Evolutionary Ecology Research, 2010, **12**: 1–22

Recombination and epistasis facilitate introgressive hybridization across reproductively isolated populations: a gamete-based simulation

Yoshinari Tanaka

Center for Environmental Risk, National Institute for Environmental Studies, Tsukuba, Japan

ABSTRACT

Questions: Why can invasive species sometimes genetically contaminate closely related indigenous species by introgressive hybridization, resisting the post-zygotic isolating mechanism? How do recombination rates and epistasis among incompatibility genes, and the number of loci affect the introgression?

Features of models: The individual based-model and gamete-based model, which tracks changes in the number of invasive genes per gamete due to selection and recombination by assuming random arrangement of genes within gamete.

Range of key variables: The recombination rate between adjacent loci ranges from 0 to 0.4. The epistatic effect between loci is measured by the exponent of the geometric function of heterozygosities representing individual fitness. It ranges from 1 (additive) to 4 (strong epistasis). The number of loci is set to 2–10 for the gamete-based model.

Conclusions: Provided that the number of loci is not very small and the fitness of the F1 hybrid is not extremely low, complete genetic replacement by introgressive hybridization is accelerated by an increase in rates of total recombination across all loci and by the epistatic fitness effect among incompatibility loci.

Keywords: Dobzhansky-Mueller model, epistasis, gamete-based model, introgressive hybridization, multi-locus underdominance, post-zygotic isolation.

INTRODUCTION

Introgressive hybridization by anthropogenic transference of exotic species brings about genetic contamination of endemic species, which, in some extreme cases, leads to genetic extinction of the endemic population by complete genetic assimilation or replacement by the invasive population (Levin *et al.*, 1995; Rieseberg and Gerber, 1995; Rhymer and Simberloff, 1996; Epifanio and Nielsen, 2000; Mallet, 2005).

Apart from ecological factors such as predation and competition by biotic invasion, hybridization of endemic species with closely related invading species or subspecies poses

Correspondence: Y. Tanaka, Center for Environmental Risk, National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, Ibaraki 305-8506, Japan. e-mail: ytanaka@nies.go.jp

two types of hazard to the endemic species. One is hybrid sterility, which inhibits normal reproduction, and the other is introgressive hybridization, which does not completely retard mating and reproduction by hybridization but instead allows exotic species or local varieties to genetically admix with, and contaminate, endemic species. From the perspective of evolutionary biology, the process of introgressive hybridization can be interpreted as breakdown of the reproductive isolation mechanisms between species (Tanaka, 2007), or breakdown of the local adaptive gene complex among intraspecific local varieties or between an artificial breed or genetically modified organism and its wild relatives (Hails and Morley, 2005; Mallet, 2005).

In this paper, I focus on the mechanism of post-zygotic isolation as a reproductive isolation mechanism responsible for whether or not introgressive hybridization is prevented, because pre-zygotic isolation, which needs secondary contacts or parapatry, is often incomplete for the species pairs that result from anthropogenic transplantation. In contrast, the post-zygotic isolating mechanism is likely to function as the final check against introgressive hybridization, because post-zygotic isolation is formed by a stochastic process and/or local adaptation, both of which can proceed in allopatry (Orr, 1995, 2001; Gavrilets, 1997a; Orr and Turelli, 2001).

Many case studies have reported a substantial risk of genetic extinction (e.g. complete loss of the purely endemic genome) by introgressive hybridization (Rhymer *et al.*, 1994; Levin *et al.*, 1995; Allendorf *et al.*, 2001). Since the genetic process of hybridization is often cryptic and not well understood, mathematical modelling approaches that are based on population genetics are needed to present a risk assessment framework based on the prediction of introgressive hybridization from measurable genetic parameters (Rhymer and Simberloff, 1996; Huxel, 1999; Hails, 2000; Colbach *et al.*, 2001; Epifanio and Philipp, 2001; Hails and Morley, 2005; Mallet, 2005; Tanaka, 2007).

Very few theoretical studies, however, have investigated the process of genetic assimilation by introgressive hybridization on the basis of population genetics or the adaptive dynamics theory in relation to reproductive isolation mechanisms (Ferdy and Austerlitz, 2002). There are several unknown aspects in the introgressive hybridization process that should be explored with population genetic models. There may be a moderate number of genes controlling reproductive isolation, which interact to generate hybrid incompatibility (Orr, 1989, 2001; Gavrilets, 1997; Gadau *et al.*, 1999; Coyne and Orr, 2004). Recombination rates are likely to influence the likelihood of introgressive hybridization occurring, because genetic mixing between partly isolated genomes must mitigate the fitness reduction of hybrid individuals and promote the introgression of exotic genes.

In response to repeated immigration and introgression by exotic individuals, the genotypic distribution of the focal population is likely to deviate from normality. The approximate hyper-geometric model can successfully simulate the dynamics of non-Gaussian genotypic distributions by disruptive selection and assortative mating (Barton, 1992; Doebeli, 1997; Shpak and Kondrashov, 1999; Barton and Shpak, 2000a, 2000b). Nonetheless, these regimes of selection and mating are not met by post-zygotic isolating mechanisms, and linkage (gamete-phase) disequilibrium caused by limited recombination and epistasis is neglected in this modelling framework by the assumption that all genotypes that comprise the same genotypic values are equiprobable.

As an alternative, I present a gamete-based simulation model that describes the genetic structure of populations by gamete-type frequencies and thus simulates the process of introgressive hybridization in a multilocus genetic system. Assuming random configuration of particular alleles in a multi-locus system within gametes, the gamete-based model

presented here – the random gamete model – aggregates all gamete types into a much smaller number of gamete classes according to the number of a specific (e.g. invasive) allele. Relevance of the random gamete assumption was tested with an individual-based simulation.

From a risk estimation perspective, we require information on quantitative relationships between measurable genetic parameters and the genetic risk of hybridization (Kanaiwa and Harada, 2000; Epifanio and Philipp, 2001; Thompson *et al.*, 2003; Hails and Morley, 2005; Mallet, 2005; Tanaka, 2007). Here I examine the effects of numbers of loci, recombination rates, strength of epistasis, and selection coefficients against incompatible or maladaptive gene combinations in terms of enhancing or reducing hybridization.

SIMULATION MODELS

To simulate the process of introgressive hybridization, two alternative models were employed: the random genetic model and the individual-based model. The random gamete model is a deterministic or an infinite population size model free from stochasticity, although based on a restrictive assumption that gametes have an equally probable configuration of alleles with a particular numerical composition of alleles. To examine the relevance of the random gamete assumption, I conducted individual-based simulations, which were free from the random gamete assumption.

Random gamete model

In the process of introgressive hybridization, linkage disequilibrium may be important in affecting hybridization, because linkage equilibrium is recovered only by recombination, and it influences the efficacy of selection purging incompatible sets of exotic genes. As a modelling framework that ensures manipulation of recombination to examine the effect of recombination on introgressive hybridization, the gamete-based model characterizes the genetic states of a population by gamete type frequencies instead of gene or genotype frequencies. However, a simulation model based on gamete type frequencies is likely to be very time consuming with standard computational power, because even with a relatively small number of loci the number of possible gamete types becomes enormous (for example, 4096 types with 10 loci with two alleles).

To balance the computational load and biological relevance of the model, I make two simplifying assumptions. One is that all loci have the same pair of alleles – the endemic allele a_0 and the exotic allele a_1 – and the endemic allele has a value of 0, whereas the exotic allele has a value of 1. Thus, the possible genotypic space is restricted from 0 to $2n$, where n is the number of loci. The gamete types possessing the same number of exotic alleles were further aggregated to a gamete class, which refers to a group of gametes that possess the same number of exotic alleles.

The second assumption is that the possible gamete types within a specific gamete class have equal probabilities; this is the random gamete assumption. In the initial phase of introgression to a purely endemic population, exotic genes in the hybridized individuals are tightly linked and must not be randomly allocated within gametes. Nonetheless, this gametic-phase disequilibrium tends to diminish with recombination across generations, and the initial phase of introgression is only a trivial part of the total process of genetic invasion by hybridization. For most of the entire process, the population may be almost balanced in

terms of introgression and recombination. Thus, the gamete types within each gamete class are approximately random if the rate of introgression per generation is small enough.

The trans-generational dynamics of gamete class frequencies can be determined if the frequency distribution of gamete classes x, $P(x|g_m, g_p)$ (in which g_m and g_p are the maternal and paternal gamete classes, respectively), produced by an individual is known from the gamete classes that were transmitted from the individual's parents and constituted the zygote of that individual.

By crossing over at meiosis before gamete formation, the parental gametes exchange a portion of their chromosomes. The recombination rate between adjacent loci is assumed to be uniformly r, and double chiasmas are disregarded for simplicity. Hence the total recombination rate, R, is simply $(n-1)r$. The present analysis postulates a single pair of chromosomes whereas extension to multiple chromosomes is relatively straightforward.

