Genetic load, eco-evolutionary feedback and extinction in metapopulations

Oluwafunmilola Olusanya[1], Ksenia A. Khudiakova[1], Himani Sachdeva[2,§ *]

[1] *Institute of Science and Technology Austria, Am Campus 1, Klosterneuburg 3400, Austria.*

[2] *Department of Mathematics, University of Vienna, Vienna 1090, Austria.*

§ Corresponding author

Abstract

Habitat fragmentation poses a significant threat to the persistence of populations, by generating increased genetic drift (and thus higher genetic load) as well as demographic stochasticity. Higher load causes population numbers to decline, which reduces the efficiency of selection and further increases load, resulting in a positive feedback which may drive entire populations to extinction. Here, we investigate this eco-evolutionary feedback in a metapopulation, focusing on how extinction thresholds depend on the total rate of deleterious mutations and their selection and dominance coefficients, the carrying capacity and growth rate of subpopulations, and the extent of migration between them. We first study the determinants of load under soft selection, where population sizes are constant, and then build upon this to analyse hard selection, where population sizes and load co-evolve. We show that in contrast to soft selection, high levels of migration are required to mitigate load and prevent extinction when selection is hard, with critical migration thresholds for metapopulation persistence increasing sharply with increasing 'hardness' of selection. Our analysis employs a combination of simulations and analytical approximations, allowing for a more comprehensive understanding of the factors influencing genetic load and extinction in fragmented landscapes.

Keywords: population fragmentation, genetic load, extinction, migration, soft selection, hard selection

Introduction

The long held belief that ecological and evolutionary processes do not influence each other has been challenged in recent years, with many studies documenting ecological and evolutionary change on contemporaneous timescales (Thompson, 1998, Hairston Jr et al., 2005). This can be the case when selection varies sharply over space and time, or when populations are too small to adapt efficiently, resulting in rapid loss in fitness (e.g., due to the accumulation of deleterious alleles), which may feed back into population demography.

* samuel.edet@imtlucca.it

Such eco-evolutionary feedback is particularly important in fragmented landscapes with many small-sized patches, where mating between genetically related individuals within patches (i.e., inbreeding) as well as stochastic changes in genetic composition (due to genetic drift) can compromise the efficiency of natural selection. As a consequence, populations become susceptible to fixing deleterious mutations, causing genetic load to increase and population numbers to decline, which can further exacerbate both genetic drift and inbreeding. The resultant positive feedback between increasing load and declining population sizes may drive populations towards extinction through a process termed *mutational meltdown* (Lynch et al., 1995b). Evidence of such extinction events has been found in plants (Matthies et al., 2004) and in vertebrates (Fagan and Holmes, 2006) and invertebrates (Saccheri et al., 1998). This kind of feedback is further aggravated by demographic and environmental stochasticity, both of which hinder adaptation to changing environments. However, despite some understanding of individual factors, a comprehensive picture of how deterministic and stochastic processes together structure genetic diversity and influence the persistence of populations remains lacking. Such an understanding is crucial for formulating effective conservation strategies that can promote long-term survival of endangered species (Lande, 1993, Frankham, 1998), particularly as human activities continue to exert significant pressures on natural environments.

Migration between subpopulations in a fragmented landscape can affect various aspects of the eco-evo feedback loop, homogenizing both genetic composition and population density across the population. Investigations of natural populations yield examples of both the beneficial and deleterious effects of migration. For instance, Finger et al. (2011) demonstrated how augmented gene flow reduced inbreeding in an isolated population of the jellyfish tree. Similarly, Land and Lacy (2000) showed that introducing eight wild-caught Texan females into a small isolated Florida panther population tripled its size, bringing it back from near extinction within just twelve years. On the flip side, there are also several examples of hybridisation-induced extinction, e.g., due to gene flow between domesticated species and their wild relatives (Todesco et al., 2016). These examples thus highlight the diverse effects of migration–ranging from 'rescue' (e.g., when migrants are introduced into genetically depauperate populations) to 'swamping' (e.g., due to gene flow between populations that are locally adapted to different environments), and more generally, the tension between inbreeding and outbreeding depression (Frankham et al., 2011, Edmands, 2007, Templeton, 1986)). This introduces complexities to practical decisions in conservation such as those concerning assisted gene flow (Aitken and Whitlock, 2013), underscoring the need for a more quantitative and theoretical understanding of gene flow in fragmented populations.

Migration can have complex effects on genetic diversity even when different populations are under *spatially uniform* selection, e.g., due to purifying selection against unconditionally deleterious alleles. Different local populations may be nearly fixed for deleterious alleles at different loci just by chance. Migration between such populations will then result in hybrid offspring of increased vigour due to masking of recessive deleterious mutations. Migration thus increases fitness variation, which has two competing effects - it prevents fixation of deleterious alleles and alleviates drift load (especially at low levels of migration), but also prevents purging of recessive alleles by increasing heterozygosity (at higher levels of migration) (Glémin et al., 2003).

The effect of migration on genetic diversity and adaptation in structured populations depends crucially on whether selection is 'soft' or 'hard', i.e., whether different local populations contribute equally to the next generation regardless of fitness, or if the contribution of fitter populations is higher (e.g., when these are

larger and send out more migrants). However, most work on hard vs. soft selection in subdivided populations assumes a rather specialised life cycle in which all (adult or juvenile) individuals across all demes join a common pool prior to reproduction, followed by uniform redistribution of zygotes back into demes (Levene, 1953, Christiansen, 1975, Ravigné et al., 2004), making it difficult to understand how *limited* dispersal between demes affects population outcomes. Moreover, models of hard selection implicitly assume that the metapopulation as a whole is under global density regulation and at carrying capacity (see, e.g., Whitlock (2002)). Thus, such models cannot be used to understand when and how the positive feedback between declining population size and increasing load might lead to local or global extinction, and to what extent this may be arrested by (limited) dispersal between populations. This therefore calls for more realistic models that allow for variation in both local and global population sizes and explicitly incorporate eco-evolutionary feedback.

Lynch et al.(1995a, 1995b) considered such a model, where load due to deleterious mutation accumulation leads to extinction. Using extensive computer simulations, they showed that populations with low effective size (typically <100) are at significant risk of extinction via mutational meltdown within only about 100 generations (Lynch et al., 1995a)– a conclusion with potential implications for, e.g., captive breeding programs. However, their analysis was limited to single randomly mating populations. Higgins and Lynch (2001) generalised this to a metapopulation with global and nearest-neighbour dispersal. Using simulations, they showed that the eventual fate of a metapopulation (as measured by the median time to extinction) depends critically on the number of demes (see also Lande et al. (2003), Ch 4, Table. 4.1) as well as on the dispersal neighbourhood. Their analysis however primarily focused on extinction times and was restricted to scenarios with relatively small patch sizes, resulting in early population extinction. Furthermore, their study was entirely simulation-based, making it hard to generalize their conclusions.

Going beyond simulations, Szép et al. (2021) derived analytical results for *polygenic* local adaptation and extinction in a metapopulation, allowing for genetic drift, demographic stochasticity and explicit eco-evo feedback between (local) population size and load. Crucially, under a polygenic model, a small increase in deleterious allele frequency at each of the very many loci influencing fitness can substantially increase total load, thus decreasing population size, which further exacerbates the effects of drift per locus. This thus generates an indirect 'eco-evolutionary' coupling between different loci, even in the absence of epistasis or linkage disequilibria. Szép et al. (2021) showed that local adaptation is more difficult under hard selection and if locally adaptive traits are more polygenic, causing populations to become extinct under much lower levels of gene flow than would be expected for a single locally adaptive locus.

This begs the question: how does gene flow mediate extinction risk under other scenarios, e.g., when selection is uniform across space, and to what extent are its effects different for selection on a few vs. many loci in this case? Using a similar theoretical model, but with spatially uniform selection, Sachdeva et al. (2022) investigated whether gene flow (in a mainland-island model) can help arrest mutational-meltdown of the island population. They found that migration can have qualitatively different effects; having a positive effect (i.e., reducing load and thus the risk of extinction) when the island population is small and isolated, and having an opposite effect (i.e., increasing load by hindering purging) when it is large and connected. They also identified the total deleterious mutation rate relative to the population growth rate as a key determinant of extinction.

However, going beyond marginal populations subject to asymmetric gene flow, it is important to understand how eco-evo feedback influences extinction risk and load in a metapopulation where migration is random and where the different islands (demes) can fix for different subsets of deleterious alleles that mask one another. We therefore extend the work of Sachdeva et al. (2022) to explore the eco-evolutionary dynamics of a metapopulation made up of a large number of demes exchanging migrants at random (i.e., under the island model). Our goal is to investigate the effect of gene flow on equilibrium load and population size, as well as the critical migration thresholds required for the persistence (i.e., for non-extinction) of the metapopulation. We analyse how these thresholds are influenced by the architecture of load (i.e., the total deleterious mutation rate and the selection and dominance coefficients of mutations) as well as features of the metapopulation landscape (carrying capacities and baseline growth rate of local demes). As we see below, a key parameter that determines the fate of the metapopulation is the total mutation rate relative to the baseline growth rate, which determines the extent to which genetic load depresses population growth and can thus be considered a proxy for the 'hardness of selection'. In addition to explicitly modeling the direct eco-evo coupling between population size and load, our theoretical approach also accounts for the indirect coupling between the dynamics of different selected loci (see above). This indirect coupling adds complexity to eco-evo dynamics and can have significant implications for population fitness and evolutionary outcomes in the metapopulation.

Model and Methods

We consider a metapopulation with a large number of demes (local populations) connected to each other via migration. Individuals are diploid and subject to deleterious mutations at L unlinked loci at a rate u per haploid locus, resulting in a genome-wide deleterious mutation rate of $2Lu$. We also allow for reverse mutations at rate v. Fitness is assumed to be multiplicative across loci. At any locus, the wild-type, heterozygote and deleterious homozygote have fitness 1, e^{-hs} and e^{-s} respectively, with h representing the dominance coefficient and s, the strength of selection against the deleterious homozygote. We take s and h to be the same across all loci, a gross simplification that is relaxed in Appendix F, SI.

In each generation, a fraction m of individuals in each deme migrate into a common pool; individuals from this pool are then evenly redistributed across all demes. We assume that all demes experience the same environment, so that the fitness of any genotype is independent of deme. We further assume a logistic model of population growth, where the population size $n_i(t+1)$ in deme i at generation $t+1$ is Poisson distributed with mean $n_i(t)e^{r_0(1-n_i(t)/K)}\overline{W_i}$. Thus, the population size of a deme after reproduction and density regulation depends on its previous size $n_i(t)$ (subsequent to migration), carrying capacity K, the baseline growth rate r_0, as well as the mean fitness $\overline{W_i}$ in deme i. Note that the carrying capacity K and the baseline growth rate r_0 are assumed to be the same across all demes. However, population sizes may differ due to stochastic variation in local population fitness $\overline{W_i}$ and demographic stochasticity.

Thus, the parameters of our model are: the number of loci L, the per locus mutation rates u and v (which we assume to be equal), the per locus strength of selection s, the dominance coefficient h, the migration rate m, the baseline growth rate r_0 and carrying capacity K per deme. As we argue below, the behaviour of this model is governed by scaled (dimensionless) parameters Km, Ks and Ku (which

represent respectively the number of migrants per deme, the per locus strength of selection relative to local drift, and the average number of deleterious mutations arising per locus per generation per deme, all defined for demes at *carrying capacity*), the dominance coefficient h, the genomewide mutation rate scaled by the baseline growth rate $(2Lu)/r_0$ (which is a measure of the 'hardness' of selection) and r_0K (which scales inversely with the strength of demographic fluctuations). The last two parameters deserve further explanation: the expected genetic load in a large population at mutation-selection balance is precisely $2Lu$, as long as alleles are not completely recessive. Thus, $(2Lu)/r_0$ represents the extent to which population growth is reduced due to mutation load in a large population. Drift typically increases load by increasing random fixation of deleterious alleles (if $Ks \lesssim 1$); thus selection can be hard even if $(2Lu)/r_0$ is not too large. Finally, the parameter r_0K represents the average number of births (or deaths) in a deme at carrying capacity. Thus, larger values of r_0K correspond to weaker demographic stochasticity.

To gain insight into how different evolutionary processes affect population outcomes, we first analyse soft selection, which corresponds to the $(2Lu)/r_0 \to 0$ limit of the full model. Under soft selection, deme sizes are independent of the degree of maladaptation, and demes are always at carrying capacity. We then consider hard selection, where genetic load is comparable to the baseline growth rate (i.e., $2Lu/r_0 \lesssim 1$), so that population size decreases with increasing maladaptation. Also, since population sizes can vary randomly across demes (due to stochastic variation in allele frequencies), larger and more fit demes contribute more to the migrant pool, while also being less influenced by incoming migrants. The hard selection model thus allows for both local extinctions (wherein a subset of demes go into a 'sink' state) or global extinction of the entire metapopulation.

Our analysis utilises simulations as well as a number of analytical approximations. We analyse the soft selection model using the diffusion approximation, which predicts the equilibrium distribution of allele frequencies at any given locus in any deme conditional on the mean allele frequency across the entire metapopulation (Wright, 1937). In using the diffusion approximation, we make a simplifying assumption that alleles at different loci evolve independently (which is valid when recombination is much faster than all evolutionary and ecological processes). However, if load is substantial, then this is no longer true: we thus study the effects of multilocus interactions on load by introducing effective migration rates into the diffusion approximation (Sachdeva, 2022).

Under hard selection, population sizes can vary between demes and also within a deme over time, requiring us to follow the joint distribution of population sizes and allele frequencies across all demes (Szép et al., 2021). However, it turns out that we cannot explicitly solve for this distribution under the diffusion approximation in the presence of mutation. Hence, to make analytical progress we use a different approach - *the semi-deterministic approximation* (see also Sachdeva et al. (2022)). This assumes that the population size of any deme is reduced with respect to carrying capacity by an amount that depends on the genetic load and is thus *determined* by load, rather than following a distribution. It further assumes that the load is that expected at mutation-selection-migration-drift equilibrium for this population size, allowing us to solve self-consistently for the population size (and load). As we discuss below, this approximation accounts for genetic drift but not demographic stochasticity, and is hence semi-deterministic.

In the following, we first introduce equations for the joint stochastic evolution of population sizes and allele frequencies, and then use this to motivate the various approximations.

