t)). \tag{6}$$

However, during the sweep, types with fewer deleterious alleles can still be generated by recombination. If one of these types establishes, it outcompetes type (i, j). Building on theory by Hartfield and Otto (2011), we describe the production and possible establishment of these types by a time-inhomogeneous branching process with immigration. Although the generation and establishment of new haplotypes is subject to strong stochasticity, we refer to this phase as to the "deterministic phase" because we model the frequency paths of the established haplotypes deterministically.

We first give a derivation for the case without unlinked deleterious material ($F = 0$), and subsequently generalize the approximation to $F > 0$.

5.2 The stochastic phase

As a first step, we determine which haplotype "rescues" the introgression process given that the process does not go extinct. For this initial phase, we again resort to the multitype branching process as defined in Eq. (1). As before, the process is initiated by a single individual of type (I, J). If $\sigma_{(I,J)} - r_{(I,J)} > 0$, type (I, J) has the chance to establish a permanent lineage of its own type. If $\sigma_{(I,J)} - r_{(I,J)} \leq 0$, type (I, J) itself will go extinct with probability 1 (ignoring the possibility of fixation by drift). However, until extinction, recombinant offspring with fewer deleterious alleles can be generated and rescue the process. In that case, to determine the "rescue type", we can consider all recombination pathways that lead to establishment of the beneficial allele and determine with which (relative) probability the various paths are realized. This idea is key for the derivation of the approximation in this section.

Throughout the analysis, the total number of recombination events from type (I, J) to any other type until extinction of type (I, J) constitutes a central quantity. This follows Serra (2006) and Serra and Haccou (2007). For $\sigma_{(I,J)} - r_{(I,J)} < 0$, we denote the corresponding probability generating function (p.g.f.) by $h(s)$. For $\sigma_{(I,J)} - r_{(I,J)} > 0$ and extinction of type (I, J), we consider the number of recombination events conditioned on extinction of type (I, J) and denote the p.g.f. by $\hat{h}(s)$. $h(s)$ and $\hat{h}(s)$ can be explicitly calculated for our model and are given by Lemma 6 in Appendix E.

As a first step, we derive an alternative expression for the survival probability of the process. To do so, we group the recombinant offspring of type (I, J) individuals into two classes: (1) individuals that found processes that survive (2) individuals that

found processes that go extinct. We denote by Y_+ and Y_- the random number of recombination events from type (I, J) to type 1 and type 2 individuals, respectively. In the lemma, we rewrite the survival probability of the process in terms of the expected number of successful recombinant lineages and an error term.

Lemma 1 *Let* $\sigma_{(I,J)} - r_{(I,J)} < 0$. *The survival probability of the process can be written as*

$$1 - Q_{(I,J)} = \left(\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)} (1 - Q_{(k,J)}) + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)} (1 - Q_{(I,k)}) \right) \frac{\frac{\mathrm{d}}{\mathrm{d}s} h(s)|_{s=1}}{r_{(I,J)}} - R_1, \tag{7}$$

where the error term R_1 is given by

$$R_1 = \frac{\partial}{\partial s_0} \left(\frac{h_1(s_0, s_1) - h_1(0, s_1)}{s_0} \right) \Bigg|_{s_0 = s_1 = 1} \tag{8}$$

with

$$h_1(s_0, s_1) = h(P_{\mathrm{success}} s_0 + (1 - P_{\mathrm{success}}) s_1) \tag{9}$$

and

$$P_{\mathrm{success}} = \frac{\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)} (1 - Q_{(k,J)}) + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)} (1 - Q_{(I,k)})}{r_{(I,J)}}. \tag{10}$$

Proof A recombinant offspring of a type (I, J) individual founds an infinite lineage with probability P_{success} and a lineage that goes extinct with probability $1 - P_{\mathrm{success}}$. According to Lemma 7 in Appendix E, the joint p.g.f of Y_+ and Y_- is given by

$$h_1(s_0, s_1) = h(P_{\mathrm{success}} s_0 + (1 - P_{\mathrm{success}}) s_1), \tag{11}$$

and we obtain for the expected number of type 1 individuals:

$$E[Y_+] = \frac{\partial}{\partial s_0} h_1(s_0, s_1)|_{s_0 = s_1 = 1} = P_{\mathrm{success}} \frac{\mathrm{d}}{\mathrm{d}s} h(s)|_{s=1}. \tag{12}$$

Now note:

$$1 - Q_{(I,J)} = P(Y_+ > 0) = E[Y_+] - R_1 \tag{13}$$

with

$$\begin{aligned} R_1 &= E[Y_+] - P(Y_+ > 0) = P(Y_+ = 2) + 2P(Y_+ = 3) + 3P(Y_+ = 4) + \dots \\ &= \frac{\partial}{\partial s_0} \left(\frac{h_1(s_0, s_1) - h_1(0, s_1)}{s_0} \right) \Bigg|_{s_0 = s_1 = 1}. \end{aligned} \tag{14}$$

I.e., if $P(Y_+ > 1) \approx 0$, the expected number of recombination events from type (I, J) individuals to individuals with fewer mutations that found a successful lineage approximates the survival probability of the process. □

Remark For $\sigma_{(I,J)} - r_{(I,J)} > 0$, we can consider the process conditioned on extinction of type (I, J). In this case, an analogous result holds if we replace $h(s)$ by $\hat{h}(s)$.

Results of a similar structure, which approximate weak recombination, appear in Iwasa et al. (2004a, Eq. (5)) and Serra and Haccou (2007, Eq. (8)).

In order to proceed, we need a formal definition of a "rescue type". Analogous to the lemma, we can then derive a recursive formula for the probability that an individual of type (i, j) rescues the process.

Definition 1 We call an (i, j) individual an (i, j) "rescue type", denoted as $(i, j, +)$, if

(1) it founds an infinite lineage of type (i, j) individuals,
(2) there is no individual in its ancestry that founds an infinite lineage of its own type.

We denote by $X_{(i,j,+)}^{(k,l)}$ the number of rescue types $(i, j, +)$ in a process which is founded by an individual of type (k, l).

We define

$$P_{(i,j)}^{(I,J)} = \text{Prob}(X_{(i,j,+)}^{(I,J)} > 0 | \text{survival of the process}). \tag{15}$$

That is, $P_{(i,j)}^{(I,J)}$ gives the probability that there exists an (i, j) rescue type conditioned on survival of the process. A priori, that does not exclude the simultaneous existence of several rescue types. For the following theorem, we again group the recombinant offspring of a type (I, J) individual into two classes: (1) individuals that found a lineage resulting in at least one individual of type $(i, j, +)$ (2) individuals that do not do that. We denote the number of recombinants of the first and second type with $Y_{(i,j,+)}$ and $Y_{(i,j,-)}$, respectively.

Theorem 2 (1) Let $\sigma_{(I,J)} - r_{(I,J)} < 0$. It holds that

$$P_{(i,j)}^{(I,J)} = \frac{\left(\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)}(1 - Q_{(k,J)})P_{(i,j)}^{(k,J)} + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)}(1 - Q_{(I,k)})P_{(i,j)}^{(I,k)}\right)\frac{\frac{d}{ds}h(s)|_{s=1}}{r_{(I,J)}} - R_1}{\left(\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)}(1 - Q_{(k,J)}) + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)}(1 - Q_{(I,k)})\right)\frac{\frac{d}{ds}h(s)|_{s=1}}{r_{(I,J)}} - R_2},$$

$$\tag{16}$$

where R_1 is defined as before and R_2 is given by

$$R_2 = \frac{\partial}{\partial s_0}\left(\frac{h_2(s_0, s_1) - h_2(0, s_1)}{s_0}\right)\Bigg|_{s_0=s_1=1} \tag{17}$$

with

$$h_2(s_0, s_1) = h(P_{(i,j,+)}s_0 + (1 - P_{(i,j,+)})s_1) \tag{18}$$

and

$$P_{(i,j,+)} = \frac{\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)}(1 - Q_{(k,J)})P_{(i,j)}^{(k,J)} + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)}(1 - Q_{(I,k)})P_{(i,j)}^{(I,k)}}{r_{(I,J)}}.$$

(19)

(2) For $\sigma_{(I,J)} - r_{(I,J)} > 0$, it holds:

$$P_{(I,J)}^{(I,J)} = \frac{1 - q_{(I,J)}}{1 - Q_{(I,J)}}$$

(20)

with

$$1 - q_{(I,J)} = \frac{\sigma_{(I,J)} - r_{(I,J)}}{1 - r_{(I,J)}},$$

(21)

where $q_{(I,J)}$ is the unconditioned probability that type (I, J) itself goes extinct. For $(i, j) \neq (I, J)$, it holds:

$$P_{(i,j)}^{(I,J)} = \left(1 - \frac{1 - q_{(I,J)}}{1 - Q_{(I,J)}}\right)$$
$$\times \frac{\left(\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)}(1 - Q_{(k,J)})P_{(i,j)}^{(k,J)} + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)}(1 - Q_{(I,k)})P_{(i,j)}^{(I,k)}\right)\frac{\frac{d}{ds}\hat{h}(s)|_{s=1}}{r_{(I,J)}} - \hat{R}_1}{\left(\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)}(1 - Q_{(k,J)}) + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)}(1 - Q_{(I,k)})\right)\frac{\frac{d}{ds}\hat{h}(s)|_{s=1}}{r_{(I,J)}} - \hat{R}_2},$$

(22)

where $\hat{h}_1$, $\hat{h}_2$, $\hat{R}_1$, and $\hat{R}_2$ are defined analogously to before (using $\hat{h}$ instead of h).

Proof We first prove part (1) of the theorem. With probability $P_{(i,j,+)}$, a recombinant offspring founds a lineage that generates at least one individual of type $(i, j, +)$. Analogous to before, we obtain

$$P(Y_{(i,j,+)} > 0) = E[Y_{(i,j,+)}] - R_2$$
$$= P_{(i,j,+)}\frac{d}{ds}h(s)|_{s=1} - R_2$$

(23)

with

$$R_2 = \frac{\partial}{\partial s_0}\left(\frac{h_2(s_0, s_1) - h_2(0, s_1)}{s_0}\right)\Bigg|_{s_0=s_1=1}.$$

(24)

It holds:

$$P_{(i,j)}^{(I,J)}(1 - Q_{(I,J)}) = P(Y_{(i,j,+)} > 0).$$

(25)

Substituting $1 - Q_{(I,J)}$ by the approximation Eq. (7) yields Eq. (16).

If $\sigma_{(I,J)} - r_{(I,J)} > 0$, type (I, J) establishes a lineage of its own type with probability

$$1 - q_{(I,J)} = \frac{\sigma_{(I,J)} - r_{(I,J)}}{1 - r_{(I,J)}} \tag{26}$$

(cf. Lemma 4). It therefore holds:

$$P_{(I,J)}^{(I,J)} = \frac{1 - q_{(I,J)}}{1 - Q_{(I,J)}}. \tag{27}$$

The probability that type (I, J) goes extinct, conditioned on survival of the process, is accordingly given by

$$P(\text{type } (I, J) \text{ goes extinct}|\text{survival of the process}) = 1 - \frac{1 - q_{(I,J)}}{1 - Q_{(I,J)}}. \tag{28}$$

We can now repeat the proof of the first part of the theorem for the process conditioned on extinction of type (I, J).   □

Remark 1 For $\sigma_{(I,J)} - r_{(I,J)} < 0$, we have $\frac{d}{ds}h(s)|_{s=1} = \frac{r_{(I,J)}}{r_{(I,J)} - \sigma_{(I,J)}}$. For $\sigma_{(I,J)} - r_{(I,J)} > 0$, we have $\frac{d}{ds}\hat{h}(s)|_{s=1} = \frac{r_{(I,J)}}{\sigma_{(I,J)} - r_{(I,J)}}$.

For $\sigma_{(I,J)} - r_{(I,J)} < 0$ not too close to zero, it is likely that only one of the few recombinant offspring of type (I, J) individuals founds an infinite lineage, and we can approximate $P(Y_+ \geq 2) \approx 0$ and consequently also $P(Y_{(i,j,+)} \geq 2) \approx 0$. This implies that the error terms R_1 and R_2 can be ignored. For $\sigma_{(I,J)} - r_{(I,J)} > 0$ and $r_{(I,J)}$ small, survival of the process is with high probability contingent on establishment of type (I, J) so that

$$P_{(i,j)}^{(I,J)} \approx \delta_{I,i}\delta_{J,j} \quad \text{with} \quad \delta_{k_1,k_2} = \begin{cases} 1 & \text{for} \quad k_1 = k_2, \\ 0 & \text{else.} \end{cases} \tag{29}$$

We can therefore formulate the following corollary:

Corollary 1 *For $\sigma_{(I,J)} - r_{(I,J)} < 0$ not too close to zero and close linkage, we can approximate*

$$P_{(i,j)}^{(I,J)} \approx \frac{P_{(i,j,+)}}{P_{\text{success}}} = \frac{\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)}(1 - Q_{(k,J)})P_{(i,j)}^{(k,J)} + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)}(1 - Q_{(I,k)})P_{(i,j)}^{(I,k)}}{\sum_{k=0}^{I-1} r_{(k,J)}^{(I,J)}(1 - Q_{(k,J)}) + \sum_{k=0}^{J-1} r_{(I,k)}^{(I,J)}(1 - Q_{(I,k)})} \tag{30}$$

with

$$P_{(k,l)}^{(i,j)} \approx \delta_{i,k}\delta_{j,l} \tag{31}$$

for $\sigma_{(i,j)} - r_{(i,j)} > 0$. In this approximation, the $P_{(i,j)}^{(I,J)}$, $i \leq I$, $j \leq J$, form a probability distribution with

$$\sum_{i,j} P_{(i,j)}^{(I,J)} = 1. \tag{32}$$

The proof for relation Eq. (32) is given in Appendix C.

The approximation [Eqs. (30) and (31)] implies that exactly one rescue type establishes in the population during the stochastic phase (types with fewer deleterious alleles can still arise later during the deterministic phase), i.e.,

$$P(X_{(i,j,+)}^{(I,J)} + X_{(k,l,+)}^{(I,J)} > 1|\text{survival}) = 0 \quad \text{for any pair} \quad (i, j), (k, l). \tag{33}$$

This assumption appears to be justified for a large parameter region with tightly linked deleterious alleles. It is also the basis for most of our analytical treatment of particular cases below. As discussed in Appendix H, the approximation is less accurate if deleterious alleles are relatively loosely linked and/or haplotypes are only slightly deleterious, but in most cases, it introduces only a small error.

We can extend the approximation to include unlinked deleterious alleles and obtain for $\sigma_{(I,J)} - r_{(I,J)} < 0$:

$$P_{(i,j;0)}^{(I,J;F)} \approx \frac{A + B}{C + D}$$

$$A = \sum_{l=0}^{F} \binom{F}{l} \left(\frac{1}{2}\right)^{F} \left\{ \sum_{i=0}^{I-1} r_{(i,J)}^{(I,J)} (1 - Q_{(i,J;l)}) P_{(i,j;0)}^{(i,J;l)} \right.$$

$$\left. + \sum_{j=0}^{J-1} r_{(I,j)}^{(I,J)} (1 - Q_{(I,j;l)}) P_{(i,j;0)}^{(I,j;l)} \right\}$$

$$B = \sum_{l=0}^{F-1} \binom{F}{l} \left(\frac{1}{2}\right)^{F} (1 - r_{(I,J)})(1 - Q_{(I,J;l)}) P_{(i,j;0)}^{(I,J;l)}$$

$$C = \sum_{l=0}^{F} \binom{F}{l} \left(\frac{1}{2}\right)^{F} \left\{ \sum_{i=0}^{I-1} r_{(i,J)}^{(I,J)} (1 - Q_{(i,J;l)}) + \sum_{j=0}^{J-1} r_{(I,j)}^{(I,J)} (1 - Q_{(I,j;l)}) \right\}$$

$$D = \sum_{l=0}^{F-1} \binom{F}{l} \left(\frac{1}{2}\right)^{F} (1 - r_{(I,J)})(1 - Q_{(I,J;l)}). \tag{34}$$

For $\sigma_{(I,J)} - r_{(I,J)} > 0$, we approximate:

$$P_{(i,j;0)}^{(I,J;F)} \approx \delta_{I,i}\delta_{J,j}. \tag{35}$$

5.3 The deterministic phase

It remains to determine whether the haplotype that establishes in the stochastic phase rises to fixation or if types with less deleterious material can establish during the sweep of the beneficial allele. In order to arrive at an approximation for the deterministic phase, we apply and extend an approach developed in Hartfield and Otto (2011). Hartfield and Otto (2011) determined the hitchhiking probability of a single deleterious allele which is closely linked to a beneficial one. For a single hitchhiker, their method can easily be adapted to our model, as shown below. In the Appendix, we further argue that the approach can be extended to a larger number of hitchhikers. Explicit results for two hitchhikers are derived in Appendix F.

For a single potential hitchhiker, assume that type $(0, 1)$ with $\sigma_{(0,1)} - r_{(0,1)} > 0$ has been introduced and established in the population. Its further growth can be well described deterministically as given by the differential equation Eq. (6). However, in the initial phase, it will on average have grown faster than the deterministic path predicts. Following Uecker and Hermisson (2011), we account for the fast initial increase by the use of an "effective initial population size" ν, which we use as an initial condition for the solution of Eq. (6) (cf. also Desai and Fisher 2007). ν is an exponentially distributed random variable with

$$P(\nu \leq \nu_0) = 1 - \exp\left(-p_{\text{est}}\nu_0\right), \tag{36}$$

where $p_{\text{est}} = 1 - q_{(0,1)}$ (cf. Eq. (26)) denotes the establishment probability of a single type $(0, 1)$ individual in a wildtype population (Uecker and Hermisson 2011 Eq. (40)). To leading order approximation, we can approximate the distribution by its mean $\bar{\nu} = 1/p_{\text{est}}$. For the relative frequency of type $(0, 1)$, it then holds (ignoring recombination):

$$x_{(0,1)}(t) = \frac{\bar{\nu} \exp\left(\sigma_{(0,1)}t\right)}{N - \bar{\nu} + \bar{\nu} \exp\left(\sigma_{(0,1)}t\right)}. \tag{37}$$

This is a good approximation up to frequency $x_{(0,1)} \approx 1 - \frac{\bar{\nu}}{N}$. At higher frequencies, the frequency will again grow faster. Individuals of type $(0, 0)$ are recurrently generated by recombination at rate $r_{(0,1)}Nx_{(0,1)}(1 - x_{(0,1)})$. As long as they are rare, their dynamics are strongly determined by stochasticity and can once again be approximated by a branching process. Their fitness depends on $x_{(0,1)}(t)$ and hence on time. The dynamics is thus described by a time-inhomogeneous branching process with birth rate 1 and death rate $1 - \sigma_{(0,0)} + x_{(0,1)}(t)\sigma_{(0,1)}$. Following Eq. (16a) in Uecker and Hermisson (2011), the establishment probability of a single individual of type $(0, 0)$ generated at time T is given by

$$p_{\text{est}}^{(0,0)}(T) = \frac{\sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,1)})}{(\sigma_{(0,0)} - \sigma_{(0,1)})(1 - x_{(0,1)}(T)) + \sigma_{(0,0)}x_{(0,1)}(T)}. \tag{38}$$

"Successful" individuals of type $(0, 0)$ are generated at rate

$$r_{(0,0)}^{(0,1)} Nx_{(0,1)}(t)(1 - x_{(0,1)}(t))p_{\text{fix}}^{(0,0)}(t). \tag{39}$$

Using this, we obtain for the probability that type $(0, 1)$ fixes in the population:

$$
\begin{aligned}
P_{\text{det}}^{((0,1)\to(0,1))} &= \exp\left[-\int_0^\infty Nr_{(0,0)}^{(0,1)} x_{(0,1)}(t)(1 - x_{(0,1)}(t)) p_{\text{fix}}^{(0,0)}(t)\mathrm{d}t \right] \\
&\approx \exp\left[-\int_{\bar{\nu}/N}^{1-\bar{\nu}/N} Nr_{(0,0)}^{(0,1)} \frac{1}{\sigma_{(0,1)}} \frac{\sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,1)})}{(\sigma_{(0,0)} - \sigma_{(0,1)})(1 - x) + \sigma_{(0,0)}x}\mathrm{d}x \right] \\
&= \left(\frac{\sigma_{(0,0)} - \frac{\bar{\nu}}{N}\sigma_{(0,1)}}{\sigma_{(0,0)} - \sigma_{(0,1)}(1 - \frac{\bar{\nu}}{N})} \right)^{-\frac{Nr_{(0,0)}^{(0,1)}\sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,1)})}{\sigma_{(0,1)}^2}},
\end{aligned}
\tag{40}
$$

where the approximation requires $\bar{\nu} < 0.5N$. For a single introgressed individual at time $t = 0$ in a large population, we can approximate $\bar{\nu}/N \approx 0$ and obtain

$$
P_{\text{det}}^{((0,1)\to(0,1))} \approx \left(\frac{\sigma_{(0,0)}}{\sigma_{(0,0)} - \sigma_{(0,1)}} \right)^{-\frac{Nr_{(0,0)}^{(0,1)}\sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,1)})}{\sigma_{(0,1)}^2}}.
\tag{41}
$$

Eq. (41) corresponds to Eq. (5) in Hartfield and Otto (2011) up to a model-specific factor of 2 in the exponent if we identify $s_a \equiv \sigma_{(0,0)}, s_d \equiv \sigma_{(0,0)} - \sigma_{(0,1)}$, and $r \equiv r_{(0,0)}^{(0,1)}$. If $N(1 - q_{(0,1)}) \approx N\sigma_{(0,1)}$ is small or if there are already other introgressed haplotypes sweeping in the population as in the generalization to more potential hitchhikers, it makes a quantitative difference whether one accounts for the fast initial increase or not, and we cannot approximate $\bar{\nu}/N \approx 0$ (see Appendix F). Alternatively, one can resort to a diffusion approach for these cases (see Hartfield and Otto 2011 and Appendix G). Note that both approaches assume that recombination is so weak that by itself, it does not influence the frequency path of type $(0, 1)$ (i.e., $\sigma_{(0,1)} \gg r_{(0,0)}^{(0,1)}$).