The probability distribution of gamete classes among gametes produced by an individual with parental gamete classes g_m and g_p is decomposed into two portions, non-recombinants and recombinants, as follows:

$$P(x|g_m, g_p) = \tfrac{1}{2}(1-R)\left\{d(x-g_m) + d(x-g_p)\right\}$$

$$+ \frac{R}{2(n-1)} \sum_c \left\{ \sum_i p(i, c, g_m)\, p(i+x-g_m, c, g_p) + \sum_j p(j, c, g_p)\, p(j+x-g_p, c, g_m) \right\}, \quad (1)$$

where $p(i, c, g)$ is the probability that the number of exotic alleles located in the upper region from the crossing-over point c in the ordered gene map of the chromosome is i, provided that the gamete class is g, and $d(z)$ is a function with a value of 1 when $z = 0$, or otherwise 0. The two summation terms in the large brackets of the second term represent the probability that a maternal or a paternal gamete, respectively, receives, by recombination with the other gamete, the specific number of exotic alleles so that the total number of exotic genes per gamete becomes x.

Based on the random gamete assumption, the above-mentioned probability is specified as

$$p(i, c, g) = \binom{n-c}{i}\binom{c}{g-i} \Big/ \binom{n}{g}, \quad (2)$$

in which i is bounded by $\max[g-c, 0] \leq i \leq \min[n-c, g]$. This is simply equivalent to the number of combinations that i distinct elements (the exotic allele) occupy among $n-c$ positions (the upper region) and $g-i$ elements occupy among c positions (the lower region), divided by the total number of combinations that g elements occupy among n positions (all loci).

Mating and selection

Specifying the mating-type frequencies in addition to the gamete class frequency distribution by specific mating types determines the frequency distribution of genotypic values among individuals in the next generation. For simplicity, mating is assumed to be random, resulting in no correlation between parental gametes within individuals and the Hardy-Weinberg equilibria.

Since the present analysis postulates the existence of post-zygotic isolation barriers, fertility or viability selection operates on individuals produced by the zygote formation, and its differential intensity is based on differences in genotypes.

As a selection scheme we employ two alternative fitness functions: the multi-locus underdominance model and the local adaptation model. The multi-locus underdominance model, the main focus of this study, postulates the existence of intrinsic factors – post-zygotic isolating barriers and epistatic fitness effect across loci, such as the Dobzhansky-Mueller incompatibilities – although the model does not follow exactly the Dobzhansky-Mueller theory. Genotypic fitness is defined in the logarithmic scale as a geometrically decreasing function of the mean heterozygosity across loci, $\bar{H}$, as follows:

$$W(g_m, g_p) = \exp\left[-c\,\{\bar{H}\,(g_m, g_p)\}^k\right], \tag{3}$$

where c and k are parameters representing the magnitude of deleterious effects and the epistatic interaction between heterozygous loci, respectively. The fitness landscape of hybrid genotypes depicted on the plane of parental gamete classes indicates that the epistatic interaction between loci substantially influences the shape of the landscape (Fig. 1). In the case of a weak epistatic interaction ($k = 1.5$ in Fig. 1b), the valley on the fitness surface is deeply concave, making shifts between the alternative fitness peaks impossible. A stronger interaction ($k = 4$ in Fig. 1a), however, makes the valley shallower and peak shifts possible if intermediate gamete classes are provided by recombination.

The F1 hybrid, which is by definition the whole heterozygote, has a fitness uniquely determined by c, $W = e^{-c}$. A k value of unity denotes additive effects of the heterozygous loci in reducing fitness, whereas larger k values mean stronger synergistic interaction between loci. The curvilinear fitness function of the number of genes has been successfully utilized to express the synergistic effects of deleterious mutations and epistatic selection on the incompatibility gene complexes (Kondrashov, 1982, 1985; Barton and Bengtsson, 1986; Barton and Shpak, 2000a; Gavrilets, 1997b). The multi-locus underdominance model parallels these deleterious mutation models if the number of heterozygous loci is interpreted as the number of deleterious combinations of genes. The reason why the mean heterozygosity across loci is used in equation (3) instead of number of heterozygous loci is that the analysis attempts to examine the net effects of the number of loci separated from the accumulation effects of multiple loci by normalizing the number of heterozygous loci with the total number of loci.

The random gamete model is not relevant to calculations of the mean heterozygosity of individuals, because the model does not specify the genotypes of individuals but instead employs an approximate description of genotypes by parental gamete classes. As an approximate derivation of fitness, an expected heterozygosity conditional to parental gamete classes is calculated as follows and assigned equally to all individuals that have the same parental gamete classes (the maternity and paternity are exchangeable):

$$\bar{H}(g_m, g_p) = 1 - n^{-1} \sum_{h=\max[g_m+g_p-n,0]}^{\min[g_m, g_p]} \left[\{n - (g_m + g_p) + 2h\}\binom{g_m}{h}\binom{n-g_m}{g_p-h}\Big/\binom{n}{g}\right]. \tag{4}$$

The number of combinations h among g_m exotic alleles coupled with other exotic alleles (making homozygous loci) and of residual $g_m - h$ exotic alleles coupled with endemic alleles (making heterozygous loci; in other words, the residual $g_p - h$ paternal exotic alleles are all coupled with a subset of $n - g_m$ maternal endemic alleles) is

$$\binom{g_m}{h}\binom{n-g_m}{g_p-h}$$

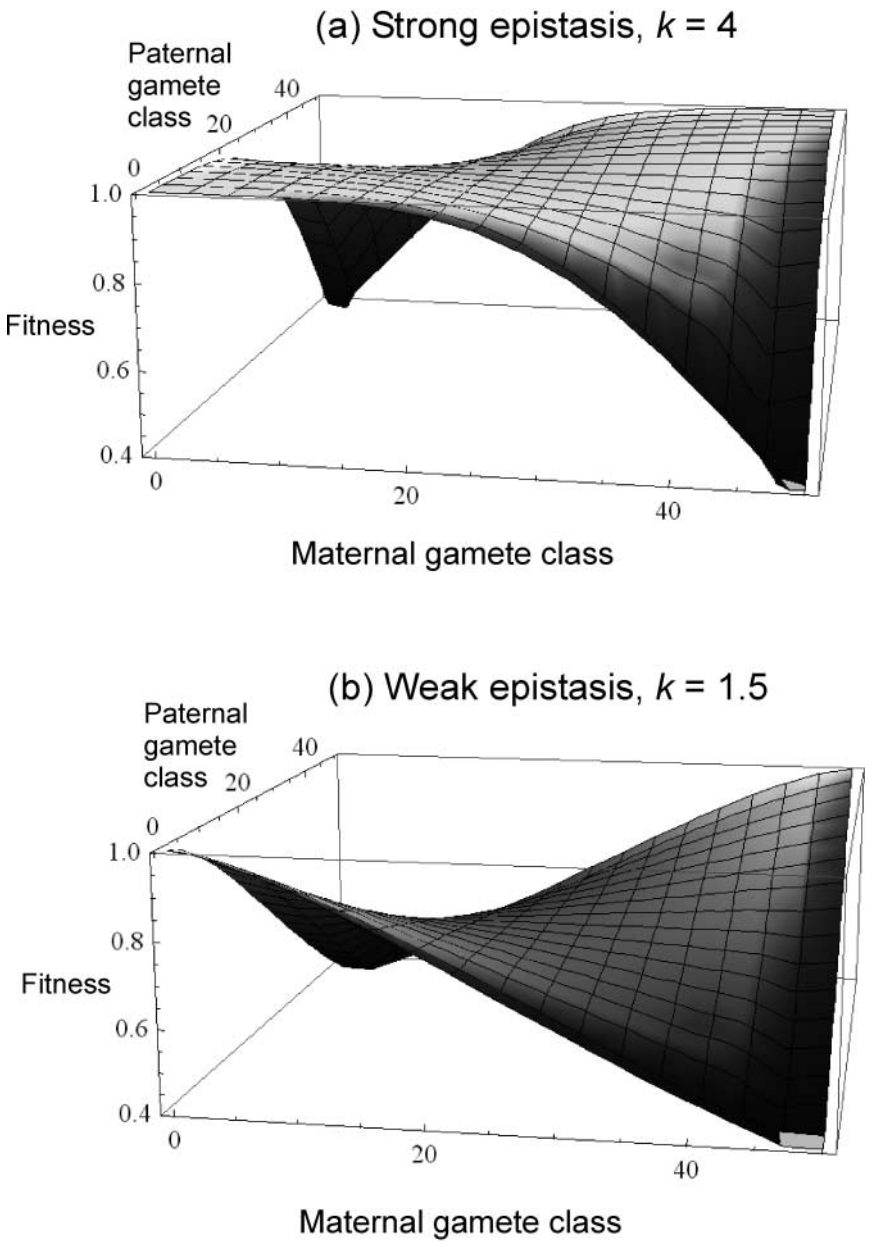


Fig. 1. Fitness landscape of the multi-locus underdominance fitness model for the cases of strong epistasis (a) $k = 4$ and weak epistasis (b) $k = 1.5$. The altitudinal axis denotes the mean offspring fitness whose parental gamete classes are specified by the vertical and the horizontal axes on the base, following equations (3) and (4) in the text. With strong epistasis, there is a peak pathway from the endemic state (the upper left corner) to the exotic state (the upper right corner), which is due to recombination. With weak epistasis, however, a deep valley separates the alternative fitness peaks. The other parameter values are $n = 50$ and $c = 1$.

for $n - g_m \geq g_p - h$. Thus, the probability that a pair of gametes of classes g_m and g_p will produce a zygote in which the number of homozygous loci for the exotic allele is h is

$$\binom{g_m}{h}\binom{n-g_m}{g_p-h}\Big/\binom{n}{g_p}.$$

Because the number of homozygous loci for the endemic allele is $(n - g_m) - (g_p - h)$ if the number of loci homozygous for the exotic allele is h, the expected heterozygosity is given by equation (4).