Evolution of allele frequencies and population sizes in continuous time

If ecological and evolutionary processes are slow relative to recombination, so that loci evolve more or less independently and allele frequencies change slowly, then one can describe the dynamics of allele frequencies, $p_{i,j}$, at locus j (where $j{=}1,2,\ldots L$) and of population size, n_i, in any deme i via continuous time equations. It is convenient to work with dimensionless parameters, which we construct by scaling population size n by carrying capacity K, i.e., $N{=}n/K$, and scaling all evolutionary rates by the baseline growth rate r_0, i.e., $S{=}s/r_0,\ U{=}u/r_0,\ M{=}m/r_0$, and time as $\tau{=}r_0\,t$ (see also Szép et al. (2021)). Then, we have:

$$\frac{dp_{i,j}}{d\tau}{=}-\frac{p_{i,j}q_{i,j}}{2}\frac{\partial R_{g,i}}{\partial p_{i,j}}{+}U_i(q_{i,j}{-}p_{i,j}){+}M\frac{\overline{N}}{N_i}\left(\frac{\overline{Np_j}}{\overline{N}}{-}p_{i,j}\right){+}\lambda_{p_{i,j}}\qquad q_{i,j}{=}1{-}p_{i,j} \tag{1a}$$

$$\frac{dN_i}{d\tau}{=}N_i(1{-}N_i{-}R_{g,i}){+}M(\overline{N}{-}N_i){+}\lambda_{N_i}\quad\text{where}\quad R_{g,i}{=}\sum_{j=1}^{L}S\,p_{i,j}(2h_{i,j}q_{i,j}{+}p_{i,j}) \tag{1b}$$

where $R_{g,i}$ is the load (scaled by r_0) in deme i, and is the sum of the load at all L loci. Further, $\overline{N}{=}\frac{1}{n_D}\sum_k N_k$ is the mean population size per deme and $\overline{Np_j}{=}\frac{1}{n_D}\sum_k p_{k,j}N_k$ the mean number of deleterious alleles (at locus j) per deme, averaged across all demes. Finally, $\lambda_{p_{i,j}}$ is a stochastic processes that captures the effect of genetic drift on allele frequencies, and satisfies $\mathbb{E}[\lambda_{p_{i,j}}(t)]{=}0$, and $\mathbb{E}[\lambda_{p_{i,j}}(\tau)\ \lambda_{p_{i',j'}}(\tau')]{-}\mathbb{E}[\lambda_{p_{i,j}}(\tau)]\ \mathbb{E}[\lambda_{p_{i',j'}}(\tau')]{=}\frac{1}{\zeta\ N_i(\tau)}p_{i,j}(\tau)q_{i,j}(\tau)\delta(\tau{-}\tau')\delta_{i,i'}\delta_{i,j'}$, where $\mathbb{E}$ denotes expectation and $\zeta{=}r_0K$. Similarly, λ_{N_i} captures the effect of demographic stochasticity (due to randomness in birth and death) on population sizes, and satisfies $\mathbb{E}[\lambda_{N_i}(t)]{=}0$, and $\mathbb{E}[\lambda_{N_i}(\tau)\ \lambda_{N_{i'}}(\tau')]{-}\mathbb{E}[\lambda_{N_i}(\tau)]\ \mathbb{E}[\lambda_{N_{i'}}(\tau')]{=}\frac{N_i(t)}{\zeta}\delta(\tau{-}\tau')\delta_{i,i'}$. Note that the strength of fluctuations is inversely proportional to $\zeta{=}r_0\ K$.

In constructing dimensionless parameters, we could have also scaled all evolutionary rates by $1/K$, yielding scaled parameters $Ks{=}\zeta S$, $Ku{=}\zeta U$ and $Km{=}\zeta M$. In the following we will use both kinds of scaling depending on the context.

Diffusion approximation for allele frequencies

Under soft selection, each deme is at carrying capacity, independent of load, so that: $N_i{=}\overline{N}{=}1$ and $\overline{Np_j}/\overline{N}{=}\overline{p_j}$ in eq. (1a). Thus, allele frequencies evolve independently of population size (which is constant) and of each other, so that the joint equilibrium distribution $\Psi\left(p_1,p_2,\ldots,p_L\right)$ of allele frequencies is simply the product of the marginal distributions $\psi\left(p_j\right)$. Moreover, following Wright (1937), we can write down an explicit expression for the allele frequency distribution $\psi(p_j)$ at any locus j in any deme, conditioned on the mean allele frequency $\overline{p_j}$ across the entire metapopulation:

$$\psi\left(p_j|\overline{p}_j\right)\propto p_i^{4Ku+4Km\overline{p}_j-1}(1{-}p_j)^{4Ku+4Km\overline{q}_j-1}e^{-2Ks(p_j^2+2hp_jq_j)};\ q_j{=}1{-}p_j,\ \overline{q}_j{=}1{-}\overline{p}_j \tag{2}$$

One can numerically integrate over this distribution to obtain the expected allele frequency $\mathbb{E}[p_j]{=}\int\psi[p_j|\overline{p}_j]\,dp_j$ as a function of $\overline{p}_j$, and then solve for $\mathbb{E}[p_j]{=}\overline{p}_j$, to obtain an explicit expression for the allele frequency distribution (by substituting the numerically determined $\overline{p}_j$ into eq. (2)). Integrating over this distribution

now yields the expected heterozygosity, load etc.

Semi-deterministic approximation for population sizes and allele frequencies

As mentioned earlier, under hard selection and in the presence of mutation, the diffusion approximation does not give an explicit formula for the joint distribution of population size and allele frequencies. Thus, we employ an approximate semi-deterministic analysis which accounts for the effect of genetic drift on allele frequencies, but neglects that of demographic fluctuations on population size (Szép et al., 2021, Sachdeva et al., 2022). This can be justified in a scenario where $\zeta{=}r_0K{\gg}1$, i.e., if population sizes are large enough for demographic fluctuations to be unimportant. Nevertheless, if individual alleles are weakly deleterious, then $\zeta S{=}Ks$ may be $\mathcal{O}(1)$, so that the effects of genetic drift on allele frequencies are still appreciable. As we see below, the semi-deterministic approximation becomes increasingly accurate as we take the limit $\zeta{=}r_0K{\to}\infty$, while holding all scaled parameters $\zeta S{=}Ks$, $\zeta U{=}Ku$, $\zeta M{=}Km$ and $LU{=}L(u/r_0)$ constant.

More concretely, the semi-deterministic analysis assumes that the equilibrium distribution of scaled population sizes $N{=}n/K$ within a deme is sharply peaked around one or more values $\{N_*\}$, which correspond to *metastable* states, e.g., a long-term 'sink' or 'large population' state. Thus, the population size is close to one or other of these $\{N_*\}$ for long periods of time, with only very occasional transitions between metastable states. This allows enough time for allele frequencies to equilibrate at a given N_*, so that genetic load within a deme can be approximated by the expectation at mutation-selection-drift-migration balance, where the strength of drift is governed by $1/N_*$. The population size N_* can then be calculated self-consistently (see eq. 3).

Let us assume that all demes in the metapopulation reach the same equilibrium N_*, so that $N_i{=}\overline{N}{=}N_*$ in eq. (1b). In this situation, migration between demes has no effect on population numbers (since no island acts as a demographic source or sink). If we further take the limit $\zeta{\to}\infty$, i.e., neglect the stochastic term in eq. (1b), then at equilibrium, we have $N_*{\sim}1{-}R_g$: in other words, population size is reduced (relative to carrying capacity) by an amount equal to the genetic load scaled by the baseline growth rate. If we further assume that load is approximately $\mathbb{E}[R_g|N_*,\overline{p}]$, the (scaled) expected load at selection-mutation-migration-drift equilibrium at that N_*, we have the following coupled equations:

$$
\begin{aligned}
N_* = 1{-}\mathbb{E}[R_g|\overline{p},N_*] \qquad \overline{p}{=}\mathbb{E}[p|\overline{p},N_*] &= \frac{\int_0^1 p\,\psi(p|\overline{p},N_*)\,dp}{\int_0^1 \psi(p|\overline{p},N_*)\,dp} \\
\mathbb{E}[R_g|\overline{p},N_*] = LS\,&\frac{\int_0^1 [p{+}(2h{-}1)pq]\,\psi(p|\overline{p},N_*)}{\int_0^1 \psi(p|\overline{p},N_*)} \\
\text{where}\quad \psi(p|\overline{p},N_*) &\propto p_i^{4N_*Ku+4N_*Km\overline{p}-1}(1{-}p_j)^{4N_*Ku+4N_*Km\overline{q}_j-1}e^{-2N_*Ks(p^2+2hpq)}
\end{aligned}
\tag{3}
$$

In order to numerically solve this set of equations, we start with an initial guess for $\overline{p}$, use this to compute N_*, $\mathbb{E}[R_g]$ and the new $\overline{p}$, and iterate this procedure until N_*, $\mathbb{E}[R_g]$ and $\overline{p}$ converge to fixed values. The final

fixed point can be sensitive to the initial guess for $\overline{p}$. Thus, we always start with a very low $\overline{p}$: this always leads to a fixed point corresponding to a 'large population' state (where $\overline{p}$ is low across the metapopulation, and $N_*{>}0$), provided such a state exists and is stable.

Taking N_* to be the same across all demes is a rather strong assumption, and will not hold close to extinction thresholds, where a significant fraction of demes is nearly extinct (while others support large populations). Our analysis is thus limited in this respect. The situation is however not too dire: as we shall see later, even in such scenarios, the semi-deterministic approximation accurately predicts the average population size in the non-extinct demes, though it fails to predict the proportion of such demes.

As discussed above, under the semi-deterministic approximation, population outcomes depend on five composite parameters. These are: the dominance coefficient h; the selection strength scaled by the carrying capacity Ks, the mutation rate scaled by the carrying capacity Ku, the migration rate scaled by the carrying capacity Km, and the genomewide mutation rate relative to the baseline growth rate, $2LU{=}2Lu/r_0$, that determines the hardness of selection (see Sachdeva et al. (2022)).

Simulations

Since multilocus individual-based simulations in metapopulations are computationally expensive, we use these in a targeted way to first, identify regimes in which multilocus associations are important under soft selection; second, to examine if the effect of such associations can be captured via simple extensions of the diffusion approximation that include effective migration rates, m_e (Sachdeva, 2022, Zwaenepoel et al., 2023). It turns out that multilocus associations only have limited effects which are captured reasonably well by approximations based on m_e. Thus, for the case of hard selection, we only perform allele frequency simulations that assume linkage and identity equilibrium between loci, i.e., neglect associations.

These simulations follow population sizes n_i and allele frequencies $p_{i,j}$ at all L loci (i.e., $j{=}1,\ldots,L$) in all n_D demes (i.e., $i{=}1,\ldots,n_D$) of the metapopulation, but not genotypes. The simulation is initialized such that each deme is initially perfectly fit (i.e., all deleterious allele frequencies are zero), with individuals accumulating deleterious mutations over time. In each generation, allele frequencies and population sizes are updated to reflect the effects of migration, mutation, selection, reproduction and stochastic events (i.e., drift and demographic stochasticity) as described in appendix A of the SI. The metapopulation is allowed to equilibrate and the equilibrium load and population size are then computed.

Code for individual-based and allele frequency simulations and for numerically evaluating analytical predictions based on eqs. 2 and 3 is available (see Olusanya et al. (2023)).

Choice of parameters.

To guide the choice of parameters, we look to broad estimates across a range of organisms. For example, the genome-wide rate of deleterious mutations per individual per generation is estimated to be of the order of $0.1{-}1$ in multicellular eukaryotes (Lynch et al., 1999), $1.2{-}1.4$ in *Drosophila* (Haag-Liautard et al., 2007), $0.25{-}2.5$ for *C. elegans* (Denver et al., 2004) and 2.2 for humans (Keightley, 2012), the latter being an unusually high value. Estimates of fitness effects of new mutations are similarly varied: while some

experimental studies have shown that a sizeable fraction ($<15\%$) of mutations are likely to be lethal (Mukai et al., 1972, Eyre-Walker et al., 2006), there is consensus that small effect mutations (with effect $\lesssim 5\%$ on quantitative traits) have higher densities (Mukai, 1964, Mukai et al., 1972, Lyman et al., 1996, Lynch et al., 1999). Dominance coefficients are harder to estimate for weakly deleterious mutations but may be ~ 0.2 for moderately deleterious variants (Charlesworth and Charlesworth, 1987). In addition, several studies point to a negative relationship between s and h, with mutations of small effect being almost additive and those of large effect being almost recessive (Simmons and Crow, 1977, Lynch et al., 1999).

Finally, indirect measures (e.g, F_{ST}) have reported moderate to high levels of gene flow in plants and mammals with typical values between closely related species being of the order $0.05-0.2$, which corresponds roughly to between $1-5$ migrants between local populations per generation. It is important to note that in populations that are on the verge of extinction, gene flow may be much less than this (Frankham et al., 2002, Casas-Marce et al., 2013, Szczecińska et al., 2016) and indeed one of the goals of this study is to identify critical levels of migration required to prevent population collapse. Thus, most of the study will focus on low to moderate levels of gene flow.

The most difficult parameter to estimate is the baseline growth rate, r_0 (see also Discussion). This enters into the composite parameter $2LU = (2Lu)/r_0$, which we use as a measure of the 'hardness' of selection (see above). We necessarily restrict ourselves to $2LU < 1$, as extinction is guaranteed at higher values of $2LU$ (regardless of gene flow).

Results

Soft selection

We first consider soft selection, where each deme is at carrying capacity K, regardless of mean fitness. If loci evolve independently, then the equilibrium allele frequency distribution $\psi[p|\overline{p}]$ at any locus in any deme depends only on the scaled parameters Ks, Ku, Km and the dominance coefficient h, and is given by eq. (2), conditional on $\overline{p}$, the mean allele frequency across all demes. As discussed above, $\overline{p}$ can be obtained self-consistently by numerically solving $\overline{p} = \int p\,\psi[p|\overline{p}]\,dp$.

The soft selection model has been analysed in earlier studies by assuming either very weak or strong selection, i.e., $Ks \ll 1$ (Whitlock, 2002), or $Ksh + Km$ at least 5 (Glémin et al., 2003, Roze, 2015). These limits are useful to consider as they yield simpler intuition and allow for explicit analytical results: for instance, when selection is much weaker than local drift (i.e., $Ks \ll 1$), probabilities of identity by descent of two or more genes are approximately the same as under the *neutral* island model, allowing one to approximate $\overline{p}$ in terms of (essentially) neutral F_{ST} (Whitlock, 2002). In this limit, selection is only efficient at the level of the population as a whole and not within demes, some of which may be locally fixed for the deleterious allele. By contrast, for large Ksh, deleterious alleles segregate at low frequencies within any deme and the allele frequency distribution is concentrated about $\overline{p}$. Now, gene flow only serves to further narrow the distribution of allele frequencies, i.e., further increase the effective size of local demes and the efficiency of selection (Glémin et al., 2003, Roze, 2015). Here, we focus on *moderately* deleterious alleles with $Ks \sim 1$ (which contribute the most to drift load in an isolated population). As we argue in appendix B

(SI), the effects of selection in this regime are more subtle and not captured by either intuitive description above.

Beyond these more conceptual issues, the concrete question we address is: how much gene flow between subpopulations is required to alleviate drift load at a single locus under soft selection, and how does this depend on the (scaled) selection strength Ks and dominance coefficient h at the locus? We then consider how load per locus is affected by multilocus associations between many *unlinked* deleterious alleles.

