Cooperative interactions among females and even more extraordinary sex ratios

Ryosuke Iritani[1], Stuart A West[2], and Jun Abe[3]

[1]RIKEN Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS), Wako, Saitama 351-0198, Japan.
[3]Faculty of Liberal Arts, Meijigakuin University, 1518 Kamikurata-cho, Totsuka, Yokohama City, Kanagawa, 244-8539, Japan
[2] Department of Zoology, University of Oxford, Oxford, UK

September 25, 2020

Abstract

Hamilton's local mate competition theory provided an explanation for extraordinary female biased sex ratios in a range of organisms. When mating takes place locally, in structured populations, a female biased sex ratio is favoured to reduce competition between related males, and to provide more mates for males. However, there are a number of wasp species where the sex ratios appear to more female biased than predicted by Hamilton's theory. We investigated theoretically the extent to which cooperative interactions between related females can interact with local mate competition to favour even more female biased sex ratios. We found that: (i) cooperative interactions between females can lead to sex ratios that are more female biased than predicted by local competition theory alone; (ii) sex ratios can be more female biased when the cooperative interactions are offspring helping parents before dispersal, rather than cooperation between siblings after dispersal. Our results can be applied to a range of organisms, and provide an explanation for the extreme sex ratio biases that have been observed in *Sclerodermus* and *Melittobia* wasps.

Key words:

- Kin selection;
- Local resource competition;
- Local mate competition;
- Local resource enhancement;
- Population structure;

Introduction

Sex ratio theory has provided one of the most productive and successful areas of evolutionary biology (Charnov 1982; Hardy 2002; West 2009). Theory predicts a number of situations in which individuals are expected to adjust the sex of their offspring in response to local conditions (Charnov 1982; Frank 1998). This theory has been applied to explain variation in the offspring sex ratio (proportion males) across a range of taxa, from malaria parasites to ants to birds (Bourke & Franks 1995; Hardy 2002; West 2009).

One of the major challenges is to explain when sex ratios are biased away equal investment in the sexes. Hamilton's (1967) local mate competition (LMC) theory provides an explanation for female biased sex ratios observed in parasitic wasps (e.g., *Scelionidae*, Waage 1982; *Alfonsiella*, Greeff 2002; *Apanteles*, Tagawa 2000; Gu & Dorn 2003; and *Nasonia*, Werren 1983; Shuker *et al.* 2006; Burton-Chellew *et al.* 2008), aphids (e.g., *Prociphilus oriens*; Yamaguchi 1985), and a number of fig wasps (Herre 1985). Specifically, Hamilton showed that if n haplodiploid females lay eggs in a patch, and that mating occurs before females disperse, then the evolutionarily stable strategy (ESS; Maynard Smith & Price 1973) is to produce an offspring sex ratio of $(n-1)/(2n)$ (Fig 1), which predicts female biased offspring sex ratios (smaller than 1/2) and becomes less biased as more females lay eggs in a patch (i.e., as n larger). Succeeding work (Taylor 1981; Bulmer 1986; Frank 1986; Taylor 1988, 1992; Frank 1998) made it clear that in Hamilton's LMC theory, selection on the female bias is mediated by the balance among three factors: (i) a benefit for reduced competition between sons, (ii) a benefit for production of more mates (daughters) for those sons ("mating bonus"; Frank 1998), and (iii) a cost for stronger local resource competition among females (LRC; Clark 1978). Hamilton's LMC theory has been extremely successful in explaining variation in the offspring sex ratio, both across species and between individuals (Taylor 1981, 1994; Avilés 1993; Gardner & West 2006; Shuker *et al.* 2004, 2006; Gardner *et al.* 2009; West 2009; Rodrigues & Gardner 2015).

However, there are a number of cases where females produce extremely female biased offspring sex ratios, which do not appear to be completely explained by LMC theory. One example is provided by *Melittobia* wasps, where females of several species produce approximately 2 % male offspring when ovipositing alone ($n = 1$), and hardly change their offspring sex ratio when more females lay eggs on a patch (larger; Fig 1A; Abe *et al.* 2003, 2014). Another example is provided by *Sclerodermus* wasps, in which multiple females can lay eggs on a host but the females still only produce 7% males (Fig 1B;

Tang *et al.* 2014; Lupi *et al.* 2017; Abdi *et al.* 2020a,b,c). These cases therefore suggest that identifying additional factors that favour female-biased sex ratio is required.

A possible explanation for the observed female-biases is that there is the potential for mutually beneficial cooperative interactions between females (Schwarz 1988; Stark 1992; Komdeur *et al.* 1997; Cronin & Schwarz 1997; Schwarz *et al.* 1998; Martins *et al.* 1999; Clutton-Brock 2002; Tang *et al.* 2014; Kapranas *et al.* 2016).For example, in presocial, allodapine bee *Exoneura bicolor*, cooperative nesting occurs among related females, which result in higher per capita reproductive outputs (Schwarz 1988; Cronin & Schwarz 1997). In this case, a more female biased offspring sex ratio can be favoured, to increase these beneficial interactions between related females, as a form of local resource enhancement (LRE; in this literature we focus on LRE provided by females). Cooperative interactions beetween females have been suggested to be important in both *Melittobia* and *Sclerodermus* wasps either (Abe *et al.* 2003, 2014; Tang *et al.* 2014; Lupi *et al.* 2017), but the extremely female-biased sex ratios under LRE in these species remain to be formally explained.

We expand existing theory to examine whether LRE can explain the extremely female biased sex ratios that have been observed in *Melittobia* and *Sclerodermus* wasps. We examine three factors that may be especially relevant to the biology of these species: (1) competitions between sons and between daughters; (2) the cooperative interactions that can occur at different times, either when laying eggs (intra-generational LRE) as in Pen & Weissing (2000) and Wild (2006), or when colonizing females help each other before competition (trans-generational LRE); (3) both females and males may disperse to different extents, hence varying the degree to which these competitive and cooperative interactions occur locally We specifically look at how sex-specific dispersal rates, the number of foundresses, and fecundity effect of LRE influence the evolution of sex ratios.

Models and Analyses

Life cycle

We assume Wright's (1931) island model of dispersal, in which the metapopulation is subdivided into an infinite number of patches each fostering n mated females. We focus on a particular female, and we denote her proportional investment of reproductive resource into sons ('sex ratio') by $x_\bullet$, the average sex ratio of the adult females in her patch in the same generation by x_0, and the average

sex ratio of adult females in the metapopulation by x. Immediately upon birth, juvenile males may disperse to an alternative patch at a rate d_{m} each, or else stay in the natal patch $(1 - d_{\mathrm{m}})$, followed by random mating on the patch, with each female mating only once but each male potentially mating many times. Males die after mating and females disperse with a probability of d_{f} each. After dispersal, breeding females compete for the limited number of breeding sites on the patch (n), after which the metapopulation is returned back to its original size and a new cycle starts (Fig 2A).

We consider two types of LRE. In the first, LRE occurs due to helping behaviors among juvenile females (before dispersal) that promotes the survival rate of all juveniles born in the same patch. We refer to this as 'intra-generational LRE' ("intra-" indicates that survival rate depends on the sex ratio, x_0, of the focal generation, $\tau = 0$, where $\tau \geq 0$ is a generic symbol to count the generations backwards ordinally in time: $\tau = 0$ for the present, $\tau = 1$ for the parental generation, and generally τ for τ-generations prior, and we refer to 'τ-th generation' henceforth). Intra-generational LRE may be relevant in species where juvenile females engage in helping behaviors before dispersal (at the latest), or equivalently, where juvenile females assist reproduction of the reproductive adults (survivorship and fecundity interpretations are mathematically equivalent). In the second model, we posit that LRE occurs due to mutual helping at the colonization stage (before competition for breeding spots), namely 'trans-generational LRE' ("trans-" indicates that the fecundity depends on the sex ratio, x_1, of the previous generation, $\tau = 1$). This applies to species where females communally colonize common patches, as in *Sclerodermus*.

LRE by natal juvenile females: intra-generational LRE

We start with our analyses for intra-generational LRE, where offspring can help members of their parents. We assume that, for a patch with the average sex ratio X, the per capita fecundity (which is the number of offspring born times their survival rate) is given by $\beta(X)$. Turning our attention to the focal patch with its inhabitants' average sex ratio x_0 (including her own sex ratio), the fecundity of the females in the patch is given by $\beta_0 := \beta(x_0)$, and the average fecundity (the total number of offspring produced per capita times their survival rate; we use ':=' to define a quantity henceforth) in the metapopulation assuming that the mutants are vanishingly rare is given by $\beta^\circ := \beta(x)$. Using a parameter α (with $0 \leq \alpha \leq 1$) which tunes the strength of LRE on fecundity (β_0), we formulate β by:

$$\beta(x_0) = K + \alpha(1 - x_0)\beta(x_0) = \frac{K}{1 - \alpha(1 - x_0)} \tag{1}$$

(see Appendix A for derivation), where K is a baseline of per capita fecundity in the absence of LRE. β_0 is larger when neighboring individuals produce more females (x_0 lower). The fecundity $\beta(x_0)$ decreases from $K/(1-\alpha)$ to K as x_0 varies from 0 to 1, and grows from K to K/x_0 as α varies from 0 to 1 (for $x_0 > 0$ fixed).

LRE by colonizing females after dispersal: trans-generational LRE

We now turn our attention to trans-generational LRE, where juvenile females of the same generation can cooperate after dispersal for communal colonization. We use the same symbol (β_0) to designate the fecundity of individuals in the focal patch, to keep the consistency with the intra-generational LRE model. We write x_τ for the average sex ratio of adult females in the focal patch in the τ-th generation, and β_τ for their fecundity (Fig 2B). We recursively define β_τ by:

$$\begin{aligned}\beta_\tau &= B(\beta_{\tau+1}, x_{\tau+1}) \\ &= K + \alpha\Big(\big(1-d_{\mathrm{f}}\big)(1-x_{\tau+1})\beta_{\tau+1} + d_{\mathrm{f}}(1-x)\overline{\beta}\Big),\end{aligned} \tag{2}$$

where α (with $0 \le \alpha \le 1$) measures the strength of LRE as before, $\big(1-d_{\mathrm{f}}\big)(1-x_{\tau+1})\beta_{\tau+1} + d_{\mathrm{f}}(1-x)\overline{\beta}$ is proportional to the density of females after female dispersal (before competition), and $\overline{\beta}$ is the metapopulation-wide average of β to be determined by assuming that it has reached a stable equilibrium value for a phenotypically monomorphic population with x. The equilibrium value for $\beta = \overline{\beta} = \beta^\circ$ is given as the solution to $\beta^\circ = B(\beta^\circ, x)$; that is:

$$\beta^\circ = \beta(x) = \frac{K}{1-\alpha(1-x)} \tag{3}$$

(see Eqn (1)), which is always locally stable for given x (i.e., β_τ converges to β° given x is fixed). Specifically, the fecundity of the focal female in the present generation $\tau = 0$ is given by:

$$\begin{aligned}\beta_0 &= B(\beta_1, x_1) \\ &= B\big(B(\beta_2, x_2), x_1\big) \\ &= B\Big(B\big(B(\beta_3, x_3), x_2\big), x_1\Big) \\ &= \dots\end{aligned} \tag{4}$$

(see Appendix A for more details), which implies that to determine (the effect of selection on) β_0, we need to consider an expected sequence of retrospective sex ratios in the focal patch, $(x_1, x_2, x_3, \ldots)$, in addition to the focal's and neighbors' sex ratios in the present generation, $(x_\bullet, x_0)$ (Lehmann 2007, 2008). LRE supplied by founding females hence generates the trans-generational kin selection effects in viscous populations, by which the impacts of biased sex ratios in the patch descend down to the reproductive success of individuals (including the focal's offspring) living in future generations, which thus in turn induces selection on the sex ratios.

Invasion fitness and the selection gradient

We can write the invasion fitness of the focal female through daughters and sons (respectively) as:

$$W_\bullet^{\mathbf{f}} := W^{\mathbf{f}}(x_\bullet, x_0, \beta_0) := \frac{(1-d_{\mathrm{f}})(1-x_\bullet)\beta_0}{(1-d_{\mathrm{f}})(1-x_0)\beta_0 + d_{\mathrm{f}}(1-x)\beta^\circ} + \frac{d_{\mathrm{f}}(1-x_\bullet)\beta_0}{(1-x)\beta^\circ}, \tag{5}$$

$$W_\bullet^{\mathbf{m}} := W^{\mathbf{m}}\left(x_\bullet, x_0, W_0^{\mathbf{f}}, \beta_0\right) := \frac{(1-d_{\mathrm{m}})x_\bullet\beta_0}{(1-d_{\mathrm{m}})x_0\beta_0 + d_{\mathrm{m}}x\beta^\circ} \underbrace{W^{\mathbf{f}}(x_0, x_0, \beta_0)}_{=:W_0^{\mathbf{f}}} + \frac{d_{\mathrm{m}}x_\bullet\beta_0}{x\beta^\circ} \tag{6}$$

(see Lehmann 2007; Gardner *et al.* 2009; see Appendix B1–3 for derivation), where the invasion subcomponent for sons (Eqn (6)) is envisioned as a function of the focal adult female's sex ratio $x_\bullet$, patch-average sex ratio x_0, the survival rate of a random female as a mate for local males $W_0^{\mathbf{f}}$ (local mating bonus; see Frank 1998, pp.199), and the average fecundity of the focal adult female β_0. Note that β_0 depends on model assumptions (intra- and trans-generational LRE).

