

**Supplementary Figures: Eco-evolutionary dynamics and collective dispersal:  
implications for salmon metapopulation robustness**

Justin D. Yeakel<sup>1,2,\*</sup>, Jean P. Gibert<sup>1</sup>, Peter A. H. Westley<sup>3</sup>, & Jonathan W. Moore<sup>4</sup>

<sup>1</sup>School of Natural Sciences, University of California, Merced, Merced CA, USA

<sup>2</sup>The Santa Fe Institute, Santa Fe NM, USA

<sup>3</sup>College of Fisheries and Ocean Sciences, University of Alaska, Fairbanks, Fairbanks AK, USA

<sup>4</sup>Earth2Oceans Research Group, Simon Fraser University, Vancouver BC, Canada

\*To whom correspondence should be addressed: [jdyeakel@gmail.com](mailto:jdyeakel@gmail.com)

## I. Gaussian approximation for the mixed trait distribution

We assume a Gaussian distribution to approximate the more complex trait distribution that results when two populations with two distinct trait distributions are linked by gene flow and reproduction. Here and throughout, we refer to the resident population as  $i$  and the straying population as  $j$ , though both populations stray simultaneously. The Gaussian approximation is assumed to have a mean that is equal to the mean of the mixed trait distribution  $(1 - m)\mu_i + m\mu_j$  with a constant variance  $\sigma^2$ , equal to the original trait variance prior to mixing. We employ this approximation because a mixed trait distribution that remixes and recombines (through reproduction) at each time-step becomes complicated very quickly, as we shall demonstrate. *Importantly, because the evolutionary framework tracks only the mean and not the variability through time, this approximation is exact with respect to the first moment.* Accordingly, the only attribute that is approximate is the assumption of normality. Below we will show that a constant variance of  $\sigma^2$  for the mixed trait distribution is a good approximation in most relevant cases, and becomes increasingly accurate when 1) generations become less overlapping (i.e. the case of semelparity), and 2) with density-dependent straying ratios.

**The trait distribution at  $t = 0$ .** At time  $t = 0$ , it is assumed that individuals from two separate populations have not yet mixed. Thus, the mean trait values for both populations are equal to the optimal trait values that accord to their respective habitats  $\mu_i(t = 0) = \theta_i$ , and that their trait values are distributed as  $X_i = x_i \sim g(x_i, \mu_i, \sigma) = \mathcal{N}(\mu_i, \sigma)$ . Here and henceforth we use upper-case to denote random variables, and lower-case to denote specific values of random variables.

**The trait distribution at  $t = 1$ .** At time  $t = 1$ , dispersal of individuals from site  $j$  to site  $i$ , and vice versa, has begun. The proportion of individuals that disperse to and from each population is given by the stray ratio  $m$ . The calculation of the mixed distribution that results from dispersal and subsequent recruitment incorporates the following processes: 1) dispersal of strays into the local population from the remote population, and vice versa; 2) recruitment of juveniles that are produced from this mixed distribution (under the assumption of random mating); 3) mortality of individuals in the adult class. These events occur in the order in which they are listed.

The mixed trait distribution for site  $i$  immediately following dispersal is simply a mixed Gaussian with weights determined by the proportion of the population that are residents  $i$  and the proportion that are strays  $j$ , such that

$$p(x_{\text{disperse}}) = w_i g(x_i, \mu_i, \sigma) + (1 - w_i) g(x_j, \mu_j, \sigma), \quad (1)$$

where the proportion of the population  $i$  composed of *resident* individuals is

$$w_i = \frac{(1 - m)N_i}{(1 - m)N_i + mN_j}, \quad (2)$$

and the proportion of population  $i$  composed of strays is  $(1 - w_i)$ . Once individuals in both habitats mix, they mate without preference. Recruits are assumed to have trait values intermediate to their parents such that  $x_r = (x_1 + x_2)/2$  where trait values  $X_1 = x_1 \sim p(x_{\text{disperse}})$  and  $X_2 = x_2 \sim p(x_{\text{disperse}})$ . We thus assume that parental traits are drawn independently from the mixed distribution  $p(x_{\text{disperse}})$ , conditional on the requirement that  $x_2 = 2x_r - x_1$ . Incorporating this condition and that traits are independently drawn from the mix of residents and strays yields the probability density for the trait value of recruits,