The local adaptation model is employed as a reference additive fitness model to allow a contrast with the peculiarities of the multi-locus underdominance model. It postulates that the environments inhabited by the exotic population are noticeably different from those of the endemic population and a quantitative trait is responsible for the local adaptation. Assuming genetic effects within and between loci to be additive and the fitness profile of the quantitative trait to depict a Gaussian curve, the fitness function of individuals that have specific parental gamete classes and are living in the invaded habitat can be expressed by the following Gaussian function:

$$W(g_m, g_p) = \exp\left[-\frac{\{(g_m + g_p)/n\}^2}{V_s}\right], \tag{5}$$

where V_s is a measure of stabilizing selection and also denotes the strength of directional selection, which is negatively associated with the values of V_s. The total of the genotypic values, $g_m + g_p$, is standardized by the number of loci n so that the total phenotypic space and individual fitness do not depend on the number of loci. The fitness decreases monotonically with the number of exotic genes from the optimum fitness for the endemic genotype, $g_m + g_p = 0$. Thus, the parameter V_s only determines the strength of natural selection on the trait and the fitness of the F1 hybrid, $W = e^{-1/V_s}$.

Simulation regime

The main purpose of simulations is to examine how recombination rate, number of loci, alternative fitness functions, and the epistatic fitness effect influence the process of hybridization. I assumed two alternative scenarios of invasion: constant invasion, in which immigration rate was constant for the whole period of the simulations, and pulse invasion (see below for explanation).

The dynamics of the gamete class frequencies at time t, which is denoted by $f_g(x, t)$, was solved numerically with iterations that involve all the processes described above. Just before reproduction, the endemic population receives immigrants from the exotic population every generation. For the constant invasion scenario, the immigration rate m per generation, defined as the proportion of new immigrants occupying the endemic population, was set at 5% for all simulations. By random mating, the genotypic distribution in the offspring generation was created by random assortment from the gamete pool, which was a mixture of gametes produced by individuals of the endemic population and by new invaders. Based on the assumption of infinite population size, the assortment is deterministic and without the perturbation that would emerge from random sampling. Thus, the frequencies of zygote types, which are characterized by parental gamete classes and denoted as $F_z(g_m, g_p, t)$, are the products of the parental gamete class frequencies, $F_z(g_m, g_p, t) = f_g(g_m, t)\ f_g(g_p, t)$.

Individuals that develop from particular zygotes are checked by selection and reproduce according to the fitness function described in the previous section. After selection within the generation, the zygote type frequencies become

$$F^*_z(g_m, g_p, t) = F_z(g_m, g_p, t)W(g_m, g_p)C^{-1}, \quad (6)$$

where C is the normalization constant, i.e. $C = \Sigma_{g_m}\Sigma_{g_p} F_z(g_m, g_p, t)W(g_m, g_p)$.

At reproduction, individuals, of which the genotypic frequencies follow $F^*_z(g_m, g_p, t)$, produce gametes by meiosis (which is subject to the recombination rule given in the section 'Random gamete model' above), generating the gamete pool for the next generation.

To monitor the extent of introgression of exotic genes to the endemic population, I evaluated the following introgression index (equivalent to the mean gene frequencies of the exotic allele across all loci) after immigration and reproduction but before selection,

$$I_g = \sum_{g=0}^{n} \left\{ \left(\frac{g}{n} \right) p(g) \right\},$$

where $p(g)$ is the frequency distribution of gametes of class g in the gamete pool.

The pulse invasion scenario assumed that immigration continued for a short period only at various rates that could be considerably higher than 5%, and afterwards it ceased. To contrast the relative importance between the immigration rates m and the duration of immigration T, a trade-off was introduced between them (i.e. T by m being equivalent to unity). This part of the assumption is intended to demonstrate the relative importance of instantaneous rate of immigration and duration of invasion for particular total numbers of immigrants.

Individual-based model

An individual-based model was designed to follow the above-mentioned simulation regimes as closely as possible. The genetic composition was described by a gamete matrix **G** with $N \times 2n$ dimensions (N = the number of individuals and n = the number of loci), in which the ith row ($1 \le i \le N$) denotes genotypes of the ith individual, and the left n columns and the right n columns respectively denote the maternal and the paternal ordered gamete types of the individual (0 = the endemic allele and 1 = the exotic allele).

Individual genotypic fitness was assigned to all individuals according to equation (3). Viability selection was operated by stochastically deciding survival or death for each individual with its fitness as used for the probability density of survival.

From those who survived, $2N$ individuals were randomly chosen with replacement to minimize stochasticity as N pairs of dams and sires (N was assumed 500 for all runs). Migration events were included in the above-mentioned sampling process by choosing exotic individuals with probability m. Meiosis occurred stochastically within individuals before gamete formation, with equiprobable recombination rates $(n-1)/R$ among the $n-1$ potential crossing over points. Coupling one of the two gametes from a dam with one of the two gametes from a sire produces an individual in the next generation.

RESULTS

The random gamete assumption is likely hard to be met for the initial phase of introgression, because considerable gamete-phase disequilibria are inevitable with a limited number of recombination events after the start of introgression. However, the comparison in gamete class frequencies between results from the random gamete simulations and the individual-based simulations indicated good accordance in the initial phase of introgression (generations 5 and 20) regardless of the total recombination rate, 0.4 or 0.2 (Fig. 1). At later generations, the two simulations exhibited poorer accordance especially for the case of high recombination rate (Fig. 1a), partly because higher uncertainties were brought about from several sources of stochasticity (i.e. migration, recombination, selection, and recruitment) with the individual-based simulation. Comparison between the individual-based simulation with a higher recombination rate and that with a lower recombination rate (between Fig. 1a and Fig. 1b) strongly indicated that the stochasticity in recombination events greatly affected the consequences of introgressive hybridization.

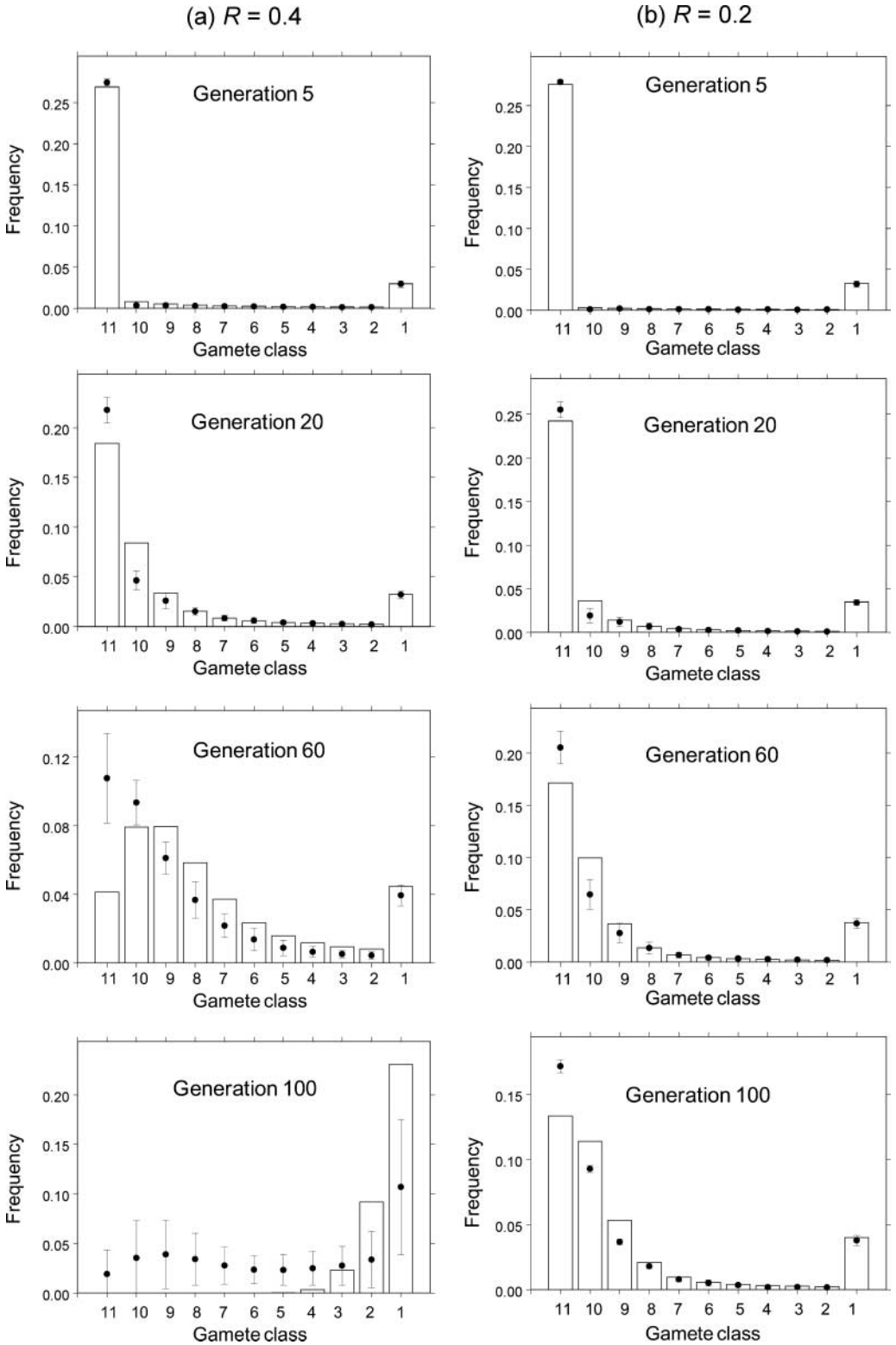
The quasi-asymptotic state, represented by introgression indices at generation 200, also showed good correspondence between the two types of simulation regardless of the number of loci, although the smaller number of loci resulted in larger uncertainties. Large discrepancies were observed only for extreme values of process rate and the dependence of quasi-asymptotic states to parameter values. Hence I conclude that the random gamete assumption is a good approximation for gamete type frequencies within gamete classes in the process of introgressive hybridization (Fig. 2).

