Approximations for the hitchhiking effect caused by the evolution of antimalarial-drug resistance

Kristan A. Schneider and
School of Life Sciences, Arizona State University, 1711 South Rural Road, Tempe, AZ 85287, USA

Department of Mathematics, University of Vienna, Nordbergstrasse 15, UZA 4, 1090 Vienna, Austria

CEMI/Biodesign Institute, Arizona State University, P. O. Box 875301, Tempe, AZ 85287-5301, USA

Yuseob Kim
School of Life Sciences, Arizona State University, 1711 South Rural Road, Tempe, AZ 85287, USA

Center for Evolutionary Medicine and Informatics, Biodesign Institute, 1001 S. McAllister Ave., Tempe, AZ 85281, USA

Kristan A. Schneider: kristan.schneider@asu.edu

Abstract

An analytically feasible, deterministic model for the spread of drug resistance among human malaria parasites, which incorporates all characteristics of the complex malaria-transmission cycle was introduced by Schneider and Kim (Theor. Popul Biol, 2010). The model accounts for the fact that only a fraction of infected hosts receive drug treatment and that hosts can be co-infected by differently many parasites. Furthermore, the model also incorporates host heterogeneity. Antimalarial-drug resistance is assumed to be caused by a single locus with two alleles—a sensitive one and a resistance one. The most important result for this model is that an analytical solution for the frequencies of a linked neutral biallelic locus exists. However, the exact solution does not admit an explicit form, and cannot straightforwardly be interpreted in terms of the model parameters. Here, we establish simple approximations for the equilibrium frequency at the neutral locus. Under the assumption that the resistant allele is initially rare—the biologically most relevant assumption in this context—and that recombination is weak, the approximations become similar to the approximations in the standard hitchhiking model. However, there are crucial differences. In particular, because of the high degree of selfing among malaria parasites in their sexual phase, a genome-wide reduction of relative heterozygosity occurs if selection is sufficiently strong. It turns out that the approximations are accurate even if the recombination rates are not small and the resistant allele is initially not very rare. The main advantage of our approximations is that they are easy to interpret in terms of model parameters. Moreover, they allow to make predictions of the size of the valley of reduced heterozygosity around the selected locus for given model parameters. Reversely, for a given reduction of heterozygosity, it is possible to identify the corresponding parameters. Moreover, we will show that incorporating host heterogeneity leads to an increased hitchhiking effect.

Correspondence to: Kristan A. Schneider, `kristan.schneider@asu.edu`.

Keywords

Selective sweep; Co-infections; Drug concentration; Relative heterozygosity; Host heterogeneity; Malaria; Plasmodium falciparum

1 Introduction

Human malaria is an infectious disease caused by parasites belonging to the genus *Plasmodium*, which is endemic in most tropical and subtropical regions in the world. Infections with *Plasmodium falciparium*, the most virulent form of human malaria, result worldwide in one to 3 million deaths per year (cf. WHO 2000; Korenromp et al. 2005). Malaria control is highly dependent on drug treatments that kill parasites in infected hosts. However, attempts to control malaria have been thwarted by rapid evolution of antimalarial-drug resistance, a fact that has been described to be a public health disaster (cf. Marsh 1998).

The limited repertoire of safe, effective, and affordable antimalarial drugs has made research on the emergence and dispersion of resistance a global health priority. Mathematical models that can use input from genetic data to investigate the dynamics of mutations associated with drug resistance are urgently needed for designing drug-deployment policies that can increase the lifespan of the available drugs. This requires a detailed understanding of population genetic processes that lead to the emergence and dispersion of drug resistance. Unfortunately, the highly complex nature of the malaria-transmission cycle as well as complex demographic and environmental factors aggravate the efforts to elaborate theoretical models.

Malaria parasites undergo a complex transmission cycle with sexual phases in the mosquito vector and asexual phases in the infected host (cf. e.g. Daily 2006; Prugnolle et al. 2009). A human host is inoculated with sporozoites by the bite of an infected Anopheles mosquito during its blood meal. In the human host the sporozoites migrate first to the liver where they differentiate into hepatic merozoites. These are released into the blood stream where some of the hepatic merozoites form gametocytes. The haploid gametocytes are extracted by a mosquito during its blood meal and immediately reproduce sexually in the mosquito's gut. Consequently, they undergo recombination. Completing the transmission cycle, this step is followed by the production of haploid sporozoites in the mosquitos's salivary glands from which they can be inoculated into a human host.

Another source of complication is that many environmental and clinical factors differ significantly across the worldwide distribution of this parasite. The transmission rate and hence the number of secondary infections varies from very low rates in parts of South America, over intermediate rates in Southeast Asia, to high rates in Africa. On the other hand the level of host-acquired immunity is much higher (which means a high number of asymptomatic infections) in most of the affected areas in Africa than in other parts of the world. Such variation in host-acquired immunity affects drug use.

Among others, two important variables that summarize the demographic and clinical setting of the particular geographic area are the (average) number of parasites (m) co-infecting a given host (the average multiplicity of infection), which is determined by transmission intensity, and the proportion (α) of infected hosts that are drug treated. This parameter depends mainly on how many hosts acquired immunity.

Several studies built population genetic models that demonstrated profound effects of m and α on the rate of drug-resistance evolution (Hastings and Mackinnon 1998; Mackinnon and

Hastings 1998; Hastings 2003,2006). However, it is still unclear which mechanisms are important to the spread of resistance, i.e., intra-host dynamics, drug half life, multiple drug treatment, migration, multiple infections, recombination, mutation etc.

Over the last 2 decades substantial advances regarding the genetic basis of antimalarial-drug resistance have been made. It is known that specific point mutations in the *dhfr* and *dhps* regions underlie resistance to *pyrimethamine* and *sulfadoxine* (cf. Cowman et al. 1988; Triglia and Cowman 1994; Brooks et al. 1994), that point mutations in mitrochandrial DNA underlie resistance to *atovaquone* (cf. Schwobel et al. 2003), and that mutations in the *pfcrt* gene are causing resistance against *Chloroquine* (CQ).

The spread of mutations causing drug resistance leads to a valley of reduced genetic variation at linked neutral regions. This removal of pre-existing variation occurs because recombination cannot effectively break the initial association with the neutral background in which the mutant first occurred, a process known as genetic hitchhiking or a selective sweep (Maynard Smith and Haigh 1974; Stephan et al. 1992; Barton 2000). For instance, Nair et al. (2003) observed a severe reduction of variation at microsatellite loci spanning over a 100 kb region surrounding the *dhfr* gene in a Southeast Asian population of *P. falciparum*. The extent of this pattern depends on how fast the favored allele increases to high frequency while meiotic recombination is constantly eroding the association between the favored allele and the surrounding chromosome segment (Kim and Stephan 2002). Selective sweeps have been initially studied for randomly mating populations of constant size, and homogenous constant selection pressures (e.g., Maynard Smith and Haigh 1974), to which we refer as the 'standard model', or standard 'hitchhiking'.

In malaria biology, detection of selective sweeps mainly contributed to confirming the location of drug resistant mutations and elucidating their mutational origins (Wootton et al. 2002; Nash et al. 2005; Mita et al. 2007; Nair et al. 2007; McCollum et al. 2008), while fewer studies attempted to relate the span of selective sweeps with the strength of drug selection (Nair et al. 2003). Recently, Schneider and Kim (2010) introduced an analytical feasible model for the spread of antimalarial-drug resistance, which allows to study genetic hitchhiking. The model covers the important characteristics of the transmission cycle, incorporates host heterogeneity, i.e., different classes of treated and untreated hosts, and accounts for the fact that hosts can be infected by differently many parasites. Moreover, the model allows simple conditions for the spread of resistance and its speed in terms of the fitness parameters and α. Hence, it is useful to find 'optimal' treatment strategies to prevent or slow down the spread of resistance (for more discussion see Schneider and Kim 2010).

Studies based on known recombination rates and the result of the standard model of selective sweeps concluded that the observed patterns of selective sweeps are compatible with the predictions (Nair et al. 2003). However, Schneider and Kim (2010) discuss in detail why the application of the standard selective-sweep model to a malaria-parasite population is highly problematic. In particular, the high degree of selfing among malaria parasites will lead to a genome-wide reduction of heterozygosity if selection is sufficiently strong (cf. also Hedrick 1980).

In the absence of reliable public-health records, a retrospective analysis of the mechanisms and parameters underlying the spread of antimalarial drug resistance may be achieved indirectly through the patterns of selective sweeps. However, the analytical solution for the valley of reduced heterozygosity in the model of Schneider and Kim (2010) is not explicit, which limits their applicability. Moreover, it is not at all straightforward to interpret these results based on the model parameters.

In this article we derive accurate approximations to the exact, analytical solutions of Schneider and Kim (2010). The approximations under the assumption that the resistant allele is initially rare and that recombination is weak are similar to but different from the usual approximation for standard hitchhiking. Unlike in standard hitchhiking, our approximations are accurate even if the recombination rates are large and the resistant allele is initially not very rare. Moreover, using the approximations, we will show that incorporating host heterogeneity will result in an increased hitchhiking effect compared to our basic model, with corresponding selection parameters, which accounts just for one class of treated and one class of untreated hosts. The simple form of the approximations will allow us to easily interpret the effect of genetic hitchhiking in terms of the model parameters. Such approximations are extremely useful to achieve applicability of the results of Schneider and Kim (2010) to real data. Moreover, we show that the approximate solutions can be easily applied to identify the range of parameters that give rise to given levels of reduced heterozygosity in a given range of recombination distances. Finally, we discuss the differences compared with standard hitchhiking and give an outlook how our results can be applied to real data.

2 The model

We consider two biallelic loci. The first locus is subject to selection with a sensitive allele A_S and a resistant allele A_R segregating. The second locus is selectively neutral with the alleles N_1 and N_2 segregating. We will use the notation and parametrization summarized in Table 1.

Let p denote the frequency of the resistant allele A_R. We have $p = p_3 + p_4$. Furthermore, we denote the frequency of the neutral allele N_1 among the sensitive and resistant haplotypes by R and Q, respectively. We therefore have $R = \frac{p_1}{p_1 + p_2} = \frac{p_1}{1 - p}$ and $Q = \frac{p_3}{p_3 + p_4} = \frac{p_3}{p}$.

Moreover, we denote the recombination rate between the two loci by r, and the vector of haplotype frequencies by $\boldsymbol{p} = (p_1, \ldots, p_4)$. We assume that each host is infected randomly and independently by exactly m haploids (parasite strains). (We assume m to be a fixed parameter until Sect. 4.2, where we assume that m follows a fixed frequency-distribution).

Hosts acquire parasite strains according to their frequencies in the mosquito phase. Hence, the configuration of infections in hosts is multinomially distributed with parameters m and $p_1, \ldots, p_4$. We assume that all haploids in a newly infected host have equal frequencies. Hence, the relative frequency of a haploid drawn from the parasite population among mosquitos is $\frac{1}{m}$. Let us denote a multi-index by $\boldsymbol{m} = (m_1, \ldots, m_4)$, and the sum over its components by $|\boldsymbol{m}| = \sum_{i=1}^{4} m_i$. The probability that a host is infected by m_i copies of haplotype i ($i = 1, \ldots, n, |\boldsymbol{m}| = m$) is given by

$$\begin{pmatrix} m \\ \boldsymbol{m} \end{pmatrix} \boldsymbol{p^m},$$

(1)

where $\begin{pmatrix} m \\ \boldsymbol{m} \end{pmatrix} = \frac{m!}{\prod_{i=1}^{4} m_i!}$ denotes the respective multinomial coefficient and, as usual, $\boldsymbol{p^m} := \prod_{i=1}^{4} p_i^{m_i}$.

After a host is infected the parasites reproduce clonally in the host. An infected host either receives drug treatment or is untreated. We assume that a fixed proportion α of infected hosts in the population is treated, whereas the remaining hosts are untreated. Thus, the probability for a host to be treated is simply α. The rate of reproduction of the haplotypes is different in treated and untreated hosts. The absolute fitness of a parasite strain is the expected number of its descendants in the host at the time of the mosquito visit. The absolute frequency (fitness) of haplotype i in an untreated host before a mosquito takes its blood meal is denoted by $\frac{m_i}{m} W_i^{(U)}$, whereas the absolute frequency of haplotype i in a treated host before a mosquito takes its blood meal is denoted by $\frac{m_i}{m} W_i^{(T)}$. [In the following, wherever it is appropriate, we use the superscript to (.) to resemble the superscripts (U) and (T).] Some of these haploids form gametocytes in male or female expressions. The frequencies of those are assumed to be proportional to the number of respective haplotypes. Furthermore, we impose that male and female gametocytes occur at the same frequencies. We assume that the number of different gametocytes taken by a mosquito during its blood meal from an infected host is proportional to its frequency in the host. Let γ denote the proportional constant, which is assumed to be the same for each mosquito. Hence, if the absolute frequencies of parasites in an infected host are $\frac{m_1}{m} W_1^{(.)}, \ldots, \frac{m_4}{m} W_4^{(.)}$, the absolute frequencies of parasites absorbed during the blood meal are $\gamma \frac{m_1}{m} W_1^{(.)}, \ldots, \gamma \frac{m_4}{m} W_4^{(.)}$. Note that this takes drug efficiency into account, because a mosquito will absorb a smaller number of parasites from a host in which drugs efficiently eliminated parasites.

In the mosquitos' guts recombination occurs immediately after the blood meal during the phase in which meiosis occurs. In the gut of a mosquito, which has taken its blood meal from a host initially infected with m_i haplotypes i ($i = 1, \ldots, 4$), the probability that a male k-gametocyte fertilizes a female l-gametocyte is

$$\frac{\gamma \frac{m_k}{m} W_k^{(.)}}{\frac{\gamma}{m} W_m^{(.)}} \cdot \frac{\gamma \frac{m_l}{m} W_l^{(.)}}{\frac{\gamma}{m} W_m^{(.)}} = \frac{m_k W_k^{(.)} m_l W_l^{(.)}}{W_m^{(.)^2}},$$

where

$$\frac{\gamma}{m} W_m^{(.)} = \frac{\gamma}{m} \sum_{k=1}^{4} m_k W_k^{(.)}$$

is the frequency of parasite haploids in the mosquito's gut. The above probability is the relative frequency of a male k-gametocyte times that of a female l-gametocyte. Therefore, the absolute frequency of pairings of a male k-gametocyte and a female l-gametocyte is, the probability of such a fertilization times the absolute numbers of parasites in the gut, i.e.,

$$\frac{\gamma^2 \frac{m_k}{m} W_k^{(.)} \frac{m_l}{m} W_l^{(.)}}{\frac{\gamma}{m} W_m^{(.)}} = \frac{\frac{\gamma}{m} m_k W_k^{(.)} m_l W_l^{(.)}}{W_m^{(.)}}. \tag{2}$$

The probability that a fertilization of a male k-gametocyte to a female l-gametocyte produces a haplotype i is denoted by $R(kl \rightarrow i)$. Therefore, the absolute frequencies of haplotype i in

the population of mosquitos that took their blood meal from treated and untreated hosts are respectively

$$p_i^{*(.)} = \sum_{|\boldsymbol{m}|=m} \binom{m}{\boldsymbol{m}} p^{\boldsymbol{m}} \sum_{k,l=1}^{4} \frac{\gamma m_k W_k^{(.)} m_l W_l^{(.)}}{m W_{\boldsymbol{m}}^{(.)}} R(kl \to i),$$

where the sum runs over all multi-indices $\boldsymbol{m} \in \mathbb{N}_+^4$ with $|\boldsymbol{m}| = m$. Therefore, the relative frequencies of haplotypes in the mosquito population become

$$p_i' = \frac{\alpha p_i^{*(T)} + (1-\alpha) p_i^{*(U)}}{\alpha \sum_{k=1}^{4} p_k^{*(T)} + (1-\alpha) \sum_{k=1}^{4} p_k^{*(U)}} = \frac{p_i^*}{\sum_{k=1}^{4} p_k^*}, \tag{3a}$$

where

$$p_i^* := \alpha p_i^{*(T)} + (1-\alpha) p_i^{*(U)}. \tag{3b}$$

For what follows, let us define the average fitness of the resistant allele among treated and untreated hosts by

$$\lambda := \alpha W_3^{(T)} + (1-\alpha) W_3^{(U)} \tag{4a}$$

and that of the sensitive by

$$\mu := \alpha W_1^{(T)} + (1-\alpha) W_1^{(U)}. \tag{4b}$$

If $p(t)$ denotes the frequency of the resistant allele at time t, it changes according to

$p(t) = \dfrac{\lambda^t p(0)}{\lambda^t p(0) + \mu^t (1 - p(0))}$, i.e., according to the standard haploid one-locus selection model (cf. Schneider and Kim 2010). Thus, the resistant allele will sweep through the population if and only if $\lambda > \mu$, otherwise it will get lost (or remain constant in frequency). Hence, in the following we will always assume $\lambda > \mu$ without further mentioning. Note that the spread of the resistant allele is independent from m. Moreover, it is obvious to calculate the time until a given level of resistance is reached (cf. Result 2 in Schneider and Kim 2010). Because of the relatively simple form of the dynamics of the resistant allele it is possible to derive simple conditions for the spread of resistance and its speed. Example 1 in Schneider and Kim (2010) directly links the fitness parameters to the spread of fixation. Furthermore, assuming that the fitnesses in treated hosts are functions of the administered drug concentration, Example 2 in Schneider and Kim (2010) illustrates the impact of the drug concentration and α on the spread of resistance and its speed. Assuming that the fitness of resistant parasites decays slower than that of the sensitive ones as a function of the drug concentration they found that resistance spreads most quickly for intermediate drug concentrations and large α.