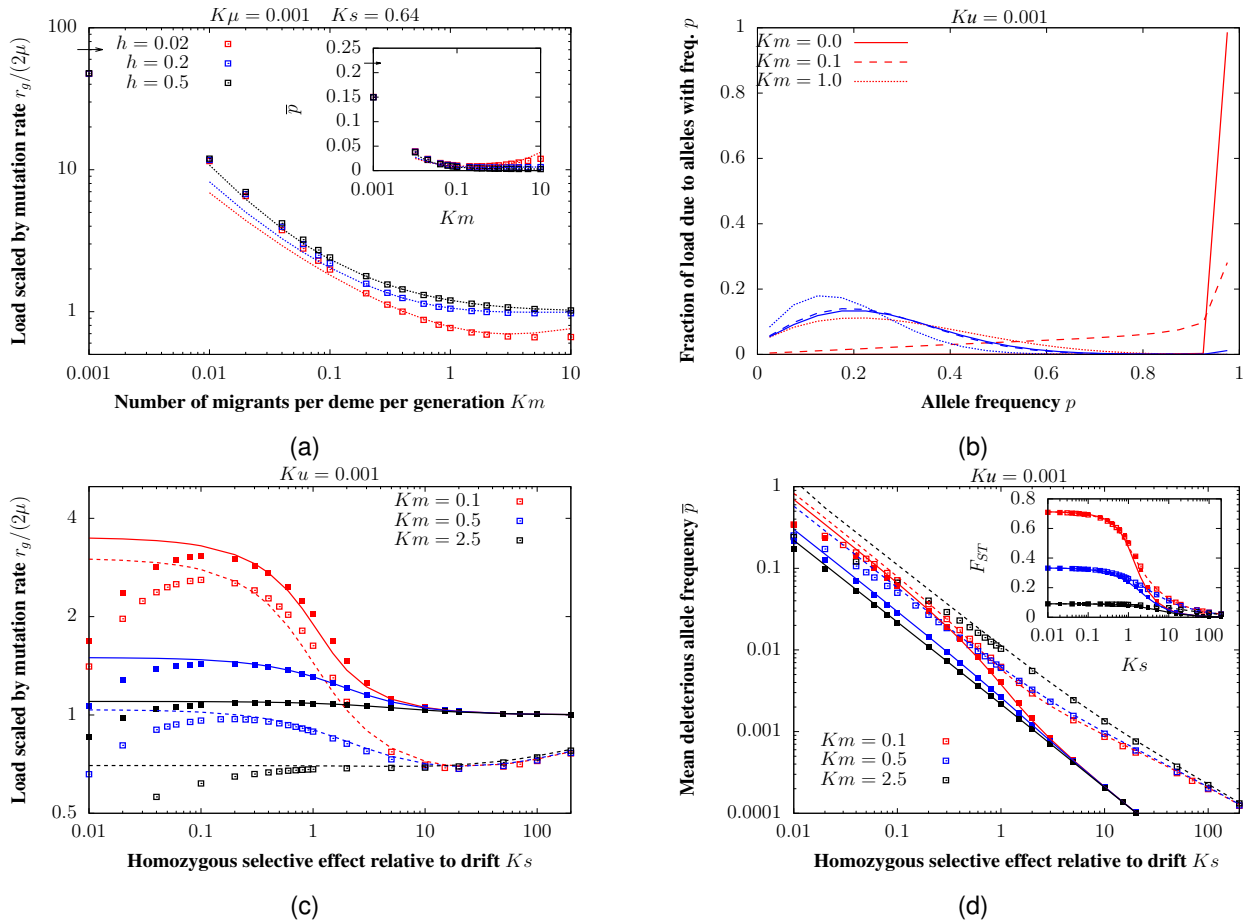

Figure 1: Mutation-selection-drift-migration equilibrium at a single locus under the infinite-island model with soft selection. (a) Main plot and inset show respectively the expected per locus load r_g (scaled by $2u$) and the mean deleterious allele frequency $\overline{p}$ vs. Km, the number of migrants per deme per generation, for various values of h, for Ks=0.64 (which is the Ks value for which load is maximum in an isolated population). Symbols depict results of the diffusion approximation (eq. (2)), while lines represent predictions of the 'moderate selection' approximation (eq. 7a, Appendix B, SI). Horizontal arrows on the y-axis represent the corresponding diffusion predictions for Km=0. (b) The fraction of total load that is due to alleles with frequency in the interval between p and $p+\Delta p$ (with Δp=0.05) for various values of Km (solid, dashed and dotted lines), for weakly selected (Ks=0.64; red) and strongly selected (Ks=6.4; blue) loci. Predictions are based on the diffusion approximation, i.e., obtained by integrating over the allele frequency distribution in eq. (2). (c) Expected per locus load r_g (scaled by $2u$) vs. Ks, the homozygous selective effect scaled by drift, for various values of Km (different colours) for h=0.5 (filled symbols) and h=0.02 (open symbols). (d) Mean deleterious allele frequency $\overline{p}$ (main plot) and the expected F_{ST} at the selected locus (inset) vs. Ks, for various Km (different colours) for h=0.5 (filled symbols) and h=0.02 (open symbols). In both (c) and (d), symbols depict results of the diffusion approximation (eq. (2)) and lines represent the approximate 'moderate selection' predictions. All plots are with Ku=0.001.

To begin with, consider the case of isolated demes (Km=0). For $Ku \ll 1$, any deme will be close to fixation for one or other allele at any locus, allowing us to use a 'fixed state' approximation (Szép et al., 2021, Barton and Olusanya, 2022). This assumes the probability of near-fixation of a wildtype allele within any deme to be proportional to: $\int_{p=0}^{1/2} dp\ p^{4Ku-1} \approx \frac{1}{4Ku}$, and the corresponding probability for the

deleterious allele to be proportional to: $\int_{p=1/2}^{1} dp\ e^{-2Ks}\ (1-p)^{4Ku-1} \approx \frac{e^{-2Ks}}{4Ku}$, regardless of h. Thus, for $Ku \ll 1$, the expected deleterious allele frequency is $\mathbb{E}[p] \approx \frac{e^{-2Ks}}{1+e^{-2Ks}}$, and the expected load r_g (scaled by the deterministic expectation $2u$) is: $\frac{G}{2u} = \frac{Ks}{2Ku}\,(\mathbb{E}[p] - (1-2h)\mathbb{E}[pq]) \approx \frac{Ks}{2Ku}\ \frac{e^{-2Ks}}{1+e^{-2Ks}}$, assuming $\mathbb{E}[pq] \approx 0$. It follows that the maximum contribution to load is from loci with $Ks \approx 0.64$, independent of h, with load per locus being $\frac{G}{2u} \approx \frac{0.139}{2Ku}$ for this value of Ks. Thus, drift can inflate load due to moderately selected loci by a factor of several hundreds or thousands in an isolated population (see also Kondrashov (1995)).

Now consider the effect of gene flow between subpopulations: fig. 1a shows load per locus (scaled by $2u$) and the expected deleterious allele frequency (inset) as a function of Km, for $Ks = 0.64$ and various values of h. Symbols depict predictions of the full diffusion framework (eq. (2)), and lines those of a simpler 'moderate selection' approximation, details of which are outlined in Appendix B, SI.

Figure 1a shows that a very low level of gene flow ($Km \sim 0.1$) is enough to dramatically reduce drift load, regardless of dominance. For instance, $r_g/(2u)$ falls from ~ 70 at $Km = 0$ to $4-5$ at $Km = 0.1$ for both nearly recessive ($h = 0.02$) and co-dominant ($h = 0.5$) deleterious alleles, for $Ku = 0.001$. In both cases, the expected allele frequency $\overline{p}$ declines from 0.22 at $Km = 0$ to ~ 0.01 at $Km = 0.1$ (inset), while heterozygosity hardly increases (from ~ 0.002 to $0.003-0.004$). Thus the reduction in load at low levels of migration is almost entirely due to the decline in the number of fixed deleterious alleles, with little to no change in the number of segregating alleles.

A further increase in migration further reduces load, though in the case of recessive alleles, this is partially offset by an increase in heterozygosity (resulting in less efficient purging). Thus, for recessive alleles, load is minimum at $Km \approx 4$ and rises again as Km increases further, approaching the level expected in a panmictic population at large Km. For the parameters shown in fig. 1a, the reduction in load due to purging at intermediate Km is rather modest and hardly visible on the scale of the plot; purging has a more substantial effect when $u/(hs)$ and Ks are smaller, and provided $h < 1/4$ (Whitlock, 2002).

Increasing gene flow also shifts the frequency spectrum of alleles that contribute to load – for $Ks \lesssim 0.1$, the predominant contribution at $Km = 0.1$ is from fixed or nearly fixed alleles; however, one migrant per deme per generation ($Km = 1$) is already enough to prevent fixation of deleterious alleles, so that load is entirely due to alleles segregating at intermediate frequencies at this higher migration level (dashed vs. dotted red curves in fig. 1b). As expected, gene flow has a much weaker effect at strongly selected loci: for $Ks = 6.4$, load is entirely due to segregating alleles regardless of Km (blue curves). Thus, gene flow only causes a minor reduction in load at strongly selected loci (fig. 1c), and may even be detrimental if it hinders purging (in the case of recessive deleterious alleles).

Moreover, even low levels of gene flow tend to 'even out' the contributions of alleles with different selection coefficients to load, so that moderately deleterious alleles no longer contribute disproportionately (as in isolated populations). For instance, with $Km = 0.1$, load is maximum for $Ks \approx 0.1$ (regardless of h), but is only about ~ 3 times larger than the deterministic expectation (fig. 1c). By contrast, in isolated populations, it may be several hundred or thousand times larger (see above). With modest levels of gene flow ($Km = 0.5$), the increase in load for intermediate Ks is weaker still, while $\gtrsim 2$ migrants per generation are enough for load to be more or less independent of Ks, as in a large, panmictic population.

One can further ask: are the processes underlying this dramatic reduction in load (even with low lev-

els of gene flow) qualitatively different for alleles with different selective effects? Expressing load as $\frac{G}{2u}=\frac{Ks}{2Ku}\,[\overline{p}-(1-2h)\,\overline{p}\overline{q}]\equiv\frac{Ks}{2Ku}\,[\overline{p}-(1-2h)\overline{p}\,\overline{q}(1-F_{ST})]$, it follows that for recessive alleles ($h<1/2$), load declines if the mean deleterious allele frequency $\overline{p}$ and/or F_{ST} at the selected locus decrease. A decline in F_{ST} (for a fixed $\overline{p}$), in turn, reflects a change in the allele frequency distribution from a more U-shaped (wherein a proportion $\overline{p}$ of demes are nearly fixed for the deleterious allele) to a unimodal distribution (wherein the deleterious allele segregates at a low frequency close to $\overline{p}$ in all demes).

Figure 1d shows $\overline{p}$ and F_{ST} vs. Ks (main plot and inset), for various values of Km, for co-dominant and nearly recessive alleles (filled vs. open symbols). Note that while selection is quite efficient in reducing the mean number of deleterious alleles already for $Ks\gtrsim 0.05$ (say, under low gene flow, i.e., $Km=0.1$), it has little effect on F_{ST} (i.e., on the probability of local fixation within demes, given $\overline{p}$) unless $Ks\gtrsim 0.5$. For example, F_{ST} is reduced by only about 15% with respect to its neutral value for alleles with $Ks=0.5$, but by $>50\%$ for $Ks=2$, for both values of h. Thus, selection for increased heterozygosity within demes (as reflected in a decrease in F_{ST}) does not markedly influence the evolutionary dynamics of alleles with $Ks\lesssim 1$ (which suffer the most severe inflation of load in isolated populations) but can be important for moderately deleterious alleles with $1\lesssim Ks\lesssim 5$ (which may still contribute substantially to load).

In Appendix B, we introduce a 'moderate selection' approximation, which allows F_{ST} to be affected by selection, but assumes that the relationship between (appropriately scaled) higher cumulants of the allele frequency distribution and F_{ST} is the same as under the neutral island model. The predictions of this approximation are shown by lines in fig. 1a, 1c and 1d, and appear to match the full diffusion (symbols) quite well for $0.5\lesssim Ks\lesssim 5$. This suggests that moderate selection essentially changes pairwise coalescence times (to different extents within and between demes), without appreciably affecting other statistical properties of the genealogy, e.g., the degree to which branching is skewed or asymmetric, at loci subject to purifying selection (see also Appendix B, SI).

Effect of multilocus associations on load. So far, we have neglected statistical associations, i.e., linkage and identity disequilibrium between deleterious alleles at different locations on the genome. These may be substantial if, for instance, different subpopulations are nearly fixed for partially recessive deleterious alleles at very many different loci, so that $F1$ individuals (offspring of migrants and residents) have higher fitness than residents, giving rise to heterosis. In such a scenario, the extent of gene flow at any locus depends not only on Km, the average number of migrants exchanged between demes, but also on the relative fitness of migrants and their descendants (which depends on all selected loci). The fitness of these descendants may be significantly higher than resident fitness when there is high differentiation between subpopulations (i.e., if migration is low or F_{ST} high), alleles are recessive (h is small) and deleterious alleles segregate at very many loci ($2Lu$ is large and Ks small).

More specifically, heterosis causes immigrant alleles to be transmitted at a higher rate to the next generations than alleles carried by residents, resulting in an *effective* migration rate that is higher than the raw migration rate (see also Ingvarsson and Whitlock (2000)), and lower per-locus load. For a genome with L unlinked loci with equal selective effects s and dominance coefficients h, and assuming $s\ll 1$, the effective

migration rate at any locus is shown by Zwaenepoel et al. (2023) to be approximately:

$$\frac{m_e}{m} \approx e^{2Ls(1-2h)\overline{p}\,\overline{q}F_{ST}} \tag{4}$$

Following Sachdeva (2022) and Zwaenepoel et al. (2023), we approximate the effects of multi-locus heterosis by assuming that allele frequencies at any locus follow the equilibrium distribution in eq. (2), but with the raw migration rate replaced by an effective migration rate which itself depends on the expected allele frequency $\overline{p}$ and F_{ST} (or equivalently, $\overline{p}$ and $\mathbb{E}[pq]$) through eq. (4). This allows us to numerically solve for $\overline{p}$ and F_{ST}, yielding the theoretical predictions (solid lines) in fig. 2.