Epistasis and strength of selection against hybrids

The magnitude of incompatibility c, which uniquely determines the strength of selection against hybrids, predominantly decided whether or not introgressive hybridization was successful (Fig. 3). Strong incompatibilities such that c values were larger than 1.5, corresponding to a 78% reduction in fitness of the F1 hybrid, did not allow introgression regardless of epistasis (k values). In contrast, very weak incompatibilities ($c = 0.5$; 40% reduction in fitness of the F1 hybrid) resulted in complete genetic replacement by introgression regardless of epistatic effects among loci.

Epistasis, a non-linear association between the number of heterozygous loci and fitness, had a marked effect on introgressive hybridization (Fig. 3). With an intermediate magnitude of incompatibility, epistatic interactions among loci noticeably affected introgressive hybridization. Stronger epistasis facilitated introgression of exotic genes in cases where incompatibility was moderately strong (65–75% reduction in fitness).

The local adaptation model, which relies on additive fitness, does not include, by definition, epistatic interactions among loci. Qualitatively the same results regarding strength of selection were obtained with the local adaptation model; strong selection against hybrids tends to prevent complete genetic replacement by introgressive hybridization (Fig. 4). However, there is a noticeable difference between the two models regarding the extent to which selection prevents introgressive hybridization. Complete genetic replacement can occur with much stronger selection against hybrids in the multi-locus underdominance model than in the local adaptation model if epistasis is large enough (compare Fig. 4 with Fig. 3). With the multi-locus underdominance model, genetic replacement is feasible even when the fitness of F1 hybrids is as low as 37% ($c = 1$), whereas


(a) R = 0.4
(b) R = 0.2
Generation 5
Generation 20
Generation 60
Generation 100
Frequency
Gamete class
11 10 9 8 7 6 5 4 3 2 1

with the local adaptation model genetic replacement is not feasible until the fitness of F1 hybrids is higher than 80%.

Recombination rates and the number of loci

The effects of recombination rates and number of loci were sharply contrasted between the two fitness models. With the multi-locus underdominance model, the recombination rate largely affected the consequences of introgression as long as the number of loci was 4 or more (Figs. 5 and 6). Figure 5 illustrates typical processes of introgression for three different total recombination rates ([a] $R = 0.05$, [b] $R = 0.2$, and [c] $R = 0.4$). High recombination rates facilitated introgressive hybridization, leading to complete replacement of genes.

More systematic calculations are given in Fig. 6. The introgression index gave a sharp, upwardly concave curve against total recombination rates, indicating the strong dependence of genetic replacement by introgressive hybridization upon recombination and the implied existence of a quasi-threshold value for the recombination rate, beyond which it allowed complete replacement of genes. In contrast, the number of loci, except when very small, had a negligible influence on introgressive hybridization.

With the local adaptation model, however, both number of loci and recombination rate appeared to have quasi-linear effects on the likelihood of genetic replacement by introgressive hybridization (Fig. 7).

Pulse invasion

The two alternative invasion scenarios – high-rate and short-term invasion versus low-rate and long-term invasion – led to contrasting results (Fig. 8). High-rate and short-term invasion succeeded in complete replacement of genes, whereas low-rate and long-term invasion did not. The asymptotic introgression index did not gradually respond to the immigration rate (or the invasion duration) but instead changed suddenly as these parameters changed (Fig. 9). Since a strict trade-off between invasion duration and immigration rate was assumed ($T_m \equiv 1$), this result strongly suggests that (maximum) immigration rate is much more important than the total number of immigrants for a long period of invasion if the multi-locus underdominance model is correct. With a lower recombination rate, higher maximum immigration rates are needed to complete the introgression.

With the local adaptation model, long-term introgression and genetic replacement do not occur as long as the invasion continues for a limited time (data not shown), since the asymmetry of the fitness landscape is persistent and the exotic genes are eventually excluded by selection.

Fig. 2. Simulated frequency distribution of gamete classes by the random gamete simulation (bars) and the individual-based simulations (solid circles) for four intermittent generations after the start of introgression. The total recombination rate is set as (a) $R = 0.4$ and (b) $R = 0.2$. The number of replicate runs is 50 in (a) and 20 in (b). The error bars from the solid circles denote standard deviations. The fitness function is the multi-locus underdominance model for both simulations. Other parameter values for both simulations are $n = 10$, $c = 1$, $k = 2$, and $m = 0.05$. The effective number of population N is 500 for all individual-based simulations.

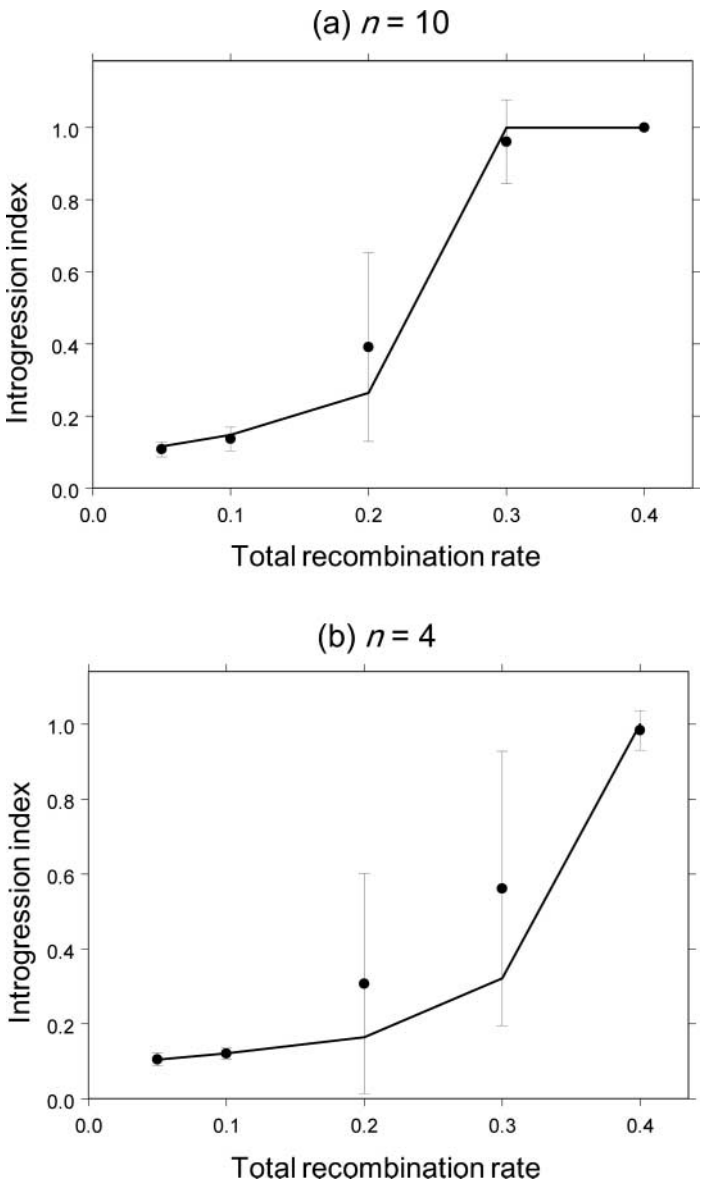


Fig. 3. Asymptotic introgression index, with the random gamete simulation (solid line) and the individual-based simulations (solid circles), plotted against total recombination rates for (a) $n = 10$ and (b) $n = 4$. The number of replicate runs is 20, and the effective number of population N is 500 for all individual-based simulations. The error bars from the solid circles denote standard deviations. The fitness function is the multi-locus underdominance model for both simulations. Other parameter values for both simulations are $c = 1$, $k = 2$, and $m = 0.05$.