We perform the neighbor-modulated fitness approach to kin selection methodology (Taylor & Frank 1996; Frank 1998; Rousset & Billiard 2000; Rousset 2004; Taylor *et al.* 2007), particularly for sex-structured populations (Taylor 1990; Taylor *et al.* 2007; Gardner *et al.* 2009) with trans-generational effects of kin selection (Lehmann 2007, 2008). We take a random juvenile female and male in the present generation each as a recipient, and adult females breeding in the τ-th generation (with $\tau = 0, 1, 2\ldots$, including the focal juveniles' mother) each as an actor.

To ease biological interpretation, we here posit that $\sigma_{\mathrm{RC}} := (1-d_{\mathrm{f}})^2$ tunes the intensity of LRC, which is mathematically the probability that the focal adult female's daughters compete for resources with a juvenile female born in the same patch (eqns (7) and A20 in Wild & Taylor 2004). Similarly, the intensity of LMC is proportional to $\sigma_{\mathrm{MC}} := (1-d_{\mathrm{m}})^2$, which is the probability that the focal adult female's sons compete for mates with a juvenile male born in the same patch. Increasing σ_{RC} (or σ_{MC})

favours less (or more) female-biased sex ratios (respectively). Also, the effect of extra daughters born locally on males' reproductive success is given by $\sigma_{\mathrm{MB}} := (1-d_{\mathrm{m}})(1-\sigma_{\mathrm{RC}})$ (i.e., local mating bonus; see Frank 1998, pp.199), which reads as the probability that males mate locally $(1-d_{\mathrm{m}})$ times the probability that the mated female(s) do not encounter local resource competition with a juvenile female born in the same patch $1-(1-d_{\mathrm{f}})^2 = 1-\sigma_{\mathrm{RC}}$.

Using these $\sigma_{\mathrm{RC}}, \sigma_{\mathrm{MC}}$ and σ_{MB}, the condition for which a slightly larger sex ratio (i.e., producing more sons than does the metapopulation average) is favoured by selection is captured by Hamilton's rule:

$$\begin{aligned} c_{\mathrm{f}} &\left(\underbrace{\frac{-1}{1-x} R_{\bullet}^{\mathbf{f}}}_{\text{daughters}} + \underbrace{\frac{\sigma_{\mathrm{RC}}}{1-x} R_{0}^{\mathbf{f}}}_{\text{LRC effect}} + \underbrace{\sum_{\tau=0}^{+\infty} \frac{1-\sigma_{\mathrm{RC}}}{\beta^{\circ}} \bullet \frac{\partial \beta_0}{\partial x_\tau} R_{\tau}^{\mathbf{f}}}_{\text{LRE on: females \& LRC}} \right) \\ &+ c_{\mathrm{m}} \left(\underbrace{\frac{1}{x} R_{\bullet}^{\mathbf{m}}}_{\text{sons}} + \underbrace{\frac{-\sigma_{\mathrm{MC}}}{x} R_{0}^{\mathbf{m}}}_{\text{LMC}} + \underbrace{\frac{-\sigma_{\mathrm{MB}}}{1-x} R_{0}^{\mathbf{m}}}_{\text{Mating bonus (MB)}} + \underbrace{\sum_{\tau=0}^{+\infty} \frac{1-\sigma_{\mathrm{MC}}}{\beta^{\circ}} \bullet \frac{\partial \beta_0}{\partial x_\tau} R_{\tau}^{\mathbf{m}}}_{\text{LRE on: males \& LMC}} + \underbrace{\sum_{\tau=0}^{+\infty} \frac{\sigma_{\mathrm{MB}}}{\beta^{\circ}} \bullet \frac{\partial \beta_0}{\partial x_\tau} R_{\tau}^{\mathbf{m}}}_{\text{LRE on MB}} \right) > 0 \end{aligned} \tag{7}$$

(see Appendix B4–7 for derivation; Lehmann 2007, 2008), where each derivative is evaluated at phenotypic neutrality ($x_\bullet = x_0 = x_1 = \cdots = x$). In Eqn (7), $R_{\bullet}^{\mathbf{f}}$ (or $R_{\bullet}^{\mathbf{m}}$) represents the regression coefficient of relatedness ('relatedness' hereafter; Michod & Hamilton 1980; Grafen 1985) for a juvenile female (or juvenile male) from the perspective of their mother each in the present generation (subscript $\tau = 0$); $R_0^{\mathbf{f}}$ (or $R_0^{\mathbf{m}}$) represents the relatedness for a random juvenile female (or juvenile male) from the perspective of a random adult female in the same patch each in the present generation; and $R_\tau^{\mathbf{f}}$ (or $R_\tau^{\mathbf{m}}$) represents the relatedness for a random juvenile female (or juvenile male) in the present from the perspective of an adult female in the τ-th generation (Taylor 1988; Bulmer 1994); c_{f} (or c_{m}) represents the class reproductive value of female (or male; Taylor 1990; Caswell 2001).

In the absence of LRE ($\alpha = 0$, thus $\sum$ terms vanishing), investing maternal reproductive resources into sons has five consequences: the decrease in daughters' success, decrease in LRC, increase in sons' success, increase in LMC, and decrease in MB, as in the previous studies (Taylor 1981). The summation ($\sum$) terms each capture the sum of LRE effects each supplied by the individuals having colonized the focal patch at time epochs $\tau = 0$ (for the intra-generational LRE), and $\tau = 1, 2, \ldots$ (for the trans-generational LRE), on the focal female's fitness (Fig 2B); this inclusive fitness effect occurs by which (i) LRE increases the number of sons and thus LMC, (ii) LRE increases the number of daughters

and LRC, and (iii) LRE increases MB for sons.

Nullifying and solving Eqn (7) for x yields a candidate ESS of sex ratio (cESS henceforth; Christiansen 1991; Takada & Kigami 1991), which we generically designate with a hat ($\widehat{x}$).

Results

No LRE

We first assess the case for $\alpha = 0$ (no LRE). By nullifying Eqn (7) for $x = \widehat{x_\emptyset}$ with $\alpha = 0$ gives:

$$\widehat{x_\emptyset} = \frac{c_{\mathrm{m}}\left(R_{\bullet}^{\mathbf{m}} - \sigma_{\mathrm{MC}} R_0^{\mathbf{m}}\right)}{c_{\mathrm{m}}\left(R_{\bullet}^{\mathbf{m}} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}}) R_0^{\mathbf{m}}\right) + c_{\mathrm{f}}\left(R_{\bullet}^{\mathbf{f}} - \sigma_{\mathrm{RC}} R_0^{\mathbf{f}}\right)} \tag{8}$$

in agreement with (Wild & Taylor 2004). This expression clarifies how the local interactions affect the adaptive value of sons and daughters, where remember that $\sigma_{\mathrm{RC}} = \left(1 - d_{\mathrm{f}}\right)^2$, $\sigma_{\mathrm{MC}} = \left(1 - d_{\mathrm{m}}\right)^2$, and $\sigma_{\mathrm{MB}} = \left(1 - d_{\mathrm{m}}\right)(1 - \sigma_{\mathrm{RC}})$. Note that for males, the total scale of competition, which includes the effect of LRC among the males' mates (i.e., females that thus received males' gametes), reads $\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}}$, which is negative when $\left(1 - d_{\mathrm{f}}\right)^2 < d_{\mathrm{m}} < 1$ (null for either $d_{\mathrm{m}} = 1$ or $d_{\mathrm{m}} = \left(1 - d_{\mathrm{f}}\right)^2$; otherwise positive). Eqn (8) is a general expression of cESS under LRC and LMC (but without LRE) when male dispersal precedes mating and subsequent female dispersal ('DMD model' in Wild & Taylor 2004). Substituting equilibrium values of the relatedness shows that $\widehat{x_\emptyset}$ exhibits overall a female- or male-bias when d_{m} is small or large (respectively) and it approaches 1/2 (Fisherian sex ratio) as n increases (Fig 3A).

As in the classical LMC theory, inserting $d_{\mathrm{m}} = 0$ (no male dispersal as in *Melittobia*; Matthews *et al.* 2009) yields $\sigma_{\mathrm{MC}} = 1 (> \sigma_{\mathrm{RC}})$ and $\sigma_{\mathrm{MB}} = 1 - \sigma_{\mathrm{RC}}$, meaning that $\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}}$ equals σ_{RC} on the denominator of Eqn (8); this subsequently supplies:

$$\widehat{x_\emptyset}\Big|_{d_{\mathrm{m}}=0} = \frac{c_{\mathrm{m}}\left(R_{\bullet}^{\mathbf{m}} - R_0^{\mathbf{m}}\right)}{c_{\mathrm{m}}\left(R_{\bullet}^{\mathbf{m}} - \sigma_{\mathrm{RC}} R_0^{\mathbf{m}}\right) + c_{\mathrm{f}}\left(R_{\bullet}^{\mathbf{f}} - \sigma_{\mathrm{RC}} R_0^{\mathbf{f}}\right)} \tag{9}$$

(Appendix B; eqn 3 in Gardner *et al.* 2009). Particularly for haploids and diploids, we get $x_\emptyset^{\mathrm{LMC}} = (n-1)/(2n)$ regardless of the female dispersal rate (Hamilton 1967; Bulmer 1986; Frank 1986; Taylor 1988; Bulmer 1994; Frank 1998; Gardner *et al.* 2009). This dispersal-invariance is due partly to the concomitant effects of producing more daughters on weaker LMC but stronger LRC with these effects

exactly canceling one another out (Taylor's (1992) cancelling principle; Wilson *et al.* 1992; Taylor 1992). For haplodiploids, Eqn (8) does depend on dispersal rate but in a negligibly minor manner (almost-invariance in female dispersal; Fig 3B, left panel).

In the case for $d_m > 0$, male biased sex ratios may occur when d_f is small, in contrast to the almost-invariance result for $d_m = 0$ (Fig 3B). Increasing $d_m > 0$ (say 0.5) leads to strongly male-biased sex ratios yet with a possibility of female-bias when d_f is relatively large (Fig 3B). When $d_m = 1$ ($\sigma_{MC} = \sigma_{MB} = 0$), only is male-bias the candidate evolutionary outcome (Fig 3B; see also SI figure in Appendix C). Overall, we find that sex ratio tends to bias towards the more dispersing sex (Bulmer & Taylor 1980; Taylor 1994; Wild & Taylor 2004).

Effects of LRE: general patterns

We now consider the consequences of LRE ($\alpha > 0$). We found three general patterns. First, both types of LRE drive the evolution of more female-biased sex ratio and less male-biased sex ratios (Fig 4A; see also SI Fig1). Second, the effect of intra-generational LRE is stronger than is trans-generational LRE (Fig 4B); more precisely, the effect of intra-generational LRE is independent of sex dispersal propensities of both sexes, while that of trans-generational LRE decreases with the dispersal rates of both sexes. Finally, $d_m = 0$ (no male dispersal) as in classic LMC theory predicts that this special case leads to 'almost-invariance results', in which sex ratio is insensitive to female dispersal rate (Hamilton 1967; Bulmer 1986; Frank 1986; Taylor 1988; Bulmer 1994; Frank 1998; Gardner *et al.* 2009).

Intra-generational LRE

We deal with general values of dispersal rates (d_f and d_m, each ranging between 0 and 1), but will make an exception for $d_m = 0$ (no male-dispersal), because the results for $d_m = 0$ are qualitatively different from the results for general values $0 < d_m \leq 1$. The other advantage of presenting the specific result for $d_m = 0$ is that this assumption gives a simple formula, comparable with the previous theoretical work (Hamilton 1967; Bulmer 1986; Frank 1986; Taylor 1988; Bulmer 1994; Frank 1998; Gardner *et al.* 2009), and applies to many species like *Melittobia* (Matthews *et al.* 2009). As such, we present the results for $d_m = 0$ and $0 < d_m \leq 1$ separately; note that $d_m = d_f = 0$ means that patches are completely isolated from each other and entails stochastic analyses (Sigmund 2010), and so we omit this possibility.

We find that Hamilton's rule (Eqn (7)), which assesses the direction of selection, is equal to:

$$\widehat{x_\emptyset} - x - \underbrace{\kappa}_{\text{scaled relatedness}} \bullet \underbrace{\frac{\alpha(1-x)}{1-\alpha(1-x)} \bullet x}_{=[\text{intra-LRE effect}]>0} > 0 \tag{10}$$

where κ, referred to as 'scaled relatedness', measures the extent to which the extra juveniles produced via LRE are likely to share the common ancestor (see Appendix B for more precise interpretation), as a function of $n, d_\mathrm{f}, d_\mathrm{m}$, in reference to the expected strength of kin competition (van Cleve 2015). The last term represents the effect of intra-generational LRE on the inclusive fitness of the focal individual. Clearly, with LRE, cESS is smaller than $\widehat{x_\emptyset}$, that is, LRE leads to more female-biased sex ratios (Fig 4A).