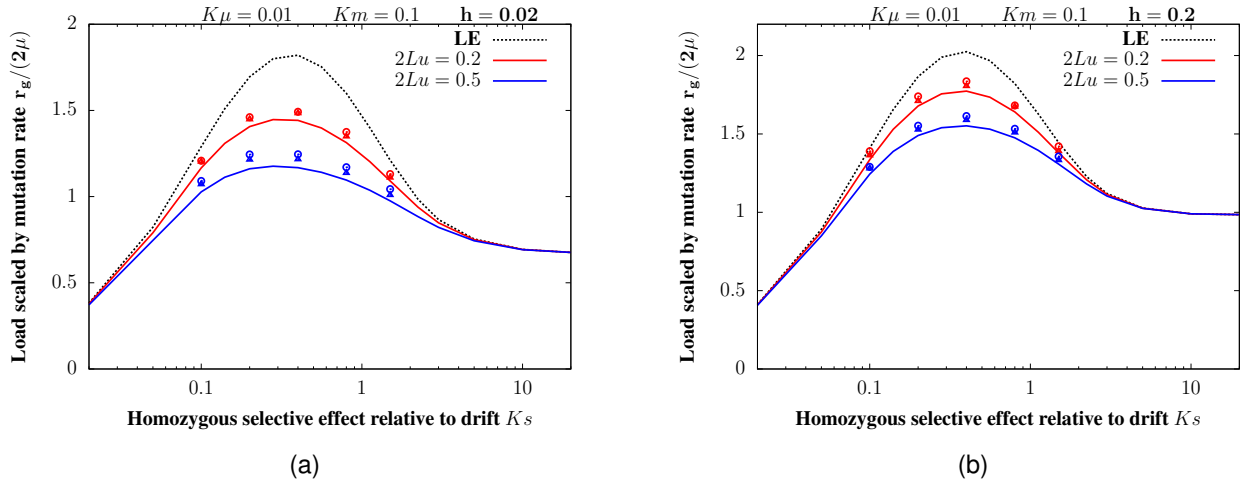

Figure 2: Effect of multi-locus associations on load. Load per locus (scaled by $2u$) vs. Ks in a population with deleterious mutations at L biallelic equal-effect loci with (a) h=0.02 and (b) h=0.2, for Km=0.1. Symbols depict results of individual-based simulations, which are carried out for two values of total mutation rate $2Lu$ and two values of L for each $2Lu$. Red circles and triangles depict respectively L=1000, u=0.0001 and L=2000, u=0.00005 both of which correspond to $2Lu$=0.2; blue circles and triangles depict L=2500, u=0.0001 and L=5000, u=0.00005, which correspond to $2Lu$=0.5. For each value of L and u, the carrying capacity K and migration rate m are chosen such that scaled parameters are Ku=0.01 and Km=0.1. For a given K, the scaled selective effect Ks is varied by varying s. Dashed curves depict single-locus predictions in the absence of multilocus associations (obtained using eq. (2)). Solid curves show predictions that account for associations via effective migration rates; these are independent of L for a fixed Lu, u/s, Ks and Km. The reduction in load per locus due to associations between loci is most significant for intermediate Ks, small h and larger Lu.

Figure 2 shows load per locus scaled by $2u$ as a function of Ks for low migration (Km=0.1), for two values of the total mutation rate ($2Lu$=0.2 in red and $2Lu$=0.5 in blue) and two dominance coefficients (h=0.02 in fig. 2a and h=0.2 in fig. 2b). For each value of $2Lu$, we further simulate with two values of L, scaling down s, u and m, and scaling up K as we increase L, so that Ks, Km, Ku and Ls remain constant. Symbols depict results of individual-based simulations for a metapopulation with 100 demes; dashed curves show single-locus predictions (obtained using eq. (2)) that do not account for multilocus heterosis; solid curves represent predictions that account for heterosis using effective migration rates: note that these depend on L only via the combination Ls (or alternatively, Lu for a given u/s). As expected, there is better agreement between simulations and theory for smaller values of s (or alternatively, larger L), for a given total mutation rate, as the expression for effective migration rate in eq. (4) becomes more accurate as $s \to 0$. As can be seen in fig. 2, load per locus is most strongly reduced by multilocus heterosis when $2Lu$ is large, and for loci with small h and intermediate Ks. Moreover, the effects of multilocus heterosis become weaker with increasing migration (which corresponds to lower F_{ST}) which reduces allele frequency differences across demes and consequently the extent of heterosis (see also Roze, 2015).

Hard selection

The analysis of soft selection in the previous section bears out 'one migrant per generation' as an approximate threshold beyond which drift load (due to random fixation of deleterious alleles) ceases to be significant (see, e.g., fig. 1a). This holds true regardless of h and even for moderately deleterious alleles that would otherwise contribute disproportionately to load. Here, we ask: what is the analogous criterion under hard selection, where feedback between population size and loadcan lead to extinction.

Under soft selection, the efficiency of selection per locus relative to local drift is governed by Ks. However, under hard selection and depending on parameter values, population sizes may be less than K, so that (scaled) selection per locus is $ns{=}NKs$ (where n and $N{=}n/K$ are the unscaled and scaled population size of a deme respectively). In general, we expect $N{\sim}1{-}R_g$ (eq. (3)), i.e., for population sizes to be reduced relative to carrying capacity by an amount equal to the total load scaled by the baseline growth rate. In the absence of drift, R_g is simply $2LU{=}2L(u/r_0)$ (which we use as a measure of the 'hardness' of selection, i.e., of the strength of the eco-evo feedback). However, drift will typically increase load (unless alleles are highly recessive) and thus depress population size. This results in a positive feedback between increase in load per locus (which is governed by per locus effects NKs and h) and a decrease in population size (which is governed by total load due to all L loci). Thus, the strength of the positive eco-evo feedback is mediated by the genetic architecture of load, i.e., the numbers, effect sizes and dominance coefficients of deleterious alleles. In the main paper, we focus on the case of equal effect loci; scenarios involving a distribution of fitness effects are explored in Appendix E of the SI.

The analysis in this section is based on the semi-deterministic approximation. We first test its accuracy by comparing against allele frequency (linkage equilibrium) simulations for different values of K, while keeping r_0 and the scaled parameters Ks, Km, Ku and $2LU{=}2L(u/r_0)$ fixed. These simulations establish that (as expected) the semi-deterministic approximation becomes increasingly accurate in the limit $r_0K{\to}\infty$, i.e., as demographic fluctuations become weaker. We then use the semi-deterministic analysis to identify critical migration thresholds below which the metapopulation becomes extinct, i.e., $N{=}n/K$ goes to zero. We explore how these thresholds depend on Ks, h and $2LU$. Finally, we also consider parameters for which the metapopulation is stable (i.e., not extinct) and explore how equilibrium load and population size per deme in this regime depend on various parameters.

Extinction thresholds with low gene flow

Figure 3a shows the mean population size per deme as a function of Km, which is the average number of migrants exchanged between any deme and the metapopulation, when demes are at carrying capacity. Deleterious mutations of equal effect occur at L loci with $2LU{=}2L(u/r_0){=}0.6$; thus selection is fairly hard. We compare scenarios where deleterious mutations have moderate ($Ks{=}1$) vs. strong ($Ks{=}10$) effects (fig. 3a vs. 3b), and also recessive vs. additive effects (red vs. blue) for each Ks. The solid lines show semi-deterministic predictions and the symbols connected by dashed lines show simulation results. In order to understand the effects of demographic stochasticity, we run simulations for various values of r_0K (where r_0 is fixed at 0.1 and K is varied), keeping all scaled parameters (Ks, Ku, Km and $2LU$) fixed. This is done by simultaneously reducing s, m, and u and increasing L, as K is increased.

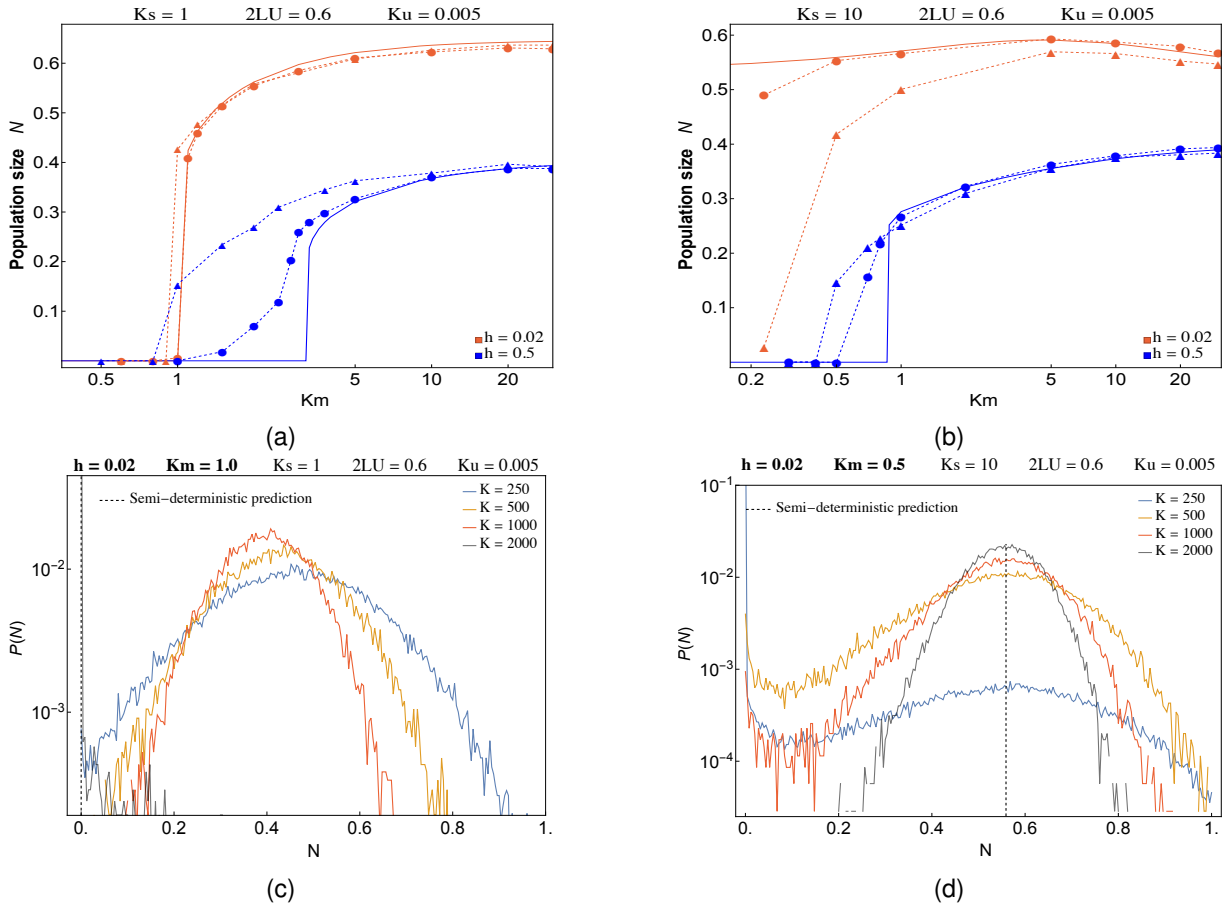

Figure 3: Comparison of simulations results with semi-deterministic predictions. Mean (scaled) population size N per deme plotted against Km for (a) $Ks{=}1$ and (b) $Ks{=}10$ for nearly recessive ($h{=}0.02$) and additive ($h{=}0.5$) alleles; other scaled parameters are: $Ku{=}0.005$ and $2LU{=}0.6$. Solid lines show predictions of the semi-deterministic approximation; symbols connected by dashed lines show results of allele frequency simulations with carrying capacity per deme K either 500 (triangles) or 3000 (circles). (c)-(d) Equilibrium distribution $P(N)$ of population sizes N, as obtained from simulations, for various values of K, for two different scaled parameter combinations for which the semi-deterministic approximation predicts extinction (in (c)) vs. stable populations (in (d)). The scaled parameters Ks, Ku, Km, $2LU{=}2L(u/r_0)$ are held constant as K is varied (by simultaneously varying s, u, L, and m). The scaled parameters have values: (c) $Ks{=}1$, $Km{=}1$, or (d) $Ks{=}10$, $Km{=}0.5$, with other parameters being $h{=}0.02$, $Ku{=}0.005$ and $2LU{=}0.6$, as above. Black vertical dashed lines represent the semi-deterministic prediction for N: this is $N{=}0$ in (c) and $N{=}0.56$ in (d). The distribution $P(N)$ becomes sharply peaked around the semi-deterministic prediction as K increases in both (c) and (d). All simulations were performed with 100 demes and $r_0{=}0.1$.

For three out of the four parameter combinations in figs. 3a and 3b, we observe a critical migration threshold (denoted by Km_c) below which the metapopulation collapses. For $Km{<}Km_c$, patches accumulate deleterious alleles and dwindle in size faster than they can be rescued by migration, leading to their extinction. A noticeable feature of figs. 3a and 3b is the discrepancy between simulations and the semi-deterministic predictions, especially for Km_c. However, as expected from our preceding discussion, this discrepancy lessens with increasing K, which corresponds to weaker demographic fluctuations. Note also that for large K, i.e., when the semi-deterministic approximation is valid, extinction thresholds resemble 'tipping points', where a slight drop in migration sets in motion a positive feedback between increasing load and decreasing population size which causes relatively large populations to collapse abruptly.

The critical threshold Km_c is larger for smaller values of Ks and larger h. In other words, higher levels of migration are required to sustain populations when selection per locus is weak relative to drift and if alleles are additive rather than recessive (since purging is more effective for recessive alleles). On the flip side, when alleles are recessive, the population size per deme N is higher and load $R_g{=}1{-}N$ correspondingly

lower at the critical threshold Km_c. For example, in Figure 3a, extinction already ensues for $R_g{\sim}0.4$ if load is due to recessive alleles, since even a slight increase in homozygosity (due to reduced migration and/or declining population sizes) causes load to increase substantially, resulting in stronger eco-evo feedback. By contrast, populations can sustain much higher levels of load *without* collapsing, when load is due to additive alleles ($R_g{\sim}1{-}N{\sim}0.75$ at Km_c; blue lines in figs. 3a and 3b) .

Interestingly, the critical migration threshold Km_c observed in simulations (with finite K and r_0K) is *lower* than the semi-deterministic prediction, all other *scaled* parameters being the same. In other words, given Ks, $2LU$, Ku, smaller patches with lower carrying capacities K (and stronger demographic fluctuations) require fewer migrants per generation to escape extinction. A related finding is that just above the extinction threshold, scaled population sizes $N{=}n/K$ are actually higher when K is smaller. These observations appear somewhat puzzling, as one might individual patches to be less stable if demographic stochasticity is stronger. However, this is counterbalanced by the fact that demographic fluctuations result in a broader distribution of deme sizes overall, allowing some demes within the metapopulation to function as demographic sources— migration from these can then rescue other demes with declining populations.

This is illustrated in fig. 3c, which shows the equilibrium distribution of deme sizes for different values of K, for $Ks{=}1$, $h{=}0.02$ and $Km{=}1$ (which is just below the semi-deterministic prediction for Km_c). When carrying capacities are small (say $K{=}250$; blue), the distribution of scaled population sizes is rather broad. For larger K (i.e., with weaker demographic fluctuations), the distribution becomes more narrow, with fewer exceptionally large patches that can potentially act as demographic sources. As a result, the entire distribution *shifts* towards smaller N. At still larger values of K and, consequently, with narrower distributions, the entire population collapses, yielding an equilibrium size $N{=}0$ (which matches the semi-deterministic prediction).