5.4 Concatenation of the stochastic and the deterministic phase

In order to determine which haplotype fixes in the population, we need to concatenate the stochastic and the deterministic phase. Let $\mathcal{A}$ be the set of all types with positive fitness. Type (k, l) establishes in the stochastic phase with probability $P_{(k,l)}^{(I,J)}$ as derived and discussed in Sect. 5.2 and hence enters the deterministic phase. We always assume that only one type does so (and it does so in a single lineage, i.e., there is only one "rescue individual"). Given establishment of type (k, l), we denote by

$$
P_{\text{det}}^{((k,l)\to(i,j))}
\tag{42}
$$

the probability that during the deterministic phase, type $(i, j) \in \mathcal{A}$ is generated and finally fixes in the population. Summing over all $(k, l) \in \mathcal{A}$ yields:

$$P(\text{"type } (i, j) \text{ fixes"}) = \sum_{(k,l)\in\mathcal{A}} P^{(I,J)}_{(k,l)} P^{((k,l)\to(i,j))}_{\det}. \tag{43}$$

If not stated otherwise, the results presented in Sect. 6 are based on Eq. (43) with $P^{(I,J)}_{(k,l)}$ obtained by Eqs. (30) and (31). The recursions are performed by a program written in the C programming language. Approximations for $P^{((0,2)\to(0,\cdot))}_{\det}$ with $\sigma_{(2,0)} - r_{(0,2)} > 0$ and $P^{((1,1)\to(\cdot,\cdot))}_{\det}$ with $\sigma_{(1,1)} - r_{(1,1)} > 0$ are derived in Appendix F. All numerical evaluation of the integrals that appear in these approximations is done in Mathematica (Wolfram Research, Champaign, USA). The accuracy of the approach and the appropriateness of the assumptions are addressed in Appendix H.

6 Application to various biological scenarios

6.1 The impact of unlinked alleles

If $I = J = 0$, the extinction probability is given by Eq. (5). For Q_0, Q_1 and Q_2, we obtain:

$$Q_0 = 1 - \sigma_0, \tag{44a}$$

$$Q_1 = 2 - \sigma_1 - \frac{1}{2}Q_0 - \sqrt{\left(2 - \sigma_1 - \frac{1}{2}Q_0\right)^2 - 2(1 - \sigma_1)}, \tag{44b}$$

$$Q_2 = 4 - 2\sigma_2 - \frac{1}{2}Q_0 - Q_1 - 2\sqrt{\left(\frac{1}{4}Q_0 + \frac{1}{2}Q_1 - 2 + \sigma_2\right)^2 - (1 - \sigma_2)}. \tag{44c}$$

How does the number of unlinked deleterious alleles impact the introgression probability if their total effect is kept constant? A comparison of $1 - Q_1$ with $\sigma_1 = \sigma_0 - 2s_{\text{del}}$ and $1 - Q_2$ with $\sigma_1 = \sigma_0 - s_{\text{del}}$ and $\sigma_2 = \sigma_0 - 2s_{\text{del}}$ yields:

$$(1 - Q_2(s_{\text{del}}))/(1 - Q_1(2s_{\text{del}})) = 1 + \mathcal{O}(s_{\text{del}}^2), \tag{45}$$

i.e., unless the deleterious effect is very strong, the establishment probability is approximately the same for both scenarios (either one deleterious allele of effect $2s_{\text{del}}$ or two deleterious alleles each of effect s_{del}). Figure 2 generalizes this result to $F > 2$. One sees that unlinked alleles significantly reduce the introgression probability. However, it is irrelevant whether there is one strongly deleterious allele or many slightly deleterious alleles. By how much do unlinked deleterious alleles of compound effect S_{del} reduce the introgression probability? Making use of the previous observation, it is sufficient to consider a single unlinked allele of effect S_{del}. A Taylor expansion yields:

$$\frac{1 - Q_1}{1 - Q_0} = 1 - \frac{2}{1 + \sigma_0}S_{\text{del}} + \mathcal{O}(S_{\text{del}}^2) \approx 1 - 2S_{\text{del}} + \mathcal{O}(S_{\text{del}}^2), \tag{46}$$
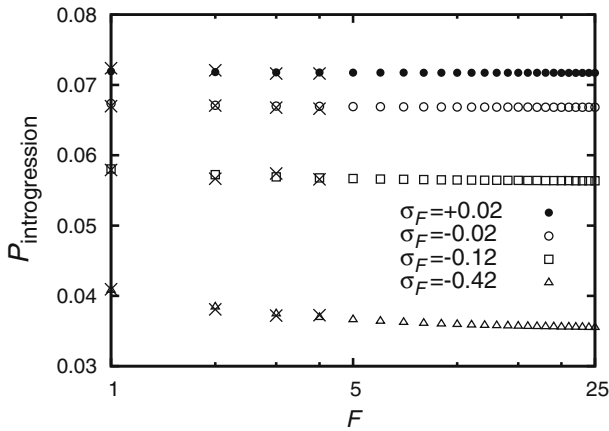
Fig. 2 The introgression probability as a function of the number of unlinked deleterious alleles. The total effect on fitness is kept constant; σ_F denotes the Malthusian fitness of a haplotype carrying the adaptive allele and F unlinked deleterious alleles. The introgression probability is approximately the same whether the effect is distributed over few strongly or many slightly deleterious alleles. The advantageous allele has Malthusian fitness $\sigma_0 = 0.08$. In the absence of deleterious alleles, it would establish with probability $1 - Q_0 = \sigma_{(0,0)} = 0.08$. The *crosses* denote simulation results. Each simulation point is the average of 10^6 introgression attempts
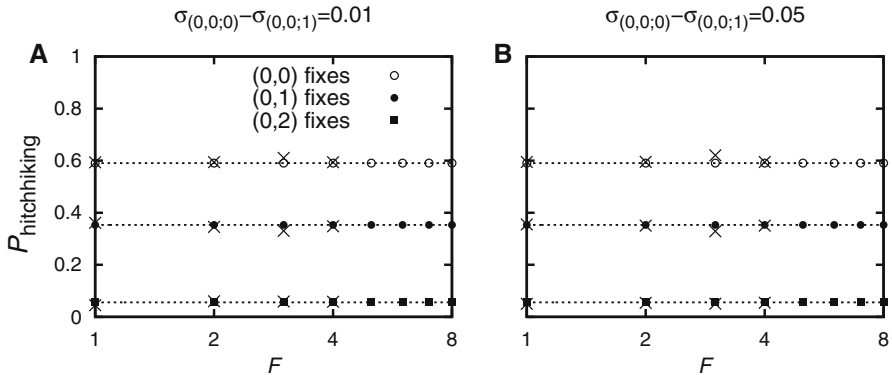

Fig. 3 The hitchhiking probability as a function of the number of unlinked deleterious alleles. The *dotted lines* correspond to the respective values for $F = 0$. Each unlinked deleterious allele reduces fitness by -0.01 (Panel **a**) and -0.05 (Panel **b**), respectively. Unlinked alleles do not visibly impact the hitchhiking probability of closely linked deleterious alleles. Parameter values are: $I = 0$, $J = 3$, $\sigma_0 = 0.075$, $\sigma_{(0,1)} = 0.07$, $\sigma_{(0,2)} = 0.05$, $\sigma_{(0,3)} = -0.015$, $N = 10{,}000$, $r_{(\cdot,\cdot)}^{(\cdot,\cdot)} = 0.0001$. The *crosses* denote simulation results. Each simulation point is the average of 2,000 successful introgression events

i.e., unlinked alleles approximately reduce the introgression probability by a factor that is independent of σ_0.

While unlinked alleles have a significant impact on the probability of adaptive gene introgression ($\sim$10–50 % in Fig. 2), they do not visibly influence the hitchhiking probability of closely linked deleterious alleles (cf. Fig. 3).

6.2 The impact of a single linked deleterious allele

In this section, we consider the impact of a single linked deleterious allele (cf. also Iwasa et al. 2004a). From Eq. (4), we obtain:

$$Q_{(0,0)} = 1 - \sigma_{(0,0)}, \tag{47a}$$

$$Q_{(0,1)} = \frac{1}{2(1 - r_{(0,1)})} \left(2 - \sigma_{(0,1)} - r_{(0,1)} Q_{(0,0)} \right. \tag{47b}$$

$$\left. - \sqrt{\left(2 - \sigma_{(0,1)} - r_{(0,1)} Q_{(0,0)}\right)^2 - 4(1 - \sigma_{(0,1)})(1 - r_{(0,1)})} \right) \tag{47c}$$

$$\approx \begin{cases} (1 - \sigma_{(0,1)}) \left(1 - \frac{\sigma_{(0,0)} - \sigma_{(0,1)}}{\sigma_{(0,1)}} r_{(0,1)} \right) & \text{for } \sigma_{(0,1)} > 0, \\ 1 + \frac{\sigma_{(0,0)}}{\sigma_{(0,1)}} r_{(0,1)} & \text{for } \sigma_{(0,1)} < 0. \end{cases} \tag{47d}$$

The approximation is a first order Taylor expansion in $r_{(0,1)}$, which yields accurate results for small $r_{(0,1)}$ if $\sigma_{(0,1)}$ is not too close to zero. Due to the assumption of single crossover only, $Q_{(0,1)}$ exactly corresponds to Q_1 for $r_{(0,1)} = 0.5$ (where $\sigma_{(0,0)} \equiv \sigma_0$ and $\sigma_{(0,1)} \equiv \sigma_1$). How does a single deleterious allele impact the probability of adaptive gene introgression? We can measure the impact by the relative reduction of the introgression probability

$$\Delta P = 1 - \frac{1 - Q_{(0,1)}}{1 - Q_{(0,0)}} = 1 - \frac{1 - Q_{(0,1)}}{\sigma_{(0,0)}}. \tag{48}$$

If ΔP is close to zero, the deleterious allele has a weak impact; if ΔP is close to one, it has a strong impact. The influence is obviously strongest for $r_{(0,1)} = 0$. If $\sigma_{(0,1)} > 0$, the maximum relative reduction in the introgression probability is given by $s_{\text{del}}/\sigma_{(0,0)}$ with $s_{\text{del}} := \sigma_{(0,0)} - \sigma_{(0,1)}$. For $\sigma_{(0,1)} \leq 0$ and $r_{(0,1)} = 0$, the advantageous allele can not introgress at all (except through fixation by drift which is not considered here). For tight linkage, we use the foregoing Taylor expansion and obtain

$$\Delta P \approx \frac{s_{\text{del}}}{\sigma_{(0,0)}} \left(1 - \frac{r_{(0,1)}}{\sigma_{(0,1)}} \right) \quad \text{for} \quad \sigma_{(0,1)} > 0, \tag{49a}$$

$$\Delta P \approx 1 + \frac{r_{(0,1)}}{\sigma_{(0,1)}} \quad \text{for} \quad \sigma_{(0,1)} < 0. \tag{49b}$$

The maximum impact is strongest for a weak beneficial mutation (where introgression is easily reduced to zero for tight linkage). The impact gets weaker with increasing recombination on the scale of $\sigma_{(0,1)}$. I.e., if either $\sigma_{(0,0)} \gg s_{\text{del}}$, or if $s_{\text{del}} \gg \sigma_{(0,0)}$, the impact declines only slowly. In order to determine the behavior for strong recombination, we perform a Taylor expansion of ΔP in s_{del}:

$$\Delta P = \frac{1}{r_{(0,1)} + \sigma_{(0,0)}(1 - r_{(0,1)})} s_{\text{del}} + \mathcal{O}\left(s_{\text{del}}^2\right). \tag{50}$$

For $r_{(0,1)} \gg \sigma_{(0,0)}$, the relative reduction becomes independent of the strength of the beneficial allele. This is in strong contrast to the behavior for the tight linkage case. The impact declines on a scale of s_{del}. It becomes irrelevant if $r_{(0,1)} \gg s_{\text{del}}$. An unlinked

deleterious allele leads to a relative reduction of $2s_{del}$, which can still be appreciable if the deleterious mutation has a strong effect.

If $\sigma_{(0,1)} - r_{(0,1)} > 0$, the deleterious allele can hitchhike to fixation. For the assessment of its hitchhiking probability, we do not follow the simple approximation (31) but give a detailed analysis of the stochastic establishment phase instead. Type $(0, 1)$ establishes with probability

$$1 - q_{(0,1)} = \frac{\sigma_{(0,1)} - r_{(0,1)}}{1 - r_{(0,1)}}, \tag{51}$$

i.e.,

$$P_{(0,1)}^{(0,1)} = \frac{\sigma_{(0,1)} - r_{(0,1)}}{1 - r_{(0,1)}} \frac{1}{1 - Q_{(0,1)}}, \tag{52}$$

and

$$P_{\text{hitchhiking}} = P_{(0,1)}^{(0,1)} P_{\text{det}}^{((0,1) \to (0,1))}. \tag{53}$$

In order to asses the relevance of the stochastic and the deterministic phases, we perform a first-order Taylor expansion of $P_{(0,1)}^{(0,1)}$ in $r_{(0,1)}$:

$$P_{(0,1)}^{(0,1)} \approx 1 - r_{(0,1)}\sigma_{(0,0)} \frac{1 - \sigma_{(0,1)}}{\sigma_{(0,1)}^2} \approx 1 - r_{(0,1)} \frac{\sigma_{(0,0)}}{\sigma_{(0,1)}^2}. \tag{54}$$

I.e., changes in $P_{(0,1)}^{(0,1)}$ occur on the scale of $r_{(0,1)} \sim \sigma_{(0,1)}^2 / \sigma_{(0,0)}$. For the deterministic phase [Eq. (41)], the scale is set by $r_{(0,1)} \sim \sigma_{(0,1)}^2 / (N\sigma_{(0,0)}s_{del})$. This allows us to distinguish two parameter regimes: if $Ns_{del} \gg 1$, the deterministic phase dominates. However, if $Ns_{del} \approx 1$ or smaller, the stochastic phase cannot be ignored. Figures 4C and D illustrate how the stochastic and deterministic phases combine to form the probability of hitchhiking for $N = 10000$ and $N = 500$, respectively. For $N = 10,000$, the behavior is dominated by the deterministic phase which decays quickly as a function of $r_{(0,1)}$. In the parameter range where $P_{\text{det}}^{((0,1) \to (0,1))}$ is appreciable, one can ignore the influence of the stochastic phase. For $N = 500$, however, $P_{\text{det}}^{((0,1) \to (0,1))}$ decays slowly, and the stochastic phase has a non-negligible impact on hitchhiking: e.g., for $r_{(0,1)} = 0.003$, we find $P_{\text{det}}^{((0,1) \to (0,1))} \approx 0.8$ and $P_{(0,1)}^{(0,1)} P_{\text{det}}^{((0,1) \to (0,1))} \approx 0.53$. Note also that if $N(1 - q_{(0,1)})$ is small, we have to account for deviations from the deterministic path. These deviations can be accounted for via the parameter $\bar{\nu}$ as in Eq. (40) or via a diffusion approach as in Hartfield and Otto (2011). In Appendix G, we give the diffusion equation adjusted to our model and compare the result to Eq. (40).

A comparison between Panels A/B with Panels C/D of Fig. 4 shows that the introgression probability changes only slightly over the depicted range of recombination, while the hitchhiking probability significantly decreases with increasing recombination distance; the scale is strongly affected by the population size.
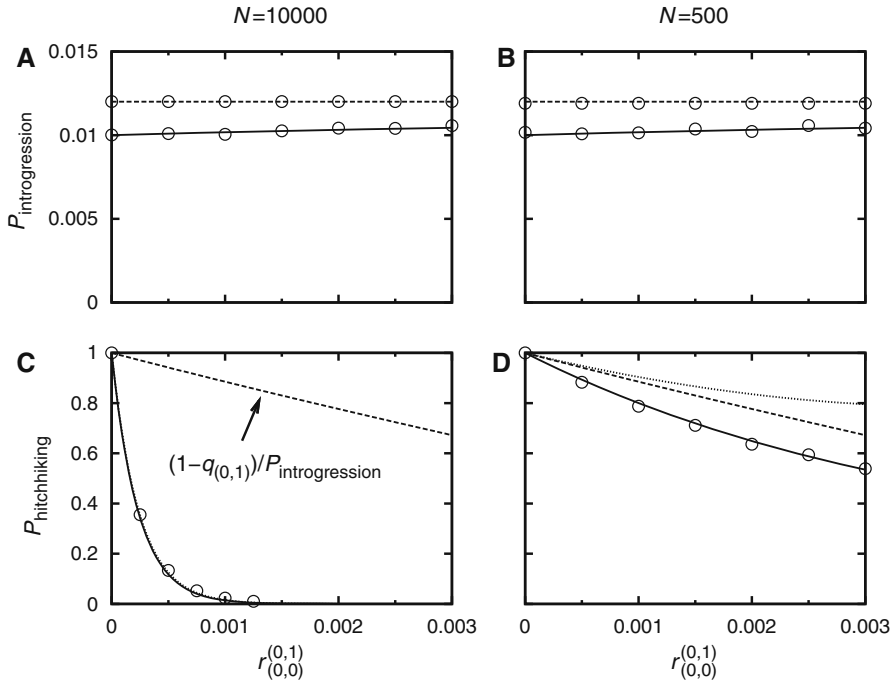
Fig. 4 Panels **a** and **b**: The introgression probability as a function of linkage. For the *solid line*: $I = 0$, $J = 1$. For the *dashed line*: $I = J = 0$. Panels **c** and **d**: The hitchhiking probability of a single deleterious allele ($I = 0$, $J = 1$). The *solid line* represents the hitchhiking probability as given by Eq. (53). The *dashed* and the *dotted lines* show the impact of the stochastic phase [Eq. (52)] and the deterministic phase [Eq. (40)], respectively. In Panel **c**, the *solid* and the *dotted line* are virtually indistinguishable; the deterministic phase dominates. In Panel **d**, the stochastic phase has a significant impact on the hitchhiking probability. Parameter values are: $N = 10,000$ (Panels **a** and **c**), $N = 500$ (Panels **b** and **d**), $\sigma_{(0,0)} = 0.012$, $\sigma_{(0,1)} = 0.01$. The *circles* denote simulation results. For Panels **a** and **b**, each simulation point is the average of 10^6 introgression attempts. For Panels **c** and **d**, each simulation point is the average of 10^4 successful introgression events

6.3 The impact of a second deleterious allele

In this section, we investigate how a second deleterious allele affects the introgression and the hitchhiking probability in dependence of the strength of selection, the genetic architecture, and linkage. The results are summarized in Figs. 5, 6, 7, 8.