We numerically evaluated the cESS to find that larger group sizes favour less female biased sex ratio and the cESS eventually approaches 1/2 (or Fisherian sex ratio) as $n \to +\infty$ (Fig 5, left panels). Increasing α leads to more female-biased sex ratios (Figs 4 and 5). As in the results for no LRE, sex ratios may be biased towards the more dispersing sex, but LRE makes the evolution of female-biased sex ratios more likely.

Example: no male-dispersal, $d_\mathrm{m} = 0$

Suppose for now $d_\mathrm{m} = 0$, and in this case, we can show that $\kappa = 1/n$ (Taylor 1992; Gardner *et al.* 2009) and therefore ploidy has no influence on the effect of LRE nor κ (Taylor 1992; Lehmann 2007). This is partly because males are fully philopatric, mating takes place in prior to female dispersal, and female dispersal allows males' and females' gametes both to disperse by the same degree, which leads to $\sigma_\mathrm{MC} - \sigma_\mathrm{MB} = \left(1 - d_\mathrm{f}\right)^2 \equiv \sigma_\mathrm{RC}$ (i.e., males and females are subject to the same degree of local competition), where '$\equiv$' is identity ('always equivalent to'). This scenario is similar to plants undergoing gametic (pollen) and zygotic (seed) dispersal when pollen dispersal is fully restricted within a patch (see Rousset 2004; Ravigné *et al.* 2006; Iritani 2020 for more details). That is, decomposing the scale of competition tells us otherwise missed fact: when $d_\mathrm{m} = 0$, the scale of competition for both sexes is equivalent (thus giving the almost-invariance result).

Varying male-dispersal, $d_\mathrm{m} > 0$

Now we tune d_m from 0 to 1 and assess its impacts upon cESS. We find that increasing d_m or d_f is likely to favour less or more female-biased sex ratios (Fig 5; respectively), and taking both to 1 leads to Fisherian sex ratio. For an intermediate male dispersal ($d_\mathrm{m} = 0.5$), male-bias is still likely

but with a possibility of switching from male- to female-bias as α or d_{f} increases. Therefore, under intra-generational LRE, the sex ratios, which could be otherwise male biased, may be biased towards female by natural selection.

Trans-generational LRE

We find that trans-generational LRE also facilitates the evolution of female-biased sex ratio (Fig 4B). Numerical estimation revealed that larger group sizes favour less female biased sex ratio and the cESS eventually approaches 1/2 (or Fisherian sex ratio) as $n \to +\infty$ (Fig 6, left panels), and increasing α leads to more female-biased sex ratio (Fig 6), as in the trans-generational LRE model. The inclusive fitness effect of trans-generational LRE decreases with d_{f}; when $d_{\mathrm{f}} = 1$ (full female dispersal), for instance, the effect of trans-generational LRE vanishes for any $\alpha > 0$ (Fig 6A, midel panel). The trans-generational LRE, which ensues after female dispersal, is sensitive to d_{f} because the probability that females can help their relatives (by remaining in the natal patch, $1 - d_{\mathrm{f}}$) decreases with d_{f}.

Example: No male-dispersal

When $d_{\mathrm{m}} = 0$, we find that Hamilton's rule reads:

$$H(x) = \widehat{x_{\emptyset}}\Big|_{d_{\mathrm{m}}=0} - x - \frac{1}{n} \bullet \underbrace{\frac{\alpha\sigma_{\mathrm{RC}}(1-x)}{1-\alpha\sigma_{\mathrm{RC}}(1-x)}x}_{=[\text{trans-LRE effect}]>0} > 0, \tag{11}$$

where the scaled relatedness is now given by $\kappa = 1/n$ as in the intra-generational model, but the last term on Eqn (11) is clearly smaller than that on Eqn (10); i.e., trans-generational LRE is weaker than is intra-generational LRE in terms of their effects on selection.

Varying male-dispersal

Varying $d_{\mathrm{m}} > 0$ turns out to give complicated form of Hamilton's rule (see Appendix B), except for the extreme case $d_{\mathrm{m}} = 1$:

$$H(x) = \widehat{x_{\emptyset}}\Big|_{d_{\mathrm{m}}=1} - x - \underbrace{\frac{c_{\mathrm{m}}R_0^{\mathbf{m}} + c_{\mathrm{f}}R_0^{\mathbf{f}}(1-\sigma_{\mathrm{RC}})}{c_{\mathrm{m}}R_{\bullet}^{\mathbf{m}} + c_{\mathrm{f}}\left(R_{\bullet}^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\right)}}_{=\kappa \text{ with } d_{\mathrm{m}}=1} \bullet \frac{\frac{\alpha}{2}\sigma_{\mathrm{RC}}(1-x)}{1-\frac{\alpha}{2}\sigma_{\mathrm{RC}}(1-x)}x > 0, \tag{12}$$

which tells us that $\alpha\sigma_{\mathrm{RC}}$ in Eqn (11) is now replaced with $\alpha\sigma_{\mathrm{MC}}/2$, with 1/2 meaning that only half of females' genes are transmitted to females (who, as opposed to males all dispersing, are likely philopatric and thus potentially contribute to the build up of trans-generational relatedness). From the expression, the scaled relatedness κ is now with reference to null (no LMC or MB as all males disperse) for males and $\sigma_{\mathrm{RC}}R_0^{\mathrm{f}}$ (LRC, which is zero when all females disperse $d_{\mathrm{f}} = 1$).

Discussion

We found that cooperative interactions between females (LRE) can lead to more female-biased sex ratios than those predicted by local mate competition (LMC) theory. Specifically, we have considered two types of LRE, and found that cooperation from offspring to their parents generation (intra-generational LRE) can lead to more female biased sex ratios than cooperation between members of the same generation (inter-generational LRE). This difference is because intra-generational LRE occurs prior to female dispersal and therefore each juvenile female has direct access to helping genetically related juveniles, thereby increasing the inclusive fitness of producing daughters over sons; in contrast, trans-generational LRE occurs after female dispersal, and dispersed females are unable to provide help to relatives, thereby reducing the inclusive fitness benefit of LRE as d_{f} increases. Our key result is therefore that LRE, alongside LMC, has the potential to generate extremely female-biased sex ratios, of the form that have been observed in nature, in species such as *Melittobia* and *Sclerodermus* wasps (Abe *et al.* 2003, 2014; Tang *et al.* 2014).

As found by previous theory, we found that, in the absence of LRE, selection in general favours the biased sex ratio towards the more dispersing sex (Bulmer & Taylor 1980; Taylor 1994; Wild & Taylor 2004), which is because kin competition between members of one sex reduces the relative value of producing that sex. In contrast, when females interact cooperatively (LRE), this increases the relative value of producing females, and natural selection thus favours less male-biased or more female-biased sex ratios (Figs 4 and 5; Taylor 1981; Emlen *et al.* 1986; Pen & Weissing 2000; Wild & Taylor 2004; Wild 2006; Wild & West 2009; but see Khwaja *et al.* 2017).

Despite the formal similarity between intra- and trans-generational LRE models, there is the quantitative difference in the consequences of dispersal rates for sex ratios. For intra-generational LRE, increasing the intensity of LRE (α) leads to more female-biased sex ratio by increasing the benefit of producing juvenile females who assist their mother (Figs 5 and 6). This selective force acts even when female dispersal rate is high, because intra-generational LRE allows juvenile females to assist their own mother before dispersal. In contrast, the trans-generational LRE predicts that increasing the intensity of LRE (α) has a weaker effect on the selection for the female-bias (Figs 4 and 6) compared to the intra-generaional LRE. The inclusive fitness effect of trans-generational LRE vanishes if females undergo complete dispersal; in other words, $d_{\mathrm{f}} = 1$ implies that cESS is independent of α for the trans-generational LRE. This result is because following complete female-dispersal, dispersed juvenile

females (the proportion d_f) do not have the access to their relatives and there are unable to engage in helping. Hence, the two models suggest that distinguishing the timing of LRE is of crucial importance for sex ratio evolution in empirical and experimental systems.

Our models provide a formal theoretical explanation for the extreme sex ratio biases that have been observed in *Sclerodermus harmand*, and several *Melittobia* wasp species. In both these cases, local mate competition is likely, but the offspring sex ratios are much more female biased than would be expected from local mate competition theory (LMC; Fig 1). We have shown in our trans-generational model that a combination of cooperative interactions between sisters (LRE) and LMC can lead to more female biased sex ratios, consistent with the empirical data. In *Sclerodermus harmandi*, females cooperate to suppress hosts and engage in brood cares (Tang *et al.* 2014; Kapranas *et al.* 2016; Lupi *et al.* 2017). In *Melittobia*, females tend to aggregate on the same hosts and cooperatively parasitise them (J. Abe, unpublished; Rosenheim 1990), females fight against the symbiont mites of host species (Okabe & Makino 2008), and their female offspring collaborate to tunnel into the materials of host nests to disperse (Deyrup *et al.* 2005), suggesting that LRE may promote female-biased sex ratios. However, in *Melittobia australica*, the quantitative discrepancy appears to be large (Figs 1 and 6). Possible factors may include multiplicativity of LRE $\beta(x) = K/\left(1 - \alpha(1-x)^{\theta}\right)$ (as in Appendix C2 and SI Fig 3; reader may want to compare this formula with Eqns (1) and (3)) which lead to even more female-biased sex ratios and LRE from juvenile females that specifically increases the production of daughters (not sons), or alternatively, we may need to consider cohesive dispersal or kin recognition among females that promotes LRE but reduces LRC. Recent field work shows that sex ratios in *Melittoiba* depend on female dispersal status, in which sex ratios produced by dispersing females are less female-biased than ones by non-dispersing females or observed in the laboratory (Fig. 1), but more female-biased than the prediction by a model incorporating dispersal status (Abe et al. in preparation). It may even be the case that *Melittobia australica* are unable to facultatively adjust sex allocation in response to local patch size due to some constraints (Shuker & West 2004; Greeff *et al.* 2020). Despite this, our models offer a useful framework to study the evolution of sex ratios under LRE and allow further development of theoretical work.

Our models are also applicable to a wide range of other organisms, in which there can be both cooperative (positive) and competitive (negative) interactions with relatives. For example: (i) mammals, birds and insects where offspring help their parents, but can also compete for breeding sites (territories; Komdeur *et al.* 1997; Clutton-Brock 2002; Stark 1992); (ii) allodapine bees where sisters

cooperate to form nests, but offspring can compete for resources (Schwarz 1988; Cronin & Schwarz 1997; Schwarz *et al.* 1998); (iii) *Diadasina distincta* bees where individuals must compete for nest sites, but nesting at higher densities can reduce parasitism (Martins *et al.* 1999). In all cases, our general prediction is that a female-biased sex ratio bias is favoured to increase cooperative interactions and reduce competitive interactions.

To conclude, our analysis suggest that intra- and trans-generational LRE from females does promote female-biased sex ratios, but their impacts upon the evolutionary outcomes differ especially in terms of the consequences of dispersal: intra-generational LRE does not depend on female dispersal rate but trans-genearational LRE does. One of the possible extensions of the present model is to study joint evolution of sex ratio and other traits under LRE (Mullon *et al.* 2016, 2018). For instance, how does joint evolution shape the association between sex-biased dispersal and sex allocation strategy, e.g., in birds and vertebrates (Frank 1990; Komdeur *et al.* 1997; Goltsman *et al.* 2005; Banks *et al.* 2008; Hjernquist *et al.* 2009)? Future studies could be directed towards more realistic modeling of LRE, by e.g., incorporating the effects of the number of adult females (n) on LRE, or working on intra-sexual LRE ('who helps whom'; Rodrigues & Gardner 2013; Rodrigues & Kokko 2016). Working with specific organisms of interest based upon multiple approaches may yield a better understanding of the evolution of sex ratios in viscous populations.

Acknowledgement

We thank Troy Day, Ian Hardy, Yoshitaka Kamimura, Asher Leeks, Joyce Tong, and anonymous reviewers for helpful comments on various versions of the manuscript; Ian Hardy, Hiroaki Yanagisawa, and Baoping Li for sharing their data; JSPS-KAKENHI (grant numbers 19K22457, 19K23768, and 20K15882 to RI), and ERC (grant number 834164 to SW) for funding. We thank Laurent Lehmann for kindly hosting RI in 2015 and his group for discussion, which helped RI develop the mathematical analyses presented here.

Author contribution

JA conceived the idea; RI carried out the mathematical analyses; JA and RI drafted the first version of the manuscript; all authors contributed to the revision.

References

Abdi, M. K., Hardy, I. C. W., Jucker, C., & Lupi, D. (2020a). Kinship effects in quasi-social parasitoids II: co-foundress relatedness and host dangerousness interactively affect host exploitation. *Biological Journal of the Linnean Society*, **130**.4, pp. 642–660. DOI: 10.1093/biolinnean/blaa047.

Abdi, M. K., Lupi, D., & Hardy, I. C. W. (2020b). Co-foundress confinement elicits kinship effects in a naturally sub-social parasitoid. *Journal of Evolutionary Biology*, DOI: 10.1111/jeb.13635.

Abdi, M. K., Lupi, D., Jucker, C., & Hardy, I. C. W. (2020c). Kinship effects in quasi-social parasitoids I: co-foundress number and relatedness affect suppression of dangerous hosts. *Biological Journal of the Linnean Society*, **130**.4, pp. 627–641. DOI: 10.1093/biolinnean/blaa046.