The effect of demographic fluctuations is, however, qualitatively different in parameter regimes where the semi-deterministic approximation predicts viable populations, i.e., non-zero equilibrium size (Figure 3d). At low K (e.g. $K{=}250$; blue in Figure 3d), population sizes show a bimodal distribution, with a small fraction of demes close to extinction and the remaining fraction peaked at a non-zero N_*. Now, as K increases, the distribution around N_* becomes narrower and smaller sub-populations become less prone to collapse, resulting in fewer nearly extinct demes. In fact, for $K{=}2000$, the distribution becomes unimodal, with all demes having sizes close to $N_*{\sim}0.56$ which coincides with the semi-deterministic prediction (dashed vertical black line). Similar plots are shown for $h{=}0.5$ in Appendix C, SI .

In what follows, we investigate how Km_c, the critical level of migration required to prevent population collapse, depends on the architecture of load, i.e., the (scaled) selection strength and dominance coefficient per locus and the total deleterious mutation rate relative to baseline growth rate. We only show predictions of the semi-deterministic approximation with the understanding that these should be accurate if r_0K is sufficiently large, but that Km_c may be somewhat lower for smaller r_0K. Importantly, we always start from a state in which the metapopulation is large and well-adapted. An alternative scenario (which we do not consider) is one where the metapopulation grows (or fails to grow) starting from just a few individuals per patch. In this case, a much higher level of migration between patches may be required to alleviate genetic Allee effects and initiate population growth (see also Sachdeva et al., 2022).

Figures 4a and 4b show that there is a non-monotonic relationship between the critical migration threshold

Km_c and the scaled strength of selection per locus, Ks. Moreover, Km_c is higher for larger values of h. This reflects the fact that load per locus is largest for alleles of moderate selective effect Ks, and for additive as opposed to recessive alleles (fig. 1c); thus the level of migration required to alleviate load and prevent extinction is also highest for intermediate values of Ks and larger h. Interestingly, the load $R_{g,c}$ just above the critical migration threshold Km_c is lower when Ks is large (dashed vs. solid lines in fig. 4e). In other words, the maximum amount of load (relative to baseline growth rate) that populations can sustain without collapsing is lower with stronger selection per locus. This is because while deleterious alleles with large values of Ks are less likely to fix, the loss in fitness when fixation does occur is higher (and the resultant eco-evo feedback between load and population size stronger) for larger Ks.

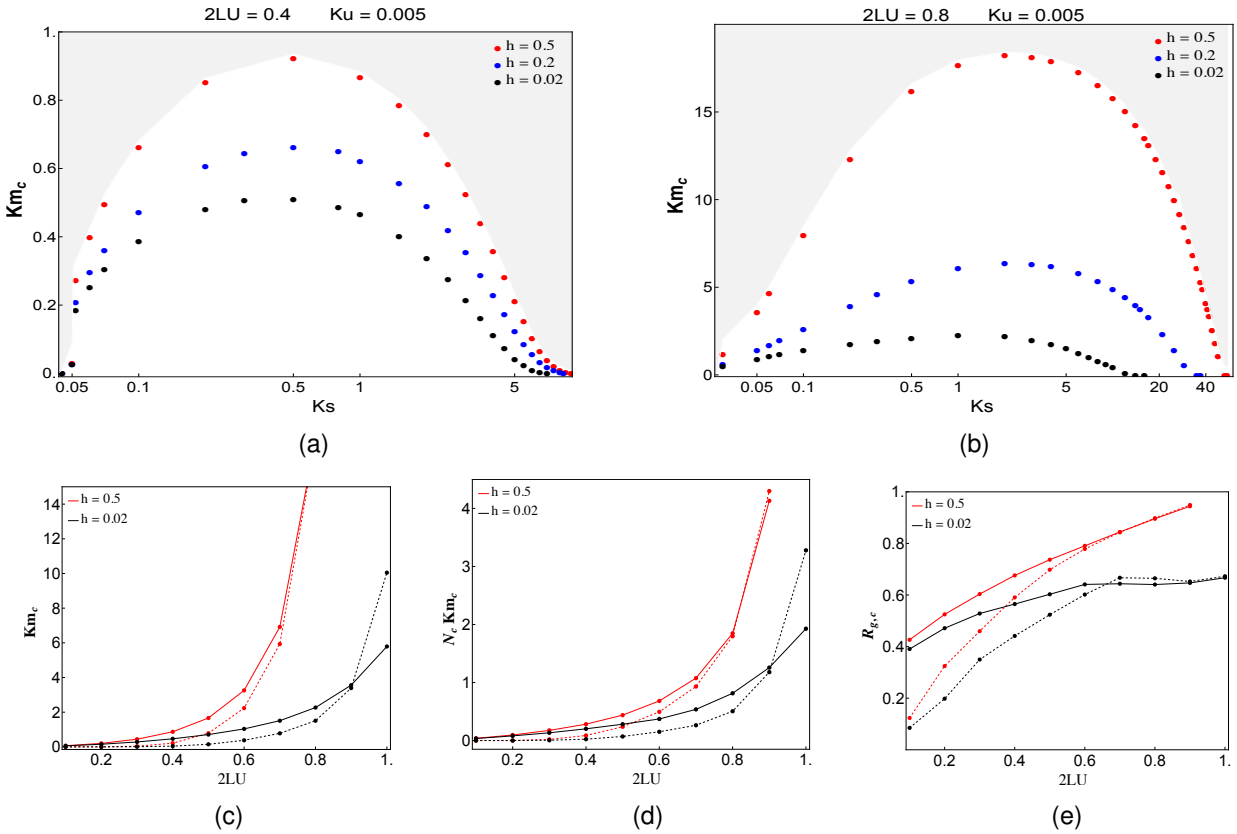

Figure 4: (a)-(b) Critical migration thresholds Km_c (below which the metapopulation goes extinct) as a function of Ks for (a) $2LU{=}(2Lu)/r_0{=}0.4$ (corresponding to moderately hard selection) and (b) $2LU{=}0.8$ (much harder selection), for $Ku{=}0.005$. Figures (c)-(e) show: (c) the critical migration threshold, Km_c (d) the number of migrants per deme , N_cKm_c, just above the critical migration threshold, and (d) the scaled load per deme, $R_{g,c}$, just above the critical migration threshold, as a function of $2LU$, which is a measure of the hardness of selection. Solid and dashed lines represent results with $Ks{=}1$ and $Ks{=}5$ respectively; different colors correspond to different values of h. All predictions are obtained from the semi-deterministic analysis (using eq. (3)), which is accurate when carrying capacity per deme is large.

A key determinant of the eco-evo feedback is $2LU{=}(2Lu)/r_0$, the total deleterious mutation rate (which determines load in the absence of drift) relative to r_0. As discussed above, $2LU$ can be viewed as a proxy for the 'hardness' of selection. Figures 4a and 4b depict Km_c for $2LU{=}0.4$ and $2LU{=}0.8$ respectively, and show that critical migration thresholds required to prevent collapse can increase dramatically as selection becomes harder. Moreover, the relatively small difference in load per locus between additive and recessive alleles that we observe at a given level of migration under soft selection (fig. 1c) can translate into rather different critical migration thresholds for the two kinds of alleles under hard selection. This is especially marked when $2LU$ is higher (fig. 4b), as a small increase in fixation probability and load per locus can translate into a substantial change in total load and population size, which further increases

fixation probabilities, resulting in very strong feedback effects.

Some care is needed in the interpretation of Km_c: this is best thought of as the minimum rate of migration (scaled by K) that allows the metapopulation to avoid collapse. It is thus equal to the number of migrants per deme that one would observe in a *hypothetical* metapopulation at carrying capacity, given this migration rate. It is not the number of migrants per deme in the actual metapopulation at or just above the critical migration threshold, which is given by N_cKm_c, where N_c is the (scaled) population size per deme at this migration threshold. Figures 4d and 4e show the average number of migrants N_cKm_c and the average (scaled) load $R_{g,c}$ per deme at the critical extinction threshold, as a function of $2LU$, for the various combinations of Ks and h. Thus, for instance, for the parameters $2LU{=}0.6$, $Ks{=}1$ and $h{=}0.5$, a minimum of $N_cKm_c \sim 0.75$ migrants per deme per generation are required for the metapopulation to avoid collapse (fig. 4d). The (scaled) load per deme in this metapopulation (just above the extinction threshold) is $R_{g,c}{\sim}0.78$ (fig. 4e), and the associated scaled population size per deme is $N_c = 1 - R_{g,c} \sim 0.22$. Thus, when demes are at a small fraction (here, $\sim 22\%$) of carrying capacity, 0.75 migrants per deme would still require a rather high rate of migration per individual– in this example, $Km_c = 0.75/0.22 \sim 3.4$ (fig. 4c), which translates into 3.4 migrants per generation between demes at carrying capacity.

Figure 4c shows that the critical migration threshold Km_c always increases with increasing $2LU$, regardless of h and Ks. Moreover, the critical number N_cKm_c of migrants per deme that is required to avoid extinction also increases with $2LU$, despite the fact that the (scaled) population size per deme N_c at the extinction threshold is lower at higher $2LU$. In other words, as selection becomes increasingly hard (either due to a higher genomewide deleterious mutation rate or lower baseline growth rate or both), each deme must exchange a larger number of migrants with other demes to avoid extinction, despite there being fewer individuals (that can potentially migrate) per deme.

Finally, we note that while $R_{g,c}$ increases with $2LU$ (fig. 4e), this is not a simple linear increase, as one would expect if load (at the critical extinction threshold) were entirely deterministic and unaffected by genetic drift. Interestingly, drift is more relevant to extinction thresholds when selection is *less* hard (i.e., $2LU$ not too large), as evident from the fact that the (scaled) load $R_{g,c}$ is then significantly higher than the deterministic expectation $2LU$. This is because populations can remain stable despite a moderate inflation of load due to drift, as long as selection is not very hard. By contrast, with very hard selection ($2LU$ close to 1), load needs to be at its minimum possible value for populations to avoid extinction. For additive or weakly recessive alleles ($1/2{<}h{<}1/4$), this minimum is $2LU$ (the deterministic expectation), while for highly recessive alleles, load can be significantly less than $2LU$ due to purging.

Load and population size under high gene flow

Having analysed extinction thresholds, we now consider the $Km{>}Km_c$ regime, where populations are stable. In this parameter regime, population sizes depend only weakly on Km, especially when alleles are recessive (fig. 3a), with load close to the deterministic prediction above Km_c (fig. 5b). This is contrary to previous findings (Whitlock, 2002) where lower gene flow is thought to reduce load, due to increased expression and consequently stronger purging of deleterious alleles in more isolated populations. To better understand our results, we plot the equilibrium allele frequency $\overline{p}$ and mean load R_g in the metapopulation

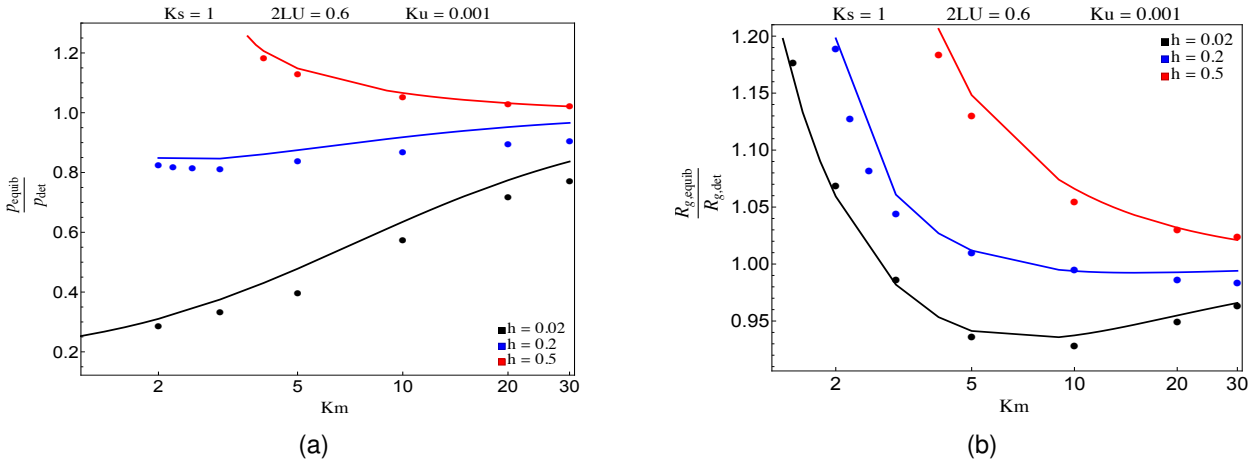

Figure 5: (a) Equilibrium frequency $\overline{p}$ of deleterious alleles and (b) equilibrium load R_g (relative to p_{det} and $R_{g,det}$, the deterministic allele frequency and load in a large undivided population), plotted against Km, for $Km{>}Km_c$. Filled circles show simulation results (using allele frequency simulations) and solid lines show results from the semi-deterministic approximation. The scaled parameters are $Ku{=}0.001$, $Ks{=}1$ and $2LU{=}0.6$; simulations are run with 100 demes with $K{=}3000$ and $r_0{=}0.1$.

as a function of Km, scaling these respectively by p_{det} and $R_{g,det}$, the corresponding (deterministic) allele frequency and load in a large undivided population (figs. 5a and 5b). The solid lines are results from the semi-deterministic approximation and colored circles represent simulation results. Here, we have obtained the deterministic p_{det} and $R_{g,det}$ in an undivided population by solving eq. (1a) at equilibrium and plugging the result into $R_{g,i}$ in eq. (1b) to get $R_{g,det}$; this is as opposed to assuming $p_{det}{\sim}u/hs$ and $R_{g,det}{\sim}2Lu$, which are overestimates when alleles are nearly fully recessive.

We see from fig. 5a that when alleles are (partially) recessive, there is a decrease in the frequency of the deleterious allele (as migration decreases) due to purging. However, fig. 5b shows that this purging effect is not strong enough to counter the increased inbreeding at lower migration rates. Thus, while increased purging does cause load to be slightly lower than the deterministic expectation for intermediate values of Km and highly recessive effects ($h{=}0.02$), this is not a significant effect. On the whole, we observe that load increases as populations become more isolated (i.e., as Km decreases), until it exceeds a critical threshold $R_{g,c}$, beyond which populations collapse. Similar figures for strongly selected alleles ($Ks{=}10$) are shown in appendix D, SI. Importantly, neglecting the fact that both purging and fixation of deleterious alleles depend on local deme sizes (which shrink with falling migration due to eco-evo feedbacks) can mislead, causing us to overestimate the significance of purging in subdivided populations.

Discussion

In this study, we have explored the varied effects of gene flow, selection and population growth on load and extinction in a metapopulation by focusing on how allele frequencies (at multiple loci) and population sizes co-evolve. We distinguish between two modes of selection: soft selection, which applies when total load is much less than the baseline growth rate (i.e., $2LU{\ll}1$), so that subpopulations are essentially at carrying capacity; and hard selection (with $2LU{\lesssim}1$), where local population sizes decline with increasing load (typically due to changes in allele frequency across multiple loci), and in turn, also affect load (e.g., via increased genetic drift). Our results thus provide insight into the intricate relationship between genetic

diversity, eco-evolutionary processes and the long-term persistence of fragmented populations.