Anyhow, the focus of this article is genetic hitchhiking. The following relations were derived in Schneider and Kim (2010).

$$\overline{W}p_i^{'}=p_i\left(\alpha W_i^{(T)}+(1-\alpha)W_i^{(U)}\right)+r_i\,p(1-p)(Q-R)\vartheta_{p,m},$$

(5a)

where

$$
\begin{aligned}
\frac{r_1}{\overline{W}} &= -r_2=-r_3=r_4=r, \\
&= p\lambda+(1-p)\mu, \\
\vartheta_{p,m} &:= \sum_{k=0}^{m-2}\binom{m-2}{k}\theta_{k,m}\,p^k(1-p)^{m-2-k},
\end{aligned}
$$

(5b)

$$\theta_{k,m}:=\alpha\theta_{k,m}^{(T)}+(1-\alpha)\theta_{k,m}^{(U)},$$

(5c)

and

$$\theta_{k,m}^{(\cdot)}:=\frac{(m-1)W_1^{(\cdot)}W_3^{(\cdot)}}{(m-1-k)W_1^{(\cdot)}+(k+1)W_3^{(\cdot)}}.$$

(5d)

It was further shown in Schneider and Kim (2010) that

$$R^{'}=R+rp(Q-R)\frac{\vartheta_{p,m}}{\mu}$$

(6a)

and

$$Q^{'}=Q-r(1-p)(Q-R)\frac{\vartheta_{p,m}}{\lambda}.$$

(6b)

From the relations (5) and (6) the equilibrium frequency of the neutral allele N_1 was calculated to be

$$\widehat{Q}^{(m)}=Q_0+\frac{r(R_0-Q_0)}{\lambda}\sum_{\tau=0}^{\infty}\frac{\vartheta_{p_\tau,m}}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^{\tau}+1}\prod_{l=0}^{\tau-1}\Lambda_{m,l},$$

(7)

where p_0, Q_0, and R_0 are the respective initial frequencies,

$$\vartheta_{p_t,m}=\left(\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^t+1\right)^{2-m}\sum_{k=0}^{m-2}\binom{m-2}{k}\theta_{k,m}\left(\frac{p_0}{1-p_0}\right)^k\left(\frac{\lambda}{\mu}\right)^{tk},$$

and

$$\Lambda_{m,t}:=1-r\frac{\vartheta_{p_t,m}}{\lambda\mu}\left(\frac{p_0\lambda^{t+1}+(1-p_0)\mu^{t+1}}{p_0\lambda^t+(1-p_0)\mu^t}\right).$$

(8)

It was mentioned by Schneider and Kim (2010) that, for practical purposes, it suffices to sum (7) until the time of quasi-fixation, for which they provided an explicit formula. However, it is not at all obvious how $\hat{Q}^{(m)}$ is influenced by the various model parameters. In particular, when studying genetic hitchhiking using (7) it seems infeasible to predict the width of the valley of reduced heterozygosity in terms of the recombinational distance. Hence, simple but accurate approximations for (7) that permit a simple interpretation in terms of the model parameters are highly desirable.

3 Approximations

We shall now calculate simple approximations for (7). First, we treat the special case $m = 2$, and later deduce the general case from it.

3.1 Equilibrium frequencies at the neutral locus for $m = 2$

The equilibrium frequency of the neutral allele N_1 is given by

$$\hat{Q}:=\hat{Q}^{(2)}=Q_0+\frac{r\vartheta(R_0 - Q_0)}{\lambda}\sum_{\tau=0}^{\infty}\frac{\prod_{l=0}^{\tau-1}\Lambda_l}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^{\tau}+1},$$

where

$$\vartheta:=\vartheta_{p_t,2} \tag{9}$$

is independent of the trajectory p_t of the resistant allele and in particular of p_0, and

$$\Lambda_l:=\Lambda_{2,l}. \tag{10}$$

In the following we will derive upper and lower bounds as well as approximations for $\hat{Q}$ which exhibit a much simpler form. This permits us to explore the effects of the various parameters on $\hat{Q}$.

First, note that Λ_l is monotone decreasing in l. To see this note that $\Lambda_l < \Lambda_{l-1}$ is equivalent to

$$1 - r\frac{\vartheta}{\lambda\mu}\left(\frac{p_0\lambda^{l+1}+(1 - p_0)\mu^{l+1}}{p_0\lambda^{l}+(1 - p_0)\mu^{l}}\right)<1 - r\frac{\vartheta}{\lambda\mu}\left(\frac{p_0\lambda^{l}+(1 - p_0)\mu^{l}}{p_0\lambda^{l-1}+(1 - p_0)\mu^{l-1}}\right).$$

This can be simplified to

$$\lambda^{l-1}\mu^{l+1}+\lambda^{l+1}\mu^{l-1}>2\lambda^{l}\mu^{l}.$$

By dividing through $\lambda^{l-1}\mu^{l-1}$ this simplifies to

$$(\lambda - \mu)^2>0.$$

Moreover, note that

$$\underline{\Lambda}:=\lim_{l\to\infty}\Lambda_l=\lim_{l\to\infty}1-r\frac{\vartheta}{\lambda\mu}\left(\frac{p_0\lambda+(1-p_0)\mu\left(\frac{\mu}{\lambda}\right)^l}{p_0+(1-p_0)\left(\frac{\mu}{\lambda}\right)^l}\right)=1-r\frac{\vartheta}{\mu}.$$

(11)

Hence, by using

$$\overline{\Lambda}:=\Lambda(0)=1-\frac{r\vartheta\left(p_0\lambda+(1-p_0)\mu\right)}{\lambda\mu}\geq\Lambda_l\geq\underline{\Lambda},$$

(12)

we obtain

$$\overline{\Lambda}^{\tau}\geq\prod_{l=0}^{\tau-1}\Lambda_l\geq\underline{\Lambda}^{\tau}.$$

Furthermore, because $\lambda>\mu$ we have

$$1-\frac{r\vartheta(p_0\lambda+(1-p_0)\mu)}{\lambda\mu}\leq1-\frac{r\vartheta(p_0\mu+(1-p_0)\mu)}{\lambda\mu}=1-\frac{r\vartheta}{\lambda}=:\tilde{\Lambda},$$

(13)

and consequently $\tilde{\Lambda}\geq\overline{\Lambda}$.

Let us define

$$Q^{(a)}:=Q_0+\frac{r\vartheta(R_0-Q_0)}{\lambda}\sum_{\tau=0}^{\infty}\frac{a^{\tau}}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^{\tau}+1}.$$

(14)

Note that by relabeling the alleles, we can assume without loss of generality that $R_0>Q_0$. Hence, (14) is monotone increasing in a. Thus, we can formulate our first result.

Result 1—*Let $Q^{(a)}$ be defined as in* (14). *Then, $Q^{(a)}$ is monotone increasing in a. Moreover, let $\underline{\Lambda}$, $\overline{\Lambda}$ and $\tilde{\Lambda}$ be defined as in* (11), (12), *and* (13), *respectively. Then*

$$\overline{Q}:=Q^{(\overline{\Lambda})}\ and\ \tilde{Q}:=Q^{(\tilde{\Lambda})}$$

are upper bounds for $\hat{Q}$, and

$$\underline{Q}:=Q^{(\underline{\Lambda})}$$

is a lower bound for $\hat{Q}$.

Although the expressions of this bounds are simpler than that of $\hat{Q}$ they are still not very explicit. Hence, we shall derive upper and lower bounds for $\overline{Q}$ and $\underline{Q}$, or more generally for $Q^{(a)}$.

First, note that $\vartheta<\lambda$. This is because it is equivalent to

$$\alpha W_3^{(T)} + (1 - \alpha)W_3^{(U)} > \frac{\alpha W_1^{(T)}W_3^{(T)}}{W_1^{(T)} + W_3^{(T)}} + \frac{(1 - \alpha)W_1^{(U)}W_3^{(U)}}{W_1^{(U)} + W_3^{(U)}}$$

which obviously holds because

$$\frac{W_1^{(T)}}{W_1^{(T)} + W_3^{(T)}} < 1 \text{ and } \frac{W_1^{(U)}}{W_1^{(U)} + W_3^{(U)}} < 1.$$

Essentially the same calculation shows $\vartheta < \mu$. Therefore, we have $\frac{\vartheta}{\lambda} < 1$ and $\frac{\vartheta}{\mu} < 1$. This combined with (11) and (13) implies $\frac{1}{2} < \underline{\Lambda} < \tilde{\Lambda} \leq 1$, since $r \in [0, \frac{1}{2}]$. Hence, in the following we will always assume $a \in [\underline{\Lambda}, \tilde{\Lambda}] \subseteq [\frac{1}{2}, 1]$.

Notice that $\tilde{Q}$ has exactly the same form as in the standard haploid hitchhiking model derived by Maynard Smith and Haigh (1974; eq. 8), with r replaced by $\frac{\vartheta r}{\lambda}$ and p_0 replaced by $\frac{\mu p_0}{\lambda(1 - p_0) + \mu p_0}$. Thus, we can interpret this as studying standard hitchhiking with an 'effective' recombination rate $\frac{\vartheta r}{\lambda}$, which is smaller than r, and an 'effective' initial frequency of $\frac{\mu p_0}{\lambda(1 - p_0) + \mu p_0}$, which is smaller than the actual initial frequency. The adjusted recombination rate leads to a more severe hitchhiking effect, whereas the adjusted initial frequency leads to a less pronounced hitchhiking effect. Combination of this factors lead to a more severe hitchhiking effect (cf. Schneider and Kim 2010).

Next, we need two definitions. For $a \in \mathbb{R}$ and $k \in \mathbb{N}^+$ the *upper Pochhammer symbol* is defined by

$$(a)_{\overline{k}} := \prod_{l=0}^{k-1}(a+l).$$

The hypergeometric function is defined by

$${}_r F_s\left(\begin{array}{c} a_1, \ldots, a_r \\ b_1, \ldots, b_s \end{array} \middle| z \right) := \sum_{k=0}^{\infty} \frac{z^k}{k!} \frac{\prod_{u=1}^{r}(a u)_{\overline{k}}}{\prod_{v=1}^{s}(b_v)_{\overline{k}}},$$

where $r, s \in \mathbb{N}, z, a_1, \ldots, a_r \in \mathbb{R}, b_1, \ldots, b_s \in \mathbb{R} \backslash \mathbb{Z}_0^-$.

With the above definition we are already able to formulate our first theorem.

Theorem 1—*Assume* $a \neq \left(\frac{\mu}{\lambda}\right)^l$ *for* $l \in \mathbb{N}$, *and let* $\csc x = \dfrac{1}{\sin x}$ *denote the cosecant of x. If* $p_0 < \dfrac{1}{2}$, *the function*

$$\psi(a) := Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda} \left[\frac{\pi \csc\left(\frac{\pi \log a}{\log \lambda - \log \mu}\right)}{(\log \lambda - \log \mu)\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \lambda - \log \mu}}} - \frac{1}{\log a} {}_2F_1\left(\begin{array}{c} 1, \frac{\log a}{\log \lambda - \log \mu} \\ 1 + \frac{\log a}{\log \lambda - \log \mu} \end{array} \middle| \frac{-p_0}{1-p_0} \right) \right]$$

(15)

is a lower bound for $Q^{(a)}$. *If* $p_0 < \dfrac{\lambda}{\lambda + \mu}$,

$$\Psi(a) := Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda} \left[\frac{\pi \csc\left(\frac{\pi \log a}{\log \lambda - \log \mu}\right)}{(\log \lambda - \log \mu)\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \lambda - \log \mu}}} - \frac{1}{a \log a} {}_2F_1\left(\begin{array}{c} 1, \frac{\log a}{\log \lambda - \log \mu} \\ 1 + \frac{\log a}{\log \lambda - \log \mu} \end{array} \middle| \frac{p_0 \mu}{(1-p_0)\lambda} \right) \right]$$

(16)

is an upper bound for $Q^{(a)}$.

Proof: Consider the function

$$g(x) = \frac{a^x}{cb^x + 1}$$

(17)

for $x > -1$ with $a \in (0, 1)$, and let

$$b := \frac{\lambda}{\mu} > 1, \quad \text{and } c := \frac{p_0}{1 - p_0}.$$

(18)

By calculating the derivative, which is given by

$$g'(x) = \frac{a^x(\log a + cb^x(\log a - \log b))}{(1 + cb^x)^2}$$

and obviously always negative, we recognize that g is monotone decreasing as a function in x. Thus, we obtain the estimate

$$\int_0^\infty g(x)dx \leq \sum_{\tau=0}^\infty g(\tau) \leq \int_0^\infty g(x-1)dx = \int_{-1}^\infty g(x)dx.$$

(19)

Since, we have

$$Q^{(a)} = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda} \sum_{\tau=0}^{\infty} g(\tau),$$

it follows from (19) that

$$\psi(a) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda} \int_0^{\infty} g(x)\mathrm{d}x \tag{20}$$

is a lower bound for $Q^{(a)}$, and that

$$\Psi(a) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda} \int_{-1}^{\infty} g(x)\mathrm{d}x \tag{21}$$

is an upper bound for $Q^{(a)}$.

We will now manipulate the above equations. First, rewrite g as

$$g(x) = \frac{e^{x\xi}}{1 + ce^{x\eta}}, \tag{22}$$

where

$$\xi := \log a \ \text{ and } \ \eta := \log b. \tag{23}$$

Assume first that $ce^{x\eta} < 1$. Hence, g is a geometric series, can be rewritten as

$$g(x) = e^{x\xi} \sum_{k=0}^{\infty} (-ce^{x\eta})^k = \sum_{k=0}^{\infty} (-c)^k e^{x(\xi + k\eta)}, \tag{24}$$

and converges absolutely for all x satisfying $ce^{x\eta} < 1$. If $ce^{x\eta} > 1$, we have $\dfrac{1}{ce^{x\eta}} < 1$ and can expand (22) into the power series

$$g(x) = \frac{e^{x\xi}}{ce^{x\eta}} \frac{1}{1 - \frac{-1}{ce^{x\eta}}} = \frac{e^{x\xi}}{ce^{x\eta}} \sum_{k=0}^{\infty} (-ce^{x\eta})^{-k} = -\sum_{k=0}^{\infty} (-c)^{-(k+1)} e^{x(-\eta(k+1) + \xi)},$$

which converges absolutely for all x satisfying $ce^{x\eta} > 1$.

Let us further define $x_0 := -\dfrac{\log c}{\log b}$. The assumption $p_0 < \dfrac{1}{2}$ guarantees $x_0 > 0$. The absolute convergence of the above series enables us to integrate by summands, i.e., for $x < x_0$ we have

$$G_1(x) = \int g(x)\mathrm{d}x = \int \sum_{k=0}^{\infty}(-c)^k e^{x(\eta k+\xi)}\mathrm{d}x = \sum_{k=0}^{\infty}\int(-c)^k e^{x(\eta k+\xi)}\mathrm{d}x \tag{25a}$$

$$\begin{aligned} &= \sum_{k=0}^{\infty}(-c)^k \frac{e^{x(\eta k+\xi)}}{\eta k+\xi} = \frac{e^{x\xi}}{\eta}\sum_{k=0}^{\infty}(-c)^k \frac{e^{x\eta k}}{k+\frac{\xi}{\eta}} \\ &= \frac{e^{x\xi}}{\xi}\sum_{k=0}^{\infty}\frac{\frac{\xi}{\eta}}{k+\frac{\xi}{\eta}}(-ce^{x\eta})^k \end{aligned} \tag{25b}$$

$$= \frac{e^{x\xi}}{\xi}\sum_{k=0}^{\infty}\frac{\left(\frac{\xi}{\eta}\right)_{\overline{k}}(1)_{\overline{k}}}{\left(1+\frac{\xi}{\eta}\right)_{\overline{k}}}\frac{(-ce^{x\eta})^k}{k!} = \frac{e^{x\xi}}{\xi}{}_2\mathrm{F}_1\left(\begin{array}{c}1,\frac{\xi}{\eta}\\1+\frac{\xi}{\eta}\end{array}\bigg|-ce^{x\eta}\right). \tag{25c}$$

Note that we have used the fact that $\frac{\xi}{\eta} \neq -k$ for all $k \in \mathbb{N}$ in (25b), which is guaranteed by the assumption $a \neq \left(\frac{\mu}{\lambda}\right)^k$ for all $k \in \mathbb{N}$.

Similarly, for $x > x_0$ we have

$$\begin{aligned} G_2(x) &= \int g(x)\mathrm{d}x = \int -\sum_{k=0}^{\infty}(-c)^{-(k+1)}e^{x(-\eta(k+1)+\xi)}\mathrm{d}x \\ &= -\sum_{k=0}^{\infty}\int(-c)^{-(k+1)}e^{x(-\eta(k+1)+\xi)}\mathrm{d}x \end{aligned} \tag{26a}$$

$$\begin{aligned} &= -\sum_{k=0}^{\infty}(-c)^{-(k+1)}\frac{e^{x(-\eta(k+1)+\xi)}}{-\eta(k+1)+\xi} \\ &= \frac{e^{x\xi}}{\eta}\sum_{k=0}^{\infty}(-c)^{-(k+1)}\frac{e^{-x\eta(k+1)}}{-(k+1)+\frac{\xi}{\eta}} \end{aligned} \tag{26b}$$

$$= \frac{-e^{x\xi}}{\eta}\sum_{k=1}^{\infty}(-c)^{-k}\frac{e^{-x\eta k}}{-k+\frac{\xi}{\eta}} = \frac{-e^{x\xi}}{\xi}\sum_{k=1}^{\infty}\frac{-\frac{\xi}{\eta}}{k-\frac{\xi}{\eta}}\left(\frac{-1}{ce^{x\eta}}\right)^k \tag{26c}$$

$$\begin{aligned} &= \frac{-e^{x\xi}}{\xi}\sum_{k=1}^{\infty}\frac{\left(-\frac{\xi}{\eta}\right)_{\overline{k}}(1)_{\overline{k}}}{\left(1-\frac{\xi}{\eta}\right)_{\overline{k}}}\frac{\left(\frac{-1}{ce^{x\eta}}\right)^k}{k!} \\ &= \frac{e^{x\xi}}{\xi}\left[1-{}_2\mathrm{F}_1\left(\begin{array}{c}1,-\frac{\xi}{\eta}\\1-\frac{\xi}{\eta}\end{array}\bigg|-\frac{1}{ce^{x\eta}}\right)\right]. \end{aligned} \tag{26d}$$

It should be mentioned that we did not need the assumption $a \neq \left(\frac{\mu}{\lambda}\right)^k$ for all $k \in \mathbb{N}$ in (26).