Abe, J., Kamimura, Y., Kondo, N., & Shimada, M. (2003). Extremely female-biased sex ratio and lethal male–male combat in a parasitoid wasp, *Melittobia australica* (Eulophidae). *Behavioral Ecology*, **14**.1, pp. 34–39. DOI: 10.1093/beheco/14.1.34.

Abe, J., Kamimura, Y., & West, S. A. (2014). Inexplicably female-biased sex ratios in *Melittobia* wasps. *Evolution*, **68**.9, pp. 2709–2717. DOI: 10.1111/evo.12451.

Avilés, L. (1993). Interdemic selection and the sex ratio: a social spider perspective. *The American Naturalist*, **142**.2, pp. 320–345. DOI: 10.1086/285540.

Banks, S. C., Knight, E. J., Dubach, J. E., & Lindenmayer, D. B. (2008). Microhabitat heterogeneity influences offspring sex allocation and spatial kin structure in possums. *Journal of Animal Ecology*, **77**.6, pp. 1250–1256. DOI: 10.1111/j.1365-2656.2008.01448.x.

Bourke, A. F. G. & Franks, N. R. (1995). *Social evolution in ants*. **62**. Princeton University Press.

Bulmer, M. G. (1986). Sex ratio theory in geographically structured populations. *Heredity*, **56**.1, pp. 69–73. DOI: 10.1038/hdy.1986.10.

Bulmer, M. G. (1994). *Evolutionary theoretical ecology*. Sinauer Associates, Sunderland, MA.

Bulmer, M. G. & Taylor, P. D. (1980). Dispersal and the sex ratio. *Nature*, **284**.5755, pp. 448–449. DOI: 10.1038/284448a0.

Burton-Chellew, M. N., Koevoets, T., Grillenberger, B. K., Sykes, E. M., Underwood, S. L., Bijlsma, K., Gadau, J., Zande, L. v. d., Beukeboom, L. W., West, S. A., & Shuker, D. M. (2008). Facultative sex ratio adjustment in natural populations of wasps: cues of local mate competition and the precision of adaptation. *The American Naturalist*, **172**.3, pp. 393–404. DOI: 10.1086/589895.

Caswell, H. (2001). *Matrix population models*. Wiley Online Library.

Charnov, E. L. (1982). *The theory of sex allocation*. **18**. Princeton University Press.

Christiansen, F. B. (1991). On conditions for evolutionary stability for a continuously varying character. *The American Naturalist*, pp. 37–50. DOI: 10.1086/285203.

Clark, A. B. (1978). Sex ratio and local resource competition in a prosimian primate. *Science*, **201**.4351, pp. 163–165. DOI: 10.1126/science.201.4351.163.

Clutton-Brock, T. H. (2002). Evolution and development of sex differences in cooperative behavior in meerkats. *Science*, **297**.5579, pp. 253–256. DOI: 10.1126/science.1071412.

Cronin, A. L. & Schwarz, M. P. (1997). Sex ratios, local fitness enhancement and eusociality in the allodapine bee *Exoneura richardsoni*. *Evolutionary Ecology*, **11**.5, pp. 567–577. DOI: 10.1007/s10682-997-1512-3.

Deyrup, L. D., Matthews, R. W., & González, J. M. (2005). Cooperative chewing in a gregariously developing parasitoid wasp, melittobia digitata dahms, is stimulated by structural cues and a pheromone in crude venom extract. *Journal of Insect Behavior*, **18**.3, pp. 293–304. DOI: 10.1007/s10905-005-3690-6.

Emlen, S. T., Emlen, J. M., & Levin, S. A. (1986). Sex-Ratio Selection in Species with Helpers-At-The-Nest. *The American Naturalist*, **127**.1, pp. 1–8. DOI: 10.1086/284463.

Frank, S. A. (1986). The genetic value of sons and daughters. *Heredity*, **56**.3, pp. 351–354. DOI: 10.1038/hdy.1986.56.

Frank, S. A. (1990). Sex allocation theory for birds and mammals. *Annual Review of Ecology and Systematics*, **21**.1, pp. 13–55. DOI: 10.1146/annurev.es.21.110190.000305.

Frank, S. A. (1998). *Foundations of social evolution*. Princeton University Press.

Gardner, A., Arce, A., & Alpedrinha, J. (2009). Budding dispersal and the sex ratio. *Journal of Evolutionary Biology*, **22**.5, pp. 1036–1045. DOI: 10.1111/j.1420-9101.2009.01719.x.

Gardner, A. & West, S. A. (2006). Demography, altruism, and the benefits of budding. *Journal of Evolutionary Biology*, **19**.5, pp. 1707–1716. DOI: 10.1111/j.1420-9101.2006.01104.x.

Goltsman, M., Kruchenkova, E. P., Sergeev, S., Johnson, P. J., & Macdonald, D. W. (2005). Effects of food availability on dispersal and cub sex ratio in the Mednyi Arctic fox. *Behavioral Ecology and Sociobiology*, **59**.2, pp. 198–206. DOI: 10.1007/s00265-005-0025-8.

Grafen, A. (1985). A geometric view of relatedness. *Oxford surveys in Evolutionary Biology*, **2**.2, pp. 28–89.

Greeff, J. M. (2002). Mating system and sex ratios of a pollinating fig wasp with dispersing males. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **269**.1507, pp. 2317–2323. DOI: 10.1098/rspb.2002.2160.

Greeff, J. M., Pentz, K., & Warren, M. (2020). The efficacy of natural selection in producing optimal sex ratio adjustments in a fig wasp species. *Proceedings of the Royal Society B: Biological Sciences*, **287**.1934, p. 20201377. DOI: 10.1098/rspb.2020.1377.

Gu, H. & Dorn, S. (2003). Mating system and sex allocation in the gregarious parasitoid Cotesia glomerata. *Animal Behaviour*, **66**.2, pp. 259–264. DOI: 10.1006/anbe.2003.2185.

Hamilton, W. D. (1967). Extraordinary sex ratios. *Science*, **156**.3774, pp. 477–488. DOI: 10.1126/science.156.3774.477.

Hardy, I. C. W. (2002). *Sex Ratios: Concepts and Research Methods*. Cambridge University Press.

Herre, E. A. (1985). Sex ratio adjustment in fig wasps. *Science*, **228**.4701, pp. 896–898. DOI: 10.1126/science.228.4701.896.

Hjernquist, M. B., Thuman Hjernquist, K. A., Forsman, J. T., & Gustafsson, L. (2009). Sex allocation in response to local resource competition over breeding territories. *Behavioral Ecology*, **20**.2, pp. 335–339. DOI: 10.1093/beheco/arp002.

Iritani, R. (2020). Gametophytic competition games among relatives: when does spatial structure select for facilitativeness or competitiveness in pollination? *Journal of Ecology*, **00**, pp. 1–13. DOI: 10.1111/1365-2745.13282.

Kapranas, A., Hardy, I. C., Tang, X., Gardner, A., & Li, B. (2016). Sex ratios, virginity, and local resource enhancement in a quasisocial parasitoid. *Entomologia Experimentalis et Applicata*, **159**.2, pp. 243–251. DOI: 10.1111/eea.12418.

Khwaja, N., Hatchwell, B. J., Freckleton, R. P., & Green, J. P. (2017). Sex allocation patterns across cooperatively breeding birds do not support predictions of the repayment hypothesis. *The American Naturalist*, **190**.4, pp. 547–556. DOI: 10.1086/693532.

Komdeur, J., Daan, S., Tinbergen, J., & Mateman, C. (1997). Extreme adaptive modification in sex ratio of the Seychelles warbler's eggs. *Nature*, **385**.6616, pp. 522–525. DOI: 10.1038/385522a0.

Lehmann, L. (2007). The evolution of trans-generational altruism: kin selection meets niche construction. *Journal of Evolutionary Biology*, **20**.1, pp. 181–189. DOI: 10.1111/j.1420-9101.2006.01202.x.

— (2008). The adaptive dynamics of niche constructing traits in spatially subdivided populations: evolving posthumous extended phenotypes. *Evolution*, **62**.3, pp. 549–566. DOI: 10.1111/j.1558-5646.2007.00291.x.

Lupi, D., Favaro, R., Jucker, C., Azevedo, C. O., Hardy, I. C., & Faccoli, M. (2017). Reproductive biology of Sclerodermus brevicornis, a European parasitoid developing on three species of invasive longhorn beetles. *Biological Control*, **105**, pp. 40–48. DOI: 10.1016/j.biocontrol.2016.11.008.

Martins, R. P., Antonini, Y., Silveira, F. A. da, & West, S. A. (1999). Seasonal variation in the sex allocation of a neotropical solitary bee. *Behavioral Ecology*, **10**.4, pp. 401–408. DOI: 10.1093/beheco/10.4.401.

Matthews, R. W., González, J. M., Matthews, J. R., & Deyrup, L. D. (2009). Biology of the Parasitoid *Melittobia* (Hymenoptera: Eulophidae). *Annual Review of Entomology*, **54**.1, pp. 251–266. DOI: 10.1146/annurev.ento.54.110807.090440.

Maynard Smith, J. & Price, G. R. (1973). The Logic of Animal Conflict. *Nature*, **246**, p. 15. DOI: 10.1038/246015a0.

Michod, R. E. & Hamilton, W. D. (1980). Coefficients of relatedness in sociobiology. *Nature*, **288**.5792, p. 694. DOI: 10.1038/288694a0.

Mullon, C., Keller, L., & Lehmann, L. (2016). Evolutionary stability of jointly evolving traits in subdivided populations. *The American Naturalist*, **188**.2, pp. 175–195. DOI: 10.1086/686900.

— (2018). Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nature Ecology & Evolution*, **2**.1, p. 132. DOI: 10.1038/s41559-017-0397-y.

Okabe, K. & Makino, S. (2008). Parasitic mites as part-time bodyguards of a host wasp. *Proceedings of the Royal Society B: Biological Sciences*, **275**.1648, pp. 2293–2297. DOI: 10.1098/rspb.2008.0586.

Pen, I. & Weissing, F. J. (2000). Sex-ratio optimization with helpers at the nest. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **267**.1443, pp. 539–543. DOI: 10.1098/rspb.2000.1034.

Ravigné, V., Olivieri, I., Martinez, S. G., & Rousset, F. (2006). Selective interactions between short-distance pollen and seed dispersal in self-compatible species. *Evolution*, **60**.11, pp. 2257–2271.

Rodrigues, A. M. M. & Gardner, A. (2013). Evolution of helping and harming in heterogeneous groups. *Evolution*, **67**.8, pp. 2284–2298. DOI: 10.1111/j.1558-5646.2012.01594.x.

Rodrigues, A. M. M. & Gardner, A. (2015). Simultaneous failure of two sex-allocation invariants: implications for sex-ratio variation within and between populations. *Proceedings of the Royal Society of London B: Biological Sciences*, **282**.1810, p. 20150570. DOI: 10.1098/rspb.2015.0570.

Rodrigues, A. M. M. & Kokko, H. (2016). Models of social evolution: can we do better to predict 'who helps whom to achieve what'? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **371**.1687.

Rosenheim, J. A. (1990). Density-dependent parasitism and the evolution of aggregated nesting in the solitary hymenoptera. *Annals of the Entomological Society of America*, **83**.3, pp. 277–286. DOI: 10.1093/aesa/83.3.277.

Rousset, F. (2004). *Genetic Structure and Selection in Subdivided Populations (MPB-40)*. Princeton University Press.

Rousset, F. & Billiard, S. (2000). A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *Journal of Evolutionary Biology*, **13**.5, pp. 814–825. DOI: 10.1046/j.1420-9101.2000.00219.x.

Schwarz, M. P. (1988). Local resource enhancement and sex ratios in a primitively social bee. *Nature*, **331**.6154, pp. 346–348. DOI: 10.1038/331346a0.

Schwarz, M., Bull, N., & Hogendoorn, K. (1998). Evolution of sociality in the allodapine bees: a review of sex allocation, ecology and evolution. *Insectes Sociaux*, **45**.4, pp. 349–368. DOI: 10.1007/s000400050095.

Shuker, D. M., Pen, I., & West, S. A. (2006). Sex ratios under asymmetrical local mate competition in the parasitoid wasp Nasonia vitripennis. *Behavioral Ecology*, **17**.3, pp. 345–352. DOI: 10.1093/beheco/arj034.

Shuker, D. M., Reece, S. E., Taylor, J. A., & West, S. A. (2004). Wasp sex ratios when females on a patch are related. *Animal Behaviour*, **68**.2, pp. 331–336. DOI: 10.1016/j.anbehav.2003.12.003.

Shuker, D. M. & West, S. A. (2004). Information constraints and the precision of adaptation: Sex ratio manipulation in wasps. *Proceedings of the National Academy of Sciences*, **101**.28, pp. 10363–10367. DOI: 10.1073/pnas.0308034101.

Sigmund, K. (2010). *The calculus of selfishness*. Princeton University Press.