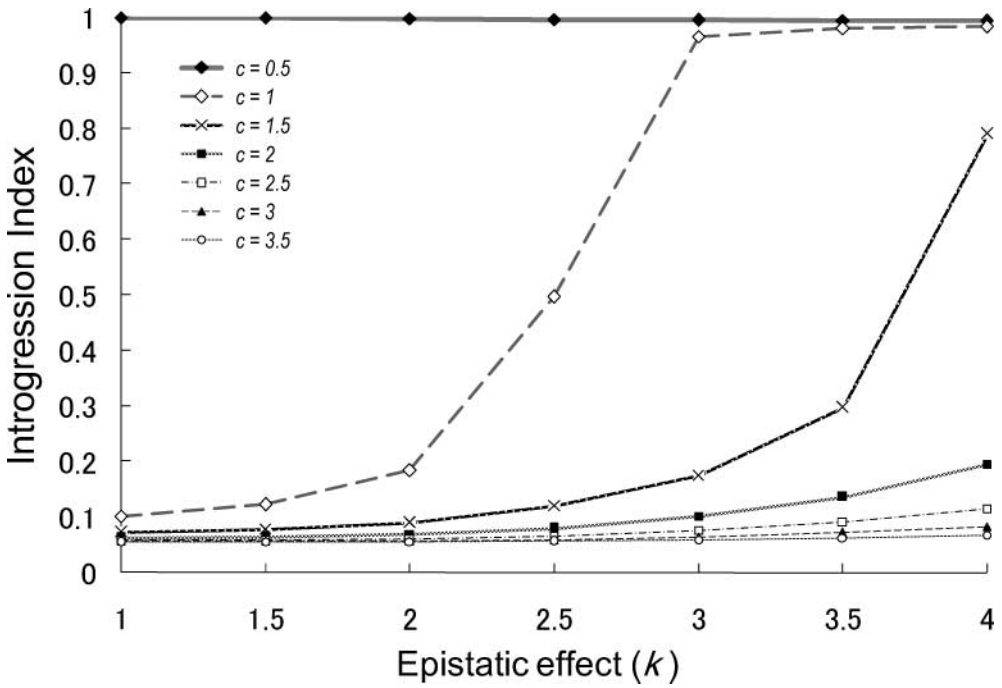


Fig. 4. Asymptotic introgression index plotted against epistatic effect with various fitness costs of F1 hybrids. The fitness function is the multi-locus underdominance model. The number of simulated generations is 200. Other parameter values are $n = 6$, $R = 0.2$, and $m = 0.05$.

DISCUSSION

The present gamete-based simulations indicated that rapid replacement of genes coding for genetic incompatibility by corresponding exotic genes was facilitated by stronger epistatic effects and higher recombination rates if individual fitness was subject to the multi-locus underdominance model. Both recombination rate and epistatic interaction between loci substantially affected the outcome of introgressive hybridization, whereas the number of loci *per se* did not influence the process of introgressive hybridization as long as the total rate of recombination was constant.

Genetic replacement by invasion might also be explained by the purely demographic effect of one-way migration. Nonetheless, the substantial effect of recombination indicates that a major portion of the simulated genetic replacement by invasion is attributable to hybridization (genetic mixing) rather than the direct demographic effects of invasion. In all cases where complete genetic replacement occurred, the frequency of the invading genes showed a sharp sigmoid curve with time. Therefore, once the population has passed the critical point of the shallow fitness valley in the final phase of genetic replacement, it quickly approaches the alternative fitness peak, which will be occupied by the purely exotic genotype. The final proliferation of the purely exotic genotype was presumably dominated by self-reproduction without hybridization, rather than by immigration or flux from admixed genotypes by recombination, because such factors were unlikely to explain the observed rapid replacement in the final phase.

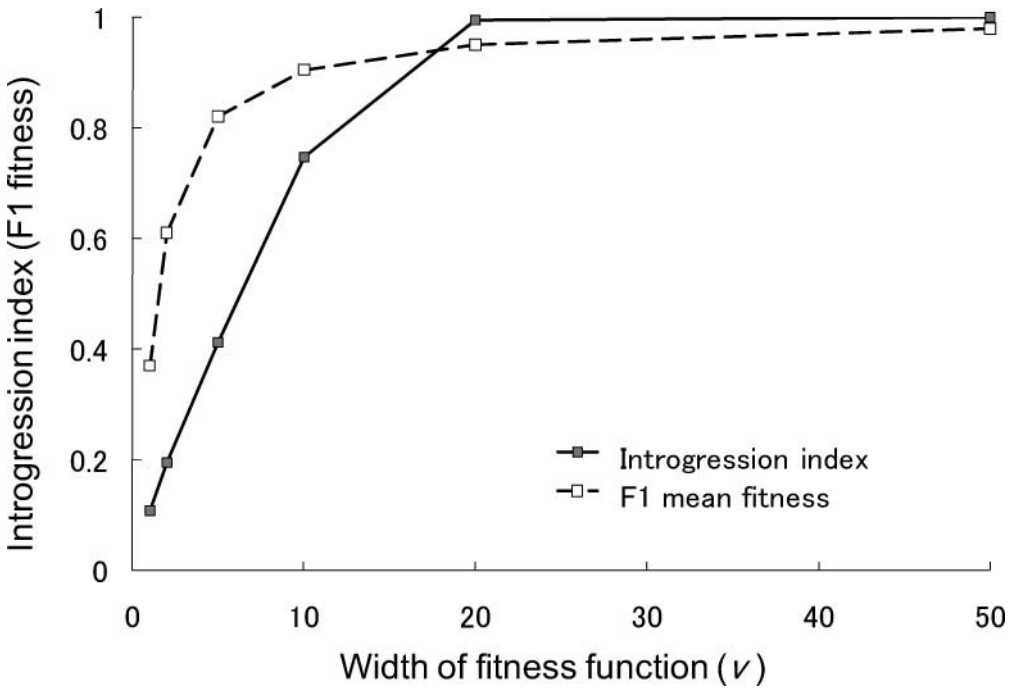


Fig. 5. Asymptotic introgression index plotted against the measure of directional selection (width of fitness function). The fitness function is the local adaptation model. The total number of simulated generations is 200. Other parameter values are $n = 10$, $R = 0.2$, and $m = 0.05$.

As indicated by the simulations, strong epistatic interactions between loci greatly facilitate introgressive hybridization. With the additive fitness effect of genes across loci (k close to 1), the endemic population rarely suffered complete genetic replacement by hybridization. With strong epistatic interactions, however, the exotic genes substantially recovered their genotypic fitness by recombination, allowing the hybridized population to move moving uphill on the fitness landscape. After reaching the shallow valley in the fitness landscape, which is formed by strong epistasis, the population can easily shift to the alternative fitness peak, owing to the weak directional forces produced by one-way migration and by flux in the purely exotic genotype from admixed genotypes by recombination.