Figures 5 and 6 consider the dependence on linkage. One sees that a strongly deleterious allele significantly affects the introgression probability even if it is only loosely linked. The impact on the hitchhiking probability is more subtle, and several cases must be distinguished. We first turn to Fig. 5, in which the second deleterious allele is on the same side of the beneficial allele as the first one. Panel C shows the behavior for $\sigma_{(0,2)} - r_{(0,2)} < 0$. In that case, applying Eq. (30) with Eq. (31), it approximately holds:

$$P_{(0,1)}^{(0,2)} \approx \frac{r_{(0,1)}^{(0,2)}(1 - Q_{(0,1)})}{r_{(0,0)}^{(0,2)}(1 - Q_{(0,0)}) + r_{(0,1)}^{(0,2)}(1 - Q_{(0,1)})} \approx \frac{r_{(0,1)}^{(0,2)}\sigma_{(0,1)}}{r_{(0,0)}^{(0,2)}\sigma_{(0,0)} + r_{(0,1)}^{(0,2)}\sigma_{(0,1)}}, \quad (55)$$
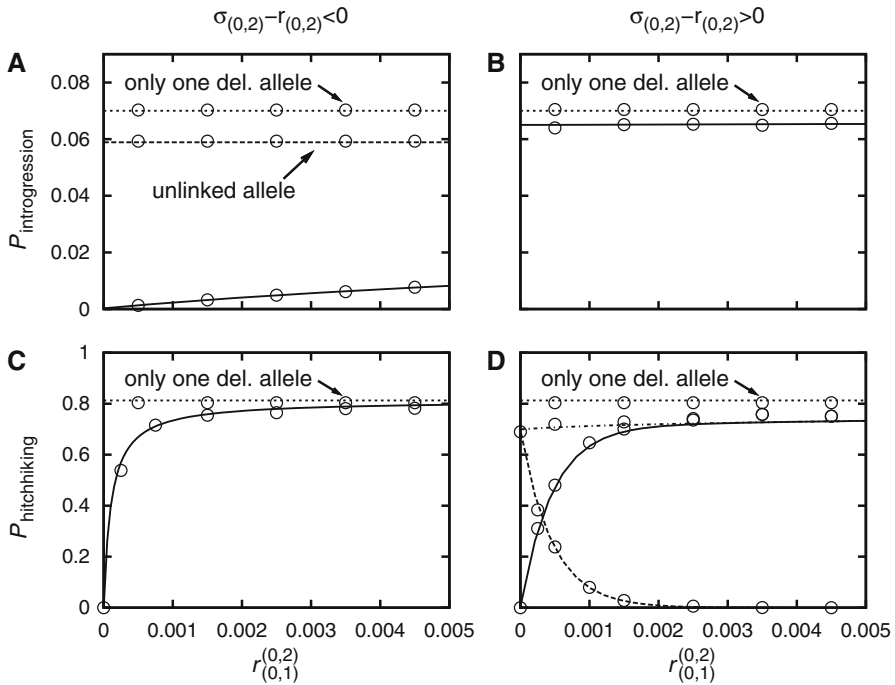
Fig. 5 Panels **a** and **b**: The introgression probability of an adaptive allele linked to two deleterious alleles. Panels **c** and **d**: The hitchhiking probability of a closely linked deleterious allele. In each Panel, the *short-dashed line* gives the respective probability in the absence of the second deleterious allele. In Panel **a**, the *long-dashed line* gives the introgression probability for one closely linked and one unlinked allele (in Panel **b**, this line would be indistinguishable from the *short-dashed line*). In Panel **d**, the *solid line* gives the probability that the haplotype $(0, 1)$ fixes. The *long-dashed line* gives the probability that the haplotype $(0, 2)$ fixes. The *dash-dotted line* gives the probability that either of the two fixes. For Panels **a** and **c**: $\sigma_{(0,2)} = -0.03$. For Panels **b** and **d**: $\sigma_{(0,2)} = 0.065$. The other parameter values are: $\sigma_{(0,0)} = 0.075$, $\sigma_{(0,1)} = 0.07$, $N = 10{,}000$, $r_{(0,0)}^{(0,1)} = 0.0001$. The *circles* denote simulation results. For Panels **a** and **b**, each simulation point it the average of 10^6 introgression attempts. For Panels, **c** and **d**, each simulation point is the average of 2,000 successful introgression events

where we have used $Q_{(0,1)} \approx 1 - \sigma_{(0,1)}$. Hence:

$$
\begin{aligned}
P_{\text{hitchhiking}}^{(0,2)} &\approx \frac{r_{(0,1)}^{(0,2)} \sigma_{(0,1)}}{r_{(0,0)}^{(0,2)} \sigma_{(0,0)} + r_{(0,1)}^{(0,2)} \sigma_{(0,1)}} P_{\text{det}}^{((0,1) \to (0,1))} \\
&\approx \frac{r_{(0,1)}^{(0,2)} \sigma_{(0,1)}}{r_{(0,0)}^{(0,2)} \sigma_{(0,0)} + r_{(0,1)}^{(0,2)} \sigma_{(0,1)}} P_{\text{hitchhiking}}^{(0,1)},
\end{aligned} \tag{56}
$$

where $P_{\text{hitchhiking}}^{(\cdot,\cdot)}$ denotes the probability that type $(0, 1)$ fixes in the population given that type $(\cdot, \cdot)$ got initially introduced. I.e., for close linkage of the second deleterious allele, the hitchhiking probability gets strongly reduced. However, as linkage gets looser, it converges quickly to its value in the absence of the second allele,
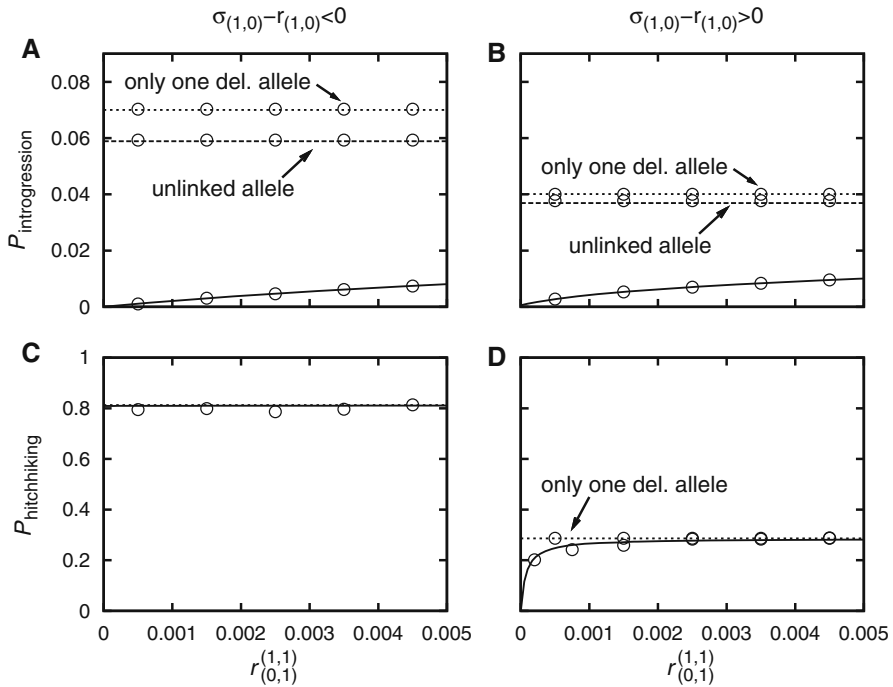
Fig. 6 Panels **a** and **b**: The introgression probability of an adaptive allele linked to two deleterious alleles. Panels **c** and **d**: The fixation probability of type $(0, 1)$. In each Panel, the *short-dashed line* gives the respective probability in the absence of the second deleterious allele ($I = 0$). In Panels **a** and **b**, the *long-dashed line* gives the introgression probability for one closely linked and one unlinked allele. For Panels **a** and **c**: $\sigma_{(0,1)} = 0.07$, $\sigma_{(1,0)} = -0.025$, $\sigma_{(1,1)} = -0.03$, i.e., type $(1, 0)$ cannot rise to fixation. For Panels **b** and **d**: $\sigma_{(0,1)} = 0.04$, $\sigma_{(1,0)} = 0.03$, $\sigma_{(1,1)} = -0.005$, i.e., either deleterious allele can fix but not both. The other parameter values are: $\sigma_{(0,0)} = 0.075$, $N = 10{,}000$, $r_{(1,0)}^{(1,1)} = 0.0001$. The *circles* denote simulation results. For Panels **a** and **b**, each simulation point it the average of 10^6 introgression attempts. For Panels **c** and **d**, each simulation point is the average of 2,000 successful introgression events

$P_{\text{det}}^{((0,1)\to(0,1))}$. Let c denote the factor by which the second deleterious allele reduces the hitchhiking probability of the first one:

$$c = \frac{P_{\text{hitchhiking}}^{(0,1)} - P_{\text{hitchhiking}}^{(0,2)}}{P_{\text{hitchhiking}}^{(0,1)}}. \tag{57}$$

By rearranging terms, we obtain

$$\frac{r_{(0,1)}^{(0,2)}}{r_{(0,0)}^{(0,2)}} = \frac{1-c}{c}\frac{\sigma_{(0,0)}}{\sigma_{(0,1)}} \quad \Leftrightarrow \quad c = \frac{1}{1 + \frac{r_{(0,1)}^{(0,2)}\sigma_{(0,1)}}{r_{(0,0)}^{(0,2)}\sigma_{(0,0)}}}. \tag{58}$$

Importantly, the strength of selection of the second deleterious allele has no effect (as long as it is strong enough for our approximation to apply), and the other selection
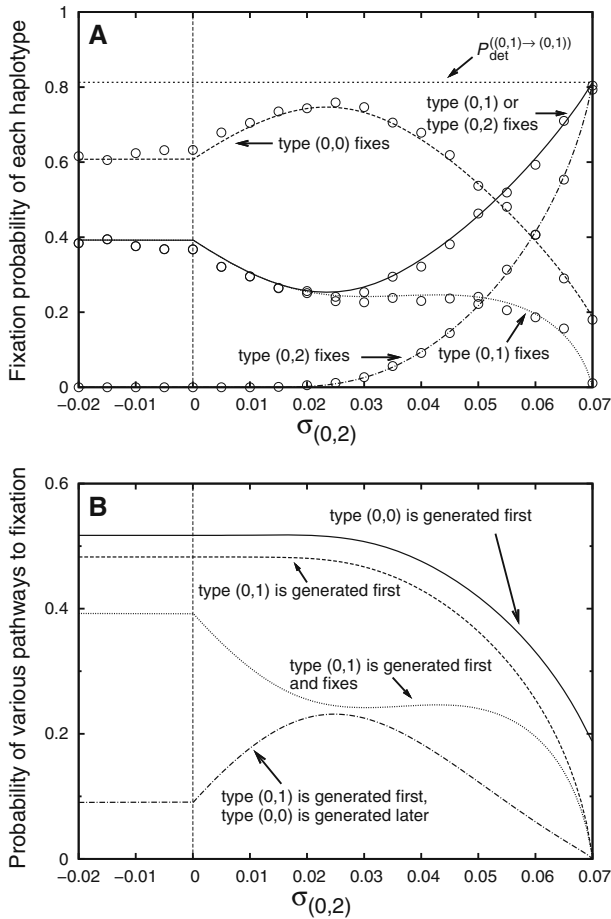
Fig. 7 Panel **a**: The hitchhiking probability of closely linked deleterious alleles as a function of the Malthusian fitness parameter $\sigma_{(0,2)}$. The *plot* shows the probabilities with which the various haplotypes fix in the population. Panel **b**: An illustration of the respective pathways of recombination and establishment that lead to fixation of types $(0, 0)$ and $(0, 1)$: by recombination, either a successful lineage of type $(0, 0)$ (*solid line*) or type $(0, 1)$ (*dashed line*) can be generated. In the first case, fixation of type $(0, 0)$ is certain. In the second case, a successful lineage of type $(0, 0)$ can still be generated later and outcompete type $(0, 1)$ (*dashed-dotted line*). This scenario is particularly likely if type $(0, 2)$ has an intermediate fitness (see text). The *dotted curves* in Panel **a** and **b** show the same probability. Parameter values are: $I = 0$, $J = 2$, $\sigma_{(0,0)} = 0.075$, $\sigma_{(0,1)} = 0.07$, $N = 10,000$, $r_{(0,0)}^{(0,1)} = r_{(0,1)}^{(0,2)} = 0.0001$. The *circles* denote simulation results. Each simulation point is the average of 2,000 successful introgression events

coefficients play only a minor role. The second deleterious allele is relevant if

$$r_{(0,1)}^{(0,2)} < r_{(0,0)}^{(0,1)} \frac{\sigma_{(0,0)}}{\sigma_{(0,1)}} \approx r_{(0,0)}^{(0,1)} \quad \text{for} \quad \sigma_{(0,0)} \gg s_{\text{del}}, \tag{59}$$

which is independent of the selection coefficients if $\sigma_{(0,0)} \gg s_{\text{del}}$. The hitchhiking probability is crucially determined by the ratio of recombination distances $r_{(0,1)}^{(0,2)}/r_{(0,0)}^{(0,1)}$.
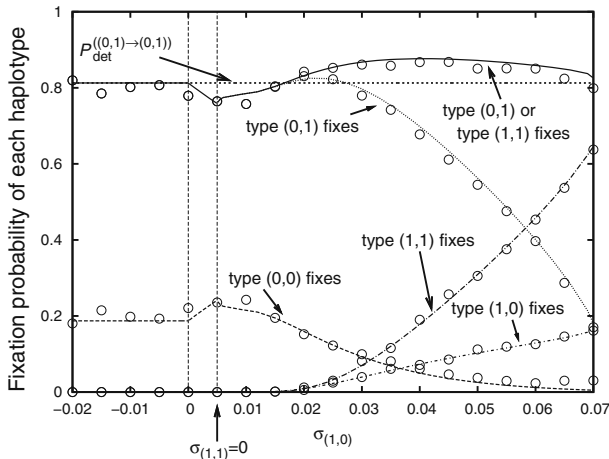
Fig. 8 The hitchhiking probability of closely linked deleterious alleles as a function of the Malthusian fitness parameter $\sigma_{(1,0)}$. The plot shows the probabilities with which the various haplotypes fix in the population. Parameter values are: $I = 1$, $J = 1$, $\sigma_{(0,0)} = 0.075$, $\sigma_{(0,1)} = 0.07$, $N = 10,000$, $r_{(0,0)}^{(0,1)} = r_{(0,0)}^{(1,0)} = 0.0001$. The *circles* denote simulation results. Each simulation point is the average of 2,000 successful introgression events

In Panel D, the second deleterious allele can hitchhike to fixation, too ($\sigma_{(0,2)} - r_{(0,2)} > 0$). The total hitchhiking probability of the closest deleterious alleles (i.e., the probability that either type $(0, 1)$ or type $(0, 2)$ fixes) is then only moderately reduced. For $r_{(0,1)}^{(0,2)}$ small, both alleles fix. With increasing recombination distance the analytical result underestimates the true hitchhiking probability (see Appendix H).

In Fig. 6, the beneficial allele is flanked by the two deleterious alleles. If one of the alleles is strongly deleterious, the hitchhiking probability of the other one is not visibly reduced (Panel C). This is because for successful introgression, the strongly deleterious allele has to recombine away very early. The situation in Panel D ($\sigma_{(0,1)} - r_{(0,1)} > 0$, $\sigma_{(1,0)} - r_{(1,0)} > 0$, but $\sigma_{(1,1)} - r_{(1,1)} < 0$) looks similar to Fig. 5C. We have

$$
P_{\text{hitchhiking}}^{(1,1)} \approx \frac{r_{(0,1)}^{(1,1)} \sigma_{(0,1)}}{r_{(1,0)}^{(1,1)} \sigma_{(1,0)} + r_{(0,1)}^{(1,1)} \sigma_{(0,1)}} P_{\text{hitchhiking}}^{(0,1)}.
\tag{60}
$$

This is formally similar to Eq. (56). There is, however, an important difference: Now, the selection coefficient of the second deleterious allele is crucial. Its influences ceases with increasing strength as $\sigma_{(1,0)}$ goes to zero. If both deleterious alleles have approximately the same effect ($\sigma_{(0,1)} \approx \sigma_{(1,0)}$), the behavior is again determined by the ratio of the recombination distances $r_{(0,1)}^{(1,1)}/r_{(1,0)}^{(1,1)}$.

Figures 7 and 8 show how the selective disadvantage of a second closely linked deleterious allele affects the hitchhiking probability. If it is on the same side of the beneficial allele as the first one (Fig. 7), the hitchhiking probability of the first one is greatly reduced unless the selective disadvantage is very slight ($\sigma_{(0,2)} \approx \sigma_{(0,1)}$). The reduction is greatest for intermediate values of the selection coefficient (see Fig. 7a).

In this parameter regime, type $(0, 2)$ significantly increases in frequency before a successful lineage of type $(0, 0)$ or $(0, 1)$ is generated. As a consequence, the time to fixation of the beneficial allele is relatively long. Even if a successful lineage of type $(0, 1)$ can establish, it is therefore likely that later, a successful lineage of type $(0, 0)$ is generated (see Fig. 7b for an illustration of this reasoning). If the beneficial allele is flanked at equal small recombination distances by two deleterious alleles as in Fig. 8, the total hitchhiking probability of the deleterious allele to the right (fixation of type $(0, 1)$ or type $(1, 1)$) is barely influenced by the presence of the second deleterious allele, irrespective of the selective disadvantage of the latter. Note, however, that recombination is weak in Fig. 8. For strong recombination and $\sigma_{(0,1)} \approx \sigma_{(1,0)}$, it is not unlikely that both a successful lineage of type $(0, 1)$ and of type $(1, 0)$ establish and coexist for a long time, making the production of a successful $(0, 0)$ recombinant very likely (see Fig. 13 in Appendix H).

6.4 The impact of several linked deleterious alleles

To start with, assume $\sigma_{(I,J)} - r_{(I,J)} > 0$. If all recombination distances are small, we can generalize the result Eq. (41) and calculate the probability that all deleterious alleles hitchhike to fixation. Analogous to the derivation of Eq. (41), we obtain for the probability that all deleterious alleles hitchhike to fixation

$$
Q_{\det}^{(I,J)} = \prod_{i=0}^{I-1} \left(\frac{\sigma_{(i,J)}}{\sigma_{(i,J)} - \sigma_{(I,J)}} \right)^{-\frac{N r_{(i,J)}^{(I,J)} \sigma_{(i,J)} (\sigma_{(i,J)} - \sigma_{(I,J)})}{\sigma_{(I,J)}^2}}
$$
$$
\times \prod_{j=0}^{J-1} \left(\frac{\sigma_{(I,j)}}{\sigma_{(I,j)} - \sigma_{(I,J)}} \right)^{-\frac{N r_{(I,j)}^{(I,J)} \sigma_{(I,j)} (\sigma_{(I,j)} - \sigma_{(I,J)})}{\sigma_{(I,J)}^2}}. \tag{61}
$$

We now turn to $\sigma_{(I,J)} - r_{(I,J)} < 0$ and study several selected scenarios which impart a general intuition. Consider first the special case $I = 0$, $\sigma_{(0,j)} - r_{(0,j)} > 0$, and $\sigma_{(0,l)} - r_{(0,l)} < 0$, $l > j$, i.e., all deleterious alleles are located at one side of the adaptive allele and at most j of them can potentially hitchhike to fixation. For tight linkage, we can again approximate $1 - Q_{(0,k)} \approx \sigma_{(0,k)}$ for $k \leq j$ and obtain (proof by induction):

$$
P_{(0,k)}^{(0,J)} \approx \frac{r_{(0,k)}^{(0,J)} \sigma_{(0,k)}}{\sum_{l=0}^{j} r_{(0,l)}^{(0,J)} \sigma_{(0,l)}}. \tag{62}
$$

I.e., the selection coefficients of the $(j + 1)$th, $(j + 2)$th, …, Jth alleles do not enter the result. Furthermore, it does not matter where the deleterious alleles beyond the $(j + 1)$th are located.