Next, assume $p_0 < \frac{1}{2}$ i.e., $x_0 > 0$. Hence, we have

$$
\begin{aligned}
\int_0^\infty g(x)\mathrm{d}x &= \int_0^{x_0} g(x)\mathrm{d}x + \int_{x_0}^\infty g(x)\mathrm{d}x \\
&= \lim_{x\to x_0-} G_1(x) - G_1(0) + \lim_{x\to+\infty} G_2(x) - \lim_{x\to x_0+} G_2(x).
\end{aligned}
$$

Since $\frac{1}{k+\frac{\xi}{\eta}} \to 0$, the Leibnitz criterium implies that

$$
G_1(x_0) = \frac{e^{-\frac{\xi \log c}{\eta}}}{\eta} \sum_{k=0}^\infty \frac{(-1)^k}{k+\frac{\xi}{\eta}}
\tag{27}
$$

converges, i.e., $-\infty < G_1(x_0) < \infty$. Furthermore, by a similar argument we obtain $-\infty < G_2(x_0) < \infty$. We also see that

$$
\begin{aligned}
|G_2(x)| &= \left| \frac{-e^{x\xi}}{\eta} \sum_{k=1}^\infty (-c)^{-k} \frac{e^{-x\eta k}}{-k+\frac{\xi}{\eta}} \right| \leq \frac{e^{x\xi}}{\eta} \sum_{k=1}^\infty \left|(-c)^{-k} \frac{e^{-x\eta k}}{-k+\frac{\xi}{\eta}}\right| \\
&= \frac{e^{x\xi}}{\eta} \sum_{k=1}^\infty \left|\frac{1}{ce^{x\eta}}\right|^k \frac{1}{k-\frac{\xi}{\eta}} \leq \frac{e^{x\xi}}{\eta} \sum_{k=1}^\infty \left|\frac{1}{ce^{x\eta}}\right|^k = \frac{e^{x\xi}}{\eta(1-ce^{x\eta})}.
\end{aligned}
$$

Hence, for $x > x_0$ this implies that the series representation (26d) of $G_2(x)$ is absolute convergent. Therefore, we obtain

$$
\lim_{x\to x_0+} G_2(x) = G_2(x_0)
\tag{28}
$$

and

$$
\lim_{x\to+\infty} G_2(x) = \frac{0}{\xi}\left[1 - {}_2F_1\left(\begin{array}{c} 1, -\frac{\xi}{\eta} \\ 1 - \frac{\xi}{\eta} \end{array}\bigg| 0\right)\right] = 0.
\tag{29}
$$

Similarly, we see that $G_1(x)$ is absolute convergent for $x_0 < x$. Therefore, we have

$$
\lim_{x\to x_0-} G_1(x) = G_1(x_0).
\tag{30}
$$

We obtain

$$\int_0^\infty g(x)\mathrm{d}x \ = \ G_1(x_0) - G_1(0) - G_2(x_0)$$

$$= \ \frac{e^{-\frac{\xi \log c}{\eta}}}{\xi}{}_2F_1\left(\begin{array}{c}1,\frac{\xi}{\eta}\\1+\frac{\xi}{\eta}\end{array}\bigg|-1\right) - \frac{1}{\xi}{}_2F_1\left(\begin{array}{c}1,\frac{\xi}{\eta}\\1+\frac{\xi}{\eta}\end{array}\bigg|-c\right) - \frac{e^{-\frac{\xi \log c}{\eta}}}{\xi}\left[1 - {}_2F_1\left(\begin{array}{c}1,-\frac{\xi}{\eta}\\1-\frac{\xi}{\eta}\end{array}\bigg|-1\right)\right].$$

(31)

If $p_0 < \frac{\lambda}{\lambda+\mu}$, we have $x_0 > -1$. The above calculations are still valid and yield

$$\int_{-1}^\infty g(x)\mathrm{d}x \ = \ G_1(x_0) - G_1(-1) - G_2(x_0)$$

$$= \ \frac{e^{-\frac{\xi \log c}{\eta}}}{\xi}{}_2F_1\left(\begin{array}{c}1,\frac{\xi}{\eta}\\1+\frac{\xi}{\eta}\end{array}\bigg|-1\right) - \frac{e^{-\xi}}{\xi}{}_2F_1\left(\begin{array}{c}1,\frac{\xi}{\eta}\\1+\frac{\xi}{\eta}\end{array}\bigg|-ce^{-\eta}\right) - \frac{e^{-\frac{\xi \log c}{\eta}}}{\xi}\left[1 - {}_2F_1\left(\begin{array}{c}1,-\frac{\xi}{\eta}\\1-\frac{\xi}{\eta}\end{array}\bigg|-1\right)\right].$$

(32)

Let us now simplify (31) and (32). Note that

$$A \ := \ {}_2F_1\left(\begin{array}{c}1,\kappa\\1+\kappa\end{array}\bigg|-1\right) + {}_2F_1\left(\begin{array}{c}1,-\kappa\\1-\kappa\end{array}\bigg|-1\right)$$

$$= \ \sum_{k=0}^\infty \frac{(1)_{\overline{k}}(\kappa)_{\overline{k}}}{(\kappa+1)_{\overline{k}}}\frac{(-1)^k}{k!} + \sum_{k=0}^\infty \frac{(1)_{\overline{k}}(-\kappa)_{\overline{k}}}{(-\kappa+1)_{\overline{k}}}\frac{(-1)^k}{k!}$$

$$= \ \sum_{k=0}^\infty \frac{(-1)^k \kappa}{\kappa+k} + \sum_{k=0}^\infty \frac{-(-1)^k \kappa}{-\kappa+k} = \sum_{k=0}^\infty \frac{2\kappa^2(-1)^k}{k^2-\kappa^2}$$

$$= \ 2\kappa^2\pi^2 \sum_{k=0}^\infty \frac{(-1)^k}{\kappa^2\pi^2-k^2\pi^2}.$$

By using the the well known series

$$\csc(z) = -\frac{1}{z} + 2z\sum_{k=0}^\infty \frac{(-1)^k}{z^2-k^2\pi^2} \ \text{ for } \frac{z}{\pi} \notin \mathbb{Z},$$

we arrive at

$$A = \pi\kappa\csc(\pi\kappa) - 1.$$

Because $a \neq \left(\frac{\mu}{\lambda}\right)^k$ for $k \in \mathbb{N}$ we have $\frac{\xi}{\eta} \notin \mathbb{Z}^+$. Because, we assume $a \in [0,1]$ and $\lambda > \mu$ we also have $a \neq \left(\frac{\mu}{\lambda}\right)^k$ for $k \in \mathbb{Z}^-$. Hence, we have $\frac{\xi}{\eta} \notin \mathbb{Z}$. Therefore, we obtain

$$\int_0^\infty g(x)\mathrm{d}x = \frac{\pi \csc(\pi \frac{\xi}{\eta})}{\eta c^{\frac{\xi}{\eta}}} - \frac{1}{\xi}{}_2\mathrm{F}_1\left(\begin{array}{c} 1, \frac{\xi}{\eta} \\ 1+\frac{\xi}{\eta} \end{array} \middle| -c \right)$$

and

$$\int_{-1}^\infty g(x)\mathrm{d}x = \frac{\pi \csc(\pi \frac{\xi}{\eta})}{\eta c^{\frac{\xi}{\eta}}} - \frac{e^{-\xi}}{\xi}{}_2\mathrm{F}_1\left(\begin{array}{c} 1, \frac{\xi}{\eta} \\ 1+\frac{\xi}{\eta} \end{array} \middle| -ce^{-\eta} \right).$$

Combining the above with (20) and (21) and the definitions of ξ, η, and c immediately yields (15) and (16).

The assumptions $p_0 < \frac{1}{2}$, or $p_0 < \frac{\lambda}{\lambda+\mu}$ hold whenever one aims to study the hitchhiking effect of a mutation that is initially rare. If the initial frequency of the mutation is not rare because one is interesting in studying mutations from standing genetic variation that become beneficial, for instance, because of a change in the environment, i.e., if one wants to study soft selective sweeps (cf. Hermisson and Pennings 2005), Theorem 1 is not applicable. The study of soft selective sweeps is relevant and hence we shall also treat the cases $p_0 \geq \frac{1}{2}$ and $p_0 \geq \frac{\lambda}{\lambda+\mu}$.

Remark 1: The proof of Theorem 1 reveals that we have to replace (15) by

$$\psi(a) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda \log a}\left[1 - {}_2\mathrm{F}_1\left(\begin{array}{c} 1, \frac{\log a}{\log \mu - \log \lambda} \\ 1+\frac{\log a}{\log \mu - \log \lambda} \end{array} \middle| -\frac{1-p_0}{p_0} \right)\right] \tag{33}$$

if $p_0 \geq \frac{1}{2}$. Moreover, we have to replace (16) by

$$\Psi(a) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda a \log a}\left[1 - {}_2\mathrm{F}_1\left(\begin{array}{c} 1, \frac{\log a}{\log \mu - \log \lambda} \\ 1+\frac{\log a}{\log \mu - \log \lambda} \end{array} \middle| -\frac{(1-p_0)\lambda}{p_0\mu} \right)\right] \tag{34}$$

if $p_0 \geq \frac{\lambda}{\lambda+\mu}$. Notably, (33) and (34) also hold if $a = \left(\frac{\lambda}{\mu}\right)^k$ for $k \in \mathbb{N}$. Furthermore, (33) and (34) are the continuations of (15) and (16) as functions in p_0.

Theorem 1 assumes $a \neq \left(\frac{\lambda}{\mu}\right)^k$ for all $k \in \mathbb{N}$. If this assumption is violated we have to make the following adjustments.

Remark 2: Assume $l \in \mathbb{N}$. If $p_0 < \frac{1}{2}$,

$$\psi\left(\left(\frac{\mu}{\lambda}\right)^l\right) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda(\log\mu - \log\lambda)}\left[\sum_{k=0}^{l-1}\frac{(-1)^k}{k-l}\left(\frac{p_0}{1-p_0}\right)^k - \left(\frac{-p_0}{1-p_0}\right)^l\log p_0\right],$$

(35)

is a lower bound for $Q^{((\frac{\mu}{\lambda})^l)}$. If $p_0 < \frac{\lambda}{\lambda+\mu}$,

$$\Psi\left(\left(\frac{\mu}{\lambda}\right)^l\right) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda(\log\mu - \log\lambda)}\left[\sum_{k=0}^{l-1}\frac{(-1)^k}{k-l}\left(\frac{\mu}{\lambda}\right)^{k-l}\left(\frac{p_0}{1-p_0}\right)^k - \left(\frac{-p_0}{1-p_0}\right)^l\log\left(\frac{p_0\mu}{p_0\mu + (1-p_0)\lambda}\right)\right]$$

(36)

is an upper bound for $Q^{((\frac{\mu}{\lambda})^l)}$. Moreover, (35) and (36) are the continuations of (15) and (16) in the limit $a \to \left(\frac{\mu}{\lambda}\right)^l$. If $p_0 \geq \frac{1}{2}$, (35) has to be replaced by (33). If $p_0 \geq \frac{\lambda}{\lambda+\mu}$, (36) has to be replaced by (34).

The proof can be found in Appendix A.1.

Let

$$\tilde{\Psi} := \Psi\left(\tilde{\Lambda}\right), \quad \overline{\Psi} := \Psi\left(\overline{\Lambda}\right), \quad \text{and} \quad \underline{\Psi} := \Psi\left(\underline{\Lambda}\right)$$

(37)

and analogously

$$\tilde{\psi} := \psi\left(\tilde{\Lambda}\right), \quad \overline{\psi} := \psi\left(\overline{\Lambda}\right), \quad \text{and} \quad \underline{\psi} := \psi\left(\underline{\Lambda}\right).$$

(38)

We immediately obtain the following corollary.

Corollary 1—*We have $\tilde{\Psi} \geq \overline{\Psi} \geq \overline{Q} \geq \hat{Q} \geq \underline{\psi}$, hence $\tilde{\Psi}$ and $\overline{\Psi}$ are upper bounds for $\hat{Q}$, and $\underline{\psi}$ is a lower bound for $\hat{Q}$. Moreover, we have $\tilde{\Psi} \geq \tilde{Q} \geq \tilde{\psi}$, $\overline{\Psi} \geq \overline{Q} \geq \overline{\psi}$, and $\underline{\Psi} \geq \underline{Q} \geq \underline{\psi}$.*

Figure 1 illustrates $\hat{Q}$ and its bounds $\underline{Q}$ and $\overline{Q}$, as well as their bounds $\overline{\Psi}$, $\underline{\Psi}$, $\overline{\psi}$, and $\underline{\psi}$ for various parameters. It becomes obvious that the upper bound $\overline{Q}$ is very close to $\hat{Q}$, whereas the lower bound $\underline{Q}$ greatly underestimates $\hat{Q}$.

It turns out that $\tilde{Q}$ and $\overline{Q}$ are almost identical unless p_0 is large. The same holds for for $\tilde{\Psi}$ and $\overline{\Psi}$. This is illustrated in Fig. 2.

Although, $\Psi(a)$ and $\psi(a)$ have closed expressions, these expressions involve the hypergeometric function. In the following we shall derive approximations for $\Psi(a)$ and $\psi(a)$ that make no use of the hypergeometric function. Our first step is the following theorem.

Theorem 2—_For $p_0 < \frac{1}{2}$ let_

$$\phi(a) := Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda}\left[\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log\mu - \log\lambda}}\left(\frac{1}{\log a} - \frac{1}{\log a + \log\lambda - \log\mu} + \frac{1}{\log a - \log\lambda + \log\mu}\right) - \frac{1}{\log a} + \frac{1}{(\log a + \log\lambda}\right.$$

(39)

and, for $p_0 < \frac{\lambda}{\lambda+\mu}$ let

$$\Phi(a) := Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda}\left[\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log\mu - \log\lambda}}\left(\frac{1}{\log a} - \frac{1}{\log a + \log\lambda - \log\mu} + \frac{1}{\log a - \log\lambda + \log\mu}\right) - \frac{1}{a\log a} + \frac{1}{a(\log a + l}\right.$$

(40)

_If $p_0 < \frac{1}{2}$, $\phi(a) \approx \psi(a)$ and if $p_0 < \frac{\lambda}{\lambda+\mu}$, $\Phi(a) \approx \Psi(a)$._

The proof can be found in Appendix A.1.

If $p_0 \geq \frac{1}{2}$, or $p_0 \geq \frac{\lambda}{\lambda+\mu}$ we have to modify the definitions of $\phi(a)$, or $\Phi(a)$, respectively.

<u>Remark 3:</u> For $p_0 \geq \frac{1}{2}$ let

$$\phi(a) := Q_0 - \frac{r\vartheta(R_0 - Q_0)(1 - p_0)}{\lambda p_0(\log\lambda - \log\mu - \log a)}$$

(41)

and, for $p_0 \geq \frac{\lambda}{\lambda+\mu}$ let

$$\Phi(a) := Q_0 - \frac{r\vartheta(R_0 - Q_0)(1 - p_0)}{\mu p_0(\log\lambda - \log\mu - \log a)}.$$

(42)

If $p_0 \geq \frac{1}{2}$, $\phi(a) \approx \psi(a)$ and if $p_0 \geq \frac{\lambda}{\lambda+\mu}$, $\Phi(a) \approx \Psi(a)$. Moreover, (41) and (42) are the continuations of (39) and (40) regarded as functions in p_0.

The proof can be found in Appendix A.1.

From Theorem 2, Remark 3, and the definitions of $\underline{\Lambda}$, $\overline{\Lambda}$, and $\tilde{\Lambda}$ we immediately obtain the following corollary.

Corollary 2—*Let* Ψ, ψ, Φ *and* ϕ *be defined as in Theorems* 1 *and* 2, *and Remarks* 1–3. *Furthermore, let*

$$\tilde{\Phi} := \Phi(\tilde{\Lambda}), \quad \overline{\Phi} := \Phi(\overline{\Lambda}), \quad \underline{\Phi} := \Phi(\underline{\Lambda}) \tag{43}$$

and

$$\tilde{\phi} := \phi(\tilde{\Lambda}), \quad \overline{\phi} := \phi(\overline{\Lambda}), \quad \underline{\phi} := \phi(\underline{\Lambda}). \tag{44}$$

Then $\tilde{\Phi} \approx \tilde{\Psi}$, $\overline{\Phi} \approx \overline{\Psi}$, $\underline{\Phi} \approx \underline{\Psi}$, $\tilde{\phi} \approx \tilde{\psi}$, $\overline{\phi} \approx \overline{\psi}$, *and* $\underline{\phi} \approx \underline{\psi}$.

The approximations provided in Corollary 2 are very accurate unless p_0 becomes too large. This is illustrated in Fig. 3.

Note that we derived the upper and lower bounds for $Q^{(a)}$ by using the estimate (19). However, we can use (19) for the standard estimate

$$\sum_{k=0}^{\infty} g(k) \approx \frac{1}{2} \int_{-1}^{\infty} g(x)\mathrm{d}x + \frac{1}{2} \int_{0}^{\infty} g(x)\mathrm{d}x. \tag{45}$$

Hence, the estimates summarized in the next corollary follow immediately.

Corollary 3—*Estimates for* $\hat{Q}$ *are given by*

$$\overline{Q}^* := \frac{1}{2}(\overline{\Psi} + \overline{\psi}) \quad \text{and} \quad \tilde{Q}^* := \frac{1}{2}(\tilde{\Psi} + \tilde{\psi}), \tag{46}$$

and

$$\overline{Q}^{\circ} := \frac{1}{2}(\overline{\Phi} + \overline{\phi}) \quad \text{and} \quad \tilde{Q}^{\circ} = \frac{1}{2}(\tilde{\Phi} + \tilde{\phi}). \tag{47}$$

The approximations of Corollary 3 are illustrated in Fig. 4. It is obvious that the above approximations are very accurate for various parameters.

Although we already arrived at relatively simple approximations for $\hat{Q}$, they are still difficult to interpret in terms of the involved parameters. In the following we will concentrate on the case in which the initial frequency p_0 of the resistant allele is small. This is the most relevant case for studying genetic hitchhiking.

3.2 Approximations for rare mutations

So far in our derivations we did not assume that the initial frequency p_0 of the resistant allele is small. However, since this is the biologically most relevant situations we shall derive further approximations for the equilibrium frequencies of the neutral alleles under the

assumption that the mutation is initially rare, i.e., $p_0 \approx 0$ and $1 - p_0 \approx 1$. Since we have $\tilde{\Lambda} \approx \overline{\Lambda}$ under this assumption we focus on estimates based on $\tilde{\Lambda}$.

For $p_0 \approx 0$ we obtain the following theorem.

Theorem 3—*Let*

$$A(a) := Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda} \left[\frac{1}{\log a} \left(\left(\frac{p_0}{1 - p_0} \right)^{\frac{\log a}{\log \mu - \log \lambda}} - 1 \right) - \frac{1 - a}{2a \log a} \right].$$
(48)

For $p_0 \approx 0$ we have $Q^\circ(a) \approx A(a)$.

The proof can be found in Appendix A.1.