In contrast to the multi-locus underdominance model, the local adaptation model predicted much more linear relationships between all the parameters examined and the extent to which introgressive hybridization replaced endemic genes. A synergistic interaction between several traits is considered to have played an important role in ecologically driven parapatric or sympatric speciation (Jones, 1998, 2003; Schemske and Bradshaw, 1999; Hawthorne and Via, 2001; Bradshaw and Schemske, 2003). Nonetheless, the additive fitness in the local adaptation model postulates that local adaptation is acquired over a relatively short time, compared with the evolution of adaptive gene complexes created by non-additive interaction between genes over a long period. Thus, the model is more appropriate for simulating the effects of release of artificial breeds or genetically modified organisms whose genetic backgrounds overlap

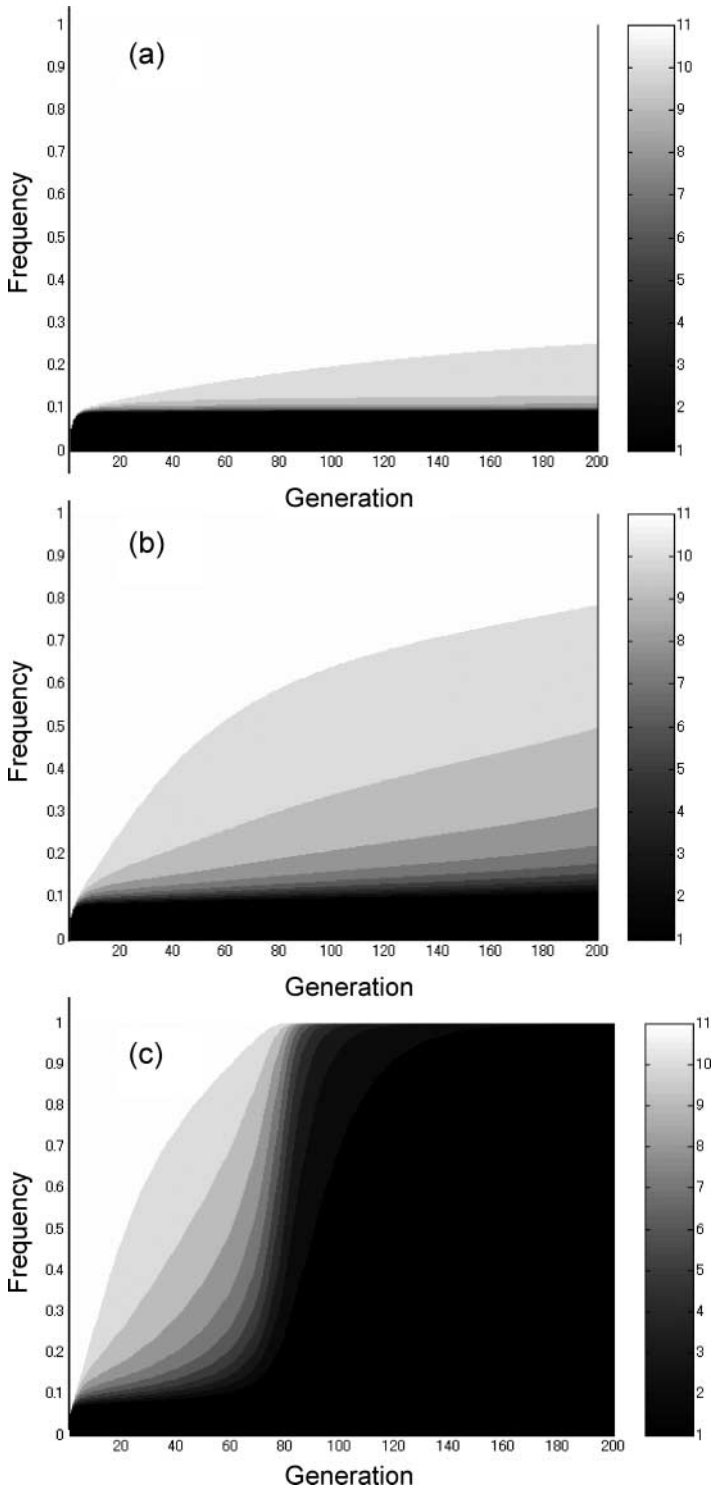


Fig. 6. Simulated introgressive hybridization when the recombination rate is (a) 0.05, (b) 0.2, or (c) 0.4. The *y*-axis represents gamete class frequencies. There are 11 gamete classes and 10 loci: the black class (1) is the purely exotic gamete, and the white class (11) is the purely endemic gamete. Other parameter values are $c = 1$ and $k = 2$.

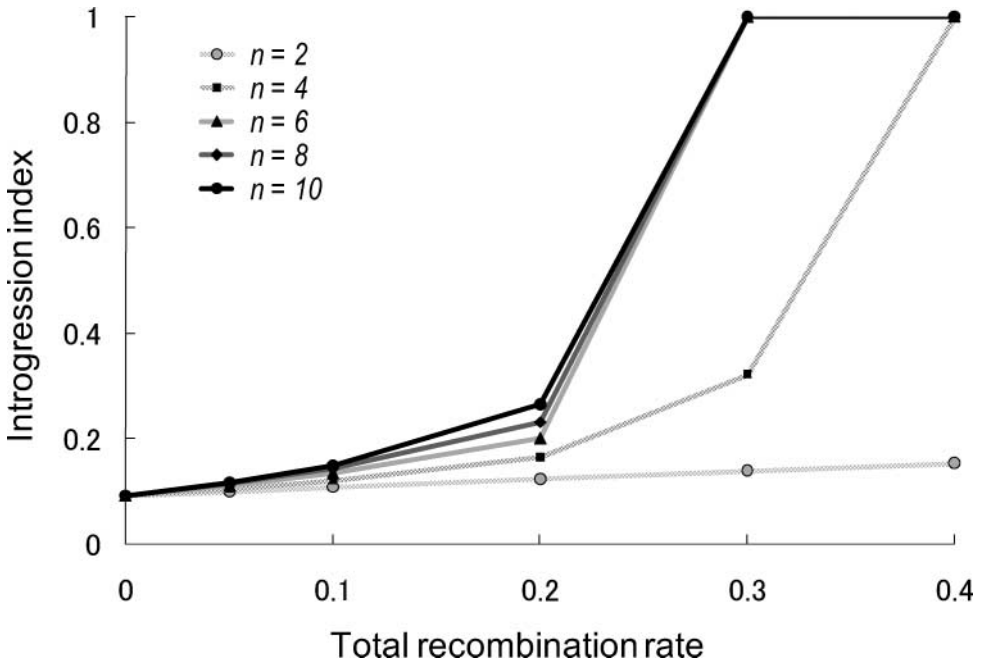


Fig. 7. Asymptotic introgression index plotted against total recombination rate with various numbers of loci. The fitness function is the multi-locus underdominance model. The total number of simulated generations is 200. Other parameter values are $c = 1$, $k = 2$, and $m = 0.05$.

with those of wild populations in nature than for simulating introgressive hybridization of a distinct species that has diverged from an ancestral species by ecological speciation.

In the case of the multi-locus underdominance model, there is a tendency for introgressive hybridization to increase abruptly with a gradual increase in the numbers of loci and recombination rates. In contrast, the local adaptation model suggests that a locally adapted population is barely replaced by another population unless the fitness reduction of the F1 hybrid is extremely small. On the other hand, partial introgression, which is characterized by intermediate values of the introgression index, is not avoidable unless natural selection against invading genes is unrealistically strong.

Therefore, it is inferred that the genetic outcome of introgressive hybridization follows two paths that are subject to the selection regime. For genes coding for genetic incompatibility subject to epistatic underdominance, introgressive hybridization can bring about an irreversible process of complete genetic replacement, depending on the number of loci and the recombination rates. Meanwhile, additive adaptive genes allow introgressive hybridization to bring about persistent but incomplete genetic contamination by exotic genes, a reversible change.

The reason why the number of loci and the recombination rate affect the process of introgressive hybridization may be explained by dilution of the fitness reduction of particular exotic genes by recombination. A common explanation applies to both the multi-locus underdominance model and the local adaptation model, although it is clearer with the latter.

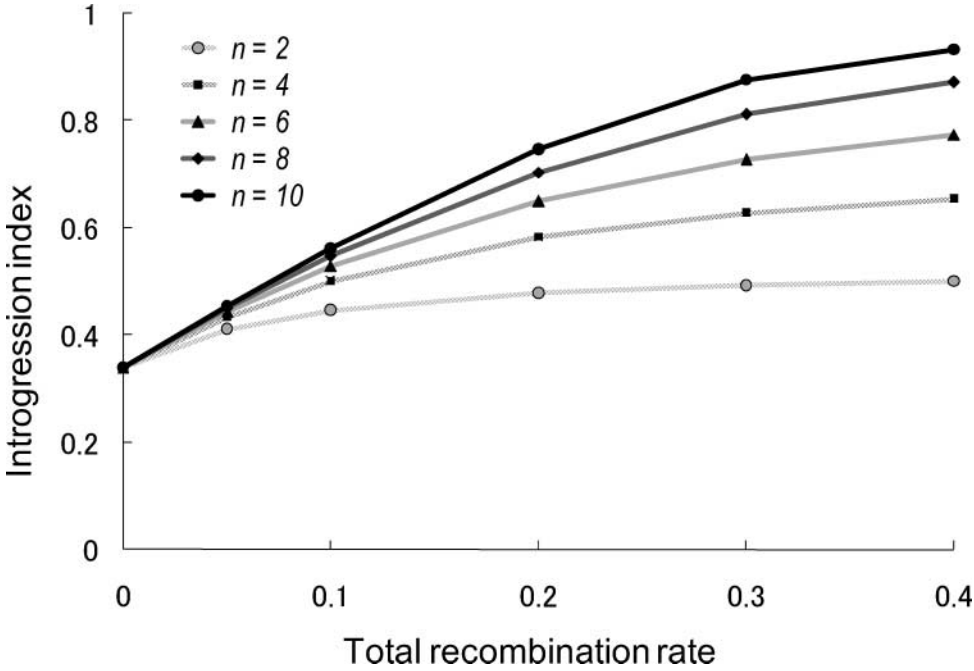


Fig. 8. Asymptotic introgression index plotted against total recombination rate with various numbers of loci. The fitness function is the local adaptation model. The total number of simulated generations is 200. Other parameter values are $v = 10$ and $m = 0.05$.