Stark, R. E. (1992). Sex ratio and maternal investment in the multivoltine large carpenter bee Xylocopa sulcatipes (Apoideh: Anthophoridae). *Ecological Entomology*, **17**.2, pp. 160–166. DOI: 10.1111/j.1365-2311.1992.tb01174.x.

Tagawa, J. (2000). Sex allocation and clutch size in the gregarious larval endoparasitoid wasp, Cotesia glomerata. *Entomologia Experimentalis et Applicata*, **97**.2, pp. 193–202. DOI: 10.1046/j.1570-7458.2000.00730.x.

Takada, T. & Kigami, J. (1991). The dynamical attainability of ESS in evolutionary games. *Journal of Mathematical Biology*, **29**.6, pp. 513–529. DOI: 10.1007/bf00164049.

Tang, X., Meng, L., Kapranas, A., Xu, F., Hardy, I. C., & Li, B. (2014). Mutually beneficial host exploitation and ultra-biased sex ratios in quasisocial parasitoids. *Nature Communications*, **5**.1. DOI: 10.1038/ncomms5942.

Taylor, P. D. (1981). Intra-sex and inter-sex sibling interactions as sex ratio determinants. *Nature*, **291**.5810, pp. 64–66. DOI: 10.1038/291064a0.

— (1988). Inclusive fitness models with two sexes. *Theoretical Population Biology*, **34**.2, pp. 145–168. DOI: 10.1016/0040-5809(88)90039-1.

— (1990). Allele-frequency change in a class-structured population. *The American Naturalist*, **135**, pp. 95–106. DOI: 10.1086/285034.

— (1992). Altruism in viscous populations—an inclusive fitness model. *Evolutionary Ecology*, **6**.4, pp. 352–356. DOI: 10.1007/bf02270971.

— (1994). Sex ratio in a stepping-stone population with sex-specific dispersal. *Theoretical Population Biology*, **45**.2, pp. 203–218. DOI: 10.1006/tpbi.1994.1011.

Taylor, P. D. & Frank, S. A. (1996). How to make a kin selection model. *Journal of Theoretical Biology*, **180**.1, pp. 27–37. DOI: 10.1006/jtbi.1996.0075.

Taylor, P. D., Wild, G., & Gardner, A. (2007). Direct fitness or inclusive fitness: how shall we model kin selection? *Journal of Evolutionary Biology*, **20**.1, pp. 301–309. DOI: 10.1111/j.1420-9101.2006.01196.x.

van Cleve, J. (2015). Social evolution and genetic interactions in the short and long term. *Theoretical Population Biology*, **103**, pp. 2–26. DOI: 10.1016/j.tpb.2015.05.002.

Waage, J. K. (1982). Sib-mating and sex ratio strategies in scelionid wasps. *Ecological Entomology*, **7**.1, pp. 103–112. DOI: 10.1111/j.1365-2311.1982.tb00648.x.

Werren, J. H. (1983). Brood Size and Sex Ratio Regulation in the Parasitic Wasp Nasonia Vitripennis (Walker) (Hymenoptera: Pteromalidae). *Netherlands Journal of Zoology*, **34**.2, pp. 123–143. DOI: 10.1163/002829684x00100.

West, S. (2009). *Sex allocation*. **48**. Princeton University Press.

Wild, G. (2006). Sex ratios when helpers stay at the nest. *Evolution*, **60**.10, pp. 2012–2022. DOI: 10.1111/j.0014-3820.2006.tb01839.x.

Wild, G. & Taylor, P. D. (2004). Kin selection models for the co-evolution of the sex ratio and sex-specific dispersal. *Evolutionary Ecology Research*, **6**.4, pp. 481–502.

Wild, G. & West, S. A. (2009). Genomic imprinting and sex allocation. *The American Naturalist*, **173**.1, E1–E14. DOI: 10.1086/593305.

Wilson, D. S., Pollock, G., & Dugatkin, L. A. (1992). Can altruism evolve in purely viscous populations? *Evolutionary Ecology*, **6**.4, pp. 331–341. DOI: 10.1007/bf02270969.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, **16**.2, p. 97. DOI: 10.1007/bf02459575.

Yamaguchi, Y. (1985). Sex ratios of an aphid subject to local mate competition with variable maternal condition. *Nature*, **318**.6045, pp. 460–462. DOI: 10.1038/318460a0.

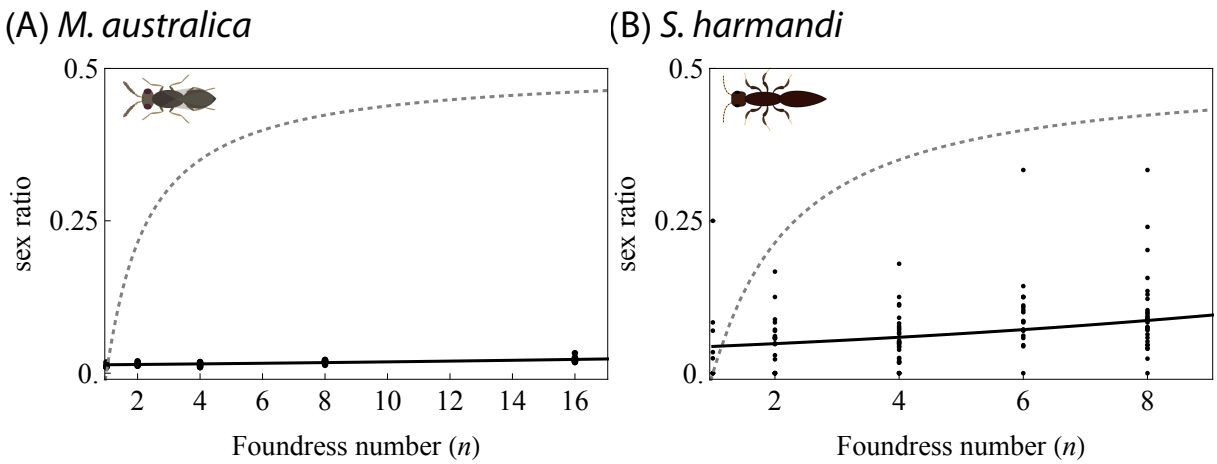


Figure 1: Extremely female-biased sex ratio in (A) *Melittobia australica* from Abe *et al.* (2003), and (B) *Sclerodermus harmandi* from Tang *et al.* (2014) and Kapranas *et al.* (2016). Both species haplodiploids. Outliers removed for (B), as in the original articles Tang *et al.* (2014) and Kapranas *et al.* (2016). Note that the horizontal axes are scaled differently. Dotted lines: Reference sex ratio give by $\big((n-1)(4n-2)\big)/\big(2n(4n-1)\big)$ (evolutionarily stable sex ratio for haplodiploids with $d_{\mathrm{f}} = 1$). Solid line in (A): predicted values by generalized linear models; in (B): shown in Tang *et al.* (2014).

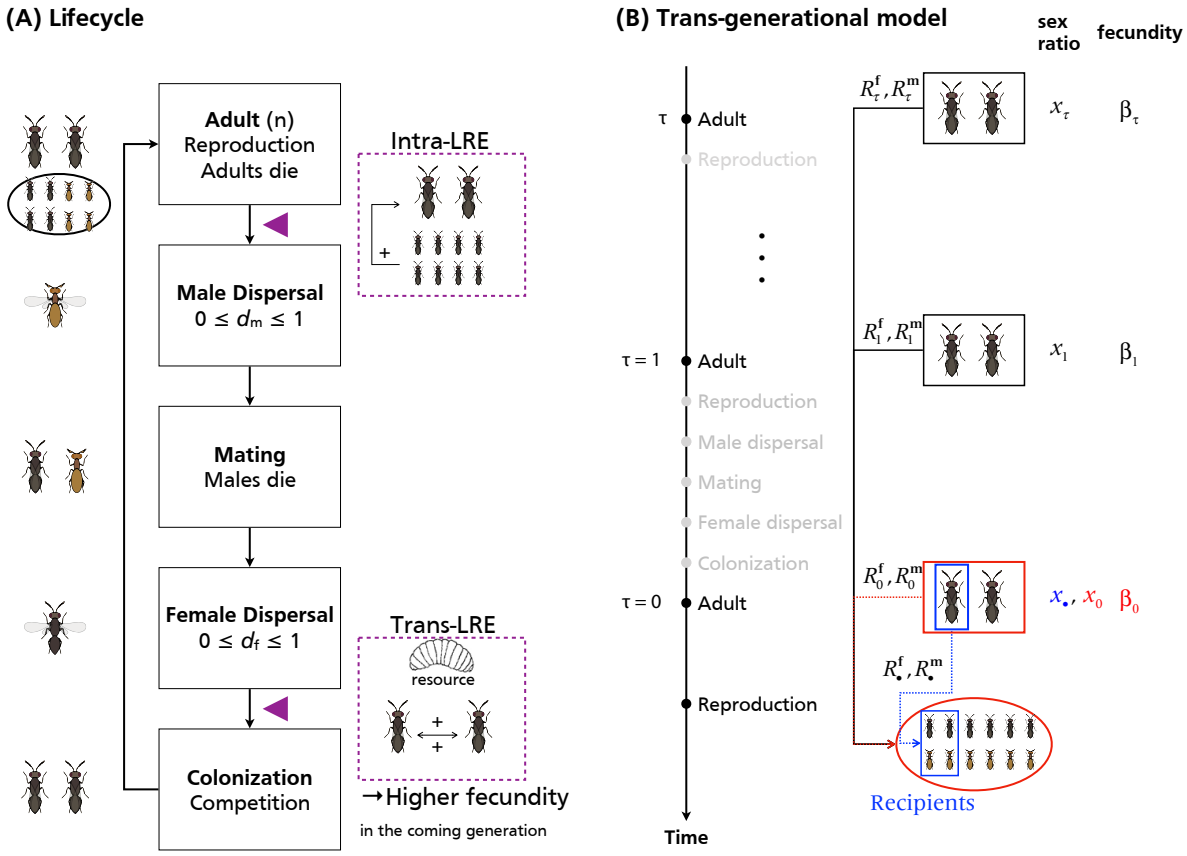


Figure 2: Schematic illustration of (A) lifecycle and (B) trans-generational relatedness. (A) Gray individuals: females. Brown individuals: males. Purple triangle: possible timing of LRE. Model species: *Melittobia australica* (but note that males of this species really are of flightless; Matthews *et al.* 2009). (B) Adult within blue box: the focal individual; juveniles within blue box: the focal individual's offspring; red: average in the patch. We count the generations backwards in time ($\tau = 0$ the present, $\tau = 1$ parental, etc). R are relatedness coefficients, each from the corresponding actor's perspective (arrows).

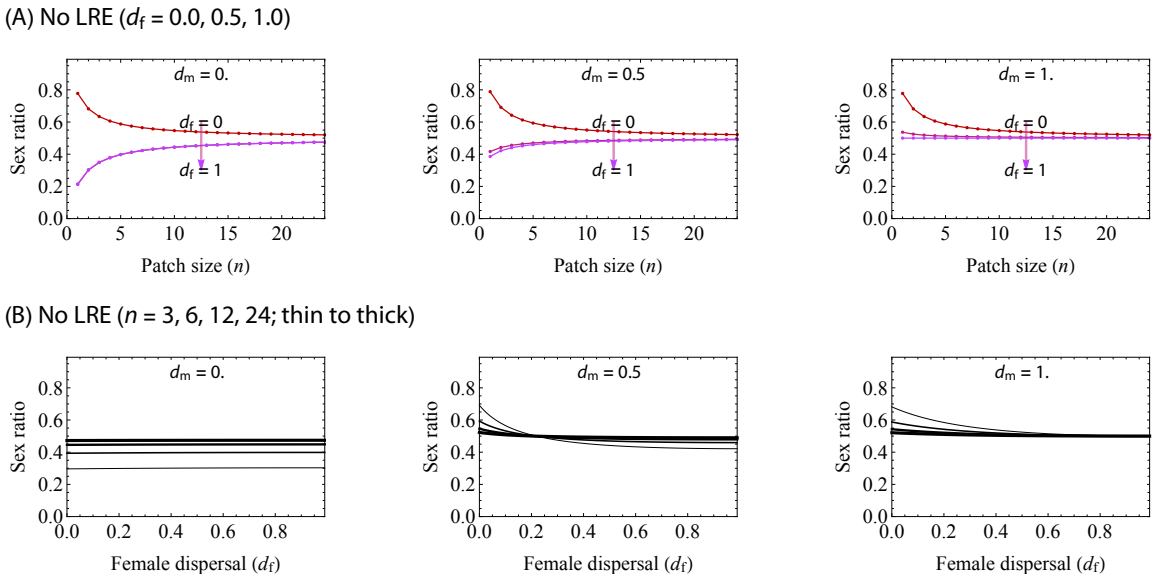


Figure 3: Evolutionary outcomes of sex ratio without LRE under haplodiploidy. (A) Sensitivity to n. Larger n monotonically tends to favour Fisherian sex ratio. Increasing d_f is likely to favour less male-bias. Note that $d_f = d_m = 0$ (red curve in the left panel) is an exceptional case in which all patches are mutually isolated, and this case therefore invalidates the present analyses (instead, entailing stochastic analyses). For (slightly) positive values for $d_f > 0$, the evolutionary outcome show very weak sensitivity to female dispersal rate (purple curves heavily overlapped in the left panel, which are visually difficult to separate from each other). (B) Sensitivity to d_f (with $d_m = 0, 0.5, 1$ from left to right panels, and $n = 2, 4, 8, 16$ from thin to thick curves). Generally, high group sizes (n) favour Fisherian sex ratio. When male dispersal is completely limited (left panel), sex ratio is almost invariant with d_f. Increasing d_m results in a shift to male-bias when female dispersal is small, and as d_f increases, female-bias is likely to be favoured by selection (middle). When male dispersal is complete ($d_m = 1$), the resulting sex ratio is male-biased and approaches Fisherian (1/2) as d_f increases. All figures produced by nullifying Eqn (7).