In accordance to out previous notation we set

$$\overline{A} := A(\overline{\Lambda}) \quad \text{and} \quad \tilde{A} := A(\tilde{\Lambda}).$$
(49)

These approximations are illustrated in Fig. 5.

Now, we shall additionally assume that $a \approx 1$.

Theorem 4—*Let $a(x) = 1 - rx$ with $x \in [\frac{\vartheta}{\lambda}, \frac{\vartheta}{\mu}]$ and*

$$B(x) := Q_0 + \frac{\vartheta(R_0 - Q_0)}{x\lambda} \left(1 - p_0^{\frac{rx}{\log \lambda - \log \mu}} \right).$$
(50)

For $p_0 \approx 0$ and $rx \approx 0$ we have $Q^\circ(a(x)) \approx B(x)$.

<u>Proof:</u> We have $\log a(x) \approx -rx$. Hence,

$$
\begin{aligned}
Q^\circ(a(x)) &\approx Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda} \left[\frac{1}{-rx} \left(\left(\frac{p_0}{1 - p_0} \right)^{\frac{rx}{\log \lambda - \log \mu}} - 1 \right) + \frac{rx}{2(1 - rx)rx} \right] \\
&= Q_0 + \frac{\vartheta(R_0 - Q_0)}{x\lambda} \left[1 - \left(\frac{p_0}{1 - p_0} \right)^{\frac{rx}{\log \lambda - \log \mu}} + \frac{rx}{2(1 - rx)} \right].
\end{aligned}
$$

If $rx \approx 0$, we can neglect $\dfrac{rx}{2(1 - rx)}$ and obtain

$$Q^\circ(a(x)) \approx Q_0 + \frac{\vartheta(R_0 - Q_0)}{x\lambda} \left(1 - \left(\frac{p_0}{1 - p_0} \right)^{\frac{rx}{\log \lambda - \log \mu}} \right).$$

Moreover, because $p_0 \approx 0$ we have $\dfrac{p_0}{1 - p_0} \approx p_0$. Thus,

$$Q^\circ(a(x)) \approx Q_0 + \frac{\vartheta(R_0 - Q_0)}{x\lambda} \left(1 - p_0^{\frac{rx}{\log \lambda - \log \mu}} \right).$$

Note that we have $a(\frac{\vartheta(p_0\lambda+(1-p_0)\mu)}{\lambda\mu})=\overline{\Lambda}$ and $a(\frac{\vartheta}{\lambda})=\tilde{\Lambda}$. Let us define

$$\overline{B}:=B\left(\frac{\vartheta(p_0\lambda+(1-p_0)\mu)}{\lambda\mu}\right) \text{ and } \tilde{B}:=B\left(\frac{\vartheta}{\lambda}\right).$$

(51)

We obtain the following corollary.

Corollary 4—*We have*

$$\tilde{B}=R_0-(R_0-Q_0)p_0^{\frac{r\vartheta}{\lambda(\log\lambda-\log\mu)}}.$$

(52)

For $p_0 \approx 0$ and $\frac{r\vartheta}{\lambda} \approx 0$, we have $\tilde{Q}^\circ \approx \tilde{B}$.

Figure 5 illustrates the approximations $\overline{A}, \tilde{A}, \overline{B}, \tilde{B}$. It is easily seen that the approximations are very accurate unless p_0 is too large ($p_0 \gtrsim 0.1$). Anyway, the approximations are still acceptably accurate for $p_0 = 0.1$.

By expanding p_0^x into a Taylor series around $x = 0$, and by setting $x=\frac{r\vartheta}{\lambda(\log\lambda-\log\mu)}$ we obtain

$$\tilde{B}=R_0-(R_0-Q_0)\sum_{k=0}^{\infty}\frac{r^k}{k!}\left(\frac{\vartheta}{\lambda(\log\lambda-\log\mu)}\right)^k \log p_0.$$

(53)

Hence, for $r \approx 0$ we can further approximate this by neglecting terms of order $O(r^2)$ or $O(r^3)$ and higher, and obtain

$$\tilde{C}=Q_0-(R_0-Q_0)\frac{r\vartheta\log p_0}{\lambda(\log\lambda-\log\mu)}$$

(54)

and

$$\tilde{D}=Q_0-(R_0-Q_0)\frac{r\vartheta\log p_0}{\lambda(\log\lambda-\log\mu)}-(R_0-Q_0)\frac{r^2}{2}\left(\frac{\vartheta}{\lambda(\log\lambda-\log\mu)}\right)^2\log p_0.$$

(55)

Figure 6 shows $\hat{Q}$ along with its approximations $\tilde{C}$ and $\tilde{D}$. It is obvious that these approximations are only accurate for very small r. Not surprisingly $\tilde{D}$ is a much better approximation than the linear approximation $\tilde{C}$.

3.3 The hitchhiking effect for general *m*

Since the sensitive allele goes extinct in the population, the equilibrium frequency of the neutral allele N_1 is given by (7), i.e., by

$$\widehat{Q}^{(m)} = Q_0 + \frac{r(R_0 - Q_0)}{\lambda} \sum_{\tau=0}^{\infty} \frac{\vartheta_{p_\tau,m}}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^\tau + 1} \prod_{l=0}^{\tau-1} \Lambda_{m,l}.$$

The above formula has the same structure as in the case $m = 2$. If we set $m = 1$ there is no recombination and hence the initial proportions of the neutral allele among genotypes with the resistant allele remain constant over time. Hence, we have $\widehat{Q}^{(1)} = Q_0^{(1)}$.

We shall now derive approximations for $\widehat{Q}^{(m)}$. As easily seen from (7), we can apply our approximations from the case $m = 2$ if we make the approximation $\prod_{l=0}^{\tau-1} \Lambda_{m,l} \approx a^\tau$ and $\vartheta_{p_\tau,m} \approx b$. We first approximate $\vartheta_{p_\tau,m}$ by a constant. From (5b) we obtain

$$\vartheta_{p_\tau,m} = \sum_{k=0}^{m-2} \binom{m-2}{k} p_\tau^k (1-p_\tau)^{m-2-k} \times \left(\frac{\alpha(m-1)W_1^{(T)}W_3^{(T)}}{(m-k-1)W_1^{(T)} + (k+1)W_3^{(T)}} + \frac{(1-\alpha)(m-1)W_1^{(U)}W_3^{(U)}}{(m-k-1)W_1^{(U)} + (k+1)W_3^{(U)}} \right).$$

Let

$$\overline{\vartheta}^{(m)} := \max_{k=0,\ldots,m-2} \left\{ \frac{\alpha(m-1)W_1^{(T)}W_3^{(T)}}{(m-k-1)W_1^{(T)} + (k+1)W_3^{(T)}} + \frac{(1-\alpha)(m-1)W_1^{(U)}W_3^{(U)}}{(m-k-1)W_1^{(U)} + (k+1)W_3^{(U)}} \right\} \tag{56}$$

and

$$\underline{\vartheta}^{(m)} := \min_{k=0,\ldots,m-2} \left\{ \frac{\alpha(m-1)W_1^{(T)}W_3^{(T)}}{(m-k-1)W_1^{(T)} + (k+1)W_3^{(T)}} + \frac{(1-\alpha)(m-1)W_1^{(U)}W_3^{(U)}}{(m-k-1)W_1^{(U)} + (k+1)W_3^{(U)}} \right\}. \tag{57}$$

Hence, we have

$$\vartheta_{p_\tau,m} \leq \overline{\vartheta}^{(m)} \sum_{k=0}^{m-2} \binom{m-2}{k} p_\tau^k (1-p_\tau)^{m-2-k} = \overline{\vartheta}^{(m)},$$

where the last equality follows from the binomial formula. Similarly, we obtain $\vartheta_{p_\tau,m} \geq \underline{\vartheta}^{(m)}$.

Therefore, we obtain

$$1 - r\frac{\overline{\vartheta}^{(m)}}{\lambda\mu}\left(\frac{p_0\lambda^{t+1} + (1-p_0)\mu^{t+1}}{p_0\lambda^t + (1-p_0)\mu^t} \right) \leq \Lambda_{m,t} \leq 1 - r\frac{\underline{\vartheta}^{(m)}}{\lambda\mu}\left(\frac{p_0\lambda^{t+1} + (1-p_0)\mu^{t+1}}{p_0\lambda^t + (1-p_0)\mu^t} \right).$$

Using the same approximation as in the case $m = 2$ yields

$$\underline{\Lambda}^{(m)} := 1 - r\frac{\overline{\vartheta}^{(m)}}{\mu} \leq \Lambda_{m,t} \leq 1 - r\frac{\underline{\vartheta}^{(m)}}{\lambda\mu}(p_0\lambda + (1-p_0)\mu) =: \overline{\Lambda}^{(m)}.$$

Hence, by replacing ϑ and a by $\overline{\vartheta}^{(m)}$ and $\overline{\Lambda}^{(m)}$, or $\underline{\vartheta}^{(m)}$ and $\underline{\Lambda}^{(m)}$, respectively, we obtain the upper and lower bounds

$$\overline{Q}^{(m)} = Q_0 + \frac{r\overline{\vartheta}^{(m)}(R_0 - Q_0)}{\lambda} \sum_{k=0}^{\infty} \frac{\overline{\Lambda}^{(m)^k}}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^k + 1}$$

and

$$\underline{Q}^{(m)} = Q_0 + \frac{r\underline{\vartheta}^{(m)}(R_0 - Q_0)}{\lambda} \sum_{k=0}^{\infty} \frac{\underline{\Lambda}^{(m)^k}}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^k + 1},$$

respectively.

It turns out that the upper and lower bounds are very inaccurate estimates for $\hat{Q}^{(m)}$. However, it is possible to find relative accurate approximations for $\hat{Q}^{(m)}$, at least if p_0 is small, which is the most important case for our current purpose. If $p_0 \approx 0$, we have $\frac{p_0}{1 - p_0} \approx 0$, so that we obtain

$$
\begin{aligned}
\vartheta_{p_\tau, m} &\approx \sum_{k=0}^{m-2} \binom{m-2}{k}\left(\frac{p_0}{1-p_0}\right)^k \left(\frac{\lambda}{\mu}\right)^{t(k)} \\
&\times \left(\frac{\alpha(m-1)W_1^{(T)}W_3^{(T)}}{(m-k-1)W_1^{(T)}+(k+1)W_3^{(T)}} + \frac{(1-\alpha)(m-1)W_1^{(U)}W_3^{(U)}}{(m-k-1)W_1^{(U)}+(k+1)W_3^{(U)}} \right) \\
&\approx \frac{\alpha(m-1)W_1^{(T)}W_3^{(T)}}{(m-1)W_1^{(T)}+W_3^{(T)}} + \frac{(1-\alpha)(m-1)W_1^{(U)}W_3^{(U)}}{(m-1)W_1^{(U)}+W_3^{(U)}} =: \tilde{\vartheta}^{(m)}.
\end{aligned}
\tag{58}
$$

From (5c) we see that $\tilde{\vartheta}^{(m)} = \theta_{0,m}$ Moreover, define

$$\tilde{\Lambda}^{(m)} := 1 - \frac{r\tilde{\vartheta}^{(m)}}{\lambda}.$$

(59)

Hence, by approximating $\vartheta_{p_\tau, m}$ by $\tilde{\vartheta}^{(m)}$ and $\Lambda_{m,l}$ by $\tilde{\Lambda}^{(m)}$ we obtain

$$\widehat{Q}^{(m)} \approx \tilde{Q}^{(m)} = Q_0 + \frac{r\tilde{\vartheta}^{(m)}(R_0 - Q_0)}{\lambda} \sum_{\tau=0}^{\infty} \frac{(\tilde{\Lambda}^{(m)})^\tau}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^\tau + 1}.$$

(60)

This approximation turns out to be sufficiently accurate unless p_0 is too large.

Clearly, $\tilde{Q}^{(m)}$ can be further approximated as in the case $m = 2$. However, note that for $m > 2$ the assumption that p_0 is small is always implicitly made because otherwise the approximation (58) for $\vartheta_{p_\tau, m}$ will be inaccurate. Hence, only those approximations that were derived under the assumption of $p_0 \approx 0$ are meaningful if $m > 2$, i.e., those of Result 4 and Corollary 4. We can therefore summarize:

Result 2—*Let $m \geq 2$ and*

$$\tilde{B}^{(m)} := R_0 - (R_0 - Q_0) p_0^{\frac{r\tilde{\partial}^{(m)}}{\lambda(\log \lambda - \log \mu)}}.$$

(61)

Then we have $\hat{Q}^{(m)} \approx \tilde{B}^{(m)}$.

Note, that the approximations become worse for large m. However, as noted in Schneider and Kim (2010), the differences in $\hat{Q}^{(m)}$ for different values of m become small for large m. Since, in reality m should be bounded by a maximum possible value it should be sufficient to assume $m < 10$. For these values Result 2 is still accurate. This is illustrated in Fig. 7. Furthermore, note that Result 2 becomes delicate for large m, because it is based on the approximation (58), which becomes very inaccurate as $m \to \infty$.

4 The general model

Schneider and Kim (2010) also presented two generalizations of their model. First, they argued that the differentiation into treated and untreated hosts is oversimplified, and that host heterogeneity should be taken into account. Host heterogeneity can reflect for instance different levels of drug concentration, drug decay, different levels of host-acquired immunity, different immune responses etc. Second, they argued that it is oversimplified to assume that each host is infected by the same number of parasites, i.e., that m is a fixed parameter, and showed how the model can be generalized to the case in which m follows a given frequency distribution.

We shall briefly summarize the two generalizations, and show how our results are generalized.

4.1 Host heterogeneity

Assume again a fixed proportion α of hosts is treated. We divide the treated and untreated host into various different discrete "treated classes" and "untreated classes". Let the proportion of infected hosts that fall into class j ($j \in \mathbb{N}$) be α_j. Let $\mathrm{U} \subseteq \mathbb{N}$ and $\mathrm{T} = \mathbb{N} \backslash \mathrm{U}$ be the sets of treated and untreated classes, respectively. Hence, we have $\sum_{j \in \mathrm{U}} \alpha_j = 1 - \alpha$ and $\sum_{j \in \mathrm{T}} \alpha_j = \alpha$. Let us denote the fitnesses of parasite haplotypes carrying the sensitive and resistant allele in hosts that fall into class j by $W_1^{(j)}$ and $W_2^{(j)}$, and $W_3^{(j)}$ and $W_4^{(j)}$, respectively.

Note that in the original formulation of the model we have

$W_k^{(U)} = \sum_{j \in \mathrm{U}} \frac{\alpha_j}{1 - \alpha} W_k^{(j)} = \mathrm{E}[W_k^{(j)} | U]$, and $W_k^{(T)} = \sum_{j \in \mathrm{T}} \frac{\alpha_j}{\alpha} W_k^{(j)} = \mathrm{E}[W_k^{(j)} | T]$. Hence, it can be regarded as the approximation that all untreated hosts are subsumed in just one class, in which $W_k^{(U)}$ is the mean fitness of the parasites among them and all treated hosts are subsumed in just one class, in which $\hat{W}_k^{(T)}$ is the mean fitness of the parasites among them.

In the model accounting for host heterogeneity the equilibrium frequency of the neutral allele N_1 is still given by (7), however the parameters λ, μ, are given by

$$\lambda = \sum_{j=0}^{\infty} \alpha_j W_3^{(j)} = \mathrm{E}\left[W_3^{(j)}\right] = (1 - \alpha)\,\mathrm{E}\left[W_3^{(j)} | \mathrm{U}\right] + \alpha \mathrm{E}\left[W_3^{(j)} | \mathrm{T}\right]$$

(62a)

and

$$\mu = \sum_{j=0}^{\infty} \alpha_j W_1^{(j)} = \mathrm{E}\left[W_1^{(j)}\right] = (1-\alpha)\,\mathrm{E}\left[W_1^{(j)}|\mathrm{U}\right] + \alpha\mathrm{E}\left[W_1^{(j)}|\mathrm{T}\right],$$

(62b)

whereas $\vartheta_{p,m}$ and $\Lambda_{m,t}$ are still given by (5b) and (8), however in (5b) $\theta_{k,m}$ has to be replaced by

$$\theta_{k,m} = (m-1)\sum_{j=0}^{\infty} \frac{\alpha_j W_1^{(j)} W_3^{(j)}}{(m-1-k)\,W_1^{(j)} + (k+1)\,W_3^{(j)}}.$$

(63)

Hence, in all approximations $\tilde{\vartheta}^{(m)}$ has to be replaced by

$$\tilde{\vartheta}^{(m)} = \sum_{j=0}^{\infty} \frac{\alpha_j\,(m-1)\,W_1^{(j)} W_3^{(j)}}{(m-1)\,W_1^{(j)} + W_3^{(j)}}.$$

(64)

We summarize:

Result 3—*Let $m \geq 2$. The equilibrium frequency of the neutral allele N_1, $\hat{Q}^{(m)}$, is approximately given by*

$$\tilde{B}^{(m)} := R_0 - (R_0 - Q_0)p_0^{\frac{r\tilde{\vartheta}^{(m)}}{\lambda(\log\lambda - \log\mu)}},$$

(65)

where λ, μ, and $\tilde{\vartheta}^{(m)}$ are given by (62a), (62b), *and* (64), *respectively.*

It was mentioned in Schneider and Kim (2010) that accounting for host heterogeneity results in a more pronounced hitchhiking compared to the basic model if the values of λ and μ coincide, or more precisely if $W_k^{(U)} = \mathrm{E}[W_k^{(j)}|\mathrm{U}]$, and $W_k^{(T)} = \mathrm{E}[W_k^{(j)}|\mathrm{T}]$. The reason is that under this assumption $\tilde{\vartheta}^{(m)}$ given by (64) is smaller than if it is given by (58), and (for fixed m) the respective approximations for the hitchhiking effect corresponds to standard hitchhiking with recombination rate $\dfrac{\tilde{\vartheta}^{(m)} r}{\lambda}$ and initial frequency $\dfrac{\mu p_0}{\lambda(1-p_0) + \mu p_0}$. We will prove this in the Sect. 5.

4.2 Number of co-infections

Now, assume that the number m of parasites that infect a host follows some probability distribution over the population. Let κ_m denote the probability that a host is infected by $m \geq 1$ parasites. Naturally, we have $\sum_{m=1}^{\infty} \kappa_m = 1$.