The direct effect of large numbers of loci and high recombination rates is to efficiently generate admixed genotypes that have partly recovered fitness in comparison with the fitness of F1 hybrids. If the recombination rate is extremely low, exotic genes must exist in most cases as the complete heterozygote (the genotype heterozygous at all loci), which has the lowest fitness. Most exotic gametes are likely to be repeatedly joined with the purely endemic gamete at zygosis because the exotic gametes are rare. The purging selection at the maximum intensity persistently acts against the exotic genes. The barrier effect of the limited recombination to resist gene flow is reinforced by epistatic selection depicted by the multi-locus underdominance model (cf. Barton and Bengtsson, 1986). In contrast, recombination across a number of loci produces various admixed genotypes, recovering the marginal fitness of particular exotic genes. The exotic genes allocated, by recombination, into purely endemic or partly exotic gametes will enjoy reduced selection pressure compared with that experienced by the exotic genes held in F1 hybrids. The gene frequency of exotic genes increases if the rate of loss is reduced, so that the input of exotic genes by immigration dominates the rate of loss by purging selection. Extremely small numbers of loci invalidate the dilution effect by recombination directed to the exotic genes, because per-locus fitness reduction is too large for exotic genes and gametes of efficiently admixed types are not produced.

In the case of the multi-locus underdominance model, when the gene frequency of exotic genes exceeds 0.5 the relative fitness of the endemic and the exotic genes is reversed owing to the symmetry of fitness, and selection favouring the exotic genes quickly accomplishes genetic replacement of the focal loci. This rapid process of genetic replacement, as indicated

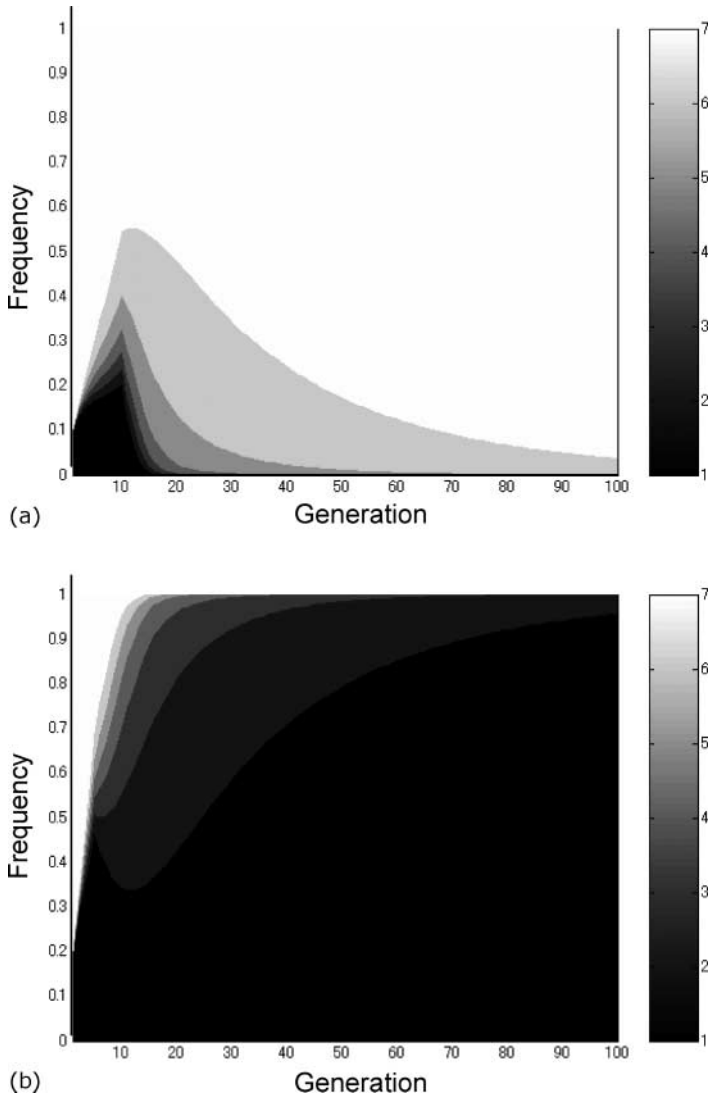


Fig. 9. Simulated introgressive hybridization by pulse invasion that is (a) low-rate and long-term ($m = 0.1$, $T = 10$) and (b) high-rate and short-term ($m = 0.2$, $T = 5$). The y-axis represents gamete class frequencies. There are 11 gamete classes and 10 loci: the black class (1) is the purely exotic gamete, and the white class (11) is the purely endemic gamete. Other parameter values are $n = 6$ and $R = 0.4$.

by the simulations, may represent an important aspect of breakdown of the post-zygotic isolating mechanism by introgressive hybridization (Tanaka, 2007).

These results imply that allopatrically diverged, closely related species are prone to human-mediated introgressive hybridization, which is characterized by long-distance movement, for three reasons. First, the derived species may not acquire efficient pre-zygotic isolation from each other due to allopatry for a long period of time, and the post-zygotic isolating mechanism may often be the only check against hybridization (Rhymer *et al.*, 1994;

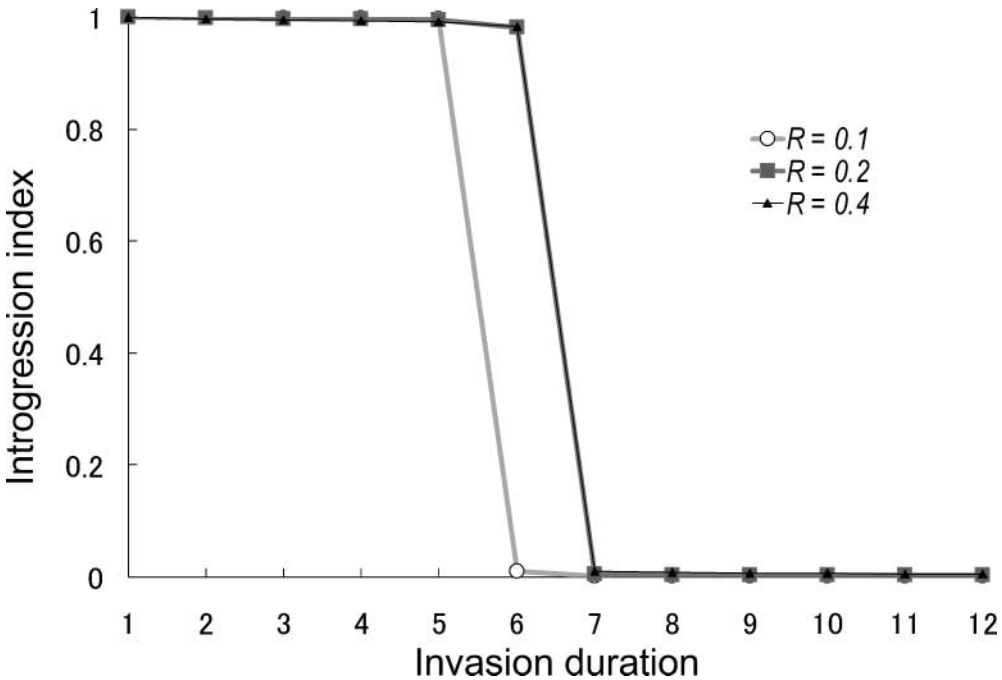


Fig. 10. Asymptotic introgression indexes plotted against invasion durations with various total recombination rates. The total number of simulated generations is 200. Parameter values are $n = 6$, $c = 1$, and $k = 2$. The migration rate m is the inverse of the invasion duration.

Rubidge and Taylor, 2004). Second, post-zygotic isolation between species is likely to occur by the accumulation of compatible combinations of mutations that have an epistatic fitness effect, as postulated by the Dobzhansky-Mueller theory and the present multi-locus underdominance model. Third, as indicated by recent molecular genetic studies, incompatibility may be controlled by multiple genes located in several regions of the genome (Orr, 2001), and the rate of recombination between incompatibility genes is therefore likely to be moderately high. In summary, allopatrically speciated taxa with large chromosome numbers are at high risk of hybridization if they encounter each other by human-mediated migration.

Caution must be paid, however, in interpreting the validity and generality of the above conclusion, because the present analysis has two inherent limitations. First, the analysis focuses on a particular isolating mechanism – post-zygotic isolation by incompatibility genes that follow the Dobzhansky-Mueller model. Focusing on the breakdown of pre-zygotic isolation, Ferdy and Austerlitz (2002) employed the approximate hypergeometric model and obtained results that contrast with those presented here (Doebeli, 1997; Shpak and Kondrashov, 1999). Quantitative characters under analysis were assumed to be associated with mating and subject to strong assortative mating. In sharp contrast to the present results, introgressive hybridization was facilitated by larger numbers of loci and fewer non-additive effects.