$$q(x_r) = \int_{-\infty}^{\infty} 2p(x_1)p(2x_r - x_1)dx_1. \quad (3)$$

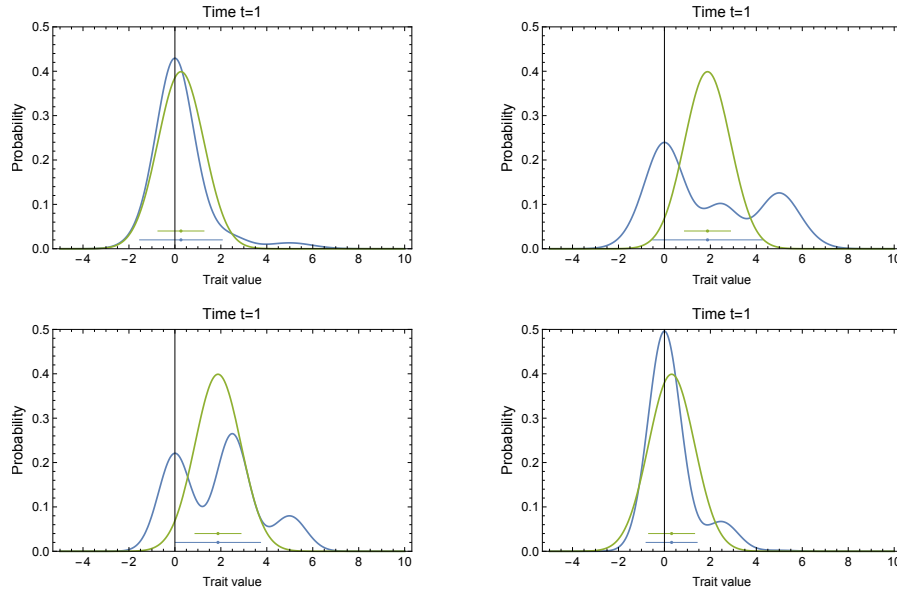


Figure S1: Clockwise from the upper-left. The calculated probability distribution function (PDF) for individuals in population  $i$  after immigration from population  $j$ , recruitment, and mortality (blue) vs. the Gaussian approximation used for the dynamical system (green), with a) low stray ratio, large straying population, overlapping generations:  $m = 0.1$ ,  $Z = 0.5$ ; b) high stray ratio, overlapping generations:  $m = 0.5$ ,  $Z = 0.5$ ; c) high stray ratio, non-overlapping generations:  $m = 0.5$ ,  $Z = 1000$ ; d) density-dependent stray ratio, non-overlapping generations:  $m = m_0(1 - N_i/(C + N_i))$ ,  $Z = 1000$ , where  $m_0 = 0.2$  and  $C = 1000$ . The mean and variability of both distributions is shown by a point (mean) and line ( $\sigma$ ) at the base of each distribution.

The mean of the recruit trait distribution is calculated to be

$$E\{q(x_r)\} = (1 - m)\mu_i + m\mu_j, \quad (4)$$

and the variance is

$$\text{Var}\{q(x_r)\} = \frac{m^2 N_j^2 (4\mu_j^2 + \sigma^2) + (m - 1)^2 N_i^2 (4\mu_i^2 + \sigma^2) - (m - 1)m N_i N_j (\mu_i^2 + 6\mu_i \mu_j + \mu_j^2 + 2\sigma^2)}{2(-m N_i + m N_j + N_i)^2}. \quad (5)$$

The relative contribution of trait values from resident parents, stray parents, and recruits depends on the numbers of each within the mixed population. The number of residents is  $(1 - m)N_i$ , and the number of strays is  $mN_j$ . However, because adult mortality occurs after reproduction, we must adjust these numbers to account for only those who survive, which is given by  $e^{-Z}$ . As  $Z \rightarrow \infty$ , the population becomes semelparous, as the adult class gives way to new recruits at each timestep. Accordingly, the number of adult residents and strays is given by  $(1 - m)N_i e^{-Z}$  and  $mN_j e^{-Z}$ , respectively. In contrast, the number of recruits is given by the combined contribution of resident and straying individuals to recruit production as  $N_r = ((1 - m)N_i R[\mu_i] + mN_j R[\mu_j])e^{(-\beta((1 - m)N_i + mN_j))}$  where  $R$  is the recruitment rate, which differs for resident vs. straying individuals (see manuscript, Eq 2).