It has long been recognized that the reduction in mean population fitness due to drift load, i.e., the stochastic fixation of weakly deleterious ($Ks{\sim}1$) mutations greatly exceeds the reduction due to strongly deleterious mutations that are essentially maintained at deterministic mutation-selection balance (Kimura et al., 1963; Kondrashov, 1995). Our results suggests that as long as selection is soft, populations have to be very strongly fragmented in order for drift load to be a real issue. As shown in fig. 1a, very little migration (as little as one migrant per ~ 10 generations, $Km{=}0.1$) between a set of interconnected patches is enough to reduce the load due to slightly deleterious mutations (i.e., those with $Ks{<}1$) by a factor of 100, bringing it close to the deterministic expectation, regardless of dominance coefficient.

At higher migration rates, load is purely due to segregating (rather than fixed) alleles, with selection against recessive alleles being most efficient at intermediate Km. This is consistent with previous work that shows that load due to partially recessive alleles is lowest at an intermediate level of migration– high enough to prevent local fixation of deleterious alleles but low enough to allow efficient purging of recessives (Whitlock (2002), Roze (2015), Zhou and Pannell (2010)). Here, we further expand previous theory on soft selection by developing two analytical approximations that describe the evolution of moderately deleterious alleles under low to moderate levels of gene flow. These include the 'moderate selection' approximation (lines in figs. 1a, 1c and 1d) that accurately captures how modest levels of selection ($0.5{<}Ks{<}5$) reduce local fixations (and thus F_{ST}) at individual loci, and an approximation based on effective migration rates that captures the effects of multi-locus heterosis on allele frequencies and load (fig. 2).

Models of soft selection are relevant for two reasons. First, some fitness components affect population size more than others (adult viability versus male mating success, say). Second, estimates of load under the soft selection model can be an upper bound to the load in real populations, answering the following question: if the effective population size stays the same, what fraction of offspring will fail to survive under the predicted burden of deleterious mutations?

In reality however, the accumulation of deleterious mutations is likely to decrease the size of the population, even leading to extinction (Kondrashov, 1995). Soft selection models do not take into account the positive feedback that ensues when the population size decreases due to genetic load, which in turn further increases genetic load (via increased drift), necessitating models with hard selection. Such feedback has been studied in earlier work on mutational meltdown (e.g., Lynch et al. 1995a,b), albeit only using simulations. Here, we develop a more comprehensive understanding using a semi-deterministic analysis which accounts for the stochastic effects of genetic drift on allele frequencies but neglects those of demographic stochasticity (Sachdeva et al., 2022).

A key result is that under hard selection, demes must exchange a much higher number of migrants per generation despite supporting fewer individuals (relative to carrying capacity)– this requires very high rates of migration per individual (figs. 4c and 4d). Further, critical migration thresholds Km_c are highly sensitive to $2LU$ (a proxy for the hardness of selection), with higher values of $2LU$ necessitating more gene flow. This is because when load is (potentially) due to many sites along the genome, then slight changes in deleterious mutation accumulation per site (e.g., due to slightly reduced gene flow and/or increased drift) can result in large changes in total load, more readily setting in motion the eco-evo feedback effects that lead to abrupt collapse and global extinction.

The parameter $2LU{=}(2Lu)/r_0$, which emerges as a crucial parameter in our analysis, captures the extent to which deleterious mutations (occurring at a genomewide rate $2Lu$) reduce the baseline growth rate r_0. Estimating $2LU$ thus requires knowledge of both the total deleterious mutation rate (e.g., from mutation accumulation experiments or genomic methods) and the baseline growth rate r_0 of a maximally fit reference population. However, r_0 is difficult to measure or even define unambiguously— for instance, one could define it either via the reproductive output of the 'least loaded' class in a specific population, or as the rate of growth associated with a hypothetical, perfectly fit genotype, with the former being (possibly) easier to measure but harder to generalise across populations. Another source of complexity is that deleterious mutation rates (or more generally, the amount of resources allocated to germline maintenance) may co-evolve with growth rates due to trade-offs between the two (Avila and Lehmann, 2023), so that composite parameters such as $2LU$ are somewhat constrained.

We also explore how the critical migration threshold Km_c for metapopulation persistence depends on selection and dominance coefficients of individual alleles, and find that Km_c (as well as N_cKm_c, the number of migrants exchanged per deme) is highest for intermediate values of Ks ($\sim$1) and for additive (as opposed to recessive) alleles (fig. 4). These results mirror those under soft selection, where load is highest for intermediate Ks and $h{\sim}1/2$ (fig. 1). The finding that populations can survive a higher degree of fragmentation (i.e., lower levels of gene flow) when load is primarily recessive appears to be at odds with verbal arguments that draw a (somewhat loose) connection between fragmentation and increased inbreeding load due to expression of recessives. Our analysis refines this argument by showing that when load is largely due to recessive alleles, mutational meltdown is triggered more 'easily' in the sense that population sizes per deme are relatively high (and load low) at the extinction threshold (fig. 4e). However, mutational meltdown is also more 'difficult' in the sense that populations can withstand much higher levels of fragmentation (without accumulating too much load), than if load were primarily additive.

Our stochastic (finite K) simulations show that close to the critical extinction threshold, local deme sizes follow a bimodal distribution (figs. 3c and 3d), with some demes teetering on the brink of extinction while others thrive with larger populations. Such bimodal distributions arise due to the stochastic nature of the model, even with uniform growth rates and carrying capacities across demes. Such bimodality can serve as an important signal of impending population collapse, where minor variations in migration rate or external factors can push some demes past the point of no return, causing local extinctions and a possible cascade towards global extinction (Scheffer et al., 2001, Drake and Griffen, 2010).

Our findings, while intuitive, can become obscured by the prevailing confusion around 'hard selection'. This term is invoked in two distinct classes of eco-evo models of subdivided populations. One class of models posits hard selection as a process wherein local population sizes decline with increasing load; these models thus explicitly track the co-evolution of population size and allele frequencies, and allow for local and global extinction (Polechová and Barton, 2015, Szép et al., 2021). By contrast, in the second class of models, hard vs. soft selection refers only to whether or not local populations contribute to the next generation in proportion to their fitness; in these models, population sizes are determined solely by some form of (local vs. global) density-dependent regulation (Christiansen, 1975, Whitlock, 2002). Thus, the latter class of models implicitly assume that the population as a whole is at carrying capacity, thereby neglecting any influence of genetic load on total population size and consequently, on the extent of drift. These two kinds of models (both ostensibly allowing for 'hard selection') yield qualitatively different

predictions. For example, we have shown that reduced gene flow is largely detrimental to populations—increasing drift load and pushing populations towards extinction. This is in contrast to previous work on hard selection in subdivided populations, which finds selection against deleterious alleles to be more efficient (due to increased purging) under conditions of low gene flow (Whitlock, 2002), but neglects any dependence of local or global population sizes on the genetic composition of the population.

What are some limitations of our study? Our theoretical framework is based on the infinite island model, which neglects explicit space. Although this is a useful simplification that ensures analytical tractability, in reality, spatial structure can strongly affect metapopulation dynamics. For instance, Bascompte and Solé (1996) argued that metapopulations embedded in space are more vulnerable to extinction than suggested by spatially implicit (e.g., infinite island) models – thus, the latter may underestimate critical thresholds and the impact of fragmentation on metapopulation persistence. This begs the question: are there specific spatial configurations more conducive to the maintenance of genetic diversity, and how can spatially and genetically explicit models inform practical decisions, e.g., the planning of migration corridors (Bennett, 1999) or assisted gene flow (Aitken and Whitlock, 2013) across continuous space?

Our study also makes various simplifying assumptions about the genetics of load. First, we neglect epistatic interactions between loci; negative epistasis, in particular, may eliminate deleterious alleles more efficiently (Kimura and Maruyama, 1966), though it has been argued that the magnitude of this effect is not sufficient to resolve the paradox of drift load (Charlesworth, 2013). Thus, generalising the present framework to incorporate epistasis could give broader insight into the relevance of epistatic fitness effects during mutational meltdown in fragmented populations. Second, we assume deleterious alleles to be unlinked, so that Hill-Robertson interference within populations is largely irrelevant. How might this change with lower rates of recombination overall or if we consider specific genomic regions with reduced or arrested recombination, e.g., inversions? On the one hand, reduced recombination should exacerbate drift load by rendering selection less efficient; on the other hand, multi-locus heterosis may be stronger. Understanding how the magnitude of these two counteracting effects depends on the extent of gene flow in subdivided populations remains an interesting direction for future work.

Finally, we have considered selection that is uniform across space. It would be interesting to extend our framework to scenarios where some loci are under spatially heterogeneous selection (as in Szép et al. (2021)) while others are subject to unconditionally deleterious mutation. While gene flow may potentially swamp local adaptation, it could also concurrently alleviate the burden of deleterious mutations. Determining the relative strengths of these contrasting effects under realistic parameter regimes is crucial for a more comprehensive understanding of evolution in fragmented, heterogeneous landscapes.

Acknowledgements.

This research was partially funded by the Austrian Science Fund (FWF) [FWF P-32896B] and DOC Fellowships of the Austrian Academy of Sciences: grants 26380 (O. Olusanya) and 26293 (K. Khudiakova). We thank Nick Barton for useful comments on the thesis chapter that led to this manuscript.

References

Aitken, S. N. and Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual review of ecology, evolution, and systematics*, 44:367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747.

Avila, P. and Lehmann, L. (2023). Life history and deleterious mutation rate coevolution. *Journal of Theoretical Biology*, 573:111598.

Barton, N. and Olusanya, O. (2022). The response of a metapopulation to a changing environment. *Philosophical Transactions of the Royal Society B*, 377(1848):20210009. https://doi.org/10.1098/rstb.2021.0009.

Bascompte, J. and Solé, R. V. (1996). Habitat fragmentation and extinction thresholds in spatially explicit models. *Journal of Animal ecology*, pages 465–473. https://doi.org/10.2307/5781.

Bennett, A. F. (1999). *Linkages in the landscape: the role of corridors and connectivity in wildlife conservation*. Number 1. Iucn.

Casas-Marce, M., Soriano, L., López-Bao, J. V., and Godoy, J. A. (2013). Genetics at the verge of extinction: insights from the iberian lynx. *Molecular ecology*, 22(22):5503–5515. https://doi.org/10.1111/mec.12498.

Charlesworth, B. (2013). Why we are not dead one hundred times over: Brief communication. *Evolution*, 67(11):3354–3361. https://doi.org/10.1111/evo.12195.

Charlesworth, D. and Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual review of ecology and systematics*, 18(1):237–268. https://doi.org/10.1146/annurev.es.18.110187.001321.

Christiansen, F. B. (1975). Hard and soft selection in a subdivided population. *The American Naturalist*, 109(965):11–16. https://doi.org/10.1086/282970.

Denver, D. R., Morris, K., Lynch, M., and Thomas, W. K. (2004). High mutation rate and predominance of insertions in the caenorhabditis elegans nuclear genome. *Nature*, 430(7000):679–682. https://doi.org/10.1038/nature02697.

Drake, J. M. and Griffen, B. D. (2010). Early warning signals of extinction in deteriorating environments. *Nature*, 467(7314):456–459. https://doi.org/10.1038/nature09389.

Edmands, S. (2007). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular ecology*, 16(3):463–475. https://doi.org/10.1111/j.1365-294X.2006.03148.x.

Eyre-Walker, A., Woolfit, M., and Phelps, T. (2006). The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics*, 173(2):891–900.

Fagan, W. F. and Holmes, E. (2006). Quantifying the extinction vortex. *Ecology letters*, 9(1):51–60. https://doi.org/10.1111/j.1461-0248.2005.00845.x.

Finger, A., Kettle, C., Kaiser-Bunbury, C., Valentin, T., Doudee, D., Matatiken, D., and Ghazoul, J. (2011). Back from the brink: potential for genetic rescue in a critically endangered tree. *Molecular Ecology*, 20(18):3773–3784. https://doi.org/10.1111/j.1365-294X.2011.05228.x.

Frankham, R. (1998). Inbreeding and extinction: island populations. *Conservation biology*, 12(3):665–675. https://doi.org/10.1046/j.1523-1739.1998.96456.x.

Frankham, R., Ballou, J. D., Eldridge, M. D., Lacy, R. C., Ralls, K., Dudash, M. R., and Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, 25(3):465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x.

Frankham, R., Briscoe, D. A., and Ballou, J. D. (2002). *Introduction to conservation genetics*. Cambridge university press.

Glémin, S., Ronfort, J., and Bataillon, T. (2003). Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics*, 165(4):2193–2212. https://doi.org/10.1093/genetics/165.4.2193.

Haag-Liautard, C., Dorris, M., Maside, X., Macaskill, S., Halligan, D. L., Charlesworth, B., and Keightley, P. D. (2007). Direct estimation of per nucleotide and genomic deleterious mutation rates in drosophila. *Nature*, 445(7123):82–85. https://doi.org/10.1038/nature05388.

Hairston Jr, N. G., Ellner, S. P., Geber, M. A., Yoshida, T., and Fox, J. A. (2005). Rapid evolution and the convergence of ecological and evolutionary time. *Ecology letters*, 8(10):1114–1127. https://doi.org/10.1111/j.1461-0248.2005.00812.x.

Higgins, K. and Lynch, M. (2001). Metapopulation extinction caused by mutation accumulation. *Proceedings of the National Academy of Sciences*, 98(5):2928–2933. https://doi.org/10.1073/pnas.031358898.

Ingvarsson, P. K. and Whitlock, M. C. (2000). Heterosis increases the effective migration rate. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 267(1450):1321–1326. https://doi.org/10.1098/rspb.2000.1145.

Keightley, P. D. (2012). Rates and Fitness Consequences of New Mutations in Humans. *Genetics*, 190(2):295–304. https://doi.org/10.1534/genetics.111.134668.

Kimura, M. and Maruyama, T. (1966). The mutational load with epistatic gene interactions in fitness. *Genetics*, 54(6):1337. https://doi.org/10.1093/genetics/54.6.1337.

Kimura, M., Maruyama, T., and Crow, J. F. (1963). The mutation load in small populations. *Genetics*, 48(10):1303–1312. https://doi.org/10.1093/genetics/48.10.1303.

Kondrashov, A. S. (1995). “contamination of the genome by very slightly deleterious mutations: Why have we not died 100 times over?”. *Journal of Theoretical Biology*, 175(4):583–594. https://doi.org/10.1006/jtbi.1995.0167.

Land, E. D. and Lacy, R. C. (2000). Introgression level achieved through florida panther genetic restoration. *Endangered Species Update*, 17(5):100–100.

Lande, R. (1993). Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist*, 142(6):911–927. https://doi.org/10.1086/285580.