As shown in Schneider and Kim (2010), the equilibrium frequency of the neutral allele N_1 is given by

$$\widehat{Q}^{(\kappa)} = Q_0 + \frac{r(R_0 - Q_0)}{\lambda} \sum_{\tau=0}^{\infty} \frac{\vartheta_{p_\tau,\kappa}}{\frac{p_0}{1-p_0}\left(\frac{\lambda}{\mu}\right)^\tau + 1} \prod_{l=0}^{\tau-1} \Lambda_{\kappa,l},$$

(66a)

with

$$\vartheta_{p_\tau,\kappa} = \sum_{m=1}^{\infty} \kappa_m \vartheta_{p_\tau,m} \quad \text{and} \quad \Lambda_{\kappa,t} = \sum_{m=1}^{\infty} \kappa_m \Lambda_{m,t},$$

(67a)

where $\vartheta_{p_\tau,m}$ and $\Lambda_{m,t}$ are given by (5b) and (8) (eventually with λ, μ, $\vartheta_{p,m}$, $\Lambda_{m,t}$, and $\theta_{k,m}$ adjusted as in Sect. 4.1 to incorporate host heterogeneity). Hence, the equilibrium frequency of the neutral allele N_1 can be approximated following the calculations of Sect. 3. We can summarize this in the following result.

Result 4—*The equilibrium frequency of the neutral allele N_1, $\hat{Q}^{(\kappa)}$ is approximately given by*

$$\tilde{Q}^{(\kappa)} \approx \tilde{B}^{(\kappa)} := R_0 - (R_0 - Q_0) p_0^{\frac{r\tilde{\vartheta}^{(\kappa)}}{\lambda(\log \lambda - \log \mu)}}$$

(68)

with $\tilde{\vartheta}_\kappa = \sum_{m=1}^{\infty} \kappa_m \tilde{\vartheta}_m$, where λ, μ, and $\tilde{\vartheta}^{(m)}$ are given by (4), and (58), respectively. If one wants to incorporate host heterogeneity, λ, μ, and $\tilde{\vartheta}^{(m)}$ are given by (62), and (64), respectively.

5 Equilibrium heterozygosity and the hitchhiking effect

From now on we assume that m follows a probability distribution as in Sect. 4.2 and we account for host heterogeneity as in Sect. 4.1 unless otherwise mentioned.

Remember that the equilibrium heterozygozity is given by

$$\widehat{H} = 2\widehat{Q}^{(\kappa)}(1 - \widehat{Q}^{(\kappa)}).$$

(69)

We have seen in the last section that the equilibrium frequency, and its upper and lower bounds and approximations are given by expressions of the form $Q_0 + (R_0 - Q_0) A(r)$, where A is a function of the recombination rate that has to be chosen appropriately. For brevity we will suppress the dependence of $\hat{Q}^{(\kappa)}$ on r unless necessary. Let us regard R_0, i.e., the initial frequency of N_1 among sensitive parasites, as a random variable and the heterozygosity as a function of R_0. Let us write

$$\widehat{H} = \widehat{H}(R_0) = 2(Q_0 + (R_0 - Q_0)A(r))(1 - Q_0 - (R_0 - Q_0)A(r)).$$

(70)

Given the initial frequency of the neutral allele N_1 is R_0, the beneficial mutation occurs initially in association with allele N_1 with probability R_0, and in association with allele N_2 with probability $(1 - R_0)$. Hence, we have $Q_0 = 1$ with probability R_0 and $Q_0 = 0$ with probability $(1 - R_0)$. Therefore, the average heterozygosity given R_0 is calculated to be

$$
\begin{aligned}
\mathrm{E}(\widehat{H}|R_0) &= 2R_0(1+(R_0-1)\,A(r))(Q_0-R_0)\,A(r)+2(1-R_0)R_0\,A(r)(1-R_0\,A(r)) \\
&= (R_0-R_0^2)(2A(r)-A^2(r)).
\end{aligned}
\tag{71}
$$

Hence, according to the theorem of total probability, the average heterozygosity is calculated to be

$$
\mathrm{E}(\widehat{H})=\mathrm{E}\left(\mathrm{E}\left(\widehat{H}|R_0\right)\right)=2\left(\mathrm{E}(R_0)-\mathrm{E}\left(R_0^2\right)\right)\,(2A(r)-A^2(r)).
\tag{72}
$$

The initial heterozygosity is given by $H_0 = 2R_0(1 - R_0)$. Hence, the fraction of the expected equilibrium heterozygosity over the initial heterozygosity is given by

$$
H(r)=\frac{\mathrm{E}(\widehat{H})}{\mathrm{E}(H_0)}=2A(r)-A^2(r),
\tag{73}
$$

which is independent of the distribution of R_0. Since $A(0) = 0$ we have $H(0) = 0$.

If we further approximate $\widehat{Q}$ by $\tilde{B}$ given by (52), we obtain

$$
H^{(\kappa)}(r) \approx \tilde{H}^{(\kappa)}=1-p_0^{\frac{2r\tilde{\vartheta}^{(\kappa)}}{\lambda(\log \lambda-\log \mu)}}.
\tag{74}
$$

From (74) it is seen that $H^{(\kappa)}$ shows a strong genome-wide reduction if κ_1 is large and selection is sufficiently strong. The approximation is illustrated in Fig. 8.

For the special case that m is constant the above reduces to $H^{(m)}(r) \approx \tilde{H}^{(m)}=1-p_0^{\frac{2r\tilde{\vartheta}^{(m)}}{\lambda(\log \lambda-\log \mu)}}$ which is increasing as a function of m. The reason is that $\tilde{\vartheta}^{(m)}$ is monotone increasing in m, either if it is given by (58) or in the case of host heterogeneity by (64).

It was mentioned in Schneider and Kim (2010) that accounting for host heterogeneity results in a more pronounced hitchhiking compared to the basic model if the values of λ and μ coincide, or, more precisely, if $W_k^{(U)}=\mathrm{E}[\,W_k^{(j)}|U]$ and $W_k^{(T)}=\mathrm{E}[\,W_k^{(j)}|T]$. The reason is that, for every m, under these assumptions $\tilde{\vartheta}^{(m)}$ given by (64) is smaller than if it is given by (58). Hence, this holds also for the corresponding values of $\tilde{\vartheta}^{(\kappa)}$, and consequently (74) is smaller if host heterogeneity is incorporated. We shall summarize this as a remark and prove it in the appendix.

Remark 4

Host heterogeneity leads to an increased hitchhiking effect, i.e., to a stronger reduction in relative heterozygosity $\tilde{H}^{(\kappa)}(r)$, compared to the basic model with only one class of treated and one class of untreated hosts with corresponding fitness parameters, i.e., $W_k^{(U)}=\mathrm{E}[\,W_k^{(j)}|U]$, and $W_k^{(T)}=\mathrm{E}[\,W_k^{(j)}|T]$.

For fixed large m the differences in $\tilde{H}^{(m)}$ become very small and vanish in the limit $m \to \infty$ because $\tilde{H}^{(m)}$ approaches the classical approximation for standard hitchhiking.

Note, that the last statement of the remark is delicate, because the approximation $\tilde{H}^{(m)}$ will be inaccurate for very large m.

We can use (73), or (74) to calculate the maximum recombination distance for which a given reduction in relative heterozygosity can be observed. This is relevant for predicting the width of the valley of relative heterozygosity for given selection parameters. By comparison of (73), or (74) with empirical data on the relative heterozygosity, it is possible to re-evaluate or validate parameter estimates (e.g., for α, λ, μ, etc.).

From (73) we obtain that the maximum recombination rate, $\hat{r}$, for which the relative heterozygosity is smaller than β by solving the equation $H(\hat{r}) = \beta$. If we solve this equation first with respect to $A(\hat{r})$, we obtain

$$A(\widehat{r}) = 1 - \sqrt{1 - \beta^2}.$$ (75)

By using the approximation according to (74), we obtain

$$\widehat{r} = \frac{\lambda \log (1 - \beta^2)(\log \lambda - \log \mu)}{2\tilde{\vartheta}^{(\kappa)} \log p_0}.$$ (76)

Figure 9 illustrates the valley of reduced heterozygosity as a function of α for different distributions κ of m. Such illustrations can be used to determine the range of parameters that lead to a given reduction in relative heterozygosity.

6 Discussion

We obtained a closed-form approximation for the expected heterozygosity shaped by genetic hitchhiking in the model of antimalarial drug-resistance evolution proposed in Schneider and Kim (2010). This model aims to capture the effect of multiple infections per host (m) and drug-treatment rate (α), which are considered the most important epidemiological parameters that characterize geographic differences in the dynamics of drug resistance (Escalante et al. 2009), as well as the complex malaria transmission cycle on the pattern of selective sweeps caused by drug-resistant mutations (Daily 2006; Prugnolle et al. 2009). Due to the model complexity the exact solution is expressed by the summation of infinite (or a very large number of) terms. Here, we provided numerous approximate solutions with varying degree of accuracy. Notably, using the assumption that the starting frequency of the resistant allele under positive directional selection is low, we could obtain a solution that is simple enough to allow clear biological interpretations regarding the effects of epidemiological parameters. Furthermore, our approximations are flexible enough to incorporate arbitrary distributions of hosts with different infection rates and/or host heterogeneity, e.g., arbitrary distributions of hosts with different drug concentrations. The latter condition, arising due to a slow decay of antimalarial drugs in the bodies of treated patients, was demonstrated to be crucial for the initiation of drug resistance evolution (Hastings et al. 2002). For more discussion on how this model can be used to predict the spread of resistance and its speed, and how it can be used to design 'optimal' treatment strategies we refer to Schneider and Kim (2010).

The mean fitness of the resistant parasites, λ, and of the sensitive parasites, μ, are crucial in the considered model. If λ and μ are not too different, $\log \lambda - \log \mu$ corresponds to the selection coefficient of the beneficial (resistant) mutation. Then, our approximation (74) is basically identical to that under the standard model of hitchhiking obtained by Maynard

Smith and Haigh (1974) except the modifying factors of the recombination rate, $\dfrac{\tilde{\vartheta}^{(\kappa)}}{\lambda}$, and of

the initial frequency, $\dfrac{\mu}{\lambda(1 - p_0) + \mu p_0}$. Therefore, the dynamics of hitchhiking unique to this

malaria model is summarized by these factors. We have $\dfrac{\tilde{\vartheta}^{(\kappa)}}{\lambda} < 1$ because $\dfrac{(m - 1)\,W_1^{(j)}}{(m - 1)\,W_1^{(j)} + W_3^{(j)}} < 1$

for all host class j and all m. The latter fraction represents the probability that a given
resistant gametocyte pairs with a sensitive gametocyte in the body of a mosquito which took
its blood meal from a j host, when the frequency of resistant allele is low. (As with the
standard hitchhiking model, the final heterozygosity is predominantly determined by the
dynamics of the resistant allele at its early stage.) If the drug in the host is very effective

($W_1^{(j)} \ll W_3^{(j)}$), it will greatly reduce $\dfrac{\tilde{\vartheta}^{(\kappa)}}{\lambda}$. Namely, the strength of selection determines the
effective rate of recombination (decay of association between beneficial and neutral allele),
unlike the case of standard hitchhiking model in which the two factors are decoupled.
Moreover, the approximation (74) will be more accurate even for larger r and p_0 compared

to the standard hitchhiking model because of the two adjustment factors $\dfrac{\tilde{\vartheta}^{(\kappa)}}{\lambda}$ and

$\dfrac{\mu}{\lambda(1 - p_0) + \mu p_0}$.

Assume m is constant. It is also obvious from the above that the number, m, of independent
parasite strains infecting a host determines the effective recombination rate. A small value of
m thus increases the hitchhiking effect. In the extreme case of $m = 1$, genetic variation in the
population is completely wiped out as parasites reproduce effectively asexually. On the
other hand, with $m \to \infty$ our approximation approaches that of standard hitchhiking. Note,

however, that in the exact solution $\dfrac{\vartheta^{(m)}}{\lambda}$ is still less than one, even with $m \to \infty$, because the
approximation leading to $\tilde{\vartheta}^{(m)}$ (Eq. 58) assumes that the frequency of the resistant allele is
low and hence that the probability that a given host is infected by two or more resistant
strains is negligible. However, if m increases to a large number, such chance is no longer
negligible. Therefore, the combined effect of strong drug pressure and the limited number of
clones in hosts can greatly amplify the effect of genetic hitchhiking beyond the level
predicted by the standard model. If m is not constant, the hitchhiking effect is more
pronounced for more left-skewed distributions κ. It will be in particular pronounced if a
large fraction of single infections occurs.

Our approximation also reveals another important departure of our model from the previous
models that assume random mating and homogeneous selective pressures. In the standard
model, the allele-frequency trajectory of a beneficial mutation is necessary and sufficient
information for predicting its hitchhiking effect on the linked neutral variation (Betancourt
et al. 2004; Chevin and Hospital 2008). For example, the speed with which the beneficial
mutation increases in the population determines the size of the genomic regions affected by
hitchhiking. However, in our model with host heterogeneity, different combinations of

parameter values (j and $W_i^{(j)}$) that specify the same allele frequency trajectory of the
resistance allele may generate different hitchhiking effects. Schneider and Kim (2010)
showed that the changes of sensitive and resistant allele frequencies are uniquely determined
by their absolute fitness, μ and λ, respectively. With host heterogeneity, the fitness is simply

the mean of $W_1^{(j)}$ or $W_3^{(j)}$ weighted by the frequencies of host classes (α_j). The modifying

factor of effective recombination rate, $\frac{\tilde{\vartheta}^{(m)}}{\lambda}$ (ore more generally $\frac{\tilde{\vartheta}^{(\kappa)}}{\lambda}$), however is not a linear function of $W_1^{(j)}$ or $W_3^{(j)}$. As a result, as shown in Remark 4, $\frac{\tilde{\vartheta}^{(m)}}{\lambda}$ (or $\frac{\tilde{\vartheta}^{(\kappa)}}{\lambda}$)) decreases as one introduces host heterogeneity while keeping λ and μ constant. This makes our simplest approximation assuming no host heterogeneity a conservative predictor of the hitchhiking effect.

Comparison of approximate solutions with empirical observation of local reductions of genetic variation around the loci of drug resistance mutations, combined with other genetic (e.g. recombination rate, the frequency change of drug resistance) and epidemiological (e.g. the mean number of independent parasite clones per host) information, will greatly advance our understanding of antimalarial drug-resistance evolution. Especially, if empirical data for the reduction in heterozygosity is available, our results can be used to determine possible ranges for parameters that are unknown and/or infeasible to measure. It should be noted that the fraction λ/μ can be easily estimated from retrospective genetic data. Since

$\log \dfrac{p_t}{1 - p_t} = t \log \dfrac{\lambda}{\mu}$, $\log \lambda/\mu$ is just the slope of the linear regression of the logarithm of the ratio of resistant over sensitive parasites measured at different time points. However, it is difficult to scale time, because the number of transmission cycles per year is difficult to quantify. Moreover, also the parameter α should be easy to measure. On the opposite, the distribution of κ of m will be difficult to estimate. However, the distribution of m, and especially single infections, lead to a genome-wide reduction in relative heterozygosity if selection is sufficiently strong. This genome-wide reduction might be used to estimate the distribution κ. However, applying our results to real data lies beyond the scope of this article and will be accomplished in a follow-up paper.

Appendix A

A.1 Bounds and approximations

Proof of Remark 2

If $a = \left(\dfrac{\mu}{\lambda}\right)^l$ for $l \in \mathbb{N}$ we just need to adapt the proof of Theorem 1. The derivation of G_1 assumed $\dfrac{\xi}{\eta} \notin \mathbb{Z}$. If this assumption is violated we have to adjust the derivation. In this case we have $-\dfrac{\xi}{\eta} = l$ and we can replace (24) by

$$
\begin{aligned}
g(x) &= \frac{e^{-xl\eta}}{ce^{x\eta}+1} = e^{-xl\eta} \sum_{k=0}^{\infty} (-ce^{x\eta})^k = \sum_{k=0}^{\infty} (-c)^k e^{x\eta(k-l)} \\
&= (-c)^l \sum_{k=0}^{\infty} (-c)^{k-l} e^{x\eta(k-l)}.
\end{aligned}
$$

Hence, we have to replace (25) in the proof of Theorem 1 by

$$
\begin{aligned}
G_1(x) &= \int g(x)\,dx = \int (-c)^l \sum_{k=0}^{\infty}(-c)^{k-l}e^{x\eta(k-l)}dx \\
&= (-c)^l \sum_{k=0}^{\infty}\int (-c)^{k-l}e^{x\eta(k-l)}dx \\
&= (-c)^l\left(\sum_{k=0}^{l-1}\int(-c)^{k-l}e^{x\eta(k-l)}dx + \int dx + \sum_{k=l+1}^{\infty}\int(-c)^{k-l}e^{x\eta(k-l)}dx\right) \\
&= (-c)^l\left(\sum_{k=0}^{l-1}(-c)^{k-l}\frac{e^{x\eta(k-l)}}{\eta(k-l)}+x+\sum_{k=l+1}^{\infty}(-c)^{k-l}\frac{e^{x\eta(k-l)}}{\eta(k-l)}\right) \\
&= (-c)^l\left(\sum_{k=0}^{l-1}(-c)^{k-l}\frac{e^{x\eta(k-l)}}{\eta(k-l)}+x-\frac{1}{\eta}\sum_{k=1}^{\infty}(-c)^{k-1}\frac{e^{x\eta k}}{k}\right) \\
&= (-c)^l\left(\sum_{k=0}^{l-1}(-c)^{k-l}\frac{e^{x\eta(k-l)}}{\eta(k-l)}+x-\frac{\log(1+ce^{x\eta})}{\eta}\right),
\end{aligned}
$$

Thus, we have

$$
G_1(-1)=(-c)^l\left(\sum_{k=0}^{l-1}\frac{(-c)^{k-l}}{\eta(k-l)}-1-\frac{\log(1+ce^{-\eta})}{\eta}\right),
\tag{77}
$$

$$
G_1(0)=(-c)^l\left(\sum_{k=0}^{l-1}\frac{(-c)^{k-l}}{\eta(k-l)}-\frac{\log(1+c)}{\eta}\right),
\tag{78}
$$

and

$$
G_1(x_0)=(-c)^l\left(\sum_{k=0}^{l-1}\frac{(-1)^{(k-l)}}{\eta(k-l)}+x_0-\frac{\log 2}{\eta}\right).
\tag{79}
$$