The proportion of individuals from the resident, straying, and recruitment classes following mortality at site  $i$  are

then  $c_i$ ,  $c_j$ , and  $c_r$  respectively, where

$$\begin{aligned} c_i &= \frac{(1-m)N_i e^{-Z}}{(1-m)N_i e^{-Z} + mN_j e^{-Z} + N_r} \\ c_j &= \frac{mN_j e^{-Z}}{(1-m)N_i e^{-Z} + mN_j e^{-Z} + N_r} \\ c_r &= \frac{N_r}{(1-m)N_i e^{-Z} + mN_j e^{-Z} + N_r}. \end{aligned} \quad (6)$$

where  $c_i + c_j + c_r = 1$ . The trait value distribution for time  $t = 1$  is then calculated as a mix of trait distributions from 1) surviving residents, 2) surviving strays, and 3) their offspring, such that

$$g_{\text{next}}(x_i) = c_i g(x_i) + c_j g(x_j) + c_r q(x_r). \quad (7)$$

The expectation of this distribution is

$$\mathbb{E}\{g_{\text{next}}(x_i)\} = (1-m)\mu_i + m\mu_j, \quad (8)$$

and the variance is

$$\begin{aligned} \text{Var}\{g_{\text{next}}(x_i)\} &= \frac{(\mu_i - \mu_j)^2 (2(m-1)^2 N_i^2 c_j^2 - (m-1)N_i c_j (2(m-1)N_i + mN_j(1-4c_i)))}{2(-mN_i + mN_j + N_i)^2} \\ &\quad + \frac{mN_j(c_i - 1)(-mN_i + 2mN_j c_i + N_i) - \sigma^2(c_i + c_j + 1)(-mN_i + mN_j + N_i)^2}{2(-mN_i + mN_j + N_i)^2}. \end{aligned} \quad (9)$$

Importantly,  $\mathbb{E}\{g_{\text{next}}(x_i)\}$  is exactly equal to the trait mean for the original mix, as well as the trait mean for recruits. Thus, a Gaussian approximation distributed as  $\mathcal{N}(w_i \mu_i + (1-w_i)\mu_j, \sigma)$  will not alter the dynamics under the assumption of constant variance in Eq 4. Moreover, the mixed distribution described above becomes more complex at the next timestep, as two multi-modal trait distributions then mix and recombine through recruitment. Although the complexity of the true distribution is increased and becomes quickly intractable, one expects the actual form to eventually approximate a Gaussian based on the central limit theorem. One could make the argument that a Gaussian approximation does not represent the multi-modality in trait values inherent to a mixed distribution, and thus underestimates variance. Although the Gaussian approximation exactly captures the expectation, we next explore to what extent the approximation does and does not capture the variability expressed in  $g(x_{\text{next}})$ .

The Gaussian approximation  $\mathcal{N}(w_i \mu_i + (1-w_i)\mu_j, \sigma)$  will generally underestimate the true variance of the mix, but not by much. If the straying ratio is low, (ca. 0.0 to 0.2), the Gaussian approximation tends to be very similar to the actual distribution (Fig. S1a). As the straying ratio increases, so does the bimodality of the actual distribution, and this magnifies our underestimate of variance (Fig. S1b). The interesting dynamics observed in our model generally occur between  $m = 0.05$  and  $m = 0.2$ , where estimates of variance are similar. Decreasing the overlap between generations also leads to greater similarity in the actual distribution and the Gaussian approximation, and the potential bias introduced by bimodality will be less for cases of non-overlapping generations, or semelparity (Fig. S1c). Most importantly, density-dependent straying increases the accuracy of our Gaussian approximation by suppressing the straying ratio (Fig. S1d).