Lande, R., Engen, S., and Saether, B.-E. (2003). *Stochastic population dynamics in ecology and conservation*. Oxford University Press, USA.

Levene, H. (1953). Genetic equilibrium when more than one ecological niche is available. *The American Naturalist*, 87(836):331–333. https://doi.org/10.1086/281792.

Lyman, R. F., Lawrence, F., Nuzhdin, S. V., and Mackay, T. F. (1996). Effects of single p-element insertions on bristle number and viability in drosophila melanogaster. *Genetics*, 143(1):277–292. https://doi.org/10.1093/genetics/143.1.277.

Lynch, M., Blanchard, J., Houle, D., Kibota, T., Schultz, S., Vassilieva, L., and Willis, J. (1999). Perspective: spontaneous deleterious mutation. *Evolution*, 53(3):645–663. https://doi.org/10.2307/2640707.

Lynch, M., Conery, J., and Burger, R. (1995a). Mutation accumulation and the extinction of small populations. *The American Naturalist*, 146(4):489–518. https://doi.org/10.1086/285812.

Lynch, M., Conery, J., and Bürger, R. (1995b). Mutational meltdowns in sexual populations. *Evolution*, 49(6):1067–1080. https://doi.org/10.2307/2410432.

Matthies, D., Bräuer, I., Maibom, W., and Tscharntke, T. (2004). Population size and the risk of local extinction: empirical evidence from rare plants. *Oikos*, 105(3):481–488. https://doi.org/10.1111/j.0030-1299.2004.12800.x.

Mukai, T. (1964). The genetic structure of natural populations of drosophila melanogaster. i. spontaneous mutation rate of polygenes controlling viability. *Genetics*, 50(1):1. https://doi.org/10.1093/genetics/50.1.1.

Mukai, T., Chigusa, S. I., Mettler, L., and Crow, J. F. (1972). Mutation rate and dominance of genes affecting viability in drosophila melanogaster. *Genetics*, 72(2):335–355. https://doi.org/10.1093/genetics/72.2.335.

Olusanya, O., Khudiakova, K., and Sachdeva, H. (2023). Genetic load, eco-evolutionary feedback and extinction in a metapopulation. https://git.ista.ac.at/oolusany/genetic-load-and-extinction-in-a-metapopulation.

Polechová, J. and Barton, N. H. (2015). Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences*, 112(20):6401–6406.

Ravigné, V., Olivieri, I., and Dieckmann, U. (2004). Implications of habitat choice for protected polymorphysms.

Roze, D. (2015). Effects of interference between selected loci on the mutation load, inbreeding depression, and heterosis. *Genetics*, 201(2):745–757. https://doi.org/10.1534/genetics.115.178533.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., and Hanski, I. (1998). Inbreeding and extinction in a butterfly metapopulation. *Nature*, 392(6675):491–494. https://doi.org/10.1038/33136.

Sachdeva, H. (2022). Reproductive isolation via polygenic local adaptation in sub-divided populations: Effect of linkage disequilibria and drift. *PLoS Genetics*, 18(9):e1010297. https://doi.org/10.1371/journal.pgen.1010297.

Sachdeva, H., Olusanya, O., and Barton, N. (2022). Genetic load and extinction in peripheral populations: the roles of migration, drift and demographic stochasticity. *Philosophical Transactions of the Royal Society B*, 377(1846):20210010. https://doi.org/10.1098/rstb.2021.0010.

Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., and Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413(6856):591–596. https://doi.org/10.1038/35098000.

Simmons, M. J. and Crow, J. F. (1977). Mutations affecting fitness in drosophila populations. *Annual review of genetics*, 11(1):49–78. https://doi.org/10.1146/annurev.ge.11.120177.000405.

Szczecińska, M., Sramko, G., Wołosz, K., and Sawicki, J. (2016). Genetic diversity and population structure of the rare and endangered plant species pulsatilla patens (l.) mill in east central europe. *PloS one*, 11(3):e0151730. https://doi.org/10.1371/journal.pone.0151730.

Szép, E., Sachdeva, H., and Barton, N. H. (2021). Polygenic local adaptation in metapopulations: A stochastic eco-evolutionary model. *Evolution*, 75(5):1030–1045. https://doi.org/10.1111/evo.14210.

Templeton, A. R. (1986). Coadaptation and outbreeding depression. *Conservation biology: the science of scarcity and diversity*.

Thompson, J. N. (1998). Rapid evolution as an ecological process. *Trends in ecology & evolution*, 13(8):329–332. https://doi.org/10.1016/S0169-5347(98)01378-0.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G., and Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, 9(7):892–908.

Whitlock, M. C. (2002). Selection, Load and Inbreeding Depression in a Large Metapopulation. *Genetics*, 160(3):1191–1202. https://doi.org/10.1093/genetics/160.3.1191.

Wright, S. (1937). The distribution of gene frequencies in populations. *Proceedings of the National Academy of Sciences of the United States of America*, 23(6):307. https://doi.org/10.1073/pnas.23.6.307.

Zhou, S.-R. and Pannell, J. R. (2010). Inbreeding depression and genetic load at partially linked loci in a metapopulation. *Genetics Research*, 92(2):127–140. https://doi.org/10.1017/S0016672310000133.

Zwaenepoel, A., Sachdeva, H., and Fraïsse, C. (2023). Polygenic barriers to gene flow: the role of dominance, haploid selection and heterogeneous genetic architectures. *bioRxiv*. https://www.biorxiv.org/content/early/2023/09/25/2023.09.24.559235.

SUPPLEMENTARY INFORMATION

Genetic load, eco-evolutionary feedback and extinction in metapopulations

Appendix A. Details of allele frequency simulations

We assume a life cycle consisting of migration between patches, mutation, followed by viability selection on adults

To account for the effect of migration on population size, we first determine the net number of migrants, $netM(t)$ (i.e., the difference between the total number of immigrants and emigrants in a given patch) and add this to the existing patch size. Under soft selection, we assume a zero net migration rate (i.e., a balance between the number of immigrants and emigrants) so as to keep the patch size fixed. However, under hard selection, the net number of migrant is assumed non-negative. Migration therefore changes the population size according to $N_i'(t)=N_i(t)$ (under soft selection) and $N_i'(t)=N_i(t)+netM(t)$ (under hard selection). Similarly, to account for the effect of migration on allele frequencies in any patch, we add the net number of migrant alleles, $netp(t)$ to the existing allele copy number in the patch and divide this by the population size $N_i'(t)$. In essence, migration changes allele frequencies according to $p_{i,j}'(t)=(N_i p_{i,j}'(t)+netp(t))/N'(t)$.

With regards to the effect of mutation, we assume equal mutation rate to and from the deleterious allele so that mutation changes allele frequency according to $p_{i,j}''(t)=p_{i,j}'(t)+\mu(1-2p_{i,j}'(t))$. However, mutation has no effect on population size (under hard selection) so that $N_i''(t)=N_i'(t)$.

Following mutation, adults in each patch mate to produce offspring that survive to be next generation parents. Under soft selection, the mating process in each patch involves sampling with replacement, $N_i(t)$ pair of individuals based on their fitness and freely recombining their gametes to form offspring gametes. This means that we do not explicitly distinguish between the male and female sexes and there is also the possibility of self mating. Under hard selection on the other hand, load and density-dependent regulation changes population size according to $N_i'''(t)=N_i''(t)e^{1-N_i''(t)-R_g}$ and demographic stochasticity (randomness in birth and death) is further imposed on the population to determine how much survivable offspring, $N_i''''(t)$ can be formed. The latter is achieved by sampling from a Poisson distribution with parameter $N_i'''(t)$. Offspring are then formed by choosing with replacement $N_i''''(t)$ pairs of parents also based on their fitness and freely recombining their gametes. For both soft and hard selection, selection changes allele frequency according to $p_{i,j}'''(t)=(p_{i,j}''(t)\ e^{-s}+q_{i,j}''(t)\ e^{-hs})/\overline{w}$ where $\overline{w}=p_{i,j}''(t)^2\ e^{-s}+2p_{i,j}''(t)\ q_{i,j}''(t)\ e^{-hs}+q_{i,j}''(t)^2$ is the mean fitness in the ith patch and $q_{i,j}''(t)=1-p_{i,j}''(t)$.

Finally, the allele frequency at the end of the generation, i.e., $p''''_{i,j}(t)$ is obtained by sampling $p'''_{i,j}(t)$ from a Binomial distribution with parameters, $N''''(t)$, $p'''_{i,j}(t)$ thus accounting for genetic drift.

Appendix B. 'Moderate selection' approximation under soft selection

Under the soft selection model, any quantity such as load, F_{ST}, etc. for a single locus can be computed from the equilibrium allele frequency distribution $\psi[p|\overline{p}]$, by first numerically solving $\overline{p} = \int p\ \psi[p|\overline{p}]\ dp$ to obtain the mean allele frequency $\overline{p|\overline{p}}$, then plugging this into the equilibrium distribution $\psi[p|\overline{p}]$, and finally integrating over $\psi[p]$ to obtain higher moments. However, it is also useful to consider an alternative approach based on moments (or cumulants) of $\psi[p]$.

In general, any moment (such as $\mathbb{E}[p]$ or $\mathbb{E}[p^2]$) will depend on higher moments, resulting in a set of recursions that is not closed. Approximations thus rely on closing this set of recursions in different ways, depending on assumptions about the relative magnitudes of Ks (or Ksh) and Km (Whitlock (2002); Glémin et al. (2003); Roze (2015)). Here, we introduce a new moment closure approximation, which applies also for recessive ($h{\sim}0$) alleles and intermediate selection coefficients ($Ks{\sim}1$).

As in the main text, let p denote the frequency of the deleterious allele at a given locus in a given deme and $\overline{p}$ the mean across all demes in the population. We denote the expected change in allele frequency per unit time by $M(p)$ and the variance of the change by $V(p)$. These are:

$$M(p) = -s(hp+(1-3h)p^2-(1-2h)p^3)+u(1-2p)+m(\overline{p}-p) \tag{1a}$$

$$V(p) = \frac{p(1-p)}{2K} \tag{1b}$$

where K is the number of individuals per deme. Let $\mathbb{E}_\psi[f(p)]$ denote the expectation of any function $f(p)$ of the allele frequency p over the frequency distribution $\psi[p]$. Under the diffusion approximation, $\mathbb{E}_\psi[f(p)]$ satisfies (see also Ohta and Kimura (1971)):

$$\frac{d\,\mathbb{E}_\psi[\,f(p)\,]}{d\,t} = \mathbb{E}_\psi\left[M(p)f'(p)+\frac{1}{2}V(p)\,f''(p)\right] \tag{2}$$

Setting $f(p)$ as p and p^2 yields the following equations for the first and second moments of the allele frequency distribution respectively (see also Whitlock (2002), Glémin et al. (2003)):

$$\frac{d\,\mathbb{E}[\,p\,]}{d\,t} = -s\left(h\mathbb{E}[p]+(1-3h)\mathbb{E}[p^2]-(1-2h)\mathbb{E}[p^3]\right)+u\left(1-2\mathbb{E}[p]\right) \tag{3a}$$

$$\begin{aligned}\frac{d\,\mathbb{E}[\,p^2\,]}{d\,t} = &-s\left(h\mathbb{E}[p^2]+(1-3h)\mathbb{E}[p^3]-(1-2h)\mathbb{E}[p^4]\right)+u\left(\mathbb{E}[p]-2\mathbb{E}[p^2]\right)\\ &+m\left(\mathbb{E}[p]\,\overline{p}-\mathbb{E}[p^2]\right)+\frac{\mathbb{E}[p]-\mathbb{E}[p^2]}{4K}\end{aligned} \tag{3b}$$

It will be useful to express the above equations in terms of appropriately scaled cumulants of the frequency

distribution:

$$F_{ST}=\frac{\mathbb{E}[p^2]-\overline{p}^2}{\overline{p}(1-\overline{p})} \qquad \gamma=\frac{\mathbb{E}[p^3]-3\overline{p}\,\mathbb{E}[p^2]+2\overline{p}^3}{\overline{p}-3\overline{p}^2+2\overline{p}^3} \qquad \kappa=\frac{\mathbb{E}[p^4]-4\overline{p}\,\mathbb{E}[p^3]+6\overline{p}^2\,\mathbb{E}[p^2]-3\overline{p}^4}{\overline{p}-4\overline{p}^2+6\overline{p}^3-3\overline{p}^4} \tag{4}$$

Here, F_{ST}, γ and κ denote respectively the variance, skew and kurtosis of the frequency distribution within any deme, scaled by the corresponding cumulant for the distribution of frequencies across all demes.

At equilibrium, moments of the frequency distribution are constant in time, i.e., $d\,\mathbb{E}[\,p\,]/dt=d\,\mathbb{E}[\,p^2\,]/dt=0$; further, $\mathbb{E}[p]=\overline{p}$. Combining equations (3) and (4), and assuming deleterious alleles to be sufficiently rare *overall* in the population that $\mathbb{O}[\overline{p}^2]$ terms can be neglected, we have at equilibrium:

$$0=\frac{u}{s}-[\,h+F_{ST}(1-3h)-\gamma(1-2h)\,]\,\overline{p} \tag{5a}$$

$$0=[1-F_{ST}(1+4Km)-4Ks(hF_{ST}+(1-3h)\gamma-(1-2h)\kappa]\overline{p} \tag{5b}$$

This *pair* of equations is under-determined as it involves *four* variables $\overline{p}$, F_{ST}, γ and κ. To obtain an approximate solution, we assume further that third and higher cumulants i.e., the skew and kurtosis are related to F_{ST} as in the *neutral* infinite-island model. Note that F_{ST} itself is not taken to be neutral (or unaffected by selection), as is assumed by Whitlock (2002), but only that:

$$\gamma=\frac{2F_{ST}^2}{1+F_{ST}} \qquad \kappa\approx\frac{6F_{ST}^3}{(1+F_{ST})(1+2F_{ST})}+\mathbb{O}(\overline{p}) \tag{6}$$

Substituting into eq. (5) and expressing load as $G=s[\overline{p}-(1-2h)(1-F_{ST})\overline{p}\,\overline{q}]\approx s\overline{p}[2h+(1-2h)F_{ST}]$, we obtain:

$$\overline{p}=\frac{u}{s}\frac{1+F_{ST}}{(1-F_{ST})(h+(1-h)F_{ST})} \qquad G=u\frac{(1+F_{ST})(2h+(1-2h)F_{ST})}{(1-F_{ST})(h+(1-h)F_{ST})} \tag{7a}$$

$$\text{where} \quad (1+F_{ST})(1+2F_{ST})(1-F_{ST}(1+4Km))-4KsF_{ST}(1-F_{ST})(h+2F_{ST}(1-h))=0 \tag{7b}$$

Equation (7b) is cubic in F_{ST} and can be solved (e.g., numerically); this solution for F_{ST} can then be substituted into eq. (7a) to obtain the average deleterious allele frequency $\overline{p}$ and load G. Thus, in essence, our approximation allows for selection that is strong enough to change F_{ST} but not the relationship between higher cumulants (γ, κ etc.) and F_{ST}.