Combining the above with (26), the definitions of ξ, η, c, (31), and (32) yields

$$
\begin{aligned}
\psi\left(\left(\tfrac{\mu}{\lambda}\right)^l\right) =\ & Q_0+\frac{r\vartheta(R_0-Q_0)}{\lambda} \\
& \times\Bigg[\frac{\left(\frac{p_0}{1-p_0}\right)^l}{\log\lambda-\log\mu}\left(\sum_{k=0}^{l-1}\frac{(-1)^k}{k-l}\left(1-\left(\frac{p_0}{1-p_0}\right)^{k-l}\right)-\log 2p_0\right) \\
& \quad -\frac{\left(\frac{p_0}{1-p_0}\right)^l}{l(\log\lambda-\log\mu)}\left(1-{}_2F_1\left(\begin{array}{c}1,l\\1+l\end{array}\bigg|-1\right)\right)\Bigg],
\end{aligned}
\tag{80}
$$

and

$$\Psi\left(\left(\frac{\mu}{\lambda}\right)^l\right) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda}$$
$$\times \left[\frac{\left(\frac{p_0}{1-p_0}\right)^l}{\log \lambda - \log \mu} \left(\sum_{k=0}^{l-1} \frac{(-1)^k}{k-l}\left(1 - \left(\frac{p_0\mu}{(1-p_0)\lambda}\right)^{k-l}\right)\right.\right.$$
$$\left. -\log\left(\frac{p_0\lambda}{p_0\mu + (1-p_0)\lambda}\right) + \log\lambda + \log\mu - \log 2\right)$$
$$\left. -\frac{\left(\frac{p_0}{1-p_0}\right)^l}{l(\log\lambda - \log\mu)}\left(1 - {}_2F_1\left(\begin{array}{c}1, l \\ 1+l\end{array}\bigg| -1\right)\right)\right],$$

(81)

respectively. Note that we have

$$1 - {}_2F_1\left(\begin{array}{c}1, l \\ 1+l\end{array}\bigg| -1\right) = 1 - \sum_{k=0}^{\infty} \frac{(1)_{\overline{k}}(l)_{\overline{k}}(-1)^k}{(l+1)_{\overline{k}}k!} = 1 - \sum_{k=0}^{\infty} \frac{l(-1)^k}{l+k}$$

(82)

$$= 1 - l(-1)^l\left(\sum_{k=1}^{\infty} \frac{(-1)^k}{k} - \sum_{k=1}^{l-1} \frac{(-1)^k}{k}\right)$$

(83)

$$= \frac{l(-1)^{2l}}{l} - l(-1)^l\left(\log 2 - \sum_{k=1}^{l-1} \frac{(-1)^k}{k}\right)$$

(84)

$$= l(-1)^l\left(\sum_{k=1}^{l} \frac{(-1)^k}{k} - \log 2\right)$$

(85)

$$= l\sum_{k=0}^{l-1} \frac{(-1)^k}{k-l} - l(-1)^l \log 2.$$

(86)

Hence, (80) and (81) simplify to

$$\psi\left(\left(\frac{\mu}{\lambda}\right)^l\right) = Q_0 + \frac{r\vartheta(R_0 - Q_0)}{\lambda}$$
$$\times \left[\frac{\left(\frac{p_0}{1-p_0}\right)^l}{\log \mu - \log \lambda} \left(\sum_{k=0}^{l-1} \frac{(-1)^k}{k-l}\left(\frac{p_0}{1-p_0}\right)^{k-l} - \log p_0\right)\right],$$

(87)

and

$$\Psi\left(\left(\frac{\mu}{\lambda}\right)^l\right) = Q_0 + \frac{r\vartheta\,(R_0 - Q_0)}{\lambda}\left[\frac{\left(\frac{p_0}{1-p_0}\right)^l}{\log\mu - \log\lambda}\left(\sum_{k=0}^{l-1}\frac{(-1)^k}{k-l}\left(\frac{p_0\mu}{(1-p_0)\lambda}\right)^{k-l} - \log\left(\frac{p_0\mu}{p_0\mu + (1-p_0)\lambda}\right)\right)\right],$$

(88)

which equal (36) and (35), respectively.

Now, let us regard $g(x)$ given by (17) as a function in x and a and write $g_a(x)$ for it. Clearly for $x \geq 0$, $g_a(x)$ is monotone increasing in a, and for $x < 0$ it is monotone decreasing. Since by Theorem 1 the integral $\int_{\tilde{x}}^{\infty} g_a(x)\mathrm{d}x$ exists for $\tilde{x} \in \{-1, 0\}$ and $a \neq \left(\frac{\mu}{\lambda}\right)^l$, it follows by the Theorems of monotone convergence and, in case $\tilde{x} = -1$, also by the theorem of dominated convergence that $\lim_{a \to a_0}\int_{\tilde{x}}^{\infty} g_a(x)\mathrm{d}x = \int_{\tilde{x}}^{\infty} g_{a_0}(x)\mathrm{d}x$. By setting $a_0 = \left(\frac{\mu}{\lambda}\right)^l$ it follows that (36) and (35) are the continuations of (16) and (15) in the limit $a \to \left(\frac{\mu}{\lambda}\right)^l$.

If $x_0 \leq 0$, we do not need the function G_1, and hence the derivations from Theorem 1 and Remark 1 hold for ψ. The same holds for Ψ if $x_0 \leq -1$. This finishes the proof.

Proof of Theorem 2

For $\kappa \neq -l(l \in \mathbb{N})$ we have

$$_2F_1\left(\begin{array}{c}1, \kappa \\ 1+\kappa\end{array}\bigg| z\right) = \sum_{k=0}^{\infty}\frac{(\kappa)_{\overline{k}}z^k}{(\kappa+1)_{\overline{k}}} = 1+\kappa\sum_{k=1}^{\infty}\frac{z^k}{\kappa+k} \approx 1+\frac{\kappa z}{\kappa+1}.$$

(89)

First assume $a \neq \left(\frac{\mu}{\lambda}\right)^l$ for $l \in \mathbb{N}$, i.e., $\frac{\xi}{\eta} \neq -l$.

Setting $\kappa = \frac{\xi}{\eta}$ and $z = -c$ gives

$$_2F_1\left(\begin{array}{c}1, \frac{\xi}{\eta} \\ 1+\frac{\xi}{\eta}\end{array}\bigg| -c\right) \approx 1 - \frac{c\xi}{\eta+\xi} = 1 - \frac{p_0\log a}{(\log a + \log\lambda - \log\mu)\,(1-p_0)},$$

(90)

whereas setting $\kappa = \frac{\xi}{\eta}$ and $z = -\frac{c}{e^\eta}$ gives

$$_2F_1\left(\begin{array}{c}1, \frac{\xi}{\eta} \\ 1+\frac{\xi}{\eta}\end{array}\bigg| -\frac{c}{e^\eta}\right) \approx 1 - \frac{c\xi}{e^\eta\,(\eta+\xi)} = 1 - \frac{p_0\mu\log a}{(\log a + \log\lambda - \log\mu)\,(1-p_0)\,\lambda}.$$

(91)

Moreover, setting $\kappa = \dfrac{-\xi}{\eta}$ and $z = -1$ yields

$$
{}_2F_1 \left(\begin{array}{c} 1, -\frac{\xi}{\eta} \\ 1 - \frac{\xi}{\eta} \end{array} \middle| -1 \right) \approx 1 - \frac{\xi}{\xi - \eta} = 1 - \frac{\log a}{\log a - \log \lambda + \log \mu},
\tag{92}
$$

whereas by setting $\kappa = \dfrac{\xi}{\eta}$ and $z = -1$ we obtain

$$
{}_2F_1 \left(\begin{array}{c} 1, \frac{\xi}{\eta} \\ 1 + \frac{\xi}{\eta} \end{array} \middle| -1 \right) \approx 1 - \frac{\xi}{\xi + \eta} = 1 - \frac{\log a}{\log a + \log \lambda - \log \mu}.
\tag{93}
$$

First, combining (93), (90), and (92) with (31), and using this approximation in (20) yields (39) by using the definitions of ξ and η. Similarly, we obtain (40) by combining (93), (91), (92), with (32) and (21).

Clearly, (39) and (40) are continuous especially also at $a = \left(\dfrac{\mu}{\lambda}\right)^l$. Since $\psi(a)$ and $\Psi(a)$ have continuations at $a = \left(\dfrac{\mu}{\lambda}\right)^l$ for ($l \in \mathbb{N}$) according to Remark 2, we have $\phi(a) \approx \psi(a)$ and $\Phi(a) \approx \Psi(a)$ for all a. This finishes the proof.

Proof of Remark 3

We obtain (41) by combining (33) with (89) for $\kappa = \dfrac{\log a}{\log \mu - \log \lambda}$ and $z = -\dfrac{1 - p_0}{p_0}$, whereas we obtain (42) by combining (34) with (89) for $\kappa = \dfrac{\log a}{\log \mu - \log \lambda}$ and $z = -\dfrac{(1 - p_0)\lambda}{p_0 \mu}$.

Clearly (39), (40), (41), and (42) are continuous functions in p_0. For $p_0 = \dfrac{1}{2}$ it is easily seen that (39) equals (41), whereas for $p_0 = \dfrac{\lambda}{\lambda + \mu}$ (40) equals (42). Hence, (41) and (42) are the continuations in p_0 of (39) and (40), respectively.

Proof of Theorem 3

Note that we have

$$
Q^\circ(a) = Q_0 + \frac{r \vartheta (R_0 - Q_0)}{\lambda} A,
\tag{94}
$$

where

$$
\begin{aligned}
A \ = \ & \left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}} \left(\frac{1}{\log a} - \frac{1}{\log \frac{a\lambda}{\mu}} + \frac{1}{\log \frac{a\mu}{\lambda}}\right) \\
& - \frac{1}{2a\log a} + \frac{p_0 \mu}{2(1-p_0)a\lambda \log \frac{a\lambda}{\mu}} - \frac{1}{2\log a} + \frac{p_0}{2(1-p_0)\log \frac{a\lambda}{\mu}} \\
= \ & \left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}} \left(\frac{1}{\log a} + \frac{1}{\log \frac{a\mu}{\lambda}}\right) - \frac{1}{\log a}\left(\frac{1+a}{2a}\right) \\
& - \left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}} \left(\frac{1}{\log \frac{a\lambda}{\mu}}\right) + \frac{p_0}{2(1-p_0)\log \frac{a\lambda}{\mu}}\left(\frac{\mu}{a\lambda}+1\right) \\
= \ & \left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}} \left(\frac{1}{\log a} + \frac{1}{\log \frac{a\mu}{\lambda}}\right) - \frac{1+a}{2a\log a} \\
& - \frac{p_0}{(1-p_0)\log \frac{a\lambda}{\mu}}\left(\left(\frac{p_0}{1-p_0}\right)^{\frac{\log \frac{a\lambda}{\mu}}{\log \mu - \log \lambda}} - 1\right) + \frac{p_0}{2(1-p_0)\log \frac{\mu}{a\lambda}}\left(\frac{\mu}{a\lambda}-1\right) \\
= \ & \frac{1}{\log a}\left(\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}} - 1\right) - \frac{p_0}{(1-p_0)\log \frac{a\lambda}{\mu}}\left(\left(\frac{p_0}{1-p_0}\right)^{\frac{\log \frac{a\lambda}{\mu}}{\log \mu - \log \lambda}} - 1\right) \\
& - \frac{1-a}{2a\log a} + \frac{1}{\log \frac{a\mu}{\lambda}}\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}} - \frac{p_0}{2(1-p_0)\log \frac{\mu}{a\lambda}}\left(\frac{\mu}{a\lambda}-1\right).
\end{aligned}
\tag{95}
$$

Furthermore, $\frac{x-1}{\log x}$ is monotone increasing in x and $\lim_{x\to 0}\frac{x-1}{\log x}=0$. Since $\lambda > \mu$ and $\frac{1}{2} \le a \le 1$, we have $0 < \frac{\mu}{a\lambda} \le 2$. Hence, we obtain

$$
0 \le \frac{\frac{\mu}{a\lambda} - 1}{\log \frac{\mu}{a\lambda}} \le \frac{1}{\log 2},
$$

and we have

$$
0 \le \frac{p_0}{2(1 - p_0)\log \frac{\mu}{a\lambda}}\left(\frac{\mu}{a\lambda} - 1\right) \le \frac{p_0}{2\log 2}(1+O(p_0)),
\tag{96}
$$

which follows from the fact that $\frac{1}{1 - p_0}$ can be written as a geometric series.

Therefore,

$$
\begin{aligned}
A \ \approx \ & \frac{1}{\log a}\left(\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}} - 1\right) - \frac{p_0}{(1-p_0)\log \frac{a\lambda}{\mu}} \\
& \times \left(\left(\frac{p_0}{1-p_0}\right)^{\frac{\log \frac{a\lambda}{\mu}}{\log \mu - \log \lambda}} - 1\right) - \frac{1-a}{2a\log a} + \frac{1}{\log \frac{a\mu}{\lambda}}\left(\frac{p_0}{1-p_0}\right)^{\frac{\log a}{\log \mu - \log \lambda}}.
\end{aligned}
\tag{97}
$$

For $C > 0$ we obtain $f(x) = \frac{C^x - 1}{x}$ is monotone increasing in x, because $f'(x) = \frac{1 - C^x + C^x \log C^x}{x^2} \ge 0$. Furthermore, we have $\lim_{x\to 0} f(x) = \log C$. By choosing $C = \left(\frac{p_0}{1 - p_0}\right)^{\frac{1}{\log \mu - \log \lambda}}$, we see that

$$\frac{p_0}{(1 - p_0) \log \frac{a\lambda}{\mu}} \left(\left(\frac{p_0}{1 - p_0} \right)^{\frac{\log \frac{a\lambda}{\mu}}{\log \mu - \log \lambda}} - 1 \right)$$

is negligible compared with

$$\frac{1}{\log a} \left(\left(\frac{p_0}{1 - p_0} \right)^{\frac{\log a}{\log \mu - \log \lambda}} - 1 \right).$$

(98)

Furthermore, also

$$\frac{1}{\log \frac{a\mu}{\lambda}} \left(\frac{p_0}{1 - p_0} \right)^{\frac{\log a}{\log \mu - \log \lambda}}$$

(99)

is negligible compared with (98). Therefore,

$$A \approx \frac{1}{\log a} \left(\left(\frac{p_0}{1 - p_0} \right)^{\frac{\log a}{\log \mu - \log \lambda}} - 1 \right) - \frac{1 - a}{2a \log a}.$$

(100)

A.2 Relative heterozygosity

Proof of Remark 4

Let $f(x, y) := \dfrac{1}{\frac{1}{x} + \frac{1}{y}}$ for $x, y \in \mathbb{R}^+$. Its Hessian matrix is calculated to be

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} \frac{-2y^2}{(x+y)^3} & \frac{2xy}{(x+y)^3} \\ \frac{2xy}{(x+y)^3} & \frac{-2x^2}{(x+y)^3} \end{pmatrix}.$$

Clearly, we have $\dfrac{-2y^2}{(x+y)^3} < 0$ and det $H = 0$, i.e., the leading minors of H are non-positive. Hence, f is concave but not strictly concave (note that $f(x, x) = x/2$). Hence, for positive random variables X and Y defined on a probability space (Ω, A, P) and a sub-σ algebra B the Jensen's inequality for higher dimensions yields

$$E[f(X, Y)|B] \leq f(E[X|B], E[Y|B]).$$

Now, choose $(\Omega, A, P) = (\mathbb{N}, P(\mathbb{N}), P)$, where $P(j) = \alpha_j$. Moreover, choose $X(j) = W_1^{(j)}, Y(j) = (m - 1)W_3^{(j)}$ and $B = \{\emptyset, U, T, \mathbb{N}\}$. Then the Jensen's, inequality gives

$$\sum_{j\in U}\frac{\alpha_j(m-1)W_1^{(j)}W_3^{(j)}}{(m-1)W_1^{(j)}+W_3^{(j)}} = (1-\alpha)\sum_{j\in U}\frac{\frac{\alpha_j}{1-\alpha}}{\frac{1}{W_1^{(j)}}+\frac{1}{(m-1)W_3^{(j)}}}=(1-\alpha)\,\mathrm{E}[\,f(X,Y)|U]$$

$$\leq (1-\alpha)f(\mathrm{E}[\,X|U],\mathrm{E}[\,Y|U])=\frac{(1-\alpha)}{\frac{1}{W_1^{(U)}}+\frac{1}{(m-1)W_3^{(U)}}}$$

$$= \frac{(1-\alpha)(m-1)W_1^{(U)}W_3^{(U)}}{(m-1)W_1^{(U)}+W_3^{(U)}}.$$

Similarly, we obtain

$$\sum_{j\in T}\frac{\alpha_j(m-1)W_1^{(j)}W_3^{(j)}}{(m-1)W_1^{(j)}+W_3^{(j)}} \leq \frac{\alpha(m-1)W_1^{(T)}W_3^{(T)}}{(m-1)W_1^{(T)}+W_3^{(T)}}.$$

Combination of the above yields

$$\sum_{j=0}^{\infty}\frac{\alpha_j(m-1)W_1^{(j)}W_3^{(j)}}{(m-1)W_1^{(j)}+W_3^{(j)}} \leq \frac{(1-\alpha)(m-1)W_1^{(U)}W_3^{(U)}}{(m-1)W_1^{(U)}+W_3^{(U)}}+\frac{\alpha(m-1)W_1^{(T)}W_3^{(T)}}{(m-1)W_1^{(T)}+W_3^{(T)}}, \tag{101}$$

i.e., that $\tilde{\vartheta}^{(m)}$ given by (64) is smaller than if it is given by (58).

Clearly in the limit $m\to\infty$ equality holds in (101). Moreover, $\tilde{\vartheta}^{(m)}\to\lambda$. Hence, in the limit case $H^{(m)}$ reduced to the approximation for standard hitchhiking.

The proof is finished by applying the argument formulated above the remark.