Because the Gaussian approximation is exact to first order, and generally provides a good approximation of realized variance, we suggest that the use of Lande's formulation for the change in mean trait values through time under the assumption of constant variance is appropriate (see Eq. 4 in main text). Because our estimate of variance is generally

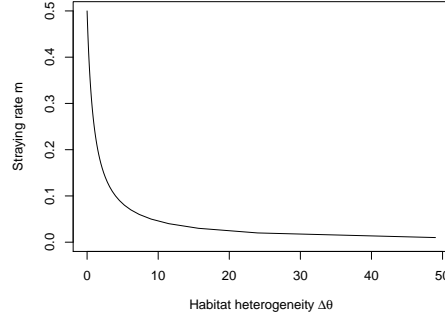


Figure S2: In some cases, habitat heterogeneity may be assumed to determine the rate of straying, if for example: 1) sites are distributed over greater spatial distances, where habitat differences are assumed to be greater between more distant sites, or 2) individuals have behaviors promoting dispersal between habitats that are more similar. To examine such cases, we use the relationship  $m = 0.5(1 + \Delta\theta)^{-1}$  where maximum straying is assumed to occur at  $m = 0.5$  (perfect mixing).

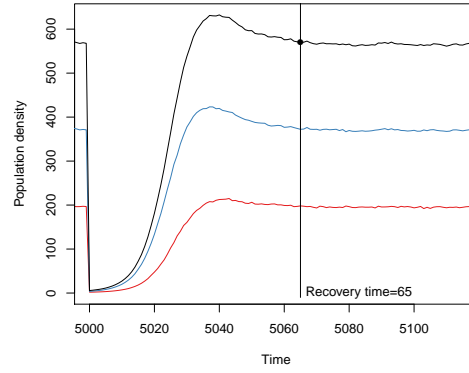


Figure S3: Example of the numerical procedure used to estimate recovery time. After a disturbance is introduced, the recovery time is calculated by measuring the point in time where  $N_T$  (in black), which is the aggregate of both populations (blue, red) settles to within one standard deviation of the new equilibrium  $N_T^*$ .

an underestimate, the rate of recruitment will be a slight overestimate (Eq. 2, main text), and the timescale of trait evolution will generally be slower (Eq. 4, main text). Changes in the rate of recruitment (Fig. S14) and heritability do not have an impact on our qualitative findings, and this suggests any bias that our approximation introduces by underestimating true variance is minimal.

## II. Additional Figures

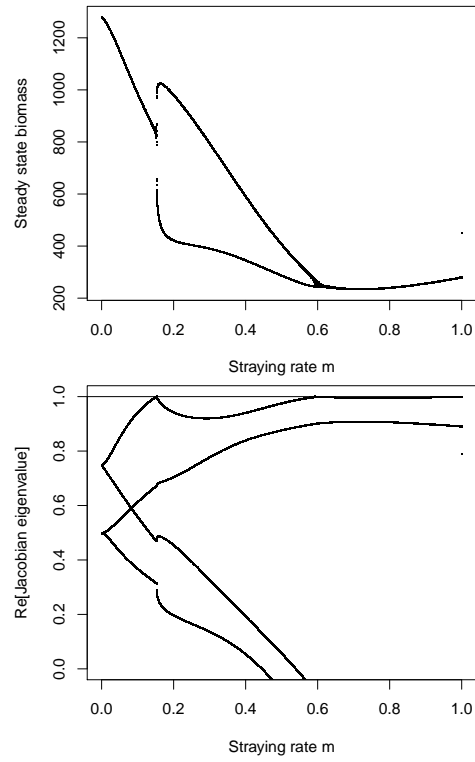


Figure S4: The real parts of the four eigenvalues for the Jacobian matrix of the 4-dimensional system. The cusp bifurcation occurs when the dominant eigenvalue crosses the unit circle at  $+1$ .