As another example, consider the special case $J = 1$, $\sigma_{(0,1)} - r_{(0,1)} > 0$, and $\sigma_{(i,j)} - r_{(i,j)} < 0$, $i > 0$. If linkage is tight, the hitchhiking probability is barely reduced by additional deleterious mutations.
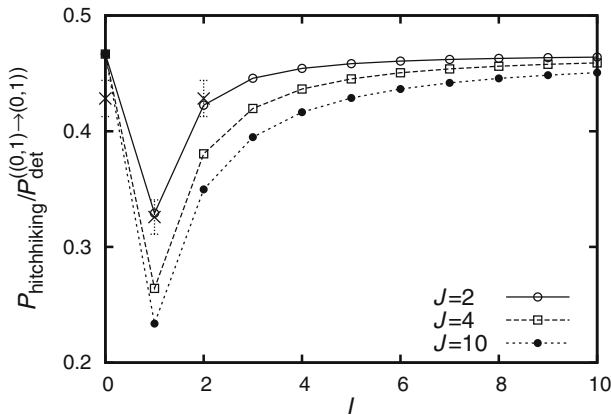
Fig. 9 The reduction of the hitchhiking probability of a closely linked deleterious allele due to the impact of other deleterious alleles. On the x axis, the number I of deleterious alleles to the left of the beneficial allele is varied. Only types $(0, 0)$ and $(0, 1)$ have positive fitness. The fixation probability of type $(0, 1)$ conditioned on fixation of the adaptative allele ($P_{\text{hitchhiking}}$) displays a pronounced dip for $I = 1$. Parameter values are: $\sigma_{(0,0)} = 0.08$, $\sigma_{(0,1)} = 0.07$, $\sigma_{(0,2)} = -0.005$, $\sigma_{(1,0)} = -0.002$, all other deleterious alleles have an effect of -0.01, $r_{(i,j)}^{(k,l)} = 0.0001$ for all k, l, i, j, $N = 10,000$. The *crosses* denote simulation results. Each simulation point is the average of 1,000 successful introgression events. The introgression probability for the scenarios shown in the plot ranges from $\approx 1.3 \cdot 10^{-7}$ ($I = J = 10$) to ≈ 0.002 ($I = 0, J = 2$), i.e., 1,000 successful introgression events correspond to $\approx 500,000$ to $6 \cdot 10^9$ introgression attempts; simulations for complex scenarios are hence impractical

Finally, let $\sigma_{(0,1)} - r_{(0,1)} > 0$, I and $J \geq 1$ arbitrary, and $\sigma_{(i,j)} - r_{(i,j)} < 0$ for $(i, j) \notin \{(0, 0), (0, 1)\}$. Figure 9 shows how additional deleterious alleles that can themselves not hitchhike to fixation can influence the hitchhiking probability of a slightly deleterious allele. The pattern can be understood by consideration of the various paths which lead to establishment of the beneficial allele: unless $I = 0$ (or $J = 1$), at least two recombination events are necessary to generate a type with positive Malthusian fitness. The position of the first successful recombination event depends on the fitness of the types that are generated by recombination. Since $\sigma_{(1,0)}$ is only slightly deleterious, the first recombination event is likely to generate this type if $I = 1$. In this case, the hitchhiking probability is strongly reduced (cf. the dip in Fig. 9). For $I > 1$, however, type $(1, 0)$ cannot be generated via a single recombination event, and the reduction is less pronounced, getting smaller with increasing I. For large I, adding more deleterious alleles to the left or the right has only a weak effect. Generally, deleterious alleles that render haplotypes strongly disfavored have to be lost as quickly as possible, and the pathway to establishment of the beneficial allele does usually not involve tunneling via more strongly deleterious haplotypes than necessary. At either side of the adaptive allele, consider the set of alleles which must be lost for establishment. As a rule of thumb, each of the two sets can be replaced by a "virtual" allele of the respective compound effect. In order to have the same effect as the set of actual alleles, this virtual allele has to be located at the same position as the deleterious allele (out of the set) which is closest to the adaptation.

7 Discussion

Gene flow between related species is frequent in nature. Although many foreign alleles burden their carrier with a selective disadvantage, exchange of genetic material between populations is often still possible. If neutral or advantageous alleles survive the fitness bottleneck caused by linked or unlinked deleterious alleles they can become permanently incorporated into the genome of the sister species. Picking up locally adaptive alleles from an indigenous species can help species to expand their range to previously uninhabitable regions (e.g., Heiser 1951; Whitney et al. 2006). Adaptive gene introgression is hence a clever evolutionary mechanism that can speed up adaptation to novel environments. Human activities create ample opportunity for hybridization between domestic animals or crop plants with their wild relatives (e.g. Fitzpatrick et al. 2010; Ellstrand et al. 1999, 2013). In this context, the introgression of alleles from genetically modified organisms (e.g., insecticide resistance genes) into weedy species is recognized as a risk that can cause permanent ecological damage. A quantitative analysis of the introgression process is essential both to assess the importance of adaptive gene introgression as an evolutionary pathway to adaptation and to estimate the ecological risks associated with unwanted hybridization.

The flow of adaptive alleles among species is hampered by the reduced fitness of inter-species crosses. This reduction in fitness is due to alleles from the donor species that are deleterious in the new environment (which we consider in this paper) and/or the new genomic background. If their compound effect outweighs the benefits of the adaptation, some deleterious alleles must be eliminated by recombination in hybrid back-crosses before the favorable allele can establish. However, closely linked slightly deleterious alleles might be dragged along to fixation. In this paper, we developed a framework to investigate the role of linked and unlinked deleterious alleles in adaptive gene introgression. The model accounts for an explicit genetic structure and describes the genetic evolution of a haploid population under the influence of selection, recombination, and drift after a single hybridization event. The analysis is based on the theory of branching processes. The early phase of spread of the advantageous allele is approximated by a reducible multitype branching process with a special structure: within the branching process approximation, offspring are either of the same type as their parent or carry fewer deleterious alleles (for similar setups see Barton and Bengtsson 1986; Demon et al. 2007; Gosh and Haccou 2010; Gosh et al. 2012a,b; Yanchukov and Proulx 2014). The fate of the recombinants that are generated after the initial establishment and until fixation of the beneficial allele is modeled by a time-inhomogeneous single-type branching process. For the analysis of the first phase, we make use of methods developed by Serra (2006) and Serra and Haccou (2007). The analysis of the second phase builds on work by Hartfield and Otto (2011). The combination of the results from both phases allows for the analytical treatment of the entire process.

The introgression probability How likely is it that the advantageous allele can establish itself in the population? For large populations, the probability of adaptive gene introgression only depends on the early phase of the spread, where the dynamics are well described by a branching process. Technically, this is similar to studies on scenarios where adaptation relies on the accumulation of new mutations via stochastic tunneling.

The gain of adaptive alleles by mutation in these scenarios, which include the crossing of fitness valleys (e.g., Weissman et al. 2009; Proulx 2011), tumor initiation (e.g., Iwasa et al. 2003, 2004a,b), or adaptation of a pathogen to a new host (e.g., Antia et al. 2003), corresponds to the removal of deleterious alleles by recombination in models of gene introgression. The survival probability of a multitype branching process is in general difficult to determine, and one has to resort to approximate formulas and numerical methods (e.g., Barton 1995; Iwasa et al. 2003, 2004b; Serra and Haccou 2007). However, in our special case, a recursive solution can be derived and readily permits the calculation of the introgression probability for any given allele configuration.

We find that both linked and unlinked deleterious alleles can significantly hamper the introgression of an adaptive allele. However, the characteristic of this barrier depends on whether linkage between the beneficial and the deleterious alleles is loose or tight. Loosely linked and unlinked deleterious alleles reduce the introgression probability by a factor that is roughly independent of the strength of the beneficial allele. In this parameter range, our results are analogous to those of Bengtsson (1985) and Barton and Bengtsson (1986), who derive a so-called gene-flow factor or barrier strength to describe the effect of a genetic barrier on the flow of a neutral marker allele (compare Appendix B). The relative reduction due to a single loosely linked allele is approximately s_{del}/r, where s_{del} is the deleterious effect and r the recombination probability (see Eq. (50), Sect. 6.2). Strongly deleterious alleles ($s_{\text{del}} > 0.05$) can have a substantial effect even if unlinked ($r = 0.5$). In agreement with Bengtsson (1985), we find that the influence of several unlinked deleterious alleles is well approximated by a single deleterious allele of the compound effect [Eq. (45)]. The simple picture of a gene-flow factor holds for $r \gg s_{\text{ben}}$ (where s_{ben} is the effect of the beneficial allele), but breaks down for tighter linkage. For small r, the relative reduction scales as $1 - r/(|s_{\text{ben}} - s_{\text{del}}|)$ and thus depends explicitly on the strength of the beneficial allele [cf. Eq. (49)].

Our results on adaptive gene introgression can also be compared with results by Barton (1995) on the reduction of the fixation probability of a new beneficial mutation due to interference with standing deleterious variation. Barton (1995) assumes that the deleterious alleles segregate under mutation-selection balance in the population when the advantageous allele appears. Consider first the limiting case where a deleterious allele segregates at a single locus with frequency u. The beneficial mutation can arise on a genome that does or does not carry the deleterious allele. Technically, the introgression probability in our model corresponds most closely to the fixation probability of the beneficial mutation given that it arises on a genome with the deleterious allele [denoted by P_u in Barton (1995)]. In that case, its fixation probability can be significantly reduced. However, the reduction due to segregating deleterious variation is generally much weaker than in our case, where both alleles enter the population via a single introgression event. This is because relative fitness of the double mutant is higher if the deleterious allele segregates in the population. A numerical comparison confirms that the result $P_u/(2\sigma_{(0,0)})$ as given by Eq. (16) and (17a) in Barton (1995) converges to $(1 - Q_{(0,1)})/\sigma_{(0,0)}$ [cf. Eq. (47c)] as the mutation rate and hence the frequency of the deleterious allele tend to zero. (Note that the results in Barton (1995) are based on a Poisson distribution of the offspring number such that the establishment probability of an isolated beneficial allele is $2\sigma_{(0,0)}$ while it is $\sigma_{(0,0)}$ in our model). Our results for the introgression probability hence represent limiting cases of the results in Barton (1995)

(P_u in the limit $u \to 0$). However, for more than one deleterious allele, the equations in Barton (1995) do not allow for an analytical solution, while this is possible for the case of introgression, as shown by our results. The total fixation probability in the segregating-alleles case is a weighted average of the cases that the beneficial allele appears on a genome with and without the deleterious allele; the weighting factor depends on the mutation rate. This can lead to widely diverging conclusions as compared to the introgression scenario. For example, for $\sigma_{(0,1)} < 0$ and complete linkage, the weighted fixation probability of the beneficial mutation is reduced by u relative to its value in absence of the deleterious allele [Barton 1995, Eq. (17b)]. In contrast, the introgression probability (as well as P_u) are zero in that case. For $\sigma_{(0,1)} > 0$, the relative reduction is at most $u(s_{\text{del}}/\sigma_{(0,0)})^2$, i.e., much smaller than for introgression ($s_{\text{del}}/\sigma_{(0,0)}$). Note that the reduction in the weighted fixation probability is caused by the recurrent generation of deleterious alleles that appear on genomes carrying the adaptation. The presence of deleterious alleles itself even slightly increases the weighted fixation probability [this term is very small and neglected in Barton (1995)]. Finally, Barton (1995) finds that for two loci flanking the beneficial mutation, the effects of the two deleterious alleles approximately multiply. This does not hold true for the different biological scenario of adaptive gene introgression.

Linked deleterious alleles can render successful introgression after a single hybridization event extremely unlikely. In order to assess whether even introgression probabilities of the order of $10^{-8} - 10^{-6}$ are still evolutionary relevant, it is helpful to compare these values to the probability of adaptation by de-novo mutations. With a point mutation probability of $\sim 10^{-8}$ and a selective advantage of 1 %, the probability that a specific mutation occurs in a specific individual and thereafter rises to fixation, is $\sim 10^{-10}$. For complex adaptations, the probability is even lower. Depending on the probability of hybridization, adaptive gene introgression can hence be a relevant evolutionary process. Hybridization rates are potentially high, and even if the success probability of each single hybridization event is low, the probability that any hybridization event is followed by adaptive gene introgression is appreciable. This consideration is particularly important in an agricultural context where (genetically modified) crops grow next to wild plants in large areas all over the world for many years. Gosh and Haccou (2010) and Gosh et al. (2012a,b) therefore suggest the so-called hazard rate as a measure for risk assessment, as the hazard rate takes both the hybridization rate and the introgression probability into account.

The hitchhiking probability Weakly deleterious alleles that are closely linked to the adaptive allele can hitchhike to fixation. We developed a framework to estimate which haplotype finally fixes in the population, depending on the alien haplotype that was originally introduced. The approach is based on a split of the process into two phases: the establishment phase of the adaptive allele and the sweep during which further deleterious alleles can be lost. What is the respective relevance of the stochastic and the deterministic phase in this scenario? In the simplest case, there are only two loci under selection: one locus with the advantageous and one locus with a deleterious allele that can hitchhike to fixation. We can then distinguish two parameter regimes: if the product of the selection coefficient and the population size, $N s_{\text{del}}$, is of order 1 or smaller, the impact of the stochastic phase is significant. The probability for the hitchhiker to survive this phase depends strongly on the selection coefficients of both the beneficial

and the deleterious allele, and on the recombination rate, but is independent of the population size $(1 - rs_{\text{ben}}/(s_{\text{ben}} - s_{\text{del}})^2$, cf. Eq. (54)). However, if the product of selection and population size is large, the stochastic phase can be ignored and the behavior is dominated by the deterministic phase. Hitchhikers will survive this phase if no haplotype without the deleterious allele can establish. Since the duration of the deterministic phase is $\sim 1/(s_{\text{ben}} - s_{\text{del}})$ and the number of successful new recombinants per generation roughly $\sim s_{\text{ben}} N r$, we see that the hitchhiking probability will strongly depend on Nr, while the effect of selection partly cancels. This is confirmed by our more precise calculations (Eq. (40)). The situation is different if additional deleterious alleles render the initial haplotype itself deleterious. In that case, establishment of the adaptation is contingent on the early loss of deleterious alleles, and depending on the allelic configuration, the stochastic establishment phase will have a strong impact on hitchhiking irrespective of the population size. To good approximation, all alleles that cause serious maladaptation and are located on the same side of the beneficial allele can be summarized to a single allele of the compound effect, reducing the dimensionality of the problem. The impact of these additional alleles fades quickly with increasing recombination distance. In contrast to introgression, unlinked alleles have no visible effect on the hitchhiking probability conditioned on successful introgression. These insights essentially generalize to more than one possible deleterious hitchhiker.

Hartfield and Otto (2011) analyze the hitchhiking probability of a single deleterious allele in the absence of other deleterious alleles. They present two approaches to the problem: a semi-deterministic approach based on branching process theory (which also serves as the basis for our analysis of the deterministic phase), and a diffusion approach. In both cases, however, they condition on establishment of type $(0, 1)$. Their analysis thus ignores aspects of the stochastic establishment phase and consequently applies only to the regime of large Ns_{del}.

Selection and recombination in the introgression process Summarizing, we can identify three fundamentally different genomic scales in units of the recombination rate that matter for adaptive gene introgression. First, deleterious alleles on the introgression haplotype affect the introgression probability of the beneficial allele across distances on the order of the deleterious selection coefficient ($r \sim s_{\text{del}}$; cf. Eq. (50)). Strongly deleterious alleles thus still matter even if they are unlinked. Importantly, the absolute strength of selection is crucial for the failure or success of introgression. For the hitchhiking probability of a single deleterious allele, we find two relevant scales that stem from the stochastic and the deterministic phases of the hitchhiking process, respectively. In the stochastic phase, this scale is set by the selection coefficient of the haplotype containing both the beneficial and the deleterious allele ($r \sim s_{\text{ben}} - s_{\text{del}}$; cf. Eq. (54)). Similar to the introgression probability it is the strength of selection that matters. In contrast, for the deterministic phase, the scale is primarily set by the inverse population size ($r \sim 1/N$). The rate of loss of the deleterious allele during the deterministic phase is also affected by selection. However, in contrast to the stochastic phase, only the ratio of selection coefficients ($s_{\text{del}}/s_{\text{ben}}$) enters and not the absolute strength [Eq. (40)]. Usually (but not always), effects from the deterministic phase are relevant already over much shorter distances than effects from the stochastic phase and will therefore dominate. Finally, a third scale for the distance r between the beneficial

allele and a deleterious hitchhiker becomes relevant if hitchhiking is compromized by further deleterious alleles in the genomic background. For cases where an additional deleterious allele at distance R to the hitchhiking allele suppresses the Malthusian fitness of that original introgression haplotype below zero, we find that the relevant scale for hitchhiking success is primarily set by the relative size of the recombination rates, i.e., $r \sim R$ [cf. Eqs. (56), (60)].

Limitations and extensions The mathematical analysis of the model has two major restrictions. The first one is the common constraint of branching process approximations for establishment probabilities of beneficial alleles: for small populations, the branching process approach underestimates the true probability. In particular, our multitype branching process can contain supercritical, critical, and subcritical types. If the population size is small, fixation of deleterious haplotypes by genetic drift can be more likely than survival of the branching process. The second restriction concerns the analysis of the hitchhiking probability. The derived approximations rely on the assumption that initially, a single haplotype establishes and starts sweeping before haplotypes with fewer deleterious alleles possibly establish. In particular when linkage is not tight, this assumption is not necessarily satisfied. However, this introduces only a small error, and the approximation usually still yields good results (cf. Fig. 12). Relaxing the assumption would require the assessment of the stochastic dynamics of two or more types that simultaneously establish in the population at different random speeds, which would severely complicate the calculations. The analysis of the deterministic phase increases in complexity with the number of positively selected haplotypes. In order to keep it tractable, we assume that at most two successful recombinant haplotypes are generated during the fixation process of the adaptive allele. While this assumption is again well justified for tight linkage, it leads to strong deviations as recombination gets stronger (cf. Fig. 13).

The model assumes a population of haploid individuals or diploids without dominance. The results for the introgression probability can be directly generalized to apply to diploids with dominance unless the beneficial allele is completely recessive: as long as introgressed alleles are rare in the population, they only appear in heterozygotes. Equation (2) therefore still applies if fitness refers to heterozygote fitness. As soon as the introgressed alleles become more frequent, both copies of an individual's chromosome might carry introgressed material. The sweep of a selectively favored haplotype is therefore strongly altered if the alleles are not co-dominant. The length and shape of this frequency path is, however, crucial for the hitchhiking probability. Our approach can be extended by this element (see Hartfield and Glémin (2014) for a single deleterious hitchhiker if there is dominance). The model furthermore assumes that selection against the deleterious alleles is independent of the genetic background. Recombination therefore increases the fitness of later generations of hybrids. However, by recombination, incompatibilities can be generated. In that case, later hybrids display a lower fitness than early generation hybrids until further recombination removes the incompatible alleles.

Beyond these genetic confinements, the second class of model extensions concerns the ecological setting. The model assumes that the alien genome arrives by long range dispersal in a panmictic population. An important extension, which would consider-

ably affect the results, is the incorporation of spatial structure. A very different situation from the one analyzed in this paper arises if dispersal is local and strong and recurrent gene flow builds up a hybrid zone. Under these circumstances, as discussed in Barton (1979), a single-locus cline does not significantly hamper the spread of a beneficial allele from one population to the other.

To summarize, we set up a minimal model in order to reveal fundamental principles that are effective in the introgression process of a favorable allele. In particular, the analysis helps to build an intuitive understanding of how deleterious alleles impact adaptive gene introgression.

Appendix A: Proof of Theorem 1

We can reinterprete the branching process as follows: An individual of type $(i, j; f)$ "dies" at rate $2 - \sigma_{(i,j;f)}$ and at death, it produces either zero or two offspring, one of which is of its own type. We now consider the embedded discrete-time process:

- With probability $\frac{1 - \sigma_{(i,j;f)}}{2 - \sigma_{(i,j;f)}}$, it has 0 offspring.
- With probability $\binom{f}{g}\left(\frac{1}{2}\right)^f \frac{1 - r_{(i,j)}}{2 - \sigma_{(i,j;f)}}$, it produces an offspring of its own type and an offspring of type $(i, j; g)$ with $g \leq f$.
- With probability $\binom{f}{g}\left(\frac{1}{2}\right)^f \frac{r_{(i,k)}^{(i,j)}}{2 - \sigma_{(i,j;f)}}$, it produces an offspring of its own type and an offspring of type $(i, k; g)$ with $g \leq f$ and $k < j$.
- With probability $\binom{f}{g}\left(\frac{1}{2}\right)^f \frac{r_{(k,j)}^{(i,j)}}{2 - \sigma_{(i,j;f)}}$, it produces an offspring of its own type and an offspring of type $(k, j; g)$ with $g \leq f$ and $k < i$.