Acknowledgments

References

Barton, NH. Genetic hitchhiking; Philos Trans R Soc Lond B Biol Sci. 2000. p. 1553-1562.http://rstb.royalsocietypublishing.org/content/355/1403/1553.abstract

Betancourt, AJ.; Kim, Y.; Orr, HA. A pseudohitchhiking model of X vs. autosomal diversity; Genetics. 2004. p. 2261-2269.http://www.genetics.org/cgi/content/abstract/168/4/2261

Brooks, DR.; Wang, P.; Read, M.; Watkins, WM.; Sims, PFG.; Hyde, JE. Sequence variation of the hydroxymethyldihydropterin pyrophosphokinase: dihydropteroate synthase gene in lines of the human malaria parasite *Plasmodium falciparum* with differing resistance to sulfadoxine; Eur J Biochem. 1994. p. 397-405.http://dx.doi.org/10.1111/j.1432-1033.1994.00397.x

Chevin, L-M.; Hospital, F. Selective sweep at a quantitative trait locus in the presence of background genetic variation; Genetics. 2008. p. 1645-1660.http://www.genetics.org/cgi/content/abstract/180/3/1645

Cowman, AF.; Morry, MJ.; Biggs, BA.; Cross, GA.; Foote, SJ. Amino acid changes linked to pyrimethamine resistance in the dihydrofolate reductase-thymidylate synthase gene of *Plasmodium falciparum*; Proc Natl Acad Sci USA. 1988. p. 9109-9113.http://www.pnas.org/content/85/23/9109.abstract

Daily, JP. Antimalarial drug therapy: the role of parasite biology and drug resistance; J Clin Pharmacol. 2006. p. 1487-1497.http://jcp.sagepub.com/cgi/content/abstract/46/12/1487

Escalante, AA.; Smith, DL.; Kim, Y. The dynamics of mutations associated with anti-malarial drug resistance in plasmodium falciparum; Trends Parasitol. 2009. p. 557-563.http://www.sciencedirect.com/science/article/B6W7G-4XJ9BS6-1/2/04d13d69d2006be02c12ef0d051cc5c6

Hastings, IM. Response to: the puzzling links between malaria transmission level and drug resistance; Trends Parasitol. 2003. p. 160-161.http://www.sciencedirect.com/science/article/B6W7G-483SMKD-1/2/680a5da4f74f112c6e1b87c270a9921d

Hastings, IM. Complex dynamics and stability of resistance to antimalarial drugs; Parasitology. 2006. p. 615-624.http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=428221&fulltextType=RA&fileId=S0031182005009790

Hastings, IM.; Mackinnon, MJ. The emergence of drug-resistant malaria; Parasitology. 1998. p. 411-417.http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=24263&fulltextType=RA&fileId=S0031182098003291

Hastings, IM.; Watkins, WM.; White, NJ. The evolution of drug-resistant malaria: the role of drug elimination half-life; Philos Trans R Soc Lond B Biol Sci. 2002. p. 505-519.http://rstb.royalsocietypublishing.org/content/357/1420/505.abstract

Hedrick, PW. Hitchhiking: a Comparison of Linkage and Partial Selfing; Genetics. 1980. p. 791-808.http://www.genetics.org/cgi/content/abstract/94/3/791

Hermisson, J.; Pennings, PS. Soft sweeps: molecular population genetics of adaptation from standing genetic variation; Genetics. 2005. p. 2335-2352.http://www.genetics.org/cgi/content/abstract/169/4/2335

Kim, Y.; Stephan, W. Detecting a local signature of genetic hitchhiking along a recombining chromosome; Genetics. 2002. p. 765-777.http://www.genetics.org/cgi/content/abstract/160/2/765

Korenromp, E.; Miller, J.; Nahlen, B.; Wardlaw, T.; Young, M. World malaria report 2005. World Health Organization (WHO); Geneva: 2005.

Mackinnon, MJ.; Hastings, IM. The evolution of multiple drug resistance in malaria parasites; Trans R Soc Tropical Med Hygiene. 1998. p. 188-195.http://www.sciencedirect.com/science/article/B75GP-4BY31N7-MS/2/bfd1e32b1cf5fd0f23e10978a356bd36

Marsh, K. Malaria disaster in africa; Lancet. 1998. p. 924-924.http://www.sciencedirect.com/science/article/B6T1B-4FWV357-JX/2/c9facab67b868b8fddd289f050bfae19

Maynard Smith, J.; Haigh, J. The hitch-hiking effect of a favourable gene; Genet Res. 1974. p. 23-35.http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=1754360&fulltextType=RA&fileId=S0016672300014634

McCollum, AM.; Basco, LK.; Tahar, R.; Udhayakumar, V.; Escalante, AA. Hitchhiking and selective sweeps of *Plasmodium falciparum* sulfadoxine and pyrimethamine resistance alleles in a population from Central Africa; Antimicrob Agents Chemother. 2008. p. 4089-4097.http://aac.asm.org/cgi/content/abstract/52/11/4089

Mita, T.; Tanabe, K.; Takahashi, N.; Tsukahara, T.; Eto, H.; Dysoley, L.; Ohmae, H.; Kita, K.; Krudsood, S.; Looareesuwan, S.; Kaneko, A.; Bjorkman, A.; Kobayakawa, T. Independent evolution of pyrimethamine resistance in *Plasmodium falciparum* isolates in Melanesia; Antimicrob Agents Chemother. 2007. p. 1071-1077.http://aac.asm.org/cgi/content/abstract/51/3/1071

Nair, S.; Nash, D.; Sudimack, D.; Jaidee, A.; Barends, M.; Uhlemann, A-C.; Krishna, S.; Nosten, F.; Anderson, TJC. Recurrent gene amplification and soft selective sweeps during evolution of multidrug resistance in malaria parasites; Mol Biol Evol. 2007. p. 562-573.http://mbe.oxfordjournals.org/cgi/content/abstract/24/2/562

Nair, S.; Williams, JT.; Brockman, A.; Paiphun, L.; Mayxay, M.; Newton, PN.; Guthmann, J-P.; Smithuis, FM.; Hien, TT.; White, NJ.; Nosten, F.; Anderson, TJC. A selective sweep driven by pyrimethamine treatment in Southeast Asian malaria parasites; Mol Biol Evol. 2003. p. 1526-1536.http://mbe.oxfordjournals.org/cgi/content/abstract/20/9/1526

Nash, D.; Nair, S.; Mayxay, M.; Newton, PN.; Guthmann, J-P.; Nosten, F.; Anderson, TJ. Selection strength and hitchhiking around two anti-malarial resistance genes; Proc R Soc B Biol Sci. 2005. p. 1153-1161.http://rspb.royalsocietypublishing.org/content/272/1568/1153.abstract

Prugnolle, F.; Durand, P.; Renaud, F.; Rousset, F. Effective size of the hierarchically structured populations of the agent of malaria: a coalescent-based model; Heredity. 2009. p. 1-7.http://dx.doi.org/10.1038/hdy.2009.127

Schneider KA, Kim Y. An analytical model for genetic hitchhiking in malaria parasites caused by drug resistance. Theor Popul Biol. 2010

Schwobel, B.; Alifrangis, M.; Salanti, A.; Jelinek, T. Different mutation patterns of atovaquone resistance to *Plasmodium falciparum* in vitro and in vivo: rapid detection of codon 268 polymorphisms in the cytochrome b as potential in vivo resistance marker; Malaria J. 2003. p. 5http://www.malariajournal.com/content/2/1/5

Stephan, W.; Wiehe, THE.; Lenz, MW. The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory; Theor Popul Biol. 1992. p. 237-254.http://www.sciencedirect.com/science/article/B6WXD-4F1Y9N0-3M/2/1245281bba0c6b542457fdd75c343edf

Triglia, T.; Cowman, AF. Primary structure and expression of the dihydropteroate synthetase gene of *Plasmodium falciparum*; Proc Natl Acad Sci USA. 1994. p. 7149-7153.http://www.pnas.org/content/91/15/7149.abstract

WHO. WHO Expert Committee on malaria; World Health Organ Tech Rep Ser. 2000. p. 1-74.http://www.genetics.org/cgi/content/abstract/160/2/765

Wootton, JC.; Feng, X.; Ferdig, MT.; Cooper, RA.; Mu, J.; Baruch, DI.; Magill, AJ. Su X-z 07 2002 Genetic diversity and chloroquine selective sweeps *in Plasmodium falciparum*; Nature. p. 320-323.http://dx.doi.org/10.1038/nature00813
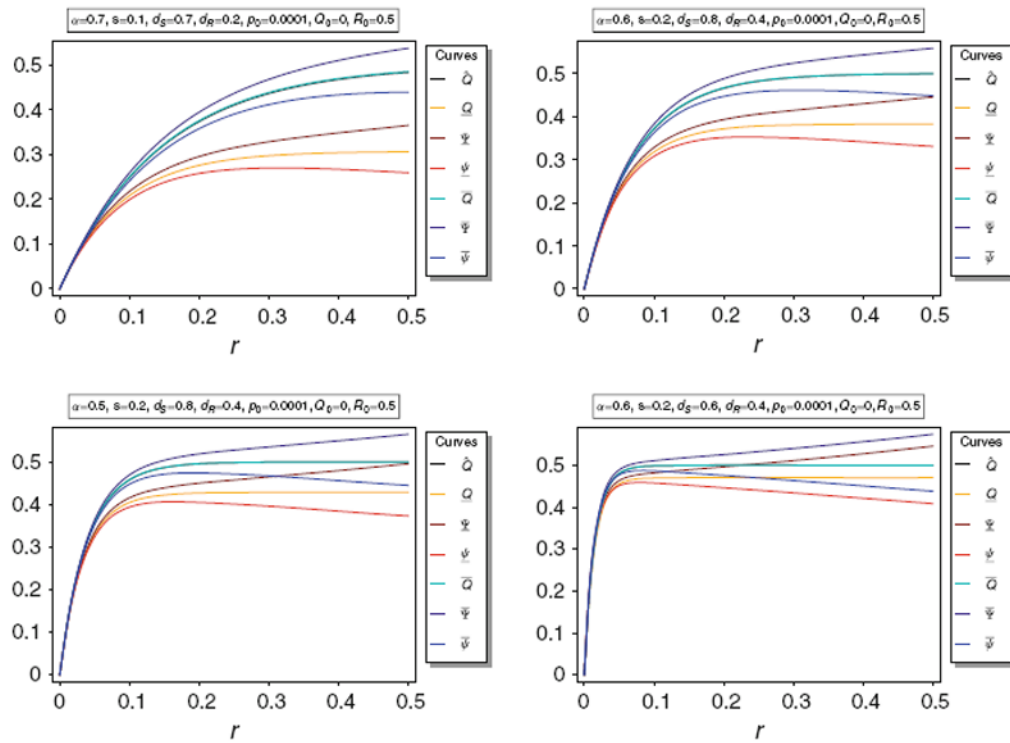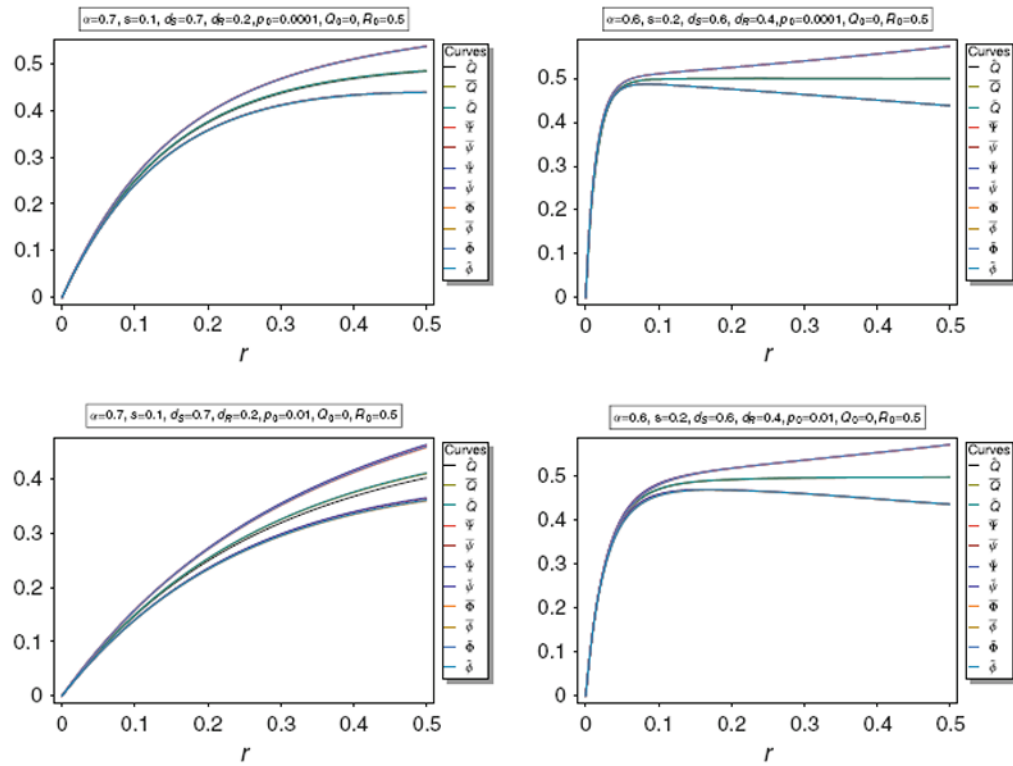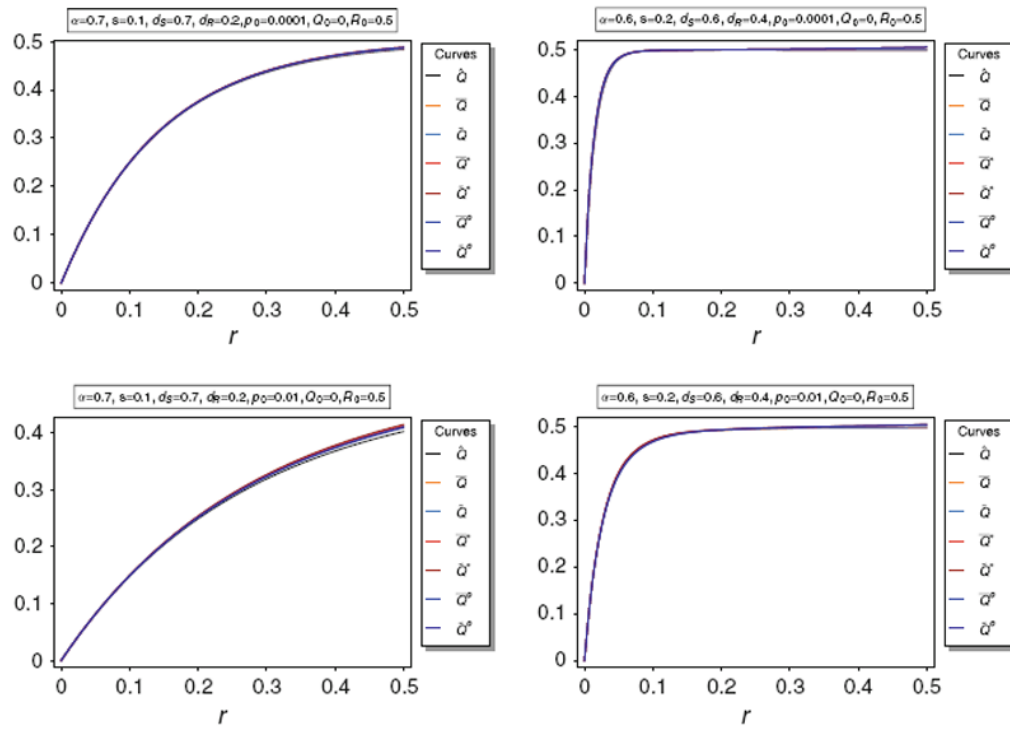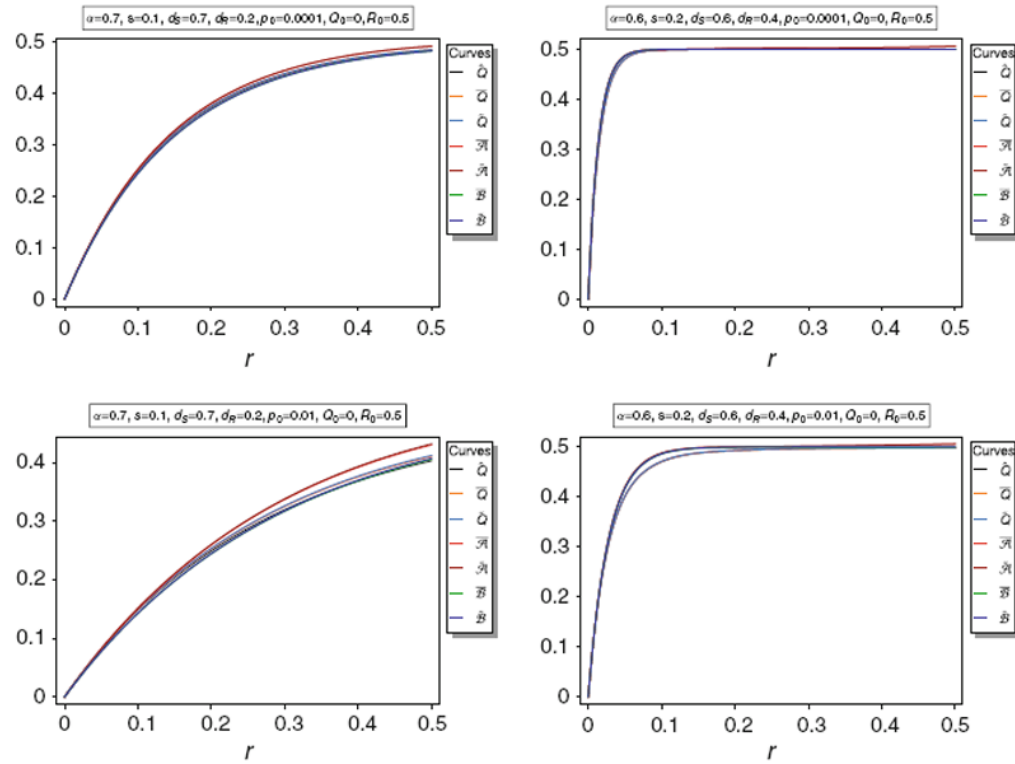
Fig. 1.
Equilibrium frequency $\hat{Q}$ of the neutral allele N_1 and various bounds for $\hat{Q}$ as a function of r for different parameter combinations. The panels show $\hat{Q}$ along with its upper and lower bounds $\overline{Q}$ and $\underline{Q}$, as well as their respective upper and lower bounds $\overline{\Psi}$ and $\overline{\psi}$, and $\underline{\Psi}$ and $\underline{\psi}$. The parameters for the various plot panels are specified in the *boxes* above the panels. In all panels $\hat{Q}$ and $\overline{Q}$ are almost identical
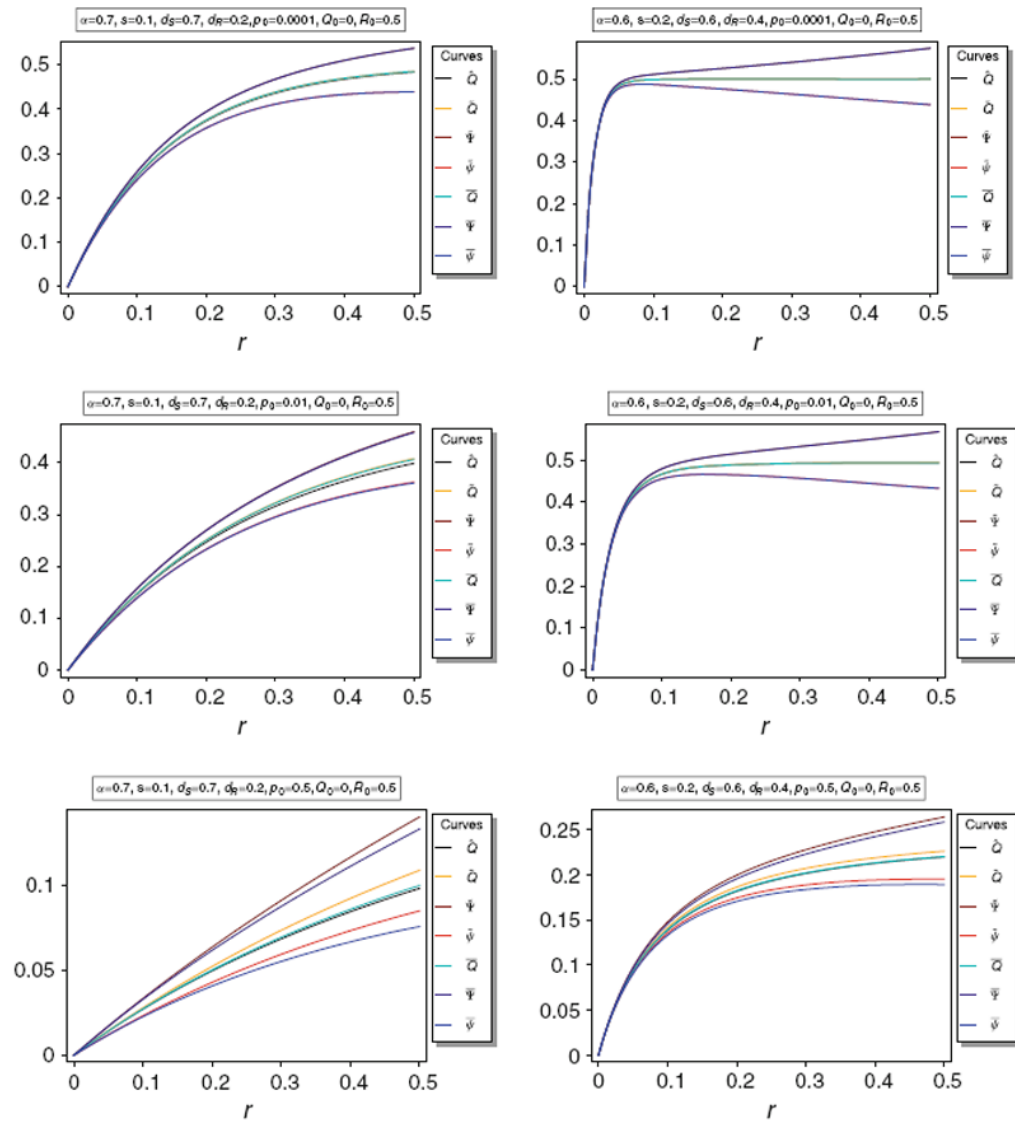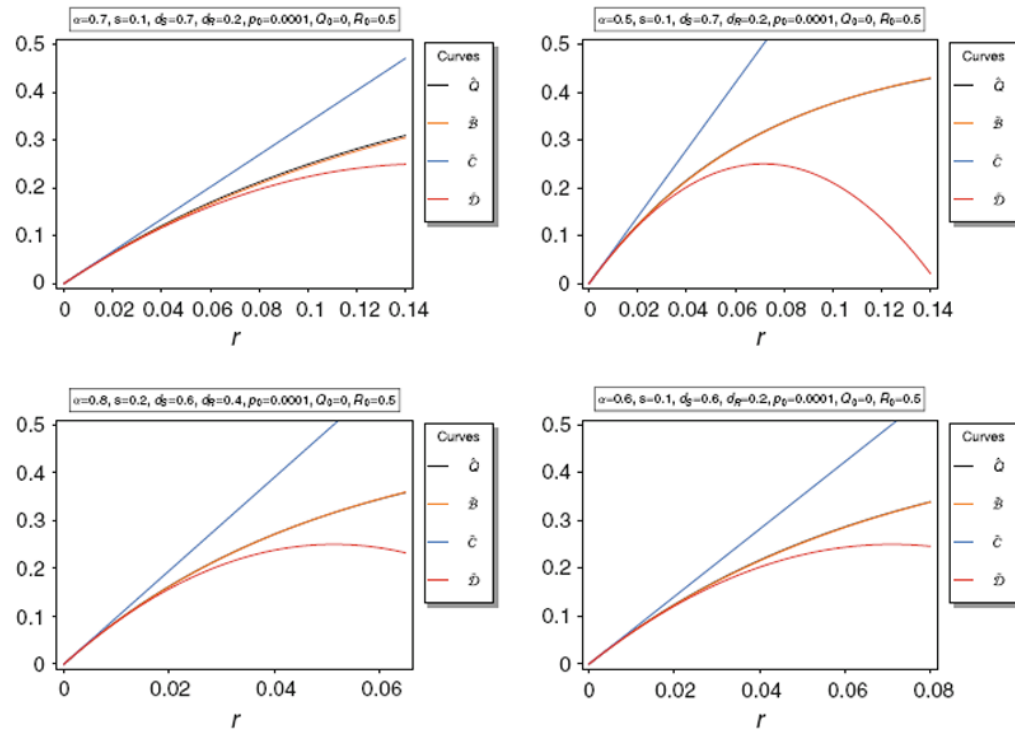
Fig. 2.
Equilibrium frequency $\hat{Q}$ of the neutral allele N_1 and various bounds of $\hat{Q}$ as a function of r for different parameter combinations. The panels show $\hat{Q}$ along with its upper bounds $\bar{Q}$ and $\tilde{Q}$, as well as their respective upper and lower bounds $\overline{\Psi}$ and $\overline{\psi}$, and $\underline{\Psi}$ and $\underline{\psi}$. The parameters for the various plot panels are specified in the *boxes* above the panels
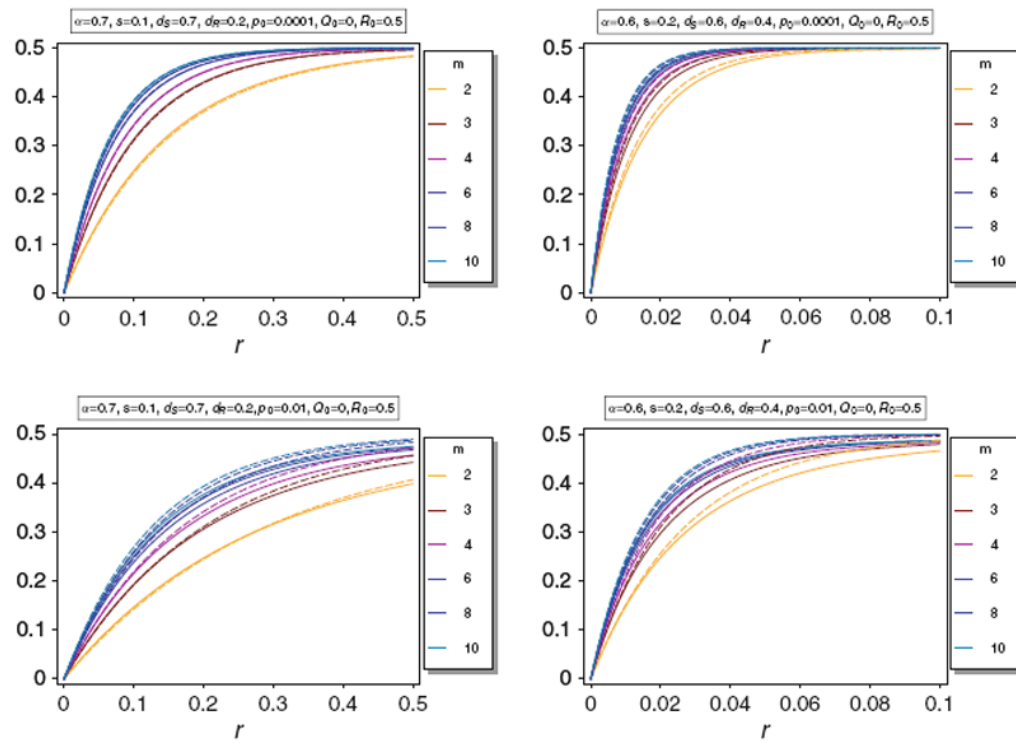
Fig. 3.