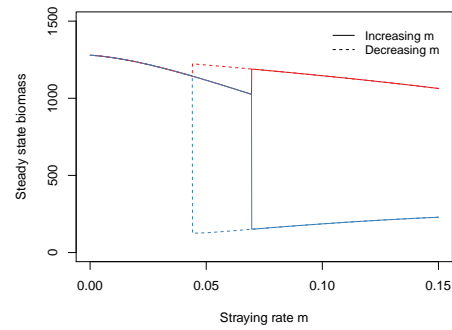


Figure S5: Increasing the straying rate results in the transition from a single steady-state for both populations to a dominant and subordinate states. If the straying rate is subsequently lowered, the single steady-state is not easily obtained, which is the hallmark of hysteresis.

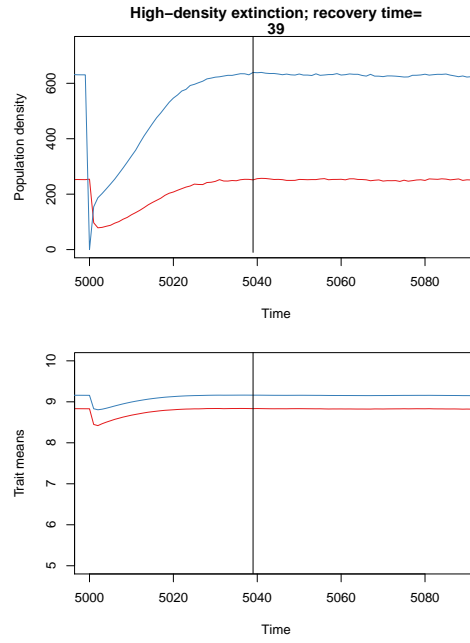


Figure S6: Extinction of high-density population with a high straying rate  $m = 0.4$  and low trait heritability  $h^2 = 0.2$  (see figure 3a). Black line marks the calculated point of recovery post-perturbation. Trait optima are  $\theta_1 = 10$  (blue population trajectory) and  $\theta_2 = 5$  (red population).

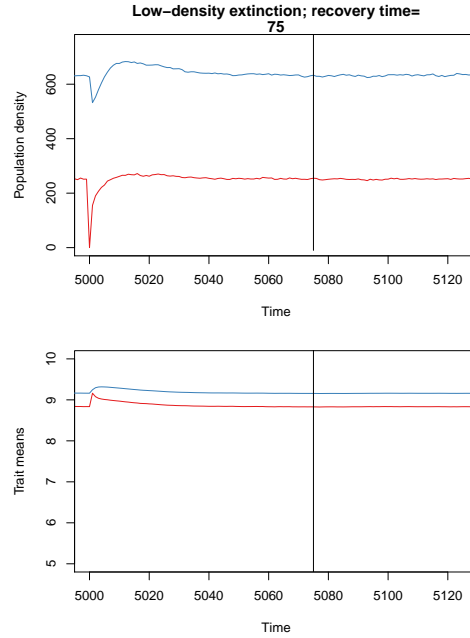


Figure S7: Extinction of low-density population with a high constant straying rate  $m = 0.4$  and low trait heritability  $h^2 = 0.2$  (see figure 3a). Black line marks the calculated point of recovery post-perturbation. Trait optima are  $\theta_1 = 10$  (blue population trajectory) and  $\theta_2 = 5$  (red population).