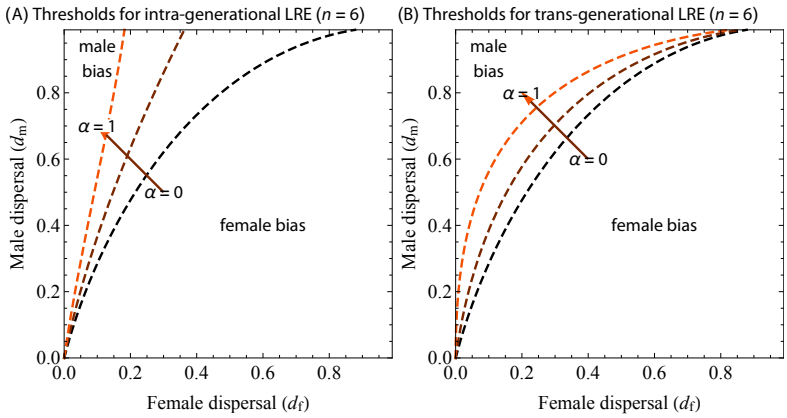


Figure 4: Threshold condition for the biased sex ratio at cESS. Each contour represents the condition for Fisherian sex ratio (1/2) to be cESS and separates the region for female- and male-biased sex ratios. Female biased sex ratios (bottom right zones) become more likely as α increases (where threshold curves plotted for $\alpha = 0, 0.5$ and 1). The contours are produced using Eqn (7) with $x = 1/2$ (Fisherian sex ratio) inserted.

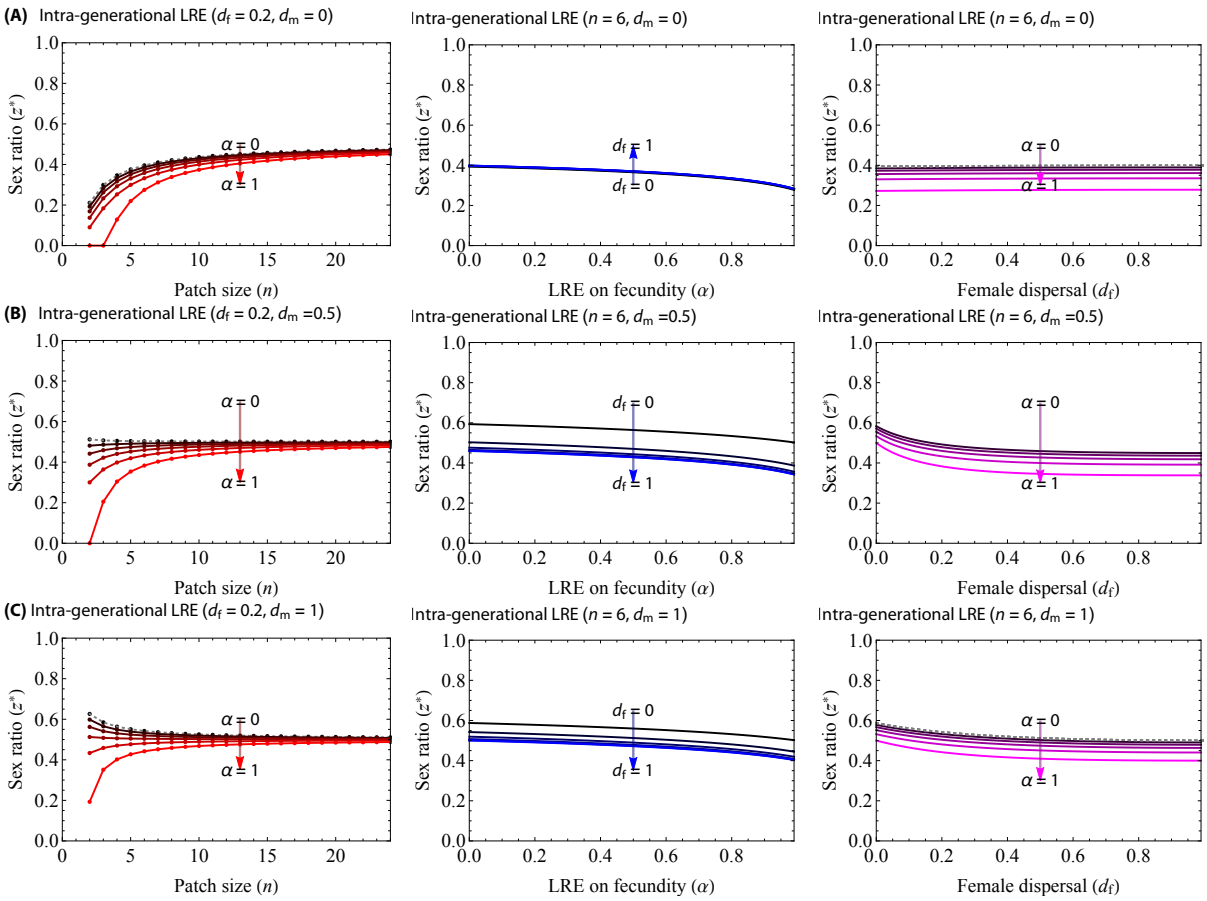


Figure 5: cESS under $d_m = 0$, for intra- generational LRE (top panels, A) and trans-generational LRE (bottom panels, B), under haplodiploidy. Gray dots: results under no LRE.

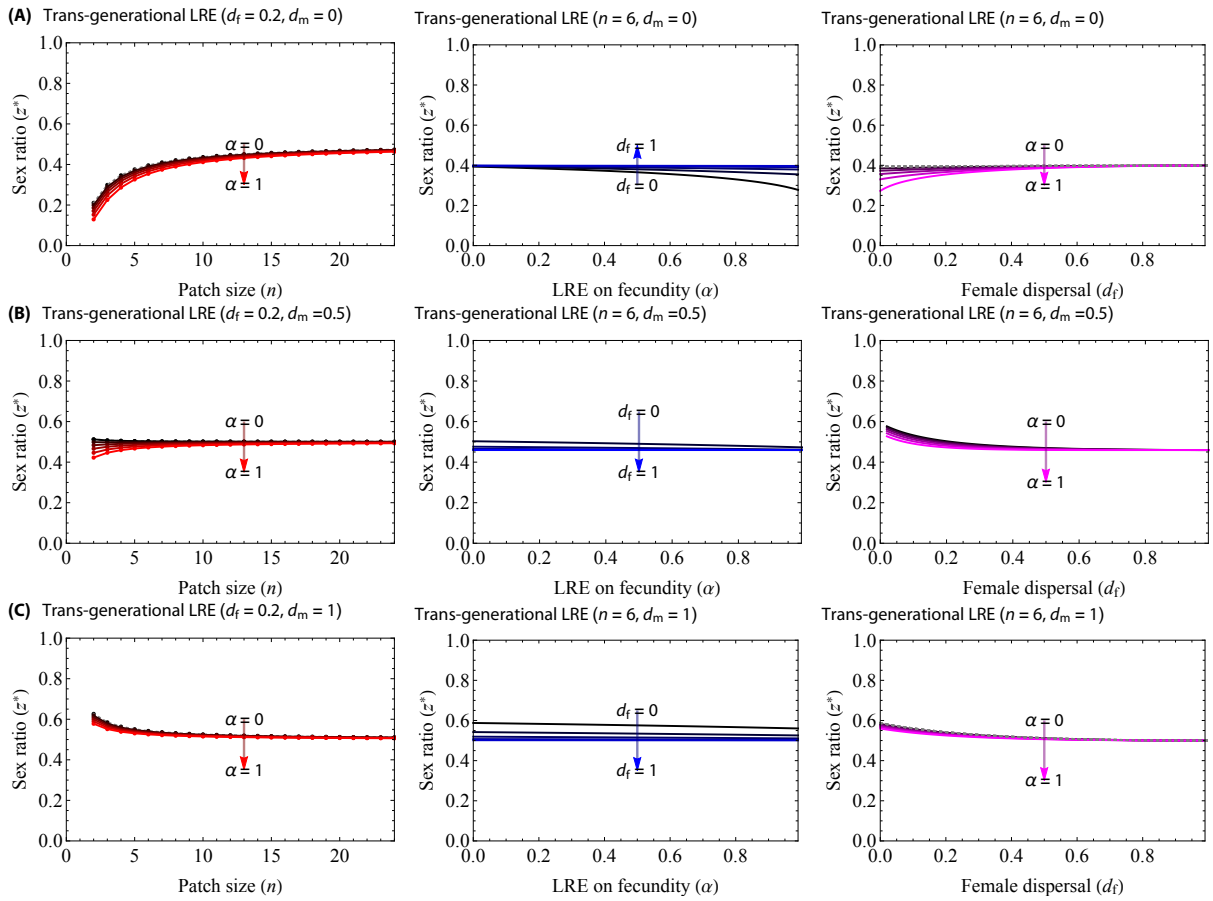


Figure 6: cESS under $d_m = 0$, for intra- generational LRE (top panels, A) and trans-generational LRE (bottom panels, B), under haplodiploidy. Gray dots: results under no LRE.

Appendix

A Fecundity mediated by LRE

A-1 Derivation

Let us for now focus on the intra-generational model. Pick a focal patch and denote the average phenotype on that patch by x_0; on that patch, the fecundity per capita may be the sum of (i) baseline fecundity (denoted by K), and (ii) effects of LRE; the latter occurs from helping behaviors of daughters, and the total number of females per patch is given by $n \bullet (1 - x_0) \bullet \beta(x_0)$, in which such an effect is equally shared by n females and thus should be divided by n (therefore n being canceled), which hence yields:

$$\beta_0 := \beta(x_0) = K + \alpha(1 - x_0)\beta(x_0) = \frac{K}{1 - \alpha(1 - x_0)}, \tag{S-1}$$

as in the main text (Eqn 1).

The same logic gives the corresponding recursion for the trans-generational model, by modifying the equation above to:

$$\begin{aligned} \beta_0 &:= K + \alpha\Big(\big(1 - d_\mathrm{f}\big)(1 - x_1)n\beta_1 + d_\mathrm{f}(1 - x)\beta^\circ n\Big)/n \\ &= K + \alpha\Big(\big(1 - d_\mathrm{f}\big)(1 - x_1)\beta_1 + d_\mathrm{f}(1 - x)\beta^\circ\Big), \end{aligned} \tag{S-2}$$

which may be reasoned as follows: the first term is a baseline fecundity; the second term overall represents the effect of LRE; the fraction $1 - d_\mathrm{f}$ of females derives from the natal patch, in a density $(1 - x_1)\beta_1 n$; d_f from other patches, in a density $(1 - x)\beta^\circ n$. Assuming that such LRE effect is shared among n individuals, we divide this term by n, obtaining the above expression. We can see that β_0 is a function of (β_1, x_1), and β_1 is a function of (β_2, x_2) (by the same reasoning), and so forth; this means that we need to deal with the nested function of β_0, based upon Lehmann (2007, 2008). With this nestedness, we henceforth write:

$$\beta_\tau = B(\beta_{\tau+1}, x_{\tau+1}) \tag{S-3}$$

for $\tau \geq 0$.

When $x_\tau = x_\bullet = x$ (for any $\tau \geq 0$; neutrality condition), we have

$$\beta^\circ = B(\beta^\circ, x), \tag{S-4}$$

which immediately gives:

$$\beta^\circ = \frac{K}{1 - \alpha(1 - x)}, \tag{S-5}$$

whereupon we recover Eqn (3) in the main text.

To guarantee β° be locally stable, we separate the timescales (holding $x_\bullet = x_\tau = x$ for all $\tau \geq 0$) by assuming the sequence of the average fecundity values $\left\{\widehat{\beta_\tau}\right\}_{\tau\geq 0}$ (with a hat put to emphasize quasi-equilibrium) is, using Eqn (S-2), determined by:

$$\widehat{\beta_\tau} = K + \alpha(1-x)\left(\left(1-d_\mathrm{f}\right)\widehat{\beta_{\tau+1}} + d_\mathrm{f}\beta^\circ\right) = B\left(\widehat{\beta_{\tau+1}}, x\right) \tag{S-6}$$

from which we have:

$$\left.\frac{\partial B\left(\widehat{\beta}, x\right)}{\partial \widehat{\beta}}\right|_{\widehat{\beta}=\beta^\circ} = \alpha(1-x)\left(1-d_\mathrm{f}\right), \tag{S-7}$$

which is obviously positive and smaller than 1, thereby guaranteeing the local stability (NB: there may be a case in which this quantity is unity but only when $x = d_\mathrm{f} = 0$ and $\alpha = 1$; the necessary condition $x = 0$ is excluded and thereby we focus on strict inequality).