Equilibrium frequency $\hat{Q}$ of the neutral allele N_1 and various bounds of $\hat{Q}$ as a function of r for different parameter combinations. The panels show $\hat{Q}$ along with its upper bounds $\bar{Q}$ and $\tilde{Q}$, as well as their respective upper and lower bounds $\bar{\Psi}$ and $\bar{\psi}$, and $\tilde{\Psi}$ and $\tilde{\psi}$, along with their approximations $\bar{\Phi}$ and $\bar{\phi}$, and $\tilde{\Phi}$ and $\tilde{\phi}$, respectively. The parameters for the various plot panels are specified in the *boxes* above the panels

Fig. 4.
Equilibrium frequency $\hat{Q}$ of the neutral allele N_1 and various bounds of $\hat{Q}$ as a function of r for different parameter combinations. The panels show $\hat{Q}$ along with its upper bounds $\overline{Q}$ and $\tilde{Q}$, as well as its approximations $\overline{Q}^*$, $\tilde{Q}^*$, $\overline{Q}^\circ$, and $\tilde{Q}^\circ$. The parameters are the same as in Fig. 3

Fig. 5.
Equilibrium frequency $\hat{Q}$ of the neutral allele N_1 and various bounds of $\hat{Q}$ as a function of r for different parameter combinations. The panels show $\hat{Q}$ along with its upper bounds $\bar{Q}$ and $\tilde{Q}$, as well as the approximations $\bar{A}, \tilde{A}, \bar{B}, \tilde{B}$. The parameters are the same as in Fig. 3

Fig. 6.
Equilibrium frequency $\hat{Q}$ of the neutral allele N_1 and various bounds of $\hat{Q}$ as a function of r for various parameter combinations. The panels show $\hat{Q}$ along with its upper bounds $\bar{Q}$ and $\tilde{Q}$, as well as the approximations $\tilde{C}$ and D.

Fig. 7.
Equilibrium frequency $\hat{Q}^{(m)}$ (*thin lines*) of the neutral allele N_1 versus the approximation $B^{(m)}$ (*thick dashed lines*) for various values of m. The parameters are as in Fig. 3
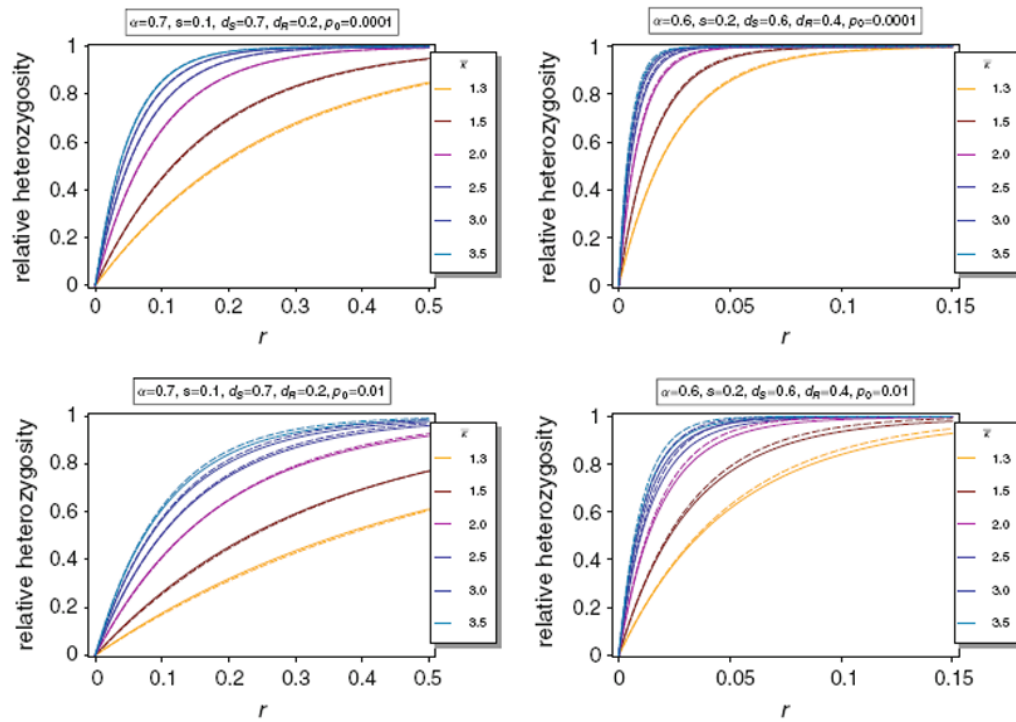
Fig. 8.
Relative expected heterozygosity $H^{(\kappa)}$ (*thin lines*) of the neutral allele N_1 versus the approximation $\tilde{H}^{(\kappa)}$ (*thick dashed lines*) for different distributions κ. In all cases we assume a truncated exponential distribution with range 1–10 and mean $\bar{\kappa}$. By truncated we mean that the probability of $m = 10$ is the probability that $m \geq 10$ for a poisson distribution with mean $\bar{\kappa}$. The parameters are as in Fig. 3
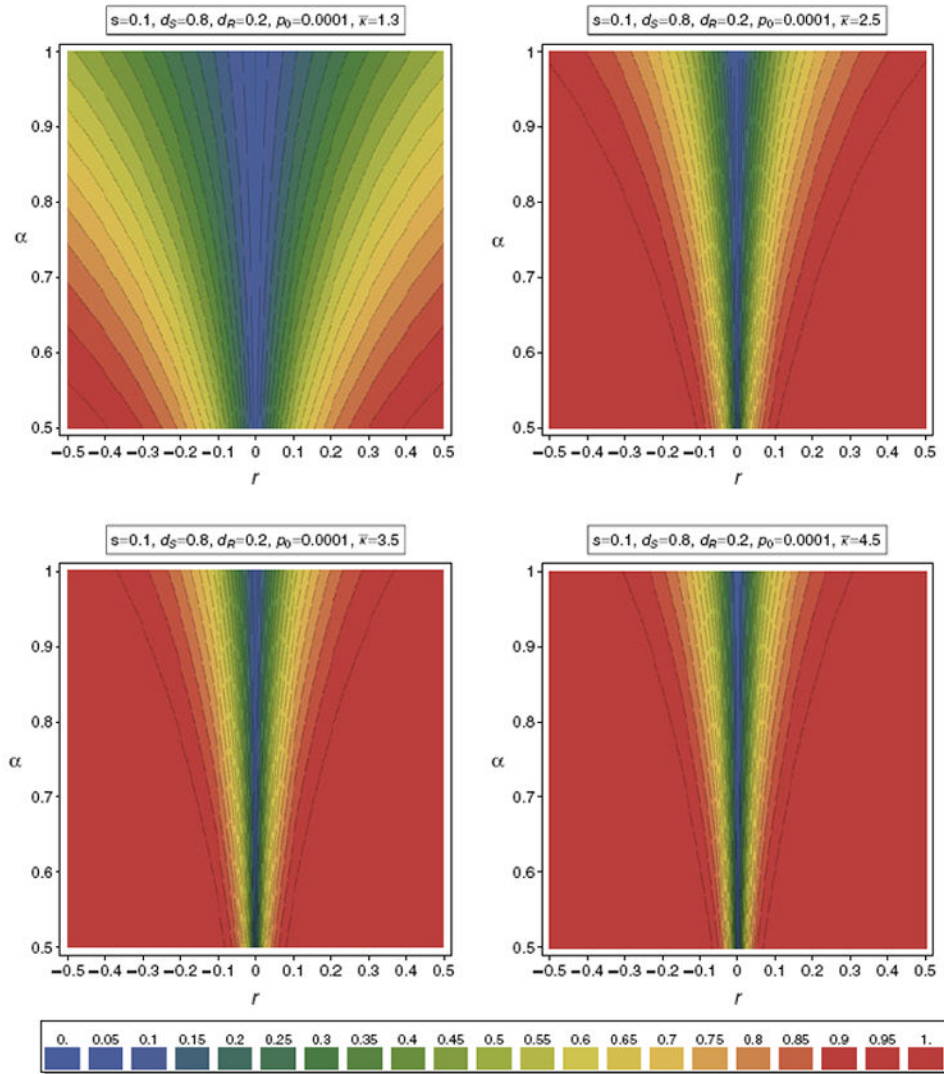
Fig. 9.
Contour plot of the valley of reduced heterozygosity for different distributions κ. In all cases we assume a truncated exponential distribution with range 1–10 and mean $\bar{\kappa}$. By truncated we mean that the probability of $m = 10$ is the probability that $m \geq 10$ for a poisson distribution with mean $\bar{\kappa}$. In all panels the same selection parameters are assumed. The parameters are summarized in the *boxes* above the panels

Table 1

Summary of notation

Haplotypes	$A_S N_1$	$A_S N_2$	$A_R N_1$	$A_R N_2$
Frequency	p_1	p_2	p_3	p_4
Fitness in	$W_1^{(U)}$	$W_2^{(U)}$	$W_3^{(U)}$	$W_4^{(U)}$
Untreated hosts	1	1	$1-s$	$1-s$
Fitness in	$W_1^{(T)}$	$W_2^{(T)}$	$W_3^{(T)}$	$W_4^{(T)}$
Treated hosts	$1-d_S$	$1-d_S$	$1-d_R$	$1-d_R$
Sensitive/resistant	Sen.	Sen.	Res.	Res.

The tables shows the notation of the four haplotypes, their frequencies, fitnesses in treated and untreated hosts, the parametrization of fitnesses that we are going to use for the illustrations in the following sections, and whether the haplotypes are sensitive or resistant. Here, s reflects metabolic costs of the resistant allele, while d_S and d_R indicate how efficiently the drug wipes out the sensitive and resistant parasites, respectively