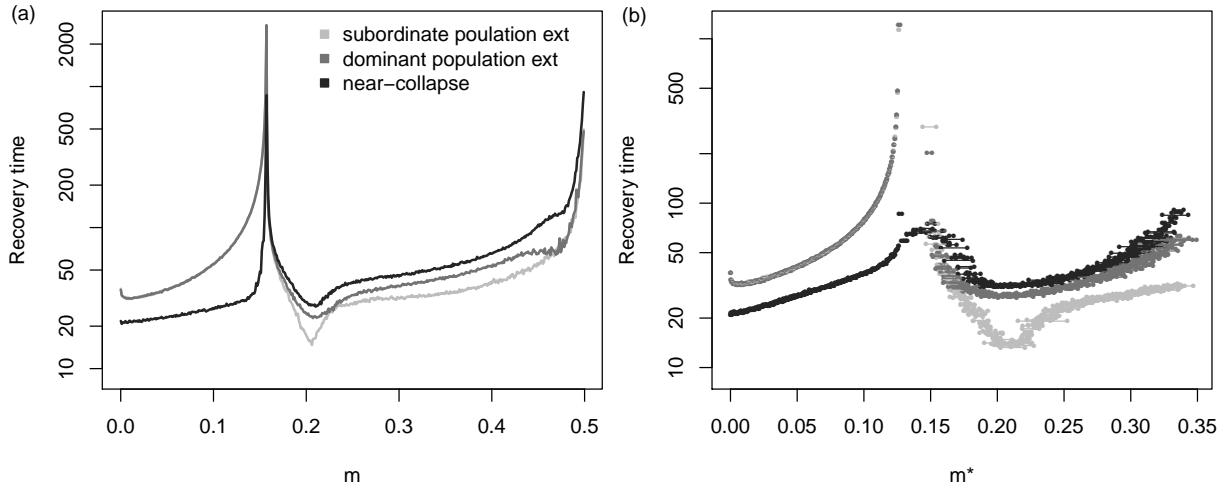


Figure S8: Recovery time of  $N_T$  following the extinction of either the low-density (light gray) or high-density (gray) population, or the near-collapse of both (dark gray) assuming (a) constant straying rates  $m$  and (b) density-dependent straying rates (evaluated at the steady-state  $m^*$ ) with trait heritability  $h^2 = 0.8$ . If  $m$  is density-dependent, in the alternative stable state regime there are two straying rates observed: one each for the low- and high-density populations, respectively, which are linked by a horizontal line.

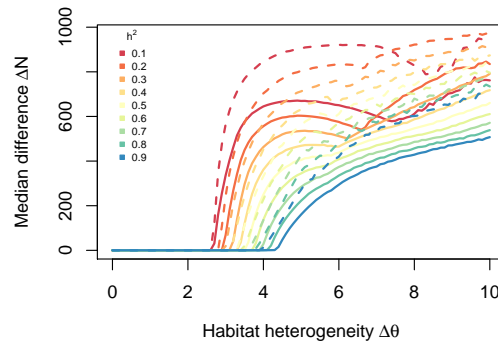


Figure S9: Median difference in population densities taken over the straying rate as a function of habitat heterogeneity  $\Delta\theta$ . Solid lines are for constant  $m$ ; dashed lines are for density-dependent  $m$ .



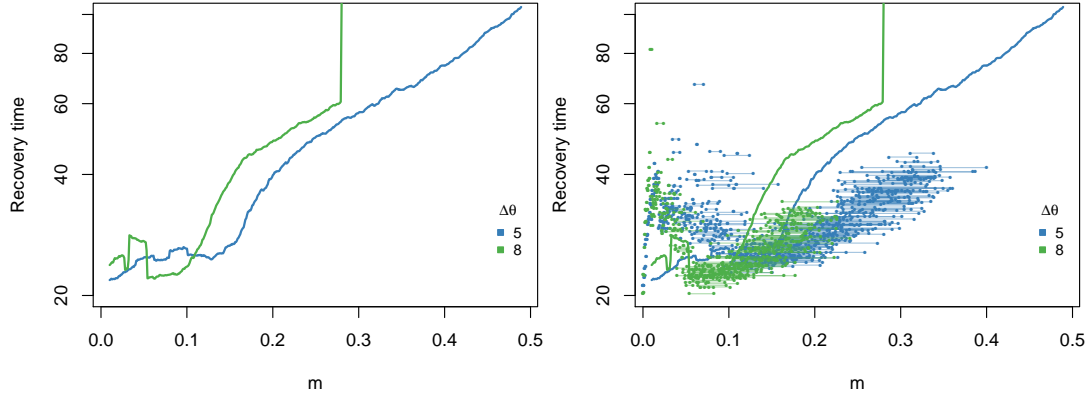


Figure S10: (a) Recovery time after near collapse of both populations as a function of straying rate  $m$  and habitat heterogeneity  $\Delta\theta$ . (b) The same as (a) but including recovery times when straying is density-dependent evaluated at  $m^*$ , shown by linked point pairs. Recovery times for systems with density-dependent straying are longer when straying is low and shorter when straying is high, mirroring the change in portfolio effects with respect to density-dependent straying shown in figure 5.