Within this scheme, the offspring generating function of an individual of type $(i, j; f)$ is given by

$$
\begin{aligned}
G^{(i,j;f)}(\mathbf{s}) &= \left(\frac{1}{2}\right)^f \frac{1 - r_{(i,j)}}{2 - \sigma_{(i,j;f)}} s_{(i,j;f)}^2 \\
&+ \sum_{g=0}^{f} \binom{f}{g}\left(\frac{1}{2}\right)^f \left\{ \sum_{k=0}^{i-1} \frac{r_{(k,j)}^{(i,j)}}{2 - \sigma_{(i,j;f)}} s_{(k,j;g)} + \sum_{k=0}^{j-1} \frac{r_{(i,k)}^{(i,j)}}{2 - \sigma_{(i,j;f)}} s_{(i,k;g)} \right\} s_{(i,j;f)} \\
&+ \sum_{g=0}^{f-1} \binom{f}{g}\left(\frac{1}{2}\right)^f \frac{1 - r_{(i,j)}}{2 - \sigma_{(i,j;f)}} s_{(i,j;g)} s_{(i,j;f)} + \frac{1 - \sigma_{(i,j;f)}}{2 - \sigma_{(i,j;f)}},
\end{aligned}
\tag{63}
$$

where $\mathbf{s}$ is a vector with elements $\{s_{(k_1,k_2;g)}; 0 \le k_1 \le I, 0 \le k_2 \le J, 0 \le g \le F\}$. Let $\mathbf{G}(\mathbf{s})$ be the vector whose components are the offspring generating functions of all possible types. According to the general theory of multitype branching processes, the extinction probability is given by the root of

$$\mathbf{G}(\mathbf{s}) = \mathbf{s} \tag{64}$$

in the unit cube, which is closest to the origin (Sewastjanow 1974 p. 115).

First note that Eq. (64) is equivalent to Eq. (3) (identifying the solution of Eq. (64) with the vector with elements $\{Q_{(k_1,k_2;g)}; 0 \le k_1 \le I, 0 \le k_2 \le J, 0 \le g \le F\}$).

Equation (64) can be solved recursively starting with type $(0, 0; 0)$, for which we obtain

$$Q_{(0,0;0)} = 1 - \sigma_{(0,0;0)} < 1. \tag{65}$$

Following Sewastjanow (1974), in each subsequent step of the recursion, there exists a unique solution < 1. It remains to show that we obtain this solution if we solve the quadratic equation for the smaller root. We use the following abbreviations:

$$\Sigma_g := \sum_{k=0}^{j-1} r_{(i,k)}^{(i,j)} Q_{(i,k;g)} + \sum_{k=0}^{i-1} r_{(k,j)}^{(i,j)} Q_{(k,j;g)}, \tag{66}$$

$$\Sigma := \left(\frac{1}{2}\right)^f \left[\sum_{g=0}^{f-1} \binom{f}{g} (1 - r_{(i,j)}) Q_{(i,j;g)} + \sum_{g=0}^{f} \binom{f}{g} \Sigma_g \right], \tag{67}$$

$$a := \left(\frac{1}{2}\right)^f (1 - r_{(i,j)}) > 0, \tag{68}$$

$$b := -(1 - \sigma_{(i,j;f)} + 1 - \Sigma), \tag{69}$$

$$c := 1 - \sigma_{(i,j;f)} \ge 0. \tag{70}$$

Note that $r \ge \Sigma_g, \forall g$.

With these abbreviations, the two roots are given by

$$q_1 = \frac{-b + \sqrt{(b^2 - 4ac)}}{2a}, \tag{71}$$

$$q_2 = \frac{-b - \sqrt{(b^2 - 4ac)}}{2a}. \tag{72}$$

We first prove that both roots are real and ≥ 0. As a preliminary step, we show that $(1 - \Sigma) \ge \left(\frac{1}{2}\right)^f (1 - r_{(i,j)})$:

$$1 - \Sigma = 1 - \left(\frac{1}{2}\right)^f \left[\sum_{g=0}^{f-1} \binom{f}{g} (1 - r_{(i,j)}) Q_{(i,j;g)} + \sum_{g=0}^{f} \binom{f}{g} \Sigma_g \right]$$

$$\ge 1 - \left(\frac{1}{2}\right)^f \left[\sum_{g=0}^{f-1} \binom{f}{g} (1 - r_{(i,j)}) + \sum_{g=0}^{f} \binom{f}{g} \Sigma_g \right]$$

$$= 1 - \left(\frac{1}{2}\right)^f \left\{ \sum_{g=0}^{f} \binom{f}{g} \left[(1 - r_{(i,j)}) + \Sigma_g\right] - (1 - r_{(i,j)}) \right\}$$

$$\geq 1 - \left\{ \sum_{g=0}^{f} \binom{f}{g} \left(\frac{1}{2}\right)^f - (1 - r_{(i,j)}) \left(\frac{1}{2}\right)^f \right\}$$

$$= (1 - r_{(i,j)}) \left(\frac{1}{2}\right)^f. \tag{73}$$

Notice: $b < 0$ since $\sigma_{(i,j;f)} \leq 1$ and $1 - \Sigma \geq (1 - r_{(i,j)}) \left(\frac{1}{2}\right)^f > 0$.

We now show that $b^2 - 4ac \geq 0$.

$$b^2 - 4ac = (1 - \sigma_{(i,j)} + 1 - \Sigma)^2 - 4(1 - \sigma_{(i,j)})(1 - r_{(i,j)}) \left(\frac{1}{2}\right)^f$$

$$\geq (1 - \sigma_{(i,j)} + 1 - \Sigma)^2 - 4(1 - \sigma_{(i,j)})(1 - \Sigma)^2$$

$$= (1 - \sigma_{(i,j)} - (1 - \Sigma))^2 \geq 0, \tag{74}$$

where we have used that $(1 - r_{(i,j)}) \left(\frac{1}{2}\right)^f \leq 1 - \Sigma$. Since $a > 0$, we conclude that both roots are positive.

It remains to prove that $q_2 \leq 1$. Note that $\mathbf{1}$ is a root of the equation since $\mathbf{G(1)} = \mathbf{1}$. We furthermore see that q_2 is a decreasing function of Σ and thus of all $Q_{(k,j)}$ and $Q_{(i,k)}$, if $c > 0$. For $c = 0$, it holds that $q_2 = 0$. It is hence clear that $q_2 \leq 1$. From Sewastjanow (1974), we even know that $q_2 < 1$ (because $Q_{(0,0;0)} < 1$). $\qquad\square$

Note that instead of considering the embedded discrete time process we could have directly resorted to the corresponding result for the extinction probability of the continuous time processes (Sewastjanow 1974, p. 116), but it seemed more illustrative in this way.

Appendix B: Approximation of the establishment probability via an effective migration rate

Bengtsson (1985) and Barton and Bengtsson (1986) determine by how much deleterious alleles hamper the flow of a neutral marker locus from one population into another, where Bengtsson (1985) mainly focuses on unlinked and Barton and Bengtsson (1986) on linked deleterious alleles. The impact is expressed via a so-called gene flow factor, which is given by the ratio of an effective migration rate m_e and the actual migration rate m (Eqs. (1), (3), and (5) in Bengtsson (1985); Eqs. (A3) and (A7) in Barton and Bengtsson (1986); the inverse of the gene flow factor gives the barrier strength to gene flow). Building on these results for the introgression of neutral marker loci, the introgression probability of a single adaptive allele with selective advantage s can be approximated by
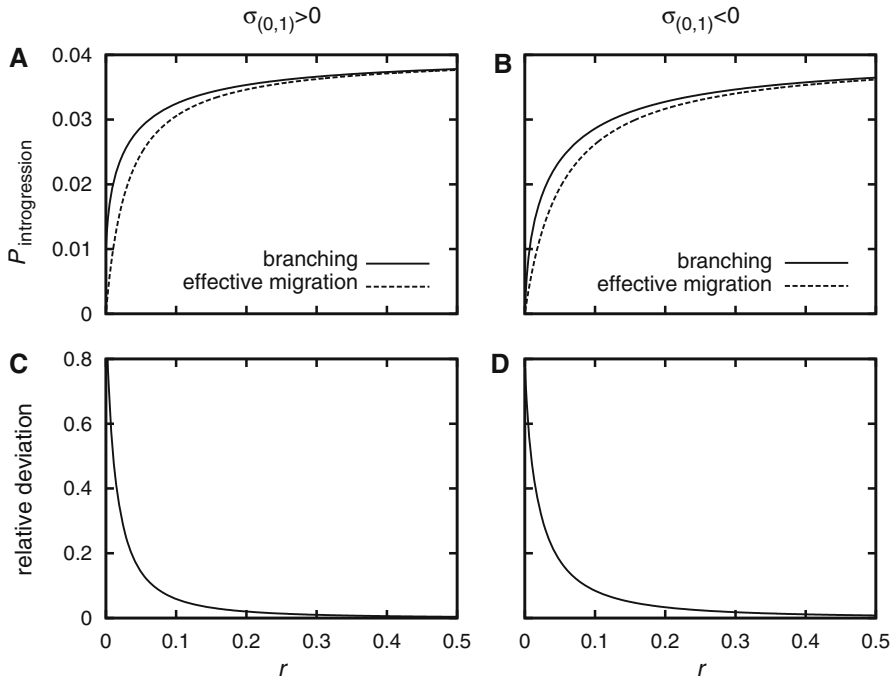
Fig. 10 Panels **a** and **b**: The introgression probability as a function of recombination. The *solid line* is based on the branching process approach, Eq. (2). The *dashed line* is obtained by Eq. (75). Panels **c** and **d**: The relative error between the approximations $(1 - Q_{(0,1)} - P^{m_{eff}}_{introgression})/(1 - Q_{(0,1)})$. The parameter values are: $I = 0$, $J = 1$, $\sigma_{(0,0)} = 0.04$, $\sigma_{(0,1)} = 0.01$ (Panels **a** and **c**) or $\sigma_{(0,1)} = -0.01$ (Panels **b** and **d**)

$$P^{m_{eff}}_{introgression} = \frac{m_e}{m} s. \tag{75}$$

Figure 10 shows how this approximation compares to the establishment probability as obtained by the branching process approach if there is just one deleterious allele. The deviation is large for tight linkage and gets small as the recombination distance between the loci increases; for lose linkage, it is negligible. This observation generalizes to scenarios with more deleterious alleles. Deviations for tight linkage are to be expected: As an example, consider the basic scenario $I = 0$, $J = 1$, and $s_{del} < \sigma_{(0,0)}$. For $r = 0$, the introgression probability of the adaptive allele equals $\sigma_{(0,0)} - s_{del}$. A neutral marker allele that is equally tightly linked, however, establishes with probability ≈ 0 (within the approximations by Bengtsson (1985) and Barton and Bengtsson (1986), its establishment probability equals zero; in a finite population, it might fix by drift). We now turn to the other extreme, i.e., when the deleterious alleles are all unlinked. We can perform a first-order Taylor expansion in s_{del} of the result for m_e/m with $I = 0$, $F = 1$ as found by Bengtsson (1985)

$$\frac{m_e}{m} = 1 - 2s_{del} + \mathcal{O}(s^2_{del}). \tag{76}$$

This is in agreement with Eq. (46). Bengtsson (1985) furthermore finds that to good approximation, the compound effect of all deleterious alleles, not their individual effects, matters. Again, this accords with our findings [cf. Sect. 6.1].

Appendix C: Proof of Eq. (32)

We prove that within approximation (30), it holds

$$\sum_{i,j} P_{(i,j)}^{(I,J)} = 1, \tag{77}$$

where $P_{(\cdot,\cdot)}^{(\cdot,\cdot)}$ is defined by the approximative formula (30). We carry out a proof by induction. Eq. (77) builds our induction hypothesis.

- Base case $(I, J) = (0, 0)$: Since $P_{(i,j)}^{(0,0)} = \delta_{i,0}\delta_{j,0}$, it holds $\sum_{i,j} P_{(i,j)}^{(0,0)} = 1$.
- Inductive step: Let the hypothesis be true for all pairs (k, m) with $k < n$ and (n, k) with $k < m$. We show that it is true for (n, m):

$$\sum_{i,j} P_{(i,j)}^{(n,m)}$$

$$= \frac{\sum_{k=0}^{n-1} r_{(k,m)}^{(n,m)} (1 - Q_{(k,m)}) \sum_{i,j} P_{(i,j)}^{(k,m)} + \sum_{k=0}^{m-1} r_{(n,k)}^{(n,m)} (1 - Q_{(n,k)}) \sum_{i,j} P_{(i,j)}^{(n,k)}}{\sum_{k=0}^{n-1} r_{(k,m)}^{(n,m)} (1 - Q_{(k,m)}) + \sum_{k=0}^{m-1} r_{(n,k)}^{(n,m)} (1 - Q_{(n,k)})}$$

$$= \frac{\sum_{k=0}^{n-1} r_{(k,m)}^{(n,m)} (1 - Q_{(k,m)}) \cdot 1 + \sum_{k=0}^{m-1} r_{(n,k)}^{(n,m)} (1 - Q_{(n,k)}) \cdot 1}{\sum_{k=0}^{n-1} r_{(k,m)}^{(n,m)} (1 - Q_{(k,m)}) + \sum_{k=0}^{m-1} r_{(n,k)}^{(n,m)} (1 - Q_{(n,k)})} = 1. \tag{78}$$

$\square$

The proof works analogously if the number of unlinked alleles is larger than 0.

Appendix D: Unlinked alleles of arbitrary effect

For the main text, we assumed that each unlinked allele has the same selection coefficient. Here, we give the generalization to arbitrary effects.

Let $\mathcal{F} = \{\mathbf{e} = (e_1, e_2, \ldots, e_F); e_i \in \{0, 1\}\}$. We define $|\mathbf{e}| = \sum_{i=1}^{F} e_i$. The set of unlinked deleterious alleles carried by an individual can be characterized by a vector in $\mathcal{F}$ where 1 and 0 at a given position denote the presence or absence of a specific deleterious allele. In this notation, the transitions of the branching process read with $|\mathbf{e}_f| = f$:

$$P((i, j; \mathbf{e}_f) \to \emptyset) = 1 - \sigma_{(i,j;\mathbf{e}_f)},$$

$$P((i, j; \mathbf{e}_f) \to \{(i, j; \mathbf{e}_f), (i, j; \mathbf{e}_g)\}) = \left(\frac{1}{2}\right)^f (1 - r_{(i,j)}) \quad \text{for} \quad \mathbf{e}_f - \mathbf{e}_g \in \mathcal{F},$$

$$P((i, j; \mathbf{e}_f) \rightarrow \{(i, j; \mathbf{e}_f), (i, k; \mathbf{e}_g)\}) = \left(\frac{1}{2}\right)^f r_{(i,k)}^{(i,j)} \quad \text{for} \quad k < j, \mathbf{e}_f - \mathbf{e}_g \in \mathcal{F},$$

$$P((i, j; \mathbf{e}_f) \rightarrow \{(i, j; \mathbf{e}_f), (k, j; \mathbf{e}_g)\}) = \left(\frac{1}{2}\right)^f r_{(k,j)}^{(i,j)} \quad \text{for} \quad k < i, \mathbf{e}_f - \mathbf{e}_g \in \mathcal{F}. \tag{79}$$

The recursive equation for the extinction probability becomes:

$$\left(\frac{1}{2}\right)^f (1 - r_{(i,j)}) Q_{(i,j;\mathbf{e}_f)}^2$$

$$+ \sum_{\mathbf{e}_g; \mathbf{e}_f - \mathbf{e}_g \in \mathcal{F}} \left(\frac{1}{2}\right)^f \left\{ \sum_{k=0}^{i-1} r_{(k,j)}^{(i,j)} Q_{(k,j;\mathbf{e}_g)} + \sum_{k=0}^{j-1} r_{(i,k)}^{(i,j)} Q_{(i,k;\mathbf{e}_g)} \right\} Q_{(i,j;\mathbf{e}_f)}$$

$$+ \sum_{\substack{\mathbf{e}_g; \mathbf{e}_f - \mathbf{e}_g \in \mathcal{F}, \\ \mathbf{e}_f - \mathbf{e}_g \neq \mathbf{0}}} \left(\frac{1}{2}\right)^f (1 - r_{(i,j)}) Q_{(i,j;\mathbf{e}_g)} Q_{(i,j;\mathbf{e}_f)} - (2 - \sigma_{(i,j;\mathbf{e}_f)}) Q_{(i,j,\mathbf{e}_f)}$$

$$+1 - \sigma_{(i,j;\mathbf{e}_f)} = 0. \tag{80}$$

The proof is analogous to before.

Similarly, we can generalize approximation Eq. (34):

$$P_{(i,j;\mathbf{0})}^{(I,J;\mathbf{e}_F)} = \frac{\sum_{\mathbf{e}_l} \left(\frac{1}{2}\right)^F A_{\mathbf{e}_l} + \sum_{\mathbf{e}_l \neq \mathbf{e}_F} \left(\frac{1}{2}\right)^F (1 - r_{(I,J)})(1 - Q_{(I,J;\mathbf{e}_l)}) P_{(i,j;\mathbf{0})}^{(I,J;\mathbf{e}_l)}}{\sum_{\mathbf{e}_l} \left(\frac{1}{2}\right)^F B_{\mathbf{e}_l} + \sum_{\mathbf{e}_l \neq \mathbf{e}_F} \left(\frac{1}{2}\right)^F (1 - r_{(I,J)})(1 - Q_{(I,J;\mathbf{e}_l)})}.$$

$$A_{\mathbf{e}_l} = \sum_{i=0}^{I-1} r_{(i,J)}^{(I,J)} (1 - Q_{(i,J;\mathbf{e}_l)}) P_{(i,j;\mathbf{0})}^{(i,J;\mathbf{e}_l)} + \sum_{j=0}^{J-1} r_{(I,j)}^{(I,J)} (1 - Q_{(I,j;\mathbf{e}_l)}) P_{(i,j;\mathbf{0})}^{(I,j;\mathbf{e}_l)}$$

$$B_{\mathbf{e}_l} = \sum_{i=0}^{I-1} r_{(i,J)}^{(I,J)} (1 - Q_{(i,J;\mathbf{e}_l)}) + \sum_{j=0}^{J-1} r_{(I,j)}^{(I,J)} (1 - Q_{(I,j;\mathbf{e}_l)}) \tag{81}$$

For $\left(\frac{1}{2}\right)^F (1 - r_{(I,J)}) - (1 - \sigma_{(I,J;F)}) > 0$, we approximate:

$$P_{(i,j;\mathbf{0})}^{(I,J;\mathbf{e}_F)} = \delta_{I,i} \delta_{J,j}. \tag{82}$$

Appendix E: General lemmata

Here we summarize some general lemmata which we use in the derivation of the hitch-hiking probability. Although the lemmata either are not new or follow immediately from known results, we briefly sketch the proofs.

We consider a multitype branching process with $N + 1$ types on $(\Omega, \mathcal{F}, P)$. The number of individuals of type i ($i = 0, 1, \ldots, N$) in generation n is $Z_n^{(i)}$. Let $Z_0^{(N)} = 1$ and $Z_0^{(i)} = 0$ if $i \neq N$, i.e., the process is started by an individual of type N; the term "generation" refers to the distance to this founding individual. Individuals of type i reproduce at rate 1 and die at rate $1 - \sigma_i$. Let the offspring of a type N individual be of type i with probability r_i. If a type N individual gives birth to an individual of type $i < N$, we call this a recombination event from type N to type i. We define $r := \sum_{k=0}^{N-1} r_k = 1 - r_N$.

Lemma 2 *The probability generating function (p.g.f.) $f(s)$ of the offspring distribution of a type N individual is given by*

$$f(s) = \frac{1 - p}{1 - ps} \quad \text{with} \quad p = \frac{1}{2 - \sigma_N}. \tag{83}$$

Proof Individuals have a geometric offspring distribution; for the random offspring number X of a type N individual, it holds:

$$P(X = k) = \left(\frac{1}{2 - \sigma_N}\right)^k \frac{1 - \sigma_N}{2 - \sigma_N}, \tag{84}$$

and therefore

$$f(s) = \sum_{k=0}^{\infty} P(X = k)s^k = \sum_{k=0}^{\infty} \left(\frac{1}{2 - \sigma_N}\right)^k \frac{1 - \sigma_N}{2 - \sigma_N} s^k$$
$$= \frac{1 - \sigma_N}{2 - \sigma_N} \frac{1}{1 - \frac{s}{2 - \sigma_N}} = \frac{1 - p}{1 - ps}. \tag{85}$$

$\square$

Lemma 3 *The joint p.g.f. of $(Z_1^{(0)}, Z_1^{(1)}, \ldots, Z_1^{(N)})$ is given by*

$$F(s_0, s_1, \ldots, s_N) = f(r_0 s_0 + r_1 s_1 + \ldots r_N s_N). \tag{86}$$

Proof We restrict to $N = 1$; the generalization to larger N is straightforward. Let $p(k)$ be the probability that an individual of type 1 has k offspring and $p(k_0, k_1)$ the probability that it has k_0 offspring of type 0 and k_1 offspring of type 1. It holds (cf. Serra 2006):

$$F(s_0, s_1) = \sum_{k_0, k_1} p(k_0, k_1) s_0^{k_0} s_1^{k_1} = \sum_{k=0}^{\infty} \sum_{k_0=0}^{k} \underbrace{p(k_0, k - k_0)}_{=p(k)\left\{r_0^{k_0}(1-r_0)^{k-k_0}\binom{k}{k_0}\right\}} s_0^{k_0} s_1^{k-k_0}$$

$$= \sum_{k=0}^{\infty} p(k) \underbrace{\sum_{k_0=0}^{k} \binom{k}{k_0} (r_0 s_0)^{k_0} (s_1(1-r_0))^{k-k_0}}_{=(r_0 s_0 + (1-r_0)s_1)^k}$$

$$= \sum_{k=0}^{\infty} p(k)(r_0 s_0 + (1 - r_0)s_1)^k = f(r_0 s_0 + (1 - r_0)s_1). \qquad (87)$$

$\square$

Lemma 4 *The extinction probability q_N of type N is given by*

$$q_N = \begin{cases} \frac{1-\sigma_N}{1-r} & \text{if } \sigma_N > r, \\ 1 & \text{if } \sigma_N \leq r. \end{cases} \qquad (88)$$

Proof The p.g.f. for the number of offspring of its own type is given by $f(r + (1-r)s)$. From the general theory of single type branching processes, it follows that q_N is the smallest root of

$$q_N = f(r + (1 - r)q_N). \qquad (89)$$

in $[0, 1]$.