A-2 Effects of selection

Intra-generational LRE

Using Eqn (S-1),

$$\left.\frac{\partial \log \beta(x_0)}{\partial x_0}\right|_{x_0=x} = -\alpha\frac{1}{1-x}, \tag{S-8}$$

and thus:

$$\left.\frac{\partial \beta_0}{\partial x_0}\right|_{x_0=x} = \left.\frac{\partial \beta(x_0)}{\partial x_0}\right|_{x_0=x} = -\alpha\frac{\beta^\circ}{1-x}. \tag{S-9}$$

Trans-generational LRE

From Eqn (S-2), similar computation gives:

$$\begin{aligned}\left.\frac{\partial \beta_0}{\partial x_1}\right|_{x_\bullet=x_\tau=x} &= \left.\frac{\partial B(\beta_1, x_1)}{\partial x_1}\right|_{x_\bullet=x_\tau=x} = -\alpha\left(1-d_\mathrm{f}\right)\beta^\circ, \\ \left.\frac{\partial \beta_0}{\partial \beta_1}\right|_{x_\bullet=x_\tau=x} &= \left.\frac{\partial B(\beta_1, x_1)}{\partial \beta_1}\right|_{x_\bullet=x_\tau=x} = \alpha\left(1-d_\mathrm{f}\right)(1-x).\end{aligned} \tag{S-10}$$

Induction yields:

$$\begin{aligned}\left.\frac{\partial\beta_0}{\partial x_\tau}\right|_{x_\bullet=x_\tau=x} &= \left.\left(\underbrace{\frac{\partial\beta_0}{\partial\beta_1}\bullet\frac{\partial\beta_1}{\partial\beta_2}\cdots\frac{\partial\beta_{\tau-2}}{\partial\beta_{\tau-1}}}_{\tau-1\text{ times}}\bullet\frac{\partial\beta_{\tau-1}}{\partial x_\tau}\right)\right|_{x_\bullet=x_\tau=x} \\ &= \Big(\alpha\big(1-d_{\mathrm{f}}\big)(1-x)\Big)^{\tau-1}\bullet\Big(-\alpha\big(1-d_{\mathrm{f}}\big)\beta^\circ\Big) \\ &= -\frac{1}{1-x}\Big(\alpha\big(1-d_{\mathrm{f}}\big)(1-x)\Big)^{\tau}\bullet\beta^\circ\end{aligned} \tag{S-11}$$

for $\tau \geq 1$. That is, the effect of selection through the expected average of sex ratios in the focal group of the τ-th generation decays geometrically in τ.

B Invasion fitness and the selection gradient: general

The reproductive success of a focal individual occurs through her daughters (winning the breeding opportunity) and her sons (winning the mating opportunity), and therefore we separate these two terms explicitly by using the class reproductive values (Taylor 1990; Caswell 2001). We use the subscripts `in` and `tr` to distinguish the two models.

B-1 Success via daughters

First of all, the focal female can produce the total number β_0 of eggs. A proportion $1-x_\bullet$ of those eggs develops into female. Therefore we have $(1-x_\bullet)\beta_0$ of daughters born to the focal. Pick one of the daughters;

- With a probability of $1-d_{\mathrm{f}}$, she stays in her natal patch, ending up with competing with, on average, $\big(1-d_{\mathrm{f}}\big)(1-x_0)\beta_0$ of juvenile females born in the same patch, plus $d_{\mathrm{f}}(1-x)\beta^\circ$ of juvenile females each born in a different patch.
- With a probability of d_{f}, she disperses to an alternative patch to compete with $(1-x)\beta^\circ$ of juvenile females.

Therefore, we have:

$$\begin{aligned}W_\bullet^{\mathbf{f}}(x_\bullet,x_0,\beta_0) &:= \big(1-d_{\mathrm{f}}\big)\frac{(1-x_\bullet)\beta_0}{\big(1-d_{\mathrm{f}}\big)(1-x_0)\beta_0+d_{\mathrm{f}}(1-x)\beta^\circ}+d_{\mathrm{f}}\frac{(1-x_\bullet)\beta_0}{(1-x)\beta^\circ} \\ &=: W_\bullet^{\mathbf{f}}\end{aligned} \tag{S-12}$$

as displayed in Eqn (5) of the main text. At phenotypic neutrality (i.e., $x_\bullet = x_0 = x$ and $\beta_0 = \beta^\circ$), we have $W_\bullet^{\mathbf{f}}(x,x,\beta^\circ) \equiv 1$ (where, by $\equiv$, we mean "always equal to" or identity).

B-2 Success via sons

The same logic gives success via sons. The focal female produces $x_\bullet \beta_0$ of sons who may compete against on average $(1-d_\mathrm{m})x_0\beta_0 + d_\mathrm{m} x\beta^\circ$ of juvenile males by staying philopatric (which occurs with a probability of $1-d_\mathrm{m}$); in which case, whether the male gamete survives to the next generation or not now relies on whether the mated female wins a breeding opportunity or not. Because mating is fully random within the patch, $W_0^\mathbf{f}(x_0, x_0, \beta_0)$ gives the proportional success of the sons' mating partners. If the sons instead disperse, which occurs with a probability of d_m, then they compete with on average $x\beta^\circ$ of males, and their mating partners' success is $W_\bullet^\mathbf{f}(x, x, \beta^\circ) = 1$. Therefore,

$$\begin{aligned} W_\bullet^\mathbf{m}\left(x_\bullet, x_0, W_0^\mathbf{f}, \beta_0\right) &:= \frac{(1-d_\mathrm{m})x_\bullet\beta_0}{(1-d_\mathrm{m})x_0\beta_0 + d_\mathrm{m} x\beta^\circ} \underbrace{\left(\frac{(1-d_\mathrm{f})(1-x_0)\beta_0}{(1-d_\mathrm{f})(1-x_0)\beta_0 + d_\mathrm{f}(1-x)\beta^\circ} + \frac{d_\mathrm{f}(1-x_0)\beta_0}{(1-x)\beta^\circ}\right)}_{=W^\mathbf{f}(x_0,x_0,\beta_0)=:W_0^\mathbf{f}} \\ &\quad + \frac{d_\mathrm{m} x_\bullet \beta_0}{x\beta^\circ} \\ &=: W_\bullet^\mathbf{m} \end{aligned} \tag{S-13}$$

as displayed in Eqn (6) of the main text.

B-3 Total fitness

The total number of genes of the focal individual transmitted to the next generation is therefore given by $W_\bullet = c_\mathrm{f} W_\bullet^\mathbf{f} + c_\mathrm{m} W_\bullet^\mathbf{m}$ (Taylor 1990; Taylor *et al.* 2007; Gardner *et al.* 2009).

B-4 Selection gradient (Hamilton's rule): preliminary

Pick the locus Ξ that encodes the sex ratio and write ξ for the genic value of an allele of the locus from a juvenile in the focal patch, with $\xi^\mathbf{f}$ for female and $\xi^\mathbf{m}$ for male, respectively. In addition, we denote the breeding value of (i) the juvenile's mother by $\widetilde{\xi}$, (ii) the average breeding value of the adult females in the same patch in the same generation by η_0, (iii) the average breeding value of all adult females in the metapopulation by ζ, and (iv) the average breeding value of the adult females in the same patch τ-generations prior to the present by η_τ for $\tau \geq 1$.

Under vanishingly small genetic variation in the metapopulation, the direction of selection can be assessed by the sign of:

$$\frac{\mathrm{d}W_\bullet}{\mathrm{d}\xi} = c_\mathrm{f}\frac{\mathrm{d}W_\bullet^\mathbf{f}}{\mathrm{d}\xi^\mathbf{f}} + c_\mathrm{m}\frac{\mathrm{d}W_\bullet^\mathbf{m}}{\mathrm{d}\xi^\mathbf{f}}, \tag{S-14}$$

evaluated at $\xi = \xi^\mathbf{f} = \xi^\mathbf{m} = \zeta$, in which c_f (or c_m) is interpreted as the probability that the focal gene lineage is found in a female (or male, respectively). If the above derivative is positive, then selection favours a slightly larger allocation to males.

Further decomposition of the partial derivatives are as follows (Lehmann 2007, 2008; Gardner *et al.* 2009):

$$\begin{aligned}\frac{\mathrm{d}W_{\bullet}^{\mathbf{f}}}{\mathrm{d}\xi^{\mathbf{f}}} &= \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_{\bullet}} \bullet \frac{\mathrm{d}x_{\bullet}}{\mathrm{d}\widetilde{\xi}} \bullet \frac{\mathrm{d}\widetilde{\xi}}{\mathrm{d}\xi^{\mathbf{f}}} + \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_0} \bullet \frac{\mathrm{d}x_0}{\mathrm{d}\eta_0} \bullet \frac{\mathrm{d}\eta_0}{\mathrm{d}\xi^{\mathbf{f}}} + \sum_{\tau=0}^{\infty} \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial \beta_0} \bullet \frac{\partial \beta_0}{\partial x_\tau} \bullet \frac{\mathrm{d}x_\tau}{\mathrm{d}\eta_\tau} \bullet \frac{\mathrm{d}\eta_\tau}{\mathrm{d}\xi^{\mathbf{f}}} \\ &= \left(\frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_{\bullet}} Q_{\bullet}^{\mathbf{f}} + \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_0} Q_0^{\mathbf{f}} + \sum_{\tau=0}^{\infty} \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial \beta_0} \bullet \frac{\partial \beta_0}{\partial x_\tau} Q_\tau^{\mathbf{f}} \right) \sigma,\end{aligned} \tag{S-15}$$

where, assuming vanishing genetic variation, we have rewritten (i) $\mathrm{d}\widetilde{\xi} / \mathrm{d}\xi^{\mathbf{f}} = Q_{\bullet}^{\mathbf{f}}$ for the consanguinity of the focal female to her own daughters, (ii) $\mathrm{d}\eta_0 / \mathrm{d}\xi^{\mathbf{f}} = Q_0^{\mathbf{f}}$ for the consanguinity of the focal female to the juvenile females born to the focal patch in the present generation, (iii) $\mathrm{d}\eta_\tau / \mathrm{d}\xi^{\mathbf{f}} = Q_\tau^{\mathbf{f}}$ for the consanguinity of a random juvenile female and an adult female breeding in the focal patch τ-generations prior to the present (for $\tau \geq 1$), and (iv) σ is an arbitrary constant representing the slope of phenotype on genotype, given by $\mathrm{d}x_{\bullet} / \mathrm{d}\widetilde{\xi} = \mathrm{d}x_0 / \mathrm{d}\eta_0 = \mathrm{d}x_\tau / \mathrm{d}\eta_\tau = \sigma$, which we hereafter rescale $\sigma = 1$.

The similar expansion yields:

$$\begin{aligned}\frac{\mathrm{d}W_{\bullet}^{\mathbf{m}}}{\mathrm{d}\xi^{\mathbf{m}}} &= \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_{\bullet}} \bullet \frac{\mathrm{d}x_{\bullet}}{\mathrm{d}\widetilde{\xi}} \bullet \frac{\mathrm{d}\widetilde{\xi}}{\mathrm{d}\xi^{\mathbf{m}}} + \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_0} \bullet \frac{\mathrm{d}x_0}{\mathrm{d}\eta_0} \bullet \frac{\mathrm{d}\eta_0}{\mathrm{d}\xi^{\mathbf{m}}} + \sum_{\tau=0}^{\infty} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}} \bullet \frac{\partial W_0^{\mathbf{f}}}{\partial \beta_0} \bullet \frac{\partial \beta_0}{\partial x_\tau} \bullet \frac{\mathrm{d}x_\tau}{\mathrm{d}\eta_\tau} \bullet \frac{\mathrm{d}\eta_\tau}{\mathrm{d}\xi^{\mathbf{m}}} + \sum_{\tau=0}^{\infty} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial \beta_0} \bullet \frac{\partial \beta_0}{\partial x_\tau} \bullet \frac{\mathrm{d}x_\tau}{\mathrm{d}\eta_\tau} \bullet \frac{\mathrm{d}\eta_\tau}{\mathrm{d}\xi^{\mathbf{m}}} \\ &= \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_{\bullet}} Q_{\bullet}^{\mathbf{m}} + \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_0} Q_0^{\mathbf{m}} + \sum_{\tau=0}^{\infty} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}} \bullet \frac{\partial W_0^{\mathbf{f}}}{\partial \beta_0} \bullet \frac{\partial \beta_0}{\partial x_\tau} Q_\tau^{\mathbf{m}} + \sum_{\tau=0}^{\infty} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial \beta_0} \bullet \frac{\partial \beta_0}{\partial x_\tau} Q_\tau^{\mathbf{m}},\end{aligned} \tag{S-16}$$

with basically the same notation as above ($^{\mathbf{f}}$ being replaced with $^{\mathbf{m}}$).