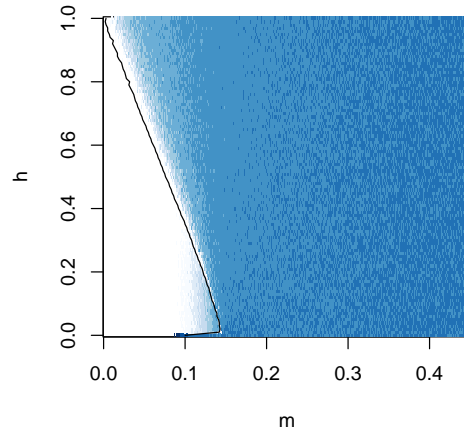


Figure S11: Portfolio effects as a function of straying rate  $m$  and trait heritability  $h^2$  when the rate of straying is  $m = 0.5(1 + \Delta\theta)^{-1}$ . Alternative steady-states emerge for low values of  $m$  (left of the cusp bifurcation, denoted by the black line), whereas a single steady-state exists for high  $m$ .

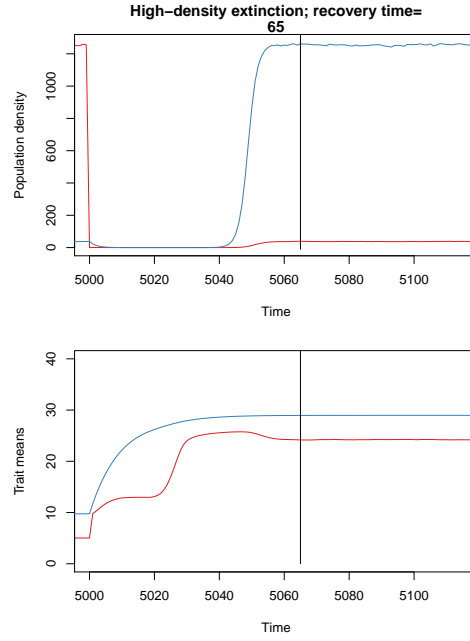


Figure S12: Population inversion where increased differences in trait optima between sites  $\Delta\theta$  corresponds to lower rates of straying  $m$ . At low rates of straying  $m = 0.02$  ( $\Delta\theta = 24$ ), extinction of the dominant population leads to slower-than-expected recovery times because the subordinate population is isolated enough to evolve towards its own trait optimum. In this case,  $m$  is less than  $m = 0.034$  (denoted by the asterisk in figure 6), such that isolation allows the subordinate population to ‘run away’ from the influence of the dominant population. This leads to a switch in subordinate/dominant states for the two populations. If  $m$  is low but greater than 0.034, isolation permits the subordinate population to ‘run away’ from the influence of the dominant population, until it is overwhelmed by the recovering dominant population, and reverts back to its previous trait mean prior to disturbance.