Lemma 5 *Let the offspring of type i be of type $j \leq i$ only. Consider the process on the set $B = \left\{ \omega : Z_n^{(N)}(\omega) \to 0 \text{ as } n \to \infty \right\}$, i.e., the set of paths were type N goes extinct. The probability of the event B is denoted by q_N. The joint p.g.f. of $(Z_1^{(0)}, Z_1^{(1)}, \ldots, Z_1^{(N)})$ in the conditioned process is given by*

$$\hat{F}(s_0, s_1, \ldots, s_N) = \frac{F(s_0, s_1, \ldots, q_N s_N)}{q_N}. \qquad (90)$$

Proof For simplicity, we stick to $N = 1$. The generalization to $N > 1$ is straightforward. It holds (cf. Athreya and Ney 1972, p. 52f) :

$$\hat{F}(s_0, s_1) = \sum_{k_0=0}^{\infty} \sum_{k_1=0}^{\infty} P(Z_1^{(0)} = k_0, Z_1^{(1)} = k_1 | B) s_0^{k_0} s_1^{k_1}$$

$$= \frac{\sum_{k_0=0}^{\infty} \sum_{k_1=0}^{\infty} P(Z_1^{(0)} = k_0, Z_1^{(1)} = k_1, B) s_0^{k_0} s_1^{k_1}}{q_1}$$

$$= \frac{\sum_{k_0=0}^{\infty} \sum_{k_1=0}^{\infty} P(Z_1^{(0)} = k_0, Z_1^{(1)} = k_1) q_1^{k_1} s_0^{k_0} s_1^{k_1}}{q_1}$$

$$= \frac{F(s_0, q_1 s_1)}{q_1}.$$
(91)

$\square$

Lemma 6 *Let the offspring of type i be of type $j \leq i$ only. Condition the process on extinction of type N. The total number of recombination events from type N individuals to other types in the whole process is determined by the p.g.f. $\hat{h}(s)$ with*

$$\hat{h}(s) = \frac{1 - psr - \sqrt{(1-psr)^2 - 4(1-p)p(1-r)}}{2pq_N(1-r)}.$$
(92)

Proof We first proof the relation

$$\hat{h}(s) = \hat{F}(s, \hat{h}(s)).$$
(93)

(cf. Serra 2006). Let $\tilde{p}(k)$ be the probability that the total number of recombination events is k and $p(k_r, k_N)$ be the probability that an individual of type N has k_N offspring of type N and k_r offspring of other types. First, note that:

$$\tilde{p}(k) = \sum_{k_N=0}^{\infty} \sum_{j=0}^{k} p(k-j, k_N) \tilde{p}(j|k_N).$$
(94)

Furthermore:

$$\sum_{k=0}^{\infty} \sum_{k_N=0}^{\infty} \sum_{j=0}^{k} p(k-j, k_N) s^k = \sum_{k_r, k_N} p(k_r, k_N) \sum_{j} \tilde{p}(j|k_N) s^{k_r+j}.$$
(95)

Using this:

$$\hat{h}(s) = \sum_{k} \tilde{p}(k) s^k = \sum_{k_r, k_N} p(k_r, k_N) \sum_{j=0}^{\infty} \tilde{p}(j|k_N) s^{j+k_r}$$

$$= \sum_{k_r, k_N} p(k_r, K_N) s_r^k \underbrace{\sum_{j=0}^{\infty} \tilde{p}(j|k_N) s^j}_{=\left(\hat{h}(s)\right)^{k_N}}.$$
(96)

$\square$

It thus holds:

$$\hat{h}(s) = \hat{F}(s, \hat{h}(s)) = \frac{F(s, q_N \hat{h}(s))}{q_N} = \frac{f(sr + q_N(1 - r)\hat{h}(s))}{q_N}$$

$$= \frac{1}{q_N} \frac{1 - p}{1 - p(sr + q_N(1 - r)\hat{h}(s))}. \tag{97}$$

This leads to a quadratic equation for $\hat{h}(s)$ which has two solutions. As it must hold that $\hat{h}(1) = 1$, we can exclude one of them and obtain Eq. (92). $\qquad\square$

If $q_N = 1$, we denote the probability generating function by $h(s)$ in the main text.

Lemma 7 *With the assumptions of the previous lemma, the joint p.g.f. of the number of recombination events from type N to type $0, 1, 2, \ldots, N - 1$ is given by*

$$h_1(s_0, s_1, \ldots, s_{N-1}) = \hat{h}(\frac{r_0}{r}s_0 + \cdots + \frac{r_{N-1}}{r}s_{N-1}). \tag{98}$$

Proof analogous to the proof of Lemma 5.

Appendix F: The deterministic phase

Appendix F1: The process is initiated by an individual of type $(0, 2)$

We now focus on three types $(0, 0)$, $(0, 1)$, and $(0, 2)$, that all have a positive selection coefficient. All other types are assumed to go extinct. A lineage of type $(0, 2)$ starts sweeping. Throughout, we assume that recombination is weak enough that it does not alter the deterministic frequency path of a type.

The deterministic frequency path is given by

$$x_{(0,2)} = \frac{\bar{v}_0 \exp(\sigma_{(0,2)}t)}{N - \bar{v}_0 + \bar{v}_0 \exp(\sigma_{(0,2)}t)} \tag{99}$$

with $\bar{v}_0 = \frac{1}{1-q_{(0,2)}}$ where $q_{(0,2)}$ is defined in Eq. (26) as $q_{(0,2)} = \frac{1-\sigma_{(0,2)}}{1-r_{(0,2)}}$. Recombination events to successful lineages occur at rate

$$\alpha_{(0,2)}(t) = \alpha_{(0,0)}(t) + \alpha_{(0,1)}(t) \tag{100}$$

with

$$\alpha_{(0,0)}(t) = Nx_{(0,2)}(t)(1 - x_{(0,2)}(t))r_{(0,0)}^{(0,2)} p_{\text{est}}^{(0,0)}(x_{(0,2)}(t)), \tag{101a}$$

$$\alpha_{(0,1)}(t) = Nx_{(0,2)}(t)(1 - x_{(0,2)}(t))r_{(0,1)}^{(0,2)} p_{\text{est}}^{(0,1)}(x_{(0,2)}(t)), \tag{101b}$$

where $p_{\text{est}}^{(0,0)}$ and $p_{\text{est}}^{(0,1)}$ are the establishment probabilities of an individual of type $(0, 0)$ and $(0, 1)$ respectively and are calculated by suitable substitutions in Eq. (38).

We assume in the following that we can set $\bar{v}_0/N \approx 0$ in the integration boundaries. Analogous to Eq. (41), the probability that no successful recombination event takes place is then given by

$$
P_{\text{det}}^{((0,2)\to(0,2))} \approx \exp\left[-\int_0^\infty \alpha_{(0,2)}(t)\mathrm{d}t\right] = \left(\frac{\sigma_{(0,0)}}{\sigma_{(0,0)} - \sigma_{(0,2)}}\right)^{-\frac{Nr_{(0,0)}^{(0,2)}\sigma_{(0,0)}(\sigma_{(0,0)}-\sigma_{(0,2)})}{\sigma_{(0,2)}^2}}
$$

$$
\times \left(\frac{\sigma_{(0,1)}}{\sigma_{(0,1)} - \sigma_{(0,2)}}\right)^{-\frac{Nr_{(0,1)}^{(0,2)}\sigma_{(0,1)}(\sigma_{(0,1)}-\sigma_{(0,2)})}{\sigma_{(0,2)}^2}}. \tag{102}
$$

The probability that no successful type $(0, 0)$ is generated up to t_R is given by

$$
\exp\left[-\int_0^{t_R} \alpha_{(0,0)}(t)\mathrm{d}t\right]
$$

$$
\approx \exp\left[-\int_0^{x_{(0,2)}(t_R)} Nr_{(0,0)}^{(0,2)} \frac{1}{\sigma_{(0,2)}} \frac{\sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,2)})}{(\sigma_{(0,0)} - \sigma_{(0,2)})(1 - x) + \sigma_{(0,0)}x}\mathrm{d}x\right]
$$

$$
= \left(\frac{\sigma_{(0,0)} - (1 - x_{(0,2)}(t_R))\sigma_{(0,2)}}{\sigma_{(0,0)} - \sigma_{(0,2)}}\right)^{-\frac{Nr_{(0,0)}^{(0,2)}\sigma_{(0,0)}(\sigma_{(0,0)}-\sigma_{(0,2)})}{\sigma_{(0,2)}\sigma_{(0,2)}}}
$$

$$
\equiv Q_{\text{det}}^{(0,0)}(x_{(0,2)}(t_R)). \tag{103}
$$

The probability that no successful type $(0, 1)$ is generated up to time t_R is given by

$$
\exp\left[-\int_0^{t_R} \alpha_{(0,1)}(t)\mathrm{d}t\right]
$$

$$
\approx \exp\left[-\int_0^{x_{(0,2)}(t_R)} Nr_{(0,1)}^{(0,2)} \frac{1}{\sigma_{(0,2)}} \frac{\sigma_{(0,1)}(\sigma_{(0,1)} - \sigma_{(0,2)})}{(\sigma_{(0,1)} - \sigma_{(0,2)})(1 - x) + \sigma_{(0,1)}x}\mathrm{d}x\right]
$$

$$
= \left(\frac{\sigma_{(0,1)} - (1 - x_{(0,2)}(t_R))\sigma_{(0,2)}}{\sigma_{(0,1)} - \sigma_{(0,2)}}\right)^{-\frac{Nr_{(0,0)}^{(0,2)}\sigma_{(0,1)}(\sigma_{(0,1)}-\sigma_{(0,2)})}{\sigma_{(0,2)}\sigma_{(0,2)}}}
$$

$$
\equiv Q_{\text{det}}^{(0,1)}(x_{(0,2)}(t_R)). \tag{104}
$$

The probability that as a first event, a successful individual of type $(0, 0)$ is generated is given by

$$
\int_0^\infty Q_{\det}^{(0,0)}(x_{(0,2)}(t_R)) Q_{\det}^{(0,1)}(x_{(0,2)}(t_R)) \alpha_{(0,0)}(t_R) \mathrm{d}t_R
$$

$$
= \int_0^\infty Q_{\det}^{(0,0)}(x_{(0,2)}(t_R)) Q_{\det}^{(0,1)}(x_{(0,2)}(t_R))
$$

$$
\times N x_{(0,2)}(t)(1 - x_{(0,2)}(t)) r_{(0,0)}^{(0,2)} p_{\mathrm{est}}^{(0,0)}(x_{(0,2)}(t)) \mathrm{d}t_R
$$

$$
= \frac{1}{\sigma_{(0,2)}} \int_0^1 Q_{\det}^{(0,0)}(x) Q_{\det}^{(0,1)}(x) N r_{(0,0)}^{(0,2)} p_{\mathrm{est}}^{(0,0)}(x) \mathrm{d}x. \tag{105}
$$

Equivalently, the probability that as a first event, a successful individual of type $(0, 1)$ is generated is given by

$$
\frac{1}{\sigma_{(0,2)}} \int_0^1 Q_{\det}^{(0,0)}(x) Q_{\det}^{(0,1)}(x) N r_{(0,1)}^{(0,2)} p_{\mathrm{est}}^{(0,1)}(x) \mathrm{d}x. \tag{106}
$$

In that case, type $(0, 1)$ might either rise to fixation, or a successful type $(0, 0)$ individual might be generated in the following process. We now calculate the probability of these two events given a type $(0, 1)$ individual was generated at time t_R. In order to calculate the probability for the generation of a successful type $(0, 0)$ individual, we need the frequencies of the three present types (wildtype, type $(0, 2)$, and type $(0, 1)$) at time t after the recombination event. We have to take into account that in its early phase of growth, the type $(0, 1)$ increases faster than predictic by the deterministic path. As before, we therefore replace its initial frequency by the mean of the effective initial population size ν. We assume that during establishment, the type $(0, 1)$ individuals exclusively replace wildtype individuals. This assumption has no visible influence on the results. We obtain for the frequencies of wildtype individuals, type $(0, 1)$ individuals, and type $(0, 2)$ individuals, respectively:

$$
\tilde{q}(t|x_{(0,2)}(t_R))
$$

$$
= \frac{1 - x_{(0,2)}(t_R) - \frac{\bar{\nu}(t_R)}{N}}{x_{(0,2)}(t_R) \exp(\sigma_{(0,2)}t) + 1 - x_{(0,2)}(t_R) - \frac{\bar{\nu}(t_R)}{N} + \frac{\bar{\nu}(t_R)}{N} \exp(\sigma_{(0,1)}t)}, \tag{107}
$$

$$
\tilde{x}_{01}(t|x_{(0,2)}(t_R))
$$

$$
= \frac{\frac{\bar{\nu}(t_R)}{N} \exp(\sigma_{(0,1)}t)}{x_{(0,2)}(t_R) \exp(\sigma_{(0,2)}t) + 1 - x_{(0,2)}(t_R) - \frac{\bar{\nu}(t_R)}{N} + \frac{\bar{\nu}(t_R)}{N} \exp(\sigma_{(0,1)}t)}, \tag{108}
$$

$$\tilde{x}_{02}(t|x_{(0,2)}(t_R))$$

$$= \frac{x_{(0,2)}(t_R)\exp(\sigma_{(0,2)}t)}{x_{(0,2)}(t_R)\exp(\sigma_{(0,2)}t) + 1 - x_{(0,2)}(t_R) - \frac{\bar{v}(t_R)}{N} + \frac{\bar{v}(t_R)}{N}\exp(\sigma_{(0,1)}t)} \quad (109)$$

with $\bar{v}(t_R) = \frac{1}{p_{\text{est}}^{(0,1)}(x_{(0,2)})(t_R))}$. The time-dependent selection coefficient of type $(0,0)$ is given by

$$\tilde{s}_{(0,0)}(t|x_{(0,2)}(t_R)) = \sigma_{(0,0)} - \sigma_{(0,1)}\tilde{x}_{01}(t|x_{(0,2)}(t_R)) - \sigma_{(0,2)}\tilde{x}_{02}(t|x_{(0,2)}(t_R)). \quad (110)$$

The establishment probability of a single individual of type $(0,0)$ that arises at time t after t_R can be calculated similarly to before and is given by

$$\tilde{p}_{\text{est}}^{(0,0)}(t|x_{(0,2)}(t_R)) = \frac{1}{A}\sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,1)})(\sigma_{(0,0)} - \sigma_{(0,2)}),$$
$$A = (\sigma_{(0,0)} - \sigma_{(0,1)})(\sigma_{(0,0)} - \sigma_{(0,2)})\tilde{q}(t|x_{(0,2)}(t_R))$$
$$+ \sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,2)})\tilde{x}_{01}(t|x_{(0,2)}(t_R))$$
$$+ \sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,1)})\tilde{x}_{02}(t|x_{(0,2)}(t_R)). \quad (111)$$

The rate of successful type $(0,0)$ individuals is

$$\tilde{\alpha}_{(0,0)}(t|x_{(0,2)}(t_R)) = Nr_{(0,0)}^{(0,1)}\tilde{q}(t|x_{(0,2)}(t_R))(1 - \tilde{q}(t|x_{(0,2)}(t_R)))\tilde{p}_{\text{est}}^{(0,0)}(t|x_{(0,2)}(t_R)). \quad (112)$$

(Remember: $r_{(0,0)}^{(0,1)} = r_{(0,0)}^{(0,2)}$.) We assume that we can ignore deviations from the deterministic frequency path for large t close to fixation of the beneficial allele. The probability that no successful recombination event takes place is then given by

$$\text{Prob(no rec}|x_{(0,2)}(t_R)) \approx \exp\left(-\int_0^\infty \tilde{\alpha}_{(0,0)}(t|x_{(0,2)}(t_R))dt\right). \quad (113)$$

The probability that type $(0,1)$ fixes is hence

$$P_{\text{det}}^{((0,2)\to(0,1))} \approx \int_0^\infty Q_{\text{det}}^{(0,0)}(x_{(0,2)}(t_R))Q_{\text{det}}^{(0,1)}(x_{(0,2)}(t_R))\alpha_{(0,1)}(t_R)$$
$$\times\text{Prob(no rec}|x_{(0,2)}(t_R))dt_R.$$
$$= \frac{1}{\sigma_{(0,2)}}\int_0^1 Q_{\text{det}}^{(0,0)}(x)Q_{\text{det}}^{(0,1)}(x)Nr_{(0,1)}^{(0,2)}p_{\text{est}}^{(0,1)}(x)\text{Prob(no rec}|x)dx.$$

$$(114)$$

The overall probability that type $(0, 0)$ fixes is obtained as

$$P_{\text{det}}^{((0,2)\to(0,0))} \approx \frac{1}{\sigma_{(0,2)}} \int_0^1 Q_{\text{det}}^{(0,0)}(x) Q_{\text{det}}^{(0,1)}(x) N r_{(0,0)}^{(0,2)} p_{\text{est}}^{(0,0)}(x) dx$$

$$+ \frac{1}{\sigma_{(0,2)}} \int_0^1 Q_{\text{det}}^{(0,0)}(x) Q_{\text{det}}^{(0,1)}(x) N r_{(0,1)}^{(0,2)} p_{\text{est}}^{(0,1)}(x)$$

$$\times (1 - \text{Prob(no rec}|x)) dx. \tag{115}$$

7.1 The process is initiated by an individual of type $(1, 1)$

We now focus on four types $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$, which all have a positive selection coefficient. An individual of type $(1, 1)$ starts sweeping. We derive an approximation which is valid if no more than two successful recombination events happen during the process, i.e., for weak recombination. We again assume that recombination does not significantly reduce the number of offspring of any individual's own type.