Another limitation arises from the unknown precision of the random gamete approximation or the equal probability assumption of all gamete types. Newly introgressed genes are not located at random in the ordered gene map of a chromosome, because they

have higher-than-equal probability of adjoining genes from a common (exotic) gamete. Persistent immigration may maintain gametic phase disequilibrium to an extent that random gamete approximation is irrelevant. Such gametic phase disequilibrium, however, disappears with crossing over every generation and may not noticeably affect the process of hybridization after the early phase of invasion. In addition, the individual-based simulation indicated that the random gamete assumption was relevant to simulate even in the early phase of introgression.

Real natural populations of endemic species that face a danger of invasion may have a seriously limited population size, whereas the present random gamete simulation assumes infinite population size. The fitness reduction by hybridization likely decreases the population size, and the size reduction likely increases immigration rate and hybridization. Thus, there may be a synergistic interaction between population decline and introgressive hybridization. The outcome of the synergistic effect of population decline and introgressive hybridization might be different between alternative fitness regimes. As for the local adaptation model, the synergistic effect likely brings about demographic extinction: the partly maladaptive genotypes created by introgressive hybridization reduce the mean fitness of the population, and the reduction of native population density increases hybridization. This scenario parallels the species range models (Garcia-Ramos and Kirkpatrick, 1997; Kirkpatrick and Barton, 1997). On the other hand, as for the multi-locus underdominance model, the exotic genes *per se* do not bring about maladaptive fitness reduction in the invaded population. The reduction of population size by F1 hybrids or partly hybridized individuals results in an inflated immigration rate, and may result in rapid replacement by the exotic gene (genetic extinction) without inducing demographic extinction (if the fitness reduction of the F1 hybrids is small enough), as depicted by the pulse invasion scenario in this study.

Bearing such limitations in mind, the present results have some implications for real conservation problems: (1) genetic extinction of purely endemic individuals is more likely if the isolation is governed mostly by a post-zygotic rather than a pre-zygotic mechanism; (2) complete replacement of compatibility genes (genetic extinction) is more likely with high recombination rates and epistatic effects; and (3) the final phase of genetic replacement is rapid and is preceded by a longer period of time characterized by a gradual increase in the number of admixed genotypes.

ACKNOWLEDGEMENTS

This work was supported by a Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science (#17510027 to Y. Tanaka). I thank Kohichi Goka, Haruki Tatsuta, and Hiroyuki Yokomizo for their helpful comments on the manuscript.

REFERENCES

Allendorf, F.W., Leary, R.F., Spruell, P. and Wenburg, J.K. 2001. The problems of hybrids: setting conservation guidelines. *Trends Ecol. Evol.*, **16**: 613–622.

Barton, N.H. and Bengtsson, B.O. 1986. The barrier to genetic exchange between hybridizing populations. *Heredity*, **56**: 357–376.

Barton, N.H. and Shpak, M. 2000a. The effect of epistasis on the structure of hybrid zones. *Genet. Res.*, **75**: 179–198.

Barton, N.H. and Shpak, M. 2000b. The stability of symmetric solutions to polygenic models. *Theor. Popul. Biol.*, **57**: 249–263.

Bradshaw, H.D. and Schemske, D.W. 2003. Allelic substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature*, **426**: 176–178.

Colbach, N., Clermont-Dauphin, C. and Meynard, J.M. 2001. GeneSys: a model of the influence of cropping system on gene escape from herbicide-tolerant rapeseed crops to rape volunteers. II. Genetic exchanges among volunteer and cropped populations in a small region. *Agric. Ecosyst. Environ.*, **83**: 255–270.

Coyne, J.A. and Orr, A. 2004. *Speciation*. Sunderland, MA: Sinauer Associates.

Doebeli, M. 1997. Genetic variation and the persistence of predator–prey interactions in the Nicholson-Bailey model. *J. Theor. Biol.*, **188**: 109–120.

Epifanio, J. and Nielsen, J. 2000. The role of hybridization in the distribution, conservation and management of aquatic species. *Rev. Fish Biol. Fish.*, **10**: 245–251.

Epifanio, J. and Philipp, D. 2001. Simulating the extinction of parental lineages from introgressive hybridization: the effects of fitness, initial proportions of parental taxa, and mate choice. *Rev. Fish Biol. Fish.*, **10**: 339–354.

Ferdy, J. and Austerlitz, F. 2002. Extinction and introgression in a community of partially cross-fertile plant species. *Am. Nat.*, **160**: 74–86.

Gadau, J., Page, R.E. and Werren, J. H. 1999. Mapping of hybrid incompatibility loci in *Nasonia. Genetics*, **153**: 1731–1741.

Garcia-Ramos, G. and Kirkpatrick, M. 1997. Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, **51**: 21–28.

Gavrilets, S. 1997a. Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.*, **12**: 307–312.

Gavrilets, S. 1997b. Hybrid zones with Dobzhansky-type epistatic selection. *Evolution*, **51**: 1027–1035.

Hails, R.S. 2000. Genetically modified plants: the debate continues. *Trends Ecol. Evol.*, **15**: 14–18.

Hails, R.S. and Morley, K. 2005. Genes invading new populations: a risk assessment perspective. *Trends Ecol. Evol.*, **20**: 245–252.

Hawthorne, D.J. and Via, S. 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature*, **412**: 904–907.

Huxel, G.R. 1999. Rapid displacement of native species by invasive species: effects of hybridization. *Biol. Conserv.*, **89**: 143–152.

Jones, C.D. 1998. The genetic basis of *Drosophila sechellia*'s resistance to a host plant taxon. *Genetics*, **149**: 1899–1908.

Jones, C.D. 2003. Genetics of egg production in *Drosophila sechellia. Heredity*, **92**: 249–256.

Kanaiwa, M. and Harada, Y. 2000. Theoretical analysis of introgression between released fish and wild fish populations considering sex differences. *Fish. Sci.*, **66**: 686–694.

Kirkpatrick, M. and Barton, N.H. 1997. Evolution of a species' range. *Am. Nat.*, **150**: 1–23.

Kondrashov, A.S. 1982. Selection against harmful mutations in large sexual and asexual populations. *Genet. Res.*, **40**: 325–332.

Kondrashov, A.S. 1985. Deleterious mutations as an evolutionary factor. II. Facultative apomixis and selfing. *Genetics*, **111**: 635–653.

Levin, D.A., Francisco-Ortega, J. and Jansen, R.K. 1995. Hybridization and the extinction of the rare plant species. *Conserv. Biol.*, **10**: 10–16.

Mallet, J. 2005. Hybridization as an invasion of the genome. *Trends Ecol. Evol.*, **20**: 229–237.

Orr, H.A. 1989. Localization of genes causing postzygotic isolation in two hybridizations involving *Drosophila pseudoobscura*. *Heredity*, **63**: 231–237.

Orr, H.A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics*, **139**: 1805–1813.

Orr, H.A. 2001. The genetics of species differences. *Trends Ecol. Evol.*, **16**: 343–350.

Orr, H.A. and Turelli, M. 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Mueller incompatibilities. *Evolution*, **55**: 1085–1094.

Rhymer, J.M. and Simberloff, D. 1996. Extinction by hybridization and introgression. *Annu. Rev. Ecol. Syst.*, **27**: 83–109.

Rhymer, J.M., Williams, M.J. and Braun, M.J. 1994. Mitochondrial analysis of gene flow between New Zealand mallards (*Anas platyrhynchos*) and grey ducks (*A. superciliosa*). *Auk*, **111**: 970–978.

Rieseberg, L.H. and Gerber, D. 1995. Hybridization in the Catalina Island mountain mahogany (*Cercocarpus traskiae*): RAPD evidence. *Conserv. Biol.*, **9**: 199–203.

Rubidge, E.M. and Taylor, E.B. 2004. Hybrid zone structure and the potential role of selection in hybridizing populations of native westslope cutthroat trout (*Oncorhynchus clarki lewisi*) and introduced rainbow trout (*O. mykiss*). *Molec. Ecol.*, **13**: 3735–3749.

Schemske, D.W. and Bradshaw, H.D. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). *Proc. Natl. Acad. Sci. USA*, **96**: 11910–11915.

Shpak, M. and Kondrashov, A.S. 1999. Applicability of the hypergeometric phenotypic model to haploid and diploid populations. *Evolution*, **53**: 600–604.

Tanaka, Y. 2007. Introgressive hybridization as the breakdown of postzygotic isolation: a theoretical perspective. *Ecol. Res.*, **22**: 929–939.

Thompson, C.J., Thompson, B.J.P., Ades, P.K., Cousens, R., Garnier-Gere, P., Landman, K. *et al*. 2003. Model-based analysis of the likelihood of genetically modified crops into wild relatives. *Ecol. Model.*, **162**: 199–209.