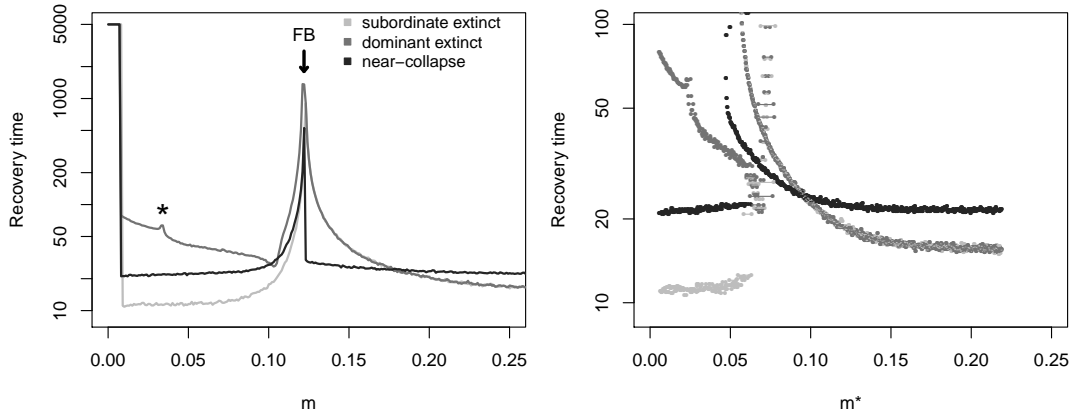


Figure S13: Recovery times for three disturbance types when the straying rate covaries with habitat heterogeneity as  $m = 0.5(1 + \Delta\theta)^{-1}$  for constant (a) and density-dependent (b) straying rates. The cusp bifurcation is not as clear in (b) because  $\Delta\theta$  is a function of the individual straying rate  $m_0$ , whereas the x-axis in (b) is the straying rate at the steady-state  $m^*$ . Despite this difference, the general behavior shown in (a) are the same in (b).

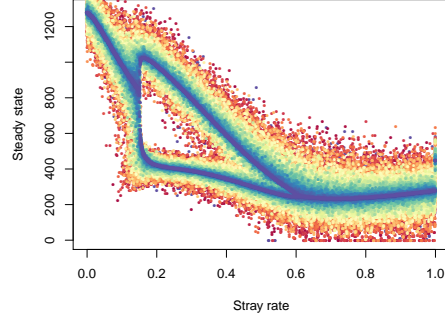


Figure S14: Steady-state densities of both populations as a function of  $m$ , where a cusp bifurcation indicates the emergence of alternative steady-states: one in a dominant state and one in a subordinate state. Steady-states for populations with symmetrical values ( $\alpha = 0$ ) in the vital rates  $r_{\max}$  and  $\beta$  are shown with cool tones. As the asymmetry among populations between sites increases ( $\alpha > 0$ ), their vital rates diverge, such that the maximal growth at sites 1 and 2 is now  $r_{\max}(1) = r_{\max}(1 + \tilde{r}v_1)$  and  $r_{\max}(2) = r_{\max}(1 + \tilde{r}v_2)$  where  $\tilde{r}v_{1,2}$  are independently drawn from  $\text{Normal}(0, \alpha)$  and  $r_{\max} = 2$ . Similarly the strength of density dependence is calculated at sites 1 and 2 as  $\beta(1) = \beta(1 + \tilde{r}v_1)$  and  $\beta(2) = \beta(1 + \tilde{r}v_2)$  where  $\tilde{r}v_{1,2}$  are independently drawn from  $\text{Normal}(0, \alpha)$  and  $\beta = 0.001$ . Steady-states for populations with increasingly asymmetric values ( $\alpha \rightarrow 0.1$ ) for vital rates  $r_{\max}$  and  $\beta$  are shown in warmer tones.

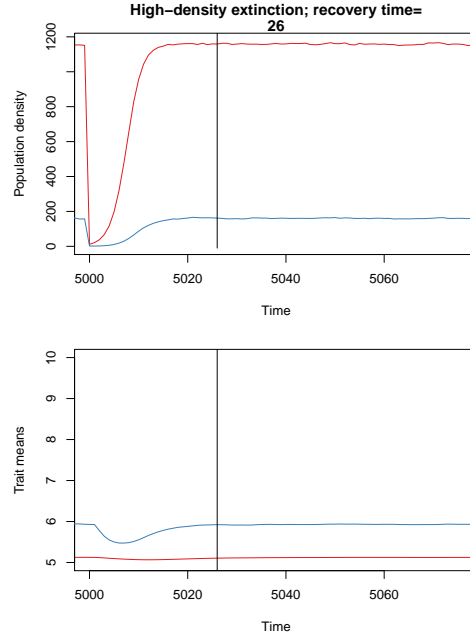


Figure S15: Near collapse of both populations with a low straying rate  $m = 0.1$  and low trait heritability  $h^2 = 0.2$  (see figure 3a). Black line marks the calculated point of recovery post-perturbation. Trait optima are  $\theta_1 = 10$  (blue population trajectory) and  $\theta_2 = 5$  (red population).