The deterministic path of type $(1, 1)$ is given by

$$x_{(1,1)} = \frac{\bar{v}_0 \exp\left(\sigma_{(1,1)}t\right)}{N - \bar{v}_0 + \bar{v}_0 \exp\left(\sigma_{(1,1)}t\right)} \tag{116}$$

with $\bar{v}_0 = \frac{1}{1-q_{(1,1)}}$, where $q_{(1,1)}$ is defined in Eq. (26) as $q_{(1,1)} = \frac{1-\sigma_{(1,1)}}{1-r_{(1,1)}}$. Recombination events that generate successful lineages occur at rate

$$\alpha_{(1,1)}(t) = \alpha_{(1,0)}(t) + \alpha_{(0,1)}(t) \tag{117}$$

with

$$\alpha_{(1,0)}(t) = N x_{(1,1)}(t)(1 - x_{(1,1)}(t)) r_{(1,0)}^{(1,1)} p_{\text{est}}^{(1,0)}(x_{(1,1)}(t)), \tag{118a}$$

$$\alpha_{(0,1)}(t) = N x_{(1,1)}(t)(1 - x_{(1,1)}(t)) r_{(0,1)}^{(1,1)} p_{\text{est}}^{(0,1)}(x_{(1,1)}(t)), \tag{118b}$$

where $p_{\text{est}}^{(1,0)}$ and $p_{\text{est}}^{(0,1)}$ are the establishment probabilities of an individual of type $(1, 0)$ and $(0, 1)$ respectively and calculated by suitable substitutions in Eq. (38). Analogous to Eq. (41), the probability that no successful recombination event takes place is given by

$$P_{\text{det}}^{((1,1)\to(1,1))} \approx \exp\left[-\int_0^\infty \alpha_{(1,1)}(t) dt\right] = \left(\frac{\sigma_{(1,0)}}{\sigma_{(1,0)} - \sigma_{(1,1)}}\right)^{-\frac{N r_{(1,0)}^{(1,1)} \sigma_{(1,0)}(\sigma_{(1,0)} - \sigma_{(1,1)})}{\sigma_{(1,1)}^2}}$$

$$\times \left(\frac{\sigma_{(0,1)}}{\sigma_{(0,1)} - \sigma_{(1,1)}}\right)^{-\frac{N r_{(0,1)}^{(1,1)} \sigma_{(0,1)}(\sigma_{(0,1)} - \sigma_{(1,1)})}{\sigma_{(1,1)}^2}}. \tag{119}$$

The probability that no successful type $(1, 0)$ is generated up to t_R is given by

$$\exp\left[-\int_0^{t_R} \alpha_{(1,0)}(t)\mathrm{d}t\right]$$

$$\approx \exp\left[-\int_0^{x_{(1,1)}(t_R)} Nr_{(1,0)}^{(1,1)} \frac{1}{\sigma_{(1,1)}} \frac{\sigma_{(1,0)}(\sigma_{(1,0)} - \sigma_{(1,1)})}{(\sigma_{(1,0)} - \sigma_{(1,1)})(1 - x) + \sigma_{(1,0)}x}\mathrm{d}x\right]$$

$$= \left(\frac{\sigma_{(1,0)} - (1 - x_{(1,1)}(t_R))\sigma_{(1,1)}}{\sigma_{(1,0)} - \sigma_{(1,1)}}\right)^{-\frac{Nr_{(1,0)}^{(1,1)}\sigma_{(1,0)}(\sigma_{(1,0)} - \sigma_{(1,1)})}{\sigma_{(1,1)}\sigma_{(1,1)}}} \equiv Q_{\mathrm{det}}^{(1,0)}(x_{(1,1)}(t_R)).$$

$$(120)$$

The probability that no successful type $(0, 1)$ is generated up to t_R is given by

$$\exp\left[-\int_0^{t_R} \alpha_{(0,1)}(t)\mathrm{d}t\right]$$

$$\approx \exp\left[-\int_0^{x_{(1,1)}(t_R)} Nr_{(0,1)}^{(1,1)} \frac{1}{\sigma_{(1,1)}} \frac{\sigma_{(0,1)}(\sigma_{(0,1)} - \sigma_{(1,1)})}{(\sigma_{(0,1)} - \sigma_{(1,1)})(1 - x) + \sigma_{(0,1)}x}\mathrm{d}x\right]$$

$$= \left(\frac{\sigma_{(0,1)} - (1 - x_{(1,1)}(t_R))\sigma_{(1,1)}}{\sigma_{(0,1)} - \sigma_{(1,1)}}\right)^{-\frac{Nr_{(0,1)}^{(1,1)}\sigma_{(0,1)}(\sigma_{(0,1)} - \sigma_{(1,1)})}{\sigma_{(1,1)}\sigma_{(1,1)}}}$$

$$\equiv Q_{\mathrm{det}}^{(0,1)}(x_{(1,1)}(t_R)).$$

$$(121)$$

The probability that as a first event, a successful type $(1, 0)$ individual is generated is given by

$$\int_0^\infty Q_{\mathrm{det}}^{(1,0)}(x_{(1,1)}(t_R))Q_{\mathrm{det}}^{(0,1)}(x_{(1,1)}(t_R))\alpha_{(1,0)}(t_R)\mathrm{d}t_R$$

$$= \int_0^\infty Q_{\mathrm{det}}^{(1,0)}(x_{(1,1)}(t_R))Q_{\mathrm{det}}^{(0,1)}(x_{(1,1)}(t_R))$$

$$\times Nx_{(1,1)}(t)(1 - x_{(1,1)}(t))r_{(1,0)}^{(1,1)}p_{\mathrm{est}}^{(1,0)}(x_{(1,1)}(t))\mathrm{d}t_R$$

$$= \frac{1}{\sigma_{(1,1)}} \int_0^1 Q_{\mathrm{det}}^{(1,0)}(x)Q_{\mathrm{det}}^{(0,1)}(x)Nr_{(1,0)}^{(1,1)}p_{\mathrm{est}}^{(1,0)}(x)\mathrm{d}x.$$

$$(122)$$

Equivalently, the probability that, as a first event, a successful type $(0, 1)$ individual is generated is given by

$$\frac{1}{\sigma_{(1,1)}} \int_0^1 Q_{\text{det}}^{(1,0)}(x) Q_{\text{det}}^{(0,1)}(x) N r_{(0,1)}^{(1,1)} p_{\text{est}}^{(0,1)}(x) \mathrm{d}x. \qquad (123)$$

We now choose without loss of generality $\sigma_{(1,0)} \leq \sigma_{(0,1)}$.

We first consider the case that as a first event, an individual of type $(0, 1)$ is generated. In that case, type $(0, 1)$ might either rise to fixation, or a successful individual of type $(0, 0)$ might be generated in the following process. We ignore the possibility that an individual of type $(1, 0)$ might be generated and temporally sweep until it goes extinct. We now calculate the probability of the two events given a successful individual of type $(0, 1)$ has been generated at time t_R. A recombination event may generate a successful type $(0, 1)$ individual at time t_R. The type $(0, 1)$ individual will rise to fixation unless a successful type $(0, 0)$ individual is generated. In order to calculate this probability, we need the frequencies of the three present types (wildtype, type $(0, 1)$, $(1, 1)$) at time t after the recombination event:

$$\tilde{q}(t | x_{(1,1)}(t_R))$$

$$= \frac{1 - x_{(1,1)}(t_R) - \frac{\bar{v}(t_R)}{N}}{x_{(1,1)}(t_R) \exp(\sigma_{(1,1)}t) + 1 - x_{(1,1)}(t_R) - \frac{\bar{v}(t_R)}{N} + \frac{\bar{v}(t_R)}{N} \exp(\sigma_{(0,1)}t)}, \qquad (124a)$$

$$\tilde{x}_{01}(t | x_{(1,1)}(t_R))$$

$$= \frac{\frac{\bar{v}(t_R)}{N} \exp(\sigma_{(0,1)}t)}{x_{(1,1)}(t_R) \exp(\sigma_{(1,1)}t) + 1 - x_{(1,1)}(t_R) - \frac{\bar{v}(t_R)}{N} + \frac{\bar{v}(t_R)}{N} \exp(\sigma_{(0,1)}t)}, \qquad (124b)$$

$$\tilde{x}_{11}(t | x_{(1,1)}(t_R))$$

$$= \frac{x_{(1,1)}(t_R) \exp(\sigma_{(1,1)}t)}{x_{(1,1)}(t_R) \exp(\sigma_{(1,1)}t) + 1 - x_{(1,1)}(t_R) - \frac{\bar{v}(t_R)}{N} + \frac{\bar{v}(t_R)}{N} \exp(\sigma_{(0,1)}t)} \qquad (124c)$$

with $\bar{v}(t_R) = \frac{1}{p_{\text{est}}^{(0,1)}(x_{(1,1)}(t_R))}$. The time-dependent selection coefficient of type $(0, 0)$ is given by

$$\tilde{s}_{(0,0)}(t | x_{(1,1)}(t_R)) = \sigma_{(0,0)} - \sigma_{(0,1)} \tilde{x}_{01}(t | x_{(1,1)}(t_R)) - \sigma_{(1,1)} \tilde{x}_{11}(t | x_{(1,1)}(t_R)). \qquad (125)$$

The establishment probability of a type $(0, 1)$ individual that arises at time t after t_R can be calculated as before and is given by

$$\tilde{p}_{\text{est}}^{(0,0)}(t | x_{(1,1)}(t_R)) = \frac{1}{A} \sigma_{(0,0)} (\sigma_{(0,0)} - \sigma_{(0,1)})(\sigma_{(0,0)} - \sigma_{(1,1)}),$$

$$A = (\sigma_{(0,0)} - \sigma_{(0,1)})(\sigma_{(0,0)} - \sigma_{(1,1)})\tilde{q}(t|x_{(1,1)}(t_R))$$
$$+ \sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(1,1)})\tilde{x}_{01}(t|x_{(1,1)}(t_R))$$
$$+ \sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(0,1)})\tilde{x}_{11}(t|x_{(1,1)}(t_R)). \qquad (126)$$

The rate of successful type $(0, 0)$ individuals is

$$\tilde{\alpha}_{(0,0)}(t|x_{(1,1)}(t_R)) = Nr_{(0,0)}^{(0,1)}\tilde{q}(t|x_{(1,1)}(t_R))\tilde{x}_{01}(t|x_{(1,1)}(t_R))\tilde{p}_{\text{est}}^{(0,0)}(t|x_{(1,1)}(t_R)). \qquad (127)$$

The probability that no successful recombination event takes place is therefore given by

$$\text{Prob(no rec}|x_{(1,1)}(t_R)) \approx \exp\left[-\int_0^\infty \tilde{\alpha}_{(0,0)}(t|x_{(1,1)}(t_R))dt\right]. \qquad (128)$$

The probability that type $(0, 1)$ fixes is therefore given by:

$$P_{\text{det}}^{((1,1)\to(0,1))} \approx \int_0^\infty Q_{\text{det}}^{(1,0)}(x_{(1,1)}(t_R))Q_{\text{det}}^{(0,1)}(x_{(1,1)}(t_R))\alpha_{(0,1)}(t_R)$$
$$\text{Prob(no rec}|x_{(1,1)}(t_R))dt_R.$$
$$= \frac{1}{\sigma_{(1,1)}}\int_0^1 Q_{\text{det}}^{(1,0)}(x)Q_{\text{det}}^{(0,1)}(x)Nr_{(0,1)}^{(1,1)}p_{\text{est}}^{(0,1)}(x)\text{Prob(no rec}|x)dx. \qquad (129)$$

Accordingly, the probability that type $(0, 0)$ fixes (given a successful individual of type $(0, 1)$ has been generated) is given by

$$\frac{1}{\sigma_{(1,1)}}\int_0^1 Q_{\text{det}}^{(1,0)}(x)Q_{\text{det}}^{(0,1)}(x)Nr_{(0,1)}^{(1,1)}p_{\text{est}}^{(0,1)}(x)(1 - \text{Prob(no rec}|x))dx. \qquad (130)$$

We now turn to the case that an individual of type $(1, 0)$ is generated first. A recombination event may generate a successful type $(1, 0)$ individual at time t_R. Type $(1, 0)$ will rise to fixation unless a successful type $(0, 0)$ or type $(0, 1)$ individual is generated. In order to calculate these probabilities, we again need the frequencies of the three present types (wildtype, type $(1, 0)$, $(1, 1)$) at time t after the recombination event:

$$\tilde{Q}(t|x_{(1,1)}(t_R))$$

$$= \frac{1 - x_{(1,1)}(t_R) - \frac{\bar{\mu}(t_R)}{N}}{x_{(1,1)}(t_R)\exp\left(\sigma_{(1,1)}t\right) + 1 - x_{(1,1)}(t_R) - \frac{\bar{\mu}(t_R)}{N} + \frac{\bar{\mu}(t_R)}{N}\exp\left(\sigma_{(1,0)}t\right)}, \quad (131a)$$

$$\tilde{X}_{10}(t|x_{(1,1)}(t_R))$$

$$= \frac{\frac{\bar{\mu}(t_R)}{N}\exp\left(\sigma_{(1,0)}t\right)}{x_{(1,1)}(t_R)\exp\left(\sigma_{(1,1)}t\right) + 1 - x_{(1,1)}(t_R) - \frac{\bar{\mu}(t_R)}{N} + \frac{\bar{\mu}(t_R)}{N}\exp\left(\sigma_{(1,0)}t\right)}, \quad (131b)$$

$$\tilde{X}_{11}(t|x_{(1,1)}(t_R))$$

$$= \frac{x_{(1,1)}(t_R)\exp\left(\sigma_{(1,1)}t\right)}{x_{(1,1)}(t_R)\exp\left(\sigma_{(1,1)}t\right) + 1 - x_{(1,1)}(t_R) - \frac{\bar{\mu}(t_R)}{N} + \frac{\bar{\mu}(t_R)}{N}\exp\left(\sigma_{(1,0)}t\right)} \quad (131c)$$

with $\bar{\mu}(t_R) = \frac{1}{p_{\text{est}}^{(1,0)}(x_{(1,1)}(t_R))}$. The time-dependent selection coefficient of a type $(0, 0)$ individual is given by

$$\tilde{S}_{(0,0)}(t|x_{(1,1)}(t_R)) = \sigma_{(0,0)} - \sigma_{(1,0)}\tilde{X}_{01}(t|x_{(1,1)}(t_R)) - \sigma_{(1,1)}\tilde{X}_{11}(t|x_{(1,1)}(t_R)). \quad (132)$$

The time-dependent selection coefficient of a type $(0, 1)$ individual is given by

$$\tilde{S}_{(0,1)}(t|x_{(1,1)}(t_R)) = \sigma_{(0,1)} - \sigma_{(1,0)}\tilde{X}_{10}(t|x_{(1,1)}(t_R)) - \sigma_{(1,1)}\tilde{X}_{11}(t|x_{(1,1)}(t_R)). \quad (133)$$

The establishment probability of a type $(0, 0)$ individual that arises at time t after t_R can be calculated as before and is given by

$$\tilde{\pi}_{\text{est}}^{(0,0)}(t|x_{(1,1)}(t_R)) = \frac{1}{A}\sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(1,0)})(\sigma_{(0,0)} - \sigma_{(1,1)}),$$

$$A = (\sigma_{(0,0)} - \sigma_{(1,0)})(\sigma_{(0,0)} - \sigma_{(1,1)})\tilde{Q}(t|x_{(1,1)}(t_R))$$
$$+ \sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(1,1)})\tilde{X}_{10}(t|x_{(1,1)}(t_R))$$
$$+ \sigma_{(0,0)}(\sigma_{(0,0)} - \sigma_{(1,0)})\tilde{X}_{11}(t|x_{(1,1)}(t_R)). \quad (134)$$

and accordingly for the establishment probability of a type $(0, 1)$ individual

$$\tilde{\pi}_{\text{est}}^{(0,1)}(t|x_{(1,1)}(t_R)) = \frac{1}{B}\sigma_{(0,1)}(\sigma_{(0,1)} - \sigma_{(1,0)})(\sigma_{(0,1)} - \sigma_{(1,1)}),$$

$$B = (\sigma_{(0,1)} - \sigma_{(1,0)})(\sigma_{(0,1)} - \sigma_{(1,1)})\tilde{Q}(t|x_{(1,1)}(t_R))$$
$$+ \sigma_{(0,1)}(\sigma_{(0,1)} - \sigma_{(1,1)})\tilde{X}_{10}(t|x_{(1,1)}(t_R))$$
$$+ \sigma_{(0,1)}(\sigma_{(0,1)} - \sigma_{(1,0)})\tilde{X}_{11}(t|x_{(1,1)}(t_R)). \quad (135)$$

The rate of successful type $(0, 0)$ individuals is

$$\tilde{\beta}_{(0,0)}(t|x_{(1,1)}(t_R)) = Nr_{(0,0)}^{(1,0)}\tilde{Q}(t|x_{(1,1)}(t_R))\tilde{X}_{10}(t|x_{(1,1)}(t_R))\tilde{\pi}_{\text{est}}^{(0,0)}(t|x_{(1,1)}(t_R)).$$
(136)

The rate of successful type $(0, 1)$ individuals is

$$\tilde{\beta}_{(0,1)}(t|x_{(1,1)}(t_R)) = Nr_{(0,1)}^{(1,1)}\tilde{Q}(t|x_{(1,1)}(t_R))\tilde{X}_{11}(t|x_{(1,1)}(t_R))\tilde{\pi}_{\text{est}}^{(0,1)}(t|x_{(1,1)}(t_R)).$$
(137)

The probability that no successful recombination event takes place is therefore given by

$$\text{Prob}_2(\text{no rec}|x_{(1,1)}(t_R)) \approx \exp\left[-\int_0^\infty \tilde{\beta}_{(0,0)}(t|x_{(1,1)}(t_R)) + \tilde{\beta}_{(0,1)}(t|x_{(1,1)}(t_R))\mathrm{d}t\right].$$
(138)

The probability that type $(1, 0)$ fixes is therefore given by:

$$\begin{aligned}
P_{\text{det}}^{((1,1)\to(1,0))} &\approx \int_0^\infty Q_{\text{det}}^{(1,0)}(x_{(1,1)}(t_R))Q_{\text{det}}^{(0,1)}(x_{(1,1)}(t_R))\alpha_{(1,0)}(t_R) \\
&\quad\times\text{Prob}_2(\text{no rec}|x_{(1,1)}(t_R))\mathrm{d}t_R. \\
&= \frac{1}{\sigma_{(1,1)}}\int_0^1 Q_{\text{det}}^{(1,0)}(x)Q_{\text{det}}^{(0,1)}(x)Nr_{(1,0)}^{(1,1)}p_{\text{est}}^{(1,0)}(x)\text{Prob}_2(\text{no rec}|x)\mathrm{d}x.
\end{aligned}$$
(139)

The probability that a successful type $(0, 0)$ individual is generated next at time τ_R, is given by

$$\frac{1}{\sigma_{(1,1)}}\int_0^1 Q_{\text{det}}^{(1,0)}(x)Q_{\text{det}}^{(0,1)}(x)Nr_{(1,0)}^{(1,1)}p_{\text{est}}^{(1,0)}(x)e^{-\int_0^{\tau_R}\tilde{\beta}_{(0,0)}(t|x)+\tilde{\beta}_{(0,1)}(t|x)\mathrm{d}t}\tilde{\beta}_{(0,0)}(\tau_R|x)\mathrm{d}x.$$
(140)

The probability that a successful type $(0, 0)$ is generated next at any time is thus

$$\frac{1}{\sigma_{(1,1)}} \int_0^1 Q_{det}^{(1,0)}(x) Q_{det}^{(0,1)}(x) N r_{(1,0)}^{(1,1)} p_{est}^{(1,0)}(x)$$

$$\times \int_0^\infty e^{-\int_0^{\tau_R} \tilde{\beta}_{(0,0)}(t|x)+\tilde{\beta}_{(0,1)}(t|x)dt} \tilde{\beta}_{(0,0)}(\tau_R|x) d\tau_R dx. \tag{141}$$

The probability that a successful type $(0, 1)$ individual is generated next, is given by

$$\frac{1}{\sigma_{(1,1)}} \int_0^1 Q_{det}^{(1,0)}(x) Q_{det}^{(0,1)}(x) N r_{(1,0)}^{(1,1)} p_{est}^{(1,0)}(x)$$

$$\times \int_0^\infty e^{-\int_0^{\tau_R} \tilde{\beta}_{(0,0)}(t|x)+\tilde{\beta}_{(0,1)}(t|x)dt} \tilde{\beta}_{(0,1)}(\tau_R|x) d\tau_R dx. \tag{142}$$

As numerical evaluation of these integrals is computationally expensive, we introduce the following approximation: We calculate the probability that at least one individual of type $(0, 0)$ is generated, ignoring type $(0, 1)$ (and vice versa), i.e.,

$$\int_0^\infty \exp\left[-\int_0^{\tau_R} \tilde{\beta}_{(0,0)}(t|x)dt\right] \tilde{\beta}_{(0,0)}(\tau_R|x)d\tau_R = 1 - \exp\left[-\int_0^\infty \tilde{\beta}_{(0,0)}(t|x)dt\right]$$

instead of

$$\int_0^\infty e^{-\int_0^{\tau_R} \tilde{\beta}_{(0,0)}(t|x)+\tilde{\beta}_{(0,1)}(t|x)dt} \tilde{\beta}_{(0,0)}(\tau_R|x)d\tau_R.$$

With this approximation, we define

$$P_{(0,0)}^{\text{upper bound}} = \frac{1}{\sigma_{(1,1)}} \int_0^1 Q_{det}^{(1,0)}(x) Q_{det}^{(0,1)}(x) N r_{(1,0)}^{(1,1)} p_{est}^{(1,0)}(x)$$

$$\times \left(1 - \exp\left[-\int_0^\infty \tilde{\beta}_{(0,0)}(t|x)dt\right]\right) dx \tag{143}$$

and equivalently $P_{(0,1)}^{\text{upper bound}}$.