From these, the selection gradient (that separates the intra- and trans-generational LRE effects)

reads:

$$
\begin{aligned}
g(x) := c_{\mathrm{f}} \frac{\mathrm{d}W_{\bullet}^{\mathbf{f}}}{\mathrm{d}\xi^{\mathbf{f}}} + c_{\mathrm{m}} \frac{\mathrm{d}W_{\bullet}^{\mathbf{m}}}{\mathrm{d}\xi^{\mathbf{m}}} = & \underbrace{c_{\mathrm{f}} \left(\frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_{\bullet}} Q_{\bullet}^{\mathbf{f}} + \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_0} Q_0^{\mathbf{f}} \right) + c_{\mathrm{m}} \left(\frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_{\bullet}} Q_{\bullet}^{\mathbf{m}} + \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_0} Q_0^{\mathbf{m}} + \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}} \frac{\partial W_0^{\mathbf{f}}}{\partial x_0} Q_0^{\mathbf{m}} \right)}_{=:g_{\emptyset}(x)} \\
& + \underbrace{\left(c_{\mathrm{f}} \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial \beta_0} \frac{\partial \beta_0}{\partial x_0} Q_0^{\mathbf{f}} + c_{\mathrm{m}} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial \beta_0} \frac{\partial \beta_0}{\partial x_0} Q_0^{\mathbf{m}} + c_{\mathrm{m}} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}} \frac{\partial W_0^{\mathbf{f}}}{\partial \beta_0} \frac{\partial \beta_0}{\partial x_0} Q_0^{\mathbf{m}} \right)}_{=:g_{\mathtt{in}}(x)} \\
& + \underbrace{\sum_{\tau=1}^{\infty} \left(c_{\mathrm{f}} \frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial \beta_0} \frac{\partial \beta_0}{\partial x_{\tau}} Q_{\tau}^{\mathbf{f}} + c_{\mathrm{m}} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial \beta_0} \frac{\partial \beta_0}{\partial x_{\tau}} Q_{\tau}^{\mathbf{m}} + c_{\mathrm{m}} \frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}} \frac{\partial W_0^{\mathbf{f}}}{\partial \beta_0} \frac{\partial \beta_0}{\partial x_{\tau}} Q_{\tau}^{\mathbf{m}} \right)}_{=:g_{\mathtt{tr}}(x)},
\end{aligned}
\tag{S-17}
$$

where $g_{\mathtt{in}}(x)$ and $g_{\mathtt{tr}}(x)$ respectively represent the subcomponent of the total selection gradient for intra- and trans-generational LRE, and $g_{\emptyset}(x)$ represents the subcomponent of the selection gradient in the absence of LRE.

B-5 Partial derivatives

Using Eqns (S-12) and (S-13), some algebraic manipulations give the following partial derivatives (each evaluated at neutrality):

$$
\begin{aligned}
\frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_{\bullet}} &= -\frac{1}{1-x}; \\
\frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial x_0} &= \frac{\left(1-d_{\mathrm{f}}\right)^2}{1-x} = \frac{\sigma_{\mathrm{RC}}}{1-x}; \\
\frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_{\bullet}} &= \frac{1}{x}; \\
\frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial x_0} &= \frac{-\left(1-d_{\mathrm{m}}\right)^2}{x} = -\frac{\sigma_{\mathrm{MC}}}{x}; \\
\frac{\partial W_{\bullet}^{\mathbf{f}}}{\partial \beta_0} &= \frac{1-\left(1-d_{\mathrm{f}}\right)^2}{\beta^{\circ}} = \frac{1-\sigma_{\mathrm{RC}}}{\beta^{\circ}}; \\
\frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}} \bullet \frac{\partial W_0^{\mathbf{f}}}{\partial \beta_0} &= \left(1-d_{\mathrm{m}}\right) \bullet \frac{1-\left(1-d_{\mathrm{f}}\right)^2}{\beta^{\circ}} = \frac{\sigma_{\mathrm{MB}}}{\beta^{\circ}}; \\
\frac{\partial W_0^{\mathbf{f}}}{\partial x_0} &= -\frac{1-\left(1-d_{\mathrm{f}}\right)^2}{1-x} = -\frac{1-\sigma_{\mathrm{RC}}}{1-x}; \\
\frac{\partial W_{\bullet}^{\mathbf{m}}}{\partial \beta_0} &= \frac{1-\left(1-d_{\mathrm{m}}\right)^2}{\beta^{\circ}} = \frac{1-\sigma_{\mathrm{MC}}}{\beta^{\circ}},
\end{aligned}
\tag{S-18}
$$

in which we have written (i) $\left(1-d_{\mathrm{f}}\right)^2=\sigma_{\mathrm{RC}}$ for the strength of local resource competition among females, (ii) $\left(1-d_{\mathrm{m}}\right)^2=\sigma_{\mathrm{MC}}$ for the strength of local mate competition among males, and (iii) $\sigma_{\mathrm{MB}}=\left(1-d_{\mathrm{m}}\right)\left(1-\left(1-d_{\mathrm{f}}\right)^2\right)$ for the strength of local mating bonus. Mathematically, these are: (i) the probability that a juvenile female competes for resource with another juvenile female born in the same patch, (ii) the probability that a juvenile male competes for mating opportunity with another juvenile male born in the same patch, and (iii) the joint probability that a male mate with a juvenile female born in the same patch and then she does escapes from competition with a juvenile female born in the same patch.

B-6 Hamilton's rule

If we divide both sides of Eqn (S-17) by $Q_\bullet>0$, we get:

$$\begin{aligned}\frac{g(x)}{Q_\bullet}=&\underbrace{c_{\mathrm{f}}\left(\frac{\partial W_\bullet^{\mathbf{f}}}{\partial x_\bullet}R_\bullet^{\mathbf{f}}+\frac{\partial W_\bullet^{\mathbf{f}}}{\partial x_0}R_0^{\mathbf{f}}\right)+c_{\mathrm{m}}\left(\frac{\partial W_\bullet^{\mathbf{m}}}{\partial x_\bullet}R_\bullet^{\mathbf{m}}+\frac{\partial W_\bullet^{\mathbf{m}}}{\partial x_0}R_0^{\mathbf{m}}+\frac{\partial W_\bullet^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}}\frac{\partial W_0^{\mathbf{f}}}{\partial x_0}R_0^{\mathbf{m}}\right)}_{\text{no LRE}}\\&+\underbrace{\sum_{\tau=0}^{\infty}\left(c_{\mathrm{f}}\frac{\partial W_\bullet^{\mathbf{f}}}{\partial\beta_0}\frac{\partial\beta_0}{\partial x_\tau}R_\tau^{\mathbf{f}}+c_{\mathrm{m}}\frac{\partial W_\bullet^{\mathbf{m}}}{\partial\beta_0}\frac{\partial\beta_0}{\partial x_\tau}R_\tau^{\mathbf{m}}+c_{\mathrm{m}}\frac{\partial W_\bullet^{\mathbf{m}}}{\partial W_0^{\mathbf{f}}}\frac{\partial W_0^{\mathbf{f}}}{\partial\beta_0}\frac{\partial\beta_0}{\partial x_\tau}R_\tau^{\mathbf{m}}\right)}_{\text{LRE effects (intra: }\tau=0\text{; trans: }\tau\geq 1)}.\end{aligned}\tag{S-19}$$

Inserting Eqn (S-18) into Eqn (S-19) without making explicit the relatedness coefficients (as functions of dispersal rates and n), we get:

$$\begin{aligned}\frac{g(x)}{Q_\bullet}=&\,c_{\mathrm{f}}\left(\frac{-1}{1-x}R_\bullet^{\mathbf{f}}+\frac{\sigma_{\mathrm{RC}}}{1-x}R_0^{\mathbf{f}}\right)+c_{\mathrm{m}}\left(\frac{1}{x}R_\bullet^{\mathbf{m}}-\frac{\sigma_{\mathrm{MC}}}{x}R_0^{\mathbf{m}}-\frac{\sigma_{\mathrm{MB}}}{1-x}R_0^{\mathbf{m}}\right)\\&+\sum_{\tau=0}^{\infty}\frac{\partial\beta_0}{\partial x_\tau}\left(c_{\mathrm{f}}\frac{1-\sigma_{\mathrm{RC}}}{\beta^\circ}R_\tau^{\mathbf{f}}+c_{\mathrm{m}}\frac{1-\sigma_{\mathrm{MC}}}{\beta^\circ}R_\tau^{\mathbf{m}}+c_{\mathrm{m}}\frac{\sigma_{\mathrm{MB}}}{\beta^\circ}R_\tau^{\mathbf{m}}\right).\end{aligned}\tag{S-20}$$

B-7 No LRE

When $\alpha=0$ (no LRE), the second line of Eqn (S-20) is null. Then Eqn (S-20) vanishes if:

$$\frac{g(x)}{Q_\bullet}=c_{\mathrm{f}}\left(\frac{-1}{1-x}R_\bullet^{\mathbf{f}}+\frac{\sigma_{\mathrm{RC}}}{1-x}R_0^{\mathbf{f}}\right)+c_{\mathrm{m}}\left(\frac{1}{x}R_\bullet^{\mathbf{m}}-\frac{\sigma_{\mathrm{MC}}}{x}R_0^{\mathbf{m}}-\frac{\sigma_{\mathrm{MB}}}{1-x}R_0^{\mathbf{m}}\right)=0,\tag{S-21}$$

which is equivalent to:

$$\begin{aligned}\frac{g(x)}{Q_\bullet}&=\frac{1}{x(1-x)}\left(c_{\mathrm{f}}\left(-xR_\bullet^{\mathbf{f}}+\sigma_{\mathrm{RC}}xR_0^{\mathbf{f}}\right)+c_{\mathrm{m}}\left((1-x)R_\bullet^{\mathbf{m}}-(1-x)\sigma_{\mathrm{MC}}R_0^{\mathbf{m}}-x\sigma_{\mathrm{MB}}R_0^{\mathbf{m}}\right)\right)\\&=0.\end{aligned}\tag{S-22}$$

Multiplying $x(1-x) > 0$ yields:

$$\begin{aligned} x(1-x)\frac{g(x)}{Q_\bullet} &= c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - \sigma_{\mathrm{MC}}R_0^{\mathbf{m}}\right) + c_{\mathrm{f}}\left(-R_\bullet^{\mathbf{f}} + \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\right)x + c_{\mathrm{m}}\left(-R_\bullet^{\mathbf{m}} + \sigma_{\mathrm{MC}}R_0^{\mathbf{m}} - \sigma_{\mathrm{MB}}R_0^{\mathbf{m}}\right)x \\ &= \left(c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\right) + c_{\mathrm{f}}\left(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\right)\right) \\ &\quad \times \left(\frac{c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - \sigma_{\mathrm{MC}}R_0^{\mathbf{m}}\right)}{c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\right) + c_{\mathrm{f}}\left(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\right)} - x\right) \\ &= 0, \end{aligned} \tag{S-23}$$

which obtains the expression for cESS (candidate ESS), of

$$\widehat{x_\emptyset} = \frac{c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - \sigma_{\mathrm{MC}}R_0^{\mathbf{m}}\right)}{c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\right) + c_{\mathrm{f}}\left(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\right)}, \tag{S-24}$$

which recovers the cESS displayed in Eqns (8) and (9) of the main text. By making the relatedness coefficients explicit with $d_{\mathrm{f}}, d_{\mathrm{m}}$ and n, we can obtain the numerical value for $\widehat{x_\emptyset}$ to reproduce Fig 3 of the main text.

From this manipulation from Eqns (S-21) to (S-23) we can see that it is of great use to scale $g(x)$ by:

$$\begin{aligned} & g(x) > 0 \\ \iff & \underbrace{\frac{x(1-x)g(x)}{Q_\bullet\left(c_{\mathrm{f}}\left(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\right) + c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\right)\right)}}_{=\widehat{x_\emptyset} - x,\ \text{when } \alpha = 0} > 0, \end{aligned} \tag{S-25}$$

as it gives a simple measure of the direction of selection. We therefore define:

$$H(x) := \frac{x(1-x)}{Q_\bullet\left(c_{\mathrm{f}}\left(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\right) + c_{\mathrm{m}}\left(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\right)\right)} \bullet g(x) \tag{S-26}$$

and refer to this as 'scaled Hamilton's rule,' presented in Results of the main text (Eqns 9 – 12).

B-8 Scaled Hamilton's rule for the intra-generational LRE

Substituting Eqns (S-10) and (S-11) into Eqn (S-20) and rescaling it to scaled Hamilton's rule $H(x)$ defined in Eqn (S-26) (but without making the relatedness coefficients explicit (while extracting $\tau = 0$

term on $\sum$) in Eqn (S-20)), we arrive at:

$$H(x) = \widehat{x_\emptyset} - x - x\frac{\alpha(1-x)}{1-\alpha(1-x)} \bullet \underbrace{\frac{c_{\mathrm{f}}R_0^{\mathbf{f}}(1-\sigma_{\mathrm{RC}}) + c_{\mathrm{m}}R_0^{\mathbf{m}}\big(1-(\sigma_{\mathrm{MC}}-\sigma_{\mathrm{MB}})\big)}{c_{\mathrm{f}}\Big(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\Big) + c_{\mathrm{m}}\big(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}}-\sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\big)}}_{=:\kappa} \tag{S-27}$$

and define κ as above.

B-9 Scaled Hamilton's rule for the trans-generational LRE

Similar computation yields:

$$\begin{aligned} H(x) &= \widehat{x_\