It is useful to juxtapose this approximation with those used in earlier work. Whitlock (2002) assumes *weak selection* such that F_{ST} is essentially neutral, i.e., equal to $1/(1+4Km)$. This is valid when $Ks\ll 1$, so that the term involving Ks in eq. (7b) (or equivalently in eq. (5b)) can be neglected. At the other extreme, Roze (2015) and Glémin et al. (2003) consider a parameter regime where the local deme size K is large enough that allele frequency distributions are essentially concentrated around the mean $\overline{p}$, which is relatively low. In practice, this means that F_{ST} is small enough that $\mathcal{O}(F_{ST}^2)$ terms in eq. (7b) can be neglected, which gives: $F_{ST}\approx\frac{1}{4Km+4Khs}$. Theirs is thus a *strong selection, strong migration* approximation; in practice, it applies for $Ksh+Km$ greater than 5. By contrast, the approximation for F_{ST} introduced above applies

also for intermediate Ks (see fig. 1d in the main text) and is thus a *moderate selection* approximation: in essence, it captures the key effect of selection which is to change coalescence times within and between demes (to *different* extents) without necessarily changing the extent to which the branching structure of genealogies is (a-)symmetric. However, this is only a rough interpretation, and more rigorous analyses are required to understand intermediate Ks regimes where neither selection nor drift can be treated as minor perturbations (to neutral or deterministic predictions respectively).

Relaxing the assumption of $\overline{p}\ll 1$. All the approximations described above assume that $\overline{p}$ is sufficiently low (i..e, deleterious alleles so far from global fixation) that $\mathcal{O}(\overline{p}^2)$ terms can be neglected. When this is no longer true (but if (6) still applies, i.e., if third and higher cumulants depend on F_{ST} as in the neutral model), equations (3), (4) and (6) together yield a cubic equation for $\overline{p}$:

$$\frac{u}{s}\left(\frac{1+F_{ST}}{1-F_{ST}}\right)=\left[F_{ST}+h(1-F_{ST})+\frac{2u}{s}\left(\frac{1+F_{ST}}{1-F_{ST}}\right)\right]\overline{p}+[1-2F_{ST}-3h(1-F_{ST})]\overline{p}^2 -(1-F_{ST})(1-2h)\overline{p}^3 \tag{8}$$

Since deleterious allele frequencies are expected to be high only for very weak selection, we can typically neglect the effects of selection on F_{ST} in this parameter regime, and simply solve equation (8) for $\overline{p}$ under the assumption $F_{ST}\approx\frac{1}{1+4Km}$.

Appendix C. Comparison of semi-deterministic predictions with simulations ($h=1/2$)

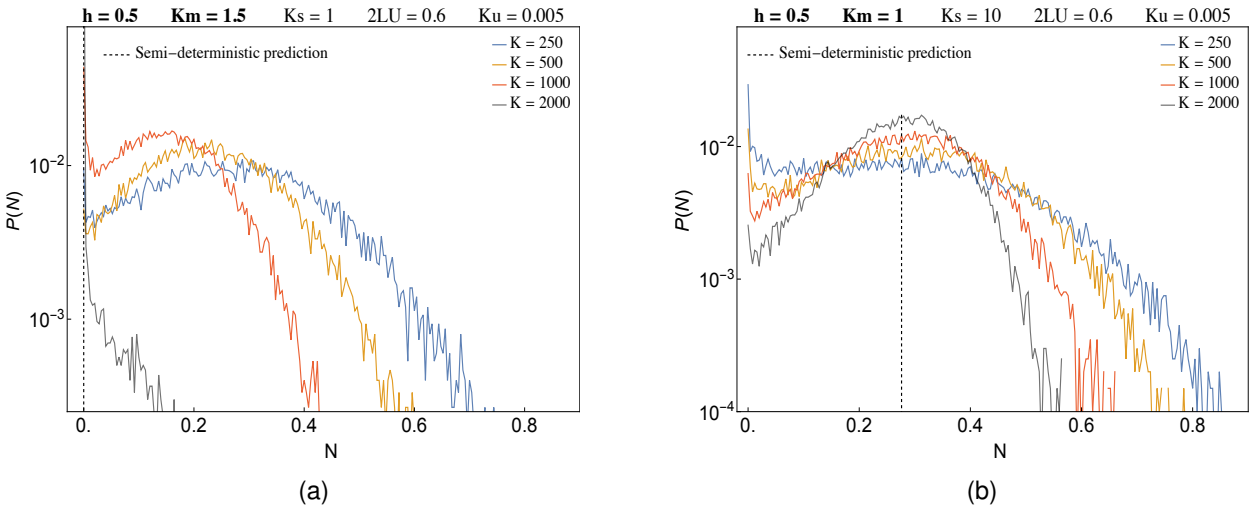

Figure 1: Comparison of simulation results with semi-deterministic predictions when load is due to additive alleles. Figures show equilibrium distribution $P(N)$ of population sizes N, as obtained from simulations, for various values of K, for two different scaled parameter combinations for which the semi-deterministic approximation predicts extinction (in (a)) vs. stable populations (in (b)). The scaled parameters Ks, Ku, Km, $2LU{=}2L(u/r_0)$ are held constant as K is varied (by simultaneously varying s, u, L, and m). The scaled parameters have values: (c) $Ks{=}1$, $Km{=}1.5$, or (d) $Ks{=}10$, $Km{=}1$, with other parameters being $h{=}0.5$, $Ku{=}0.005$ and $2LU{=}0.6$, as above. Black vertical dashed lines represent the semi-deterministic prediction for N: this is $N{=}0$ in (c) and $N{=}0.276$ in (d). The distribution $P(N)$ becomes sharply peaked around the semi-deterministic prediction as K increases in both (c) and (d). All simulations were performed with 100 demes and $r_0{=}0.1$.

Here, we demonstrate the convergence of simulation results to the semi-deterministic prediction with increasing K more generally, by considering load due to additive ($h{=}1/2$) alleles, in contrast to figures 3c and 3d in the main text, which show results for nearly recessive ($h{=}0.02$) alleles. As in the main text, we consider two different parameter combinations for which the semi-deterministic approximation predicts global extinction (fig. 1a) vs. stable populations with $N_*{=}0.276$ (fig. 1b). As in the case of recessive alleles (figures 3c and 3d in the main text), we find that the distribution of scaled population sizes $P(N)$ becomes narrower with increasing K (which corresponds to weakening demographic fluctuations). In fig. 1a, this causes the mean population size to decrease (since there are fewer large islands that can act as demographic sources), with the entire population nearly extinct at $K{=}2000$ (as also predicted by the semi-deterministic analysis). By contrast, in fig. 1b the distribution of deme sizes gets concentrated around a finite N_* (which matches the semi-deterministic prediction) with increasing K.

Appendix D. Effect of gene flow on equilibrium allele frequency and equilibrium mean load for $Ks{=}10$.

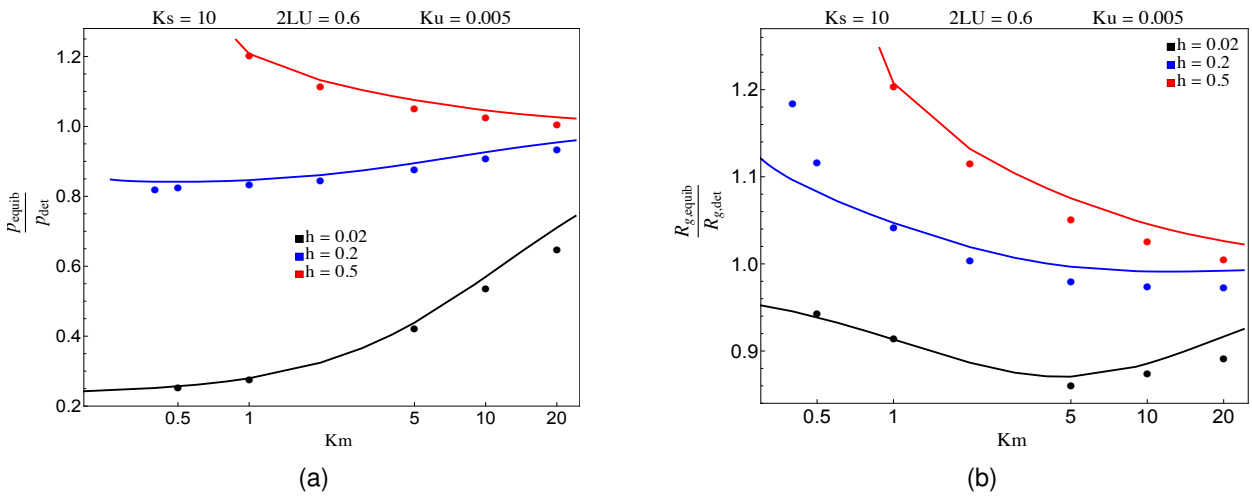

Figure 2: (a) Equilibrium frequency $\overline{p}$ of deleterious alleles and (b) equilibrium load R_g (relative to p_{det} and $R_{g,det}$, the deterministic allele frequency and load in a large undivided population), plotted against Km, for $Km{>}Km_c$. Filled circles show simulation results (using allele frequency simulations) and solid lines show results from the semi-deterministic approximation. The scaled parameters are $Ku{=}0.001$, $Ks{=}1$ and $2LU{=}0.6$; simulations are run with 100 demes with $K{=}3000$ and $r_0{=}0.1$.

Appendix E. Extinction thresholds with a distribution of selection and dominance coefficients

Here, we explore the impact of gene flow on the mean population size in the metapopulation when we have a distribution of selection and dominance coefficients. We only consider a simple distribution, where a fraction a of loci are additive ($h{=}0.5$) and have scaled selection coefficients Ks_A and the remaining fraction, $(1{-}a)$ are nearly recessive ($h{=}0.02$), with selection coefficients Ks_R.

We see from fig. 3a - 3c that independent of the value of a, when additive alleles are nearly neutral and recessive alleles are non-neutral, increasing the selective effect of the recessive alleles, i.e., making

them more strongly deleterious has little or no effect on the critical migration threshold above which the metapopulation survives. This also holds for the case where the additive alleles are mildly deleterious and occupy a higher proportion of loci (dashed lines in fig. 3c).

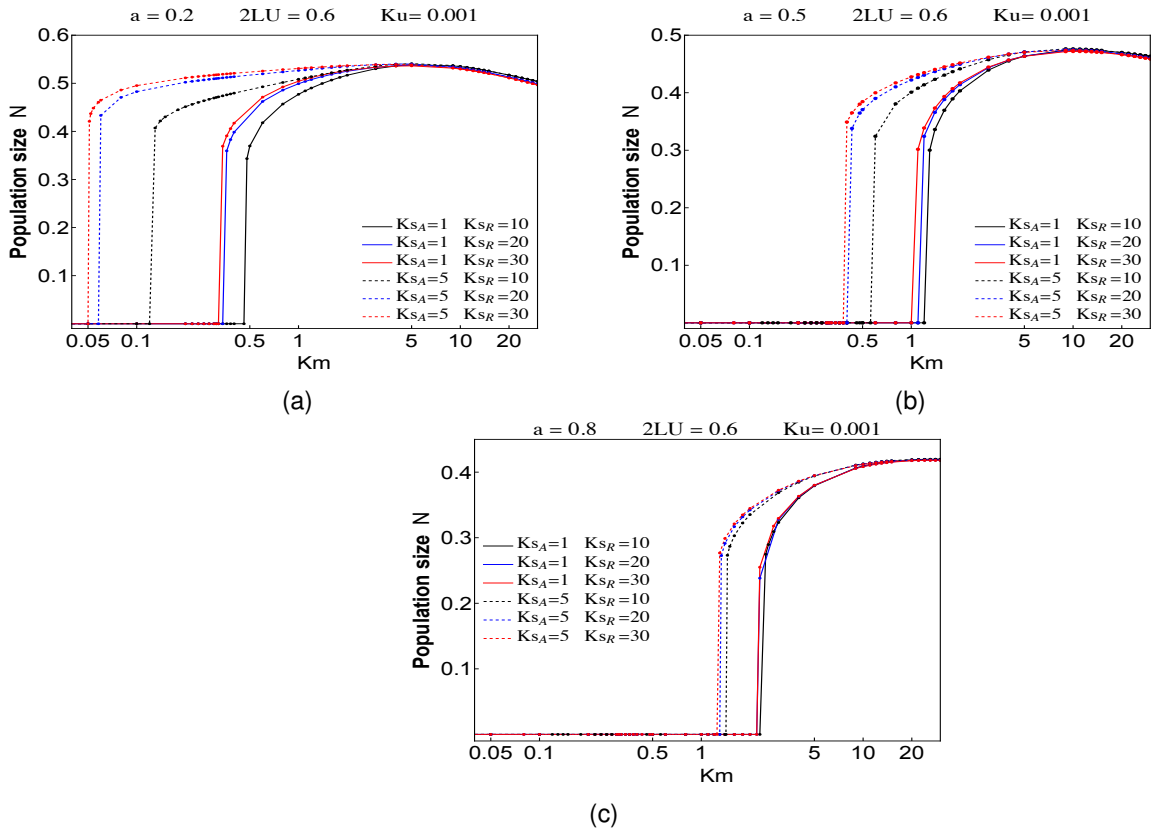

Figure 3: Average population size across the metapopulation plotted against Km with (a.) a=0.2 i.e., 20% of loci are additive and 80% are recessive (b.) a=0.5 i.e., with equal proportion of additive and recessive alleles (c.) a=0.8 i.e., with 80% of loci being additive and 20% recessive.

On the other hand, when additive alleles occupy a lower or equal proportion of loci as the recessive alleles, we see a somewhat different dynamics. We observe a (slightly) higher critical migration threshold when additive and recessive alleles are mildly deleterious (black dashed line in fig. 3a and 3b). As the recessive alleles become moderately deleterious, the threshold reduces and increasing the selective effect further makes little or no effect on the threshold migration rate (compare blue and red dashed lines in both fig. 3b and 3c). Finally, for all Ks_A, Ks_R combinations considered, we see the critical migration threshold at least doubling as a increases.

References

Glémin, S., Ronfort, J., and Bataillon, T. (2003). Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics*, 165(4):2193–2212. https://doi.org/10.1093/genetics/165.4.2193.

Ohta, T. and Kimura, M. (1971). Linkage disequilibrium between two segregating nucleotide sites under the steady flux of mutations in a finite population. *Genetics*, 68(4):571. https://doi.org/10.1093/genetics/68.4.571.

Roze, D. (2015). Effects of interference between selected loci on the mutation load, inbreeding depression, and heterosis. *Genetics*, 201(2):745–757. https://doi.org/10.1534/genetics.115.178533.

Whitlock, M. C. (2002). Selection, Load and Inbreeding Depression in a Large Metapopulation. *Genetics*, 160(3):1191–1202. https://doi.org/10.1093/genetics/160.3.1191.