In order to make the fixation probabilities of the various types sum up to one, we subsequently introduce a normalization factor, which overall leads to the approximation

$$P_{(0,0)}^{\text{upper bound}} \times \frac{\frac{1}{\sigma_{(1,1)}} \int_0^1 Q_{\text{det}}^{(1,0)}(x) Q_{\text{det}}^{(0,1)}(x) N r_{(1,0)}^{(1,1)} p_{\text{est}}^{(1,0)}(x)(1 - \text{Prob}_2(\text{no rec}|x)) dx}{P_{(0,0)}^{\text{upper bound}} + P_{(0,1)}^{\text{upper bound}}}$$

(144)

for the probability that type $(0, 0)$ is generated next at any time.

For the probability that type $(0, 1)$ is generated next at any time, we approximate accordingly

$$P_{(0,1)}^{\text{upper bound}} \times \frac{\frac{1}{\sigma_{(1,1)}} \int_0^1 Q_{\text{det}}^{(1,0)}(x) Q_{\text{det}}^{(0,1)}(x) N r_{(1,0)}^{(1,1)} p_{\text{est}}^{(1,0)}(x)(1 - \text{Prob}_2(\text{no rec}|x)) dx}{P_{(0,0)}^{\text{upper bound}} + P_{(0,1)}^{\text{upper bound}}}$$

(145)

After establishment of type $(0, 1)$, a successful individual of type $(0, 0)$ might still be generated. We ignore this probability. This will be a good approximation when the probability of three successful recombination events is low, i.e., in particular when recombination is low.

The probability that type $(0, 0)$ rises to fixation is then approximately given by the sum of Eq. (130) and Eq. (144). And the probability that type $(0, 0)$ rises to fixation is approximately given by the sum of Eq. (129) and Eq. (145).

Appendix G: Diffusion with killing for $I = 0, J = 1$

Hartfield and Otto (2011) derive a diffusion equation for the probability that type $(0, 1)$ fixes in the population conditioned on fixation of the beneficial allele. We here give the equation adjusted to our model (for the derivation, we refer to Hartfield and Otto (2011)). We define:

$$S_{(0,0)} = N\sigma_{(0,0)},$$ (146a)
$$S_{(0,1)} = N\sigma_{(0,1)},$$ (146b)
$$\rho_{(0,0)}^{(0,1)} = N r_{(0,0)}^{(0,1)}.$$ (146c)

The scaled version of Eq. (38) reads:

$$\pi(p) = \frac{S_{(0,0)}(S_{(0,0)} - S_{(0,1)})}{(S_{(0,0)} - S_{(0,1)})(1 - p) + S_{(0,0)}p},$$ (147)

where p is the relative frequency of type $(0, 1)$. The mean change in p over a time step measured in N generations is the same in both models. The variance in change of p is
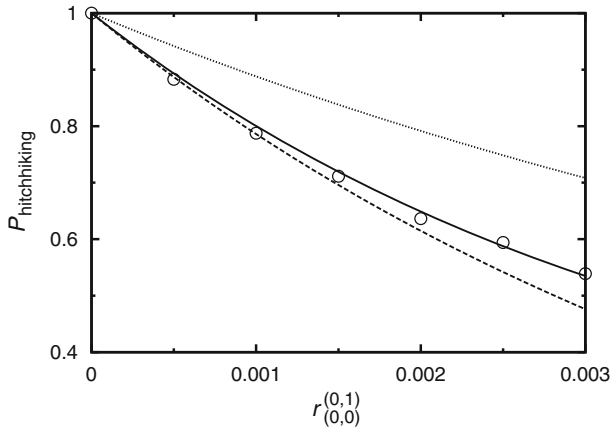
Fig. 11 The hitchhiking probability of a linked deleterious allele as a function of the recombination probability. The *solid line* is based on Eq. (53) as in Fig. 4. The *dashed line* shows $P_{(0,1)}^{(0,1)} P^*(1/N)$ where $P_{(0,1)}^{(0,1)}$ and P^* are given by Eq. (52) and the solution of Eq. (148), respectively. The *dotted line* represents $P^*(1/N)$. The parameter values are the same as for Fig. 4D. For $N = 10,000$ as in Fig. 4C, both results are basically indistinguishable

$V(p) = 2p(1 - p)$ in our model (vs $V(p) = p(1 - p)$ in the Wright-Fisher model). Following Hartfield and Otto (2011), we denote by $P^*(p)$ the conditional probability that type $(0, 1)$ fixes if its current frequency is p. We obtain:

$$\frac{\mathrm{d}^2 P^*(p)}{\mathrm{d}p^2} + \sigma_{(0,1)} \frac{1 + e^{-pS_{(0,1)}}}{1 - e^{-pS_{(0,1)}}} \frac{\mathrm{d}P^*(p)}{\mathrm{d}p} - \rho_{(0,0)}^{(0,1)} \pi(p) P^*(p) = 0. \tag{148}$$

The differential equation is solved with the boundary conditions

$$P^*(1) = 1, \tag{149a}$$

$$\left. \frac{\mathrm{d}P^*(p)}{\mathrm{d}p} \right|_{p=0} = 0. \tag{149b}$$

Figure 11 compares the results for the hitchhiking probability as obtained via Eq. (148) and via Eq. (53). Note that the diffusion equation is conditioned on establishment of type $(0, 1)$, assuming an absolute rate of increase of $\sigma_{(0,1)}$, while $\bar{\nu}$ is based on an absolute rate of increase of $\sigma_{(0,1)} - r_{(0,1)}$.

Appendix H: Accuracy of the approximations

Appendix H.1: The hitchhiking probability: the stochastic phase

Our approximation (43) is based on the assumption that only one "rescue individual" is generated. Here, we investigate the limits of this assumption. As an example, we

consider the case $I = 0$, $J = 2$ with $\sigma_{(0,2)} < 0$ in more detail. We denote by $\tilde{h}_1(s_0, s_1)$ the joint p.g.f. for the number of "successful" (Y_+) and "unsuccessful" (Y_-) recombination events until the extinction of type $(0, 2)$ conditioned on survival of the process. A successful recombination event means that the recombinant founds an infinite lineage. It holds:

$$
\begin{aligned}
\tilde{h}_1(s_0, s_1) &= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} P(Y_+ = k_1, Y_- = k_2 | \text{survival}) s_0^{k_1} s_1^{k_2} \\
&= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \frac{P(Y_+ = k_1, Y_- = k_2, \text{survival})}{1 - Q_{(0,2)}} s_0^{k_1} s_1^{k_2} \\
&= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \frac{P(Y_+ = k_1, Y_- = k_2, \text{survival})}{1 - Q_{(0,2)}} s_0^{k_1} s_1^{k_2} \\
&= \frac{h_1(s_0, s_1)}{1 - Q_{(0,2)}} - \sum_{k_2=0}^{\infty} \frac{P(Y_+ = 0, Y_- = k_2)}{1 - Q_{(0,2)}} s_1^{k_2}
\end{aligned}
\tag{150}
$$

with $h_1(s_0, s_1)$ given by Eq. (11) and Lemma 6. The probability that more than one successful recombination event takes place is given by

$$
\begin{aligned}
P(Y_+ > 1 | \text{survival}) &= 1 - P(Y_+ = 1 | \text{survival}) \\
&= 1 - \frac{\partial}{\partial s_1} \tilde{h}_1(s_1, s_2)|_{s_1=0, s_2=1} \\
&= 1 - \frac{P_{\text{success}}}{1 - Q_{(0,2)}} h'(1 - P_{\text{success}}).
\end{aligned}
\tag{151}
$$

The smaller $P(Y_+ > 1 | \text{survival})$, the better is the assumption fulfilled. Panels A and B of Fig. 12 show the probability $P(Y_+ > 1 | \text{survival})$ as a function of recombination and fitness of type $(0, 2)$, respectively. Note that $P(Y_+ > 1 | \text{survival})$ displays a maximum as a function of the recombination probability (not shown): For low recombination, an offspring is a recombinant type with low probability. For large recombination, the number of offspring which are themselves of type $(0, 2)$ gets significantly reduced by recombination, leading to a faster extinction of type $(0, 2)$; this in turn leads to few recombination events even if the recombination probability is high. Panels C and D of Fig. 12 show the hitchhiking probabilities corresponding to the scenarios of Panels A and B. The deviation between the theory and simulations is smaller than expected from $P(Y_+ > 1 | \text{survival})$. For $\sigma_{(0,2)}$ close to zero, type $(0, 2)$ can survive for a long time until it goes extinct and $P(Y_+ > 1 | \text{survival})$ is large. However, successful recombination events are rare. This implies that, within the multitype branching process, the time between the generation of the first and second rescue individuals can be large. Assume that a rescue individual of type $(0, 1)$ is generated first. Until the generation of the second rescue individual, type $(0, 1)$ has already reached such a high frequency that the multitype branching process description has become obsolete. I.e., even if the
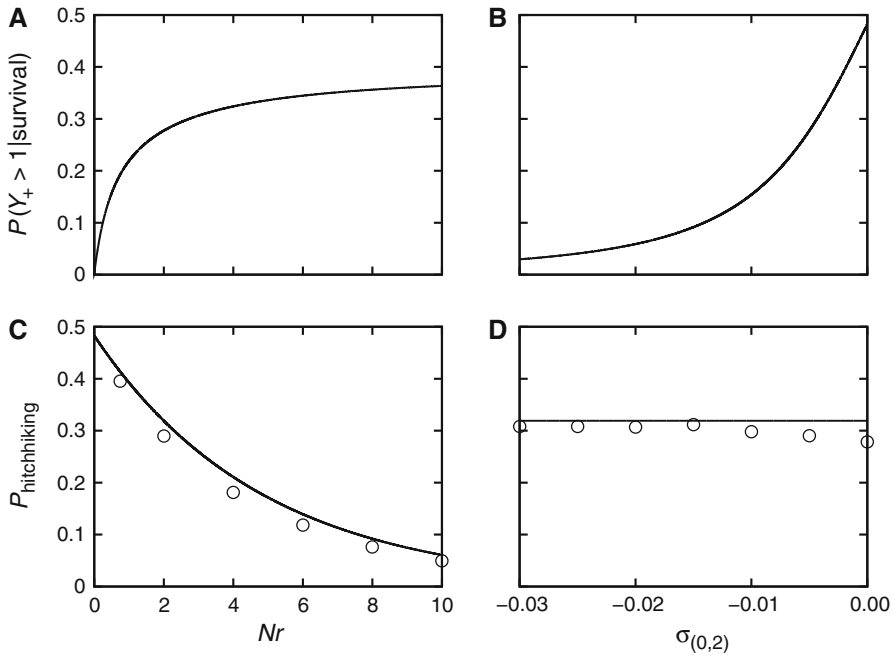
Fig. 12 Panels **a** and **b**: The probability that more than one successful recombination event takes place, $P(Y_+ > 1|\text{survival})$ [see Eq. (151)] as a function of r and $\sigma_{(0,2)}$, respectively. Panels **c** and **d**: The hitchhiking probability of the deleterious allele as a function of r and $\sigma_{(0,2)}$. Parameter values are: $I = 0$, $J = 2$, $\sigma_{(0,0)} = 0.075$, $\sigma_{(0,1)} = 0.07$, $N = 10{,}000$, $r_{(0,0)}^{(0,2)} = r_{(0,1)}^{(0,2)} = r$, $\sigma_{(0,2)} = -0.005$ (Panels **a** and **c**), $Nr = 2$ (Panels **b** and **d**). The *circles* denote simulation results. Each simulation point is the average of $2 \cdot 10^4$ successful introgression events

multitype branching process predicts the generation of a second rescue individual of type $(0, 0)$, the fixation of type $(0, 0)$ is not certain in the full model.

Appendix H.2: The hitchhiking probability: the deterministic phase

Our results for $P_{\text{det}}^{(1,1)\to(0,0)}$ and $P_{\text{det}}^{(1,1)\to(0,1)}$ rely on the assumption that no more than two successful recombination events happen until fixation of the adaptive allele (and additionally on an approximation of two integrals, see Eqs. (144) and (145). This assumption is only justified if recombination is weak. In particular, if recombination is strong, it is likely that a lineage of type $(0, 1)$ individuals and a lineage of type $(1, 0)$ individuals segregate simultaneously in the population and recombine to found a successful lineage of type $(0, 0)$ individuals. In this case, our approximation will strongly overestimate the fixation probability of type $(0, 1)$ (assuming that $\sigma_{(0,1)} > \sigma_{(1,0)}$) and underestimate the fixation probability of type $(0, 0)$. This can be clearly seen in Fig. 13. Note that for increasing recombination, the fixation probability of type $(0, 1)$ attains a maximum and decreases then in favor of fixation of type $(0, 0)$ due to the described mechanism (not shown). This behavior is not covered by our theory. The prediction for fixation of type $(1, 0)$ or $(1, 1)$ remains highly accurate though.
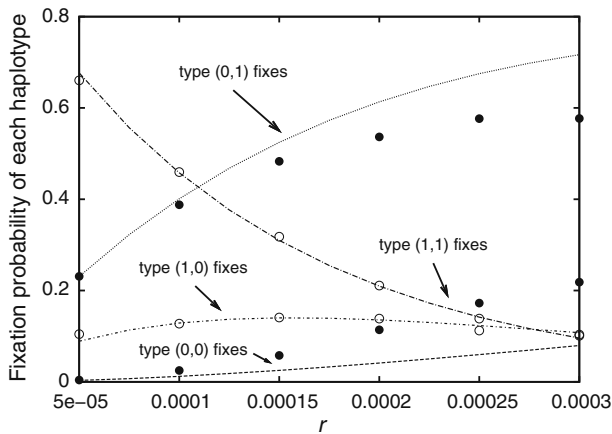
Fig. 13 The hitchhiking probability of linked alleles as a function of recombination. Parameter values are: $I = J = 1$, $\sigma_{(0,0)} = 0.075$, $\sigma_{(0,1)} = 0.07$, $\sigma_{(1,0)} = 0.06$, $N = 10,000$, $r_{(0,1)}^{(1,1)} = r_{(1,0)}^{(1,1)}$. The *circles* denote simulation results. Each simulation point is the average of 2,000 successful introgression events

Appendix H.3: The hitchhiking probability: concatenation of the stochastic and the deterministic phase

Our theory is based on a clear separation of the process into two phases. In some cases, however, the two phases blend together.

In Fig. 5, Panel d, deviations between theory and simulations arise for increasing recombination. This has two reasons: First, the approximation $P_{(0,2)}^{(0,2)} \approx 1$ becomes worse as recombination increases. Second (and more importantly here), it is not unlikely that type $(0, 1)$ and type $(0, 2)$ establish simultaneously while our approximation assumes that type $(0, 1)$ establishes after type $(0, 2)$. A slight shift in the time of establishment as introduced by our approach can lead to appreciable deviations of the modeled frequency paths of both types from their true paths. This, in turn, has a non-negligible impact on the assessment of the hitchhiking probability.

References

Abi-Rached L et al (2011) The shaping of modern human immune systems by multiregional admixture with archaic humans. Science 334(89):89–94

Adams JR, Leonard JA, Waits LP (2003) Widespread occurence of a domestic dog mitochondrial DNA haplotype in southeastern US coyotes. Mol Ecol 12(2):541–546

Antia R, Regoes RR, Koella JC, Bergstrom CT (2003) The role of evolution in the emergence of infectious diseases. Nature 426:658–661

Arnold MJ (2004) Transfer and origin of adaptations through natural hybridization: Were Anderson and Stebbins right? Plant Cell 16:562–570

Arnold ML, Martin NH (2009) Adaptation by introgression. J Biol 8:82

Arnold ML, Bulger MR, Burke JM, Hempel AL, Williams JH (1999) Natural hybridization: How low can you go and still be important? Ecology 80(2):371–381

Athreya KB, Ney PE (1972) Branch Processes. Springer, Berlin

Barton NH (1995) Linkage and the limits to natural selection. Genetics 140:821–841

Barton NH (1979) Gene flow past a cline. Heredity 43(3):333–339

Barton N, Bengtsson BO (1986) The barrier to genetic exchange between hybridising populations. Heredity 56:357–376

Beaumont M et al (2001) Genetic diversity and introgression in the Scottish wildcat. Mol Ecol 10:319–336

Bengtsson BO (1985) The flow of genes through a genetic barrier. In: Greenwood JJ, Harvey PH, Slatkin M (eds) Evolution essays in honour of John Maynard Smith. Cambridge University Press, Cambridge, pp 31–42

Demon I, Haccou P, van den Bosch F (2007) Introgression of resistance genes between populations: a model study of insecticide resistance in *Bemisia tabaci*. Theor Popul Biol 72:292–304

Desai MM, Fisher DS (2007) Beneficial mutation-selection balance and the effect of linkage on positive selection. Genetics 176:1759–1798

Ellstrand NC, Prentice HC, Hancock JF (1999) Gene flow and introgression from domesticated plants into their wild relatives. Annu Rev Ecol Sytematics 30:539–563

Ellstrand NC et al (2013) Introgression of crop alleles into wild or weedy populations. Annu Rev Ecol Evol Syst 44:325–345

Fitzpatrick BM et al (2010) Rapid spread of invasive genes into a threatend native species. Proc Natl Acad Sci 107(8):3606–3610

Galassi M, Davies J, Theiler J, Gough B, Jungman G, Alken P, Booth M, Rossi F (2009) GNU Scientific Library Reference Manual, 3rd edn. Network Theory Ltd., Bristol

Gosh A, Haccou P (2010) Quantifying stochastic introgression processes with hazard rates. Theor Popul Biol 77:171–180

Gosh A, Meirmans PG, Haccou P (2012a) Quantifying introgression risk with realistic population genetics. Proc R Soc B 279:4747–4754

Gosh A, Serra MC, Haccou P (2012b) Quantifying time-inhomogeneous stochastic introgression processes with hazard rates. Theor Popul Biol 81:253–263

Gottelli D et al (1994) Molecular genetics of the most endangered canid: the ethopian wolf Canis simensis. Mol Ecol 3(4):301–312

Hartfield M, Glémin S (2014) Hitchhiking of deleterious alleles and the cost of adaptation in partially selfing species. Genetics p. doi:10.1534/genetics.113.158196

Hartfield M, Otto SP (2011) Recombination and hitchhiking of deleterious alleles. Evolution 65(9):2421–2434

Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing genetic variation as sources of adaptive variation. Mol Ecol 22:4606–4618

Heiser CB Jr (1951) Hybridization in the annual sunflowers: *Helianthus annuus × H. debilis* Var. cucumerifolius. Evolution 5(1):42–51

Iwasa Y, Michor F, Nowak MA (2003) Evolutionary dynamics of escape from biomedical intervention. Proc R Soc B 270:2573–2578

Iwasa Y, Michor F, Nowak MA (2004a) Evolutionary dynamics of invasion and escape. J Theor Biol 226:205–214

Iwasa Y, Michor F, Nowak MA (2004b) Stochastic tunnels in evolutionary dynamics. Genetics 166(3):1571–1579

Lindner CR, Taha I, Seiler GJ, Snow AA, Rieseberg LH (1998) Long-term introgression of crop genes into wild sunflower populations. Theor App Genet 96:339–347

Mallet J (2005) Hybridization as an invasion of the genome. Trends Ecol Evol 20(5):229–237

Miller W, et al. (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. Proc Natl Acad Sci E2382–E2390

Proulx SR (2011) The rate of multi-step evolution in moran and wright-fisher populations. Theor Popul Biol 80:197–207

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. Annu Rev Ecol Systematics 27:83–109

Schwenk K, Brede N, Streit B (2008) Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. Philos Trans R Soc B 363:2805–2811

Serra MC (2006) On the waiting time to escape. J Appl Prob 43:296–302

Serra MC, Haccou P (2007) Dynamics of escape mutations. Theor Popul Biol 72:167–178

Sewastjanow BA (1974) Verzweigungsprozesse. Akademie-Verlag, Berlin

Snow AA (2002) Transgenic crops—why gene flow matters. Nat Biotechnol 20:542

Snow AA et al (2003) A Bt transgene reduces herbivory and enhances fecundity in wild sunflowers. Ecol Appl 13(2):279–286

Stewart CN, Halfhill MD, Warwick SI (2003) Genetic modification: transgene introgression from genetically modified crops to their wild relatives. Nat Rev Genet 4:806–817

The Heliconius Genome Consortium (2012) Buttlerfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487:95–98

Uecker H, Hermisson J (2011) On the fixation process of a beneficial mutation in a variable environment. Genetics 188:915–930

Weissman DB, Desai MM, Fisher DS, Fisher MW (2009) The rate at which asexual populations cross fitness valleys. Theor Popul Biol 75:286–300

Whitney KD, Randell RA, Rieseberg LH (2006) Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. Am Nat 167(6):794–807

Yanchukov A, Proulx SR (2014) Migration-selection balance at multiple loci and selection on dominance and recombination. PLoS One 9(2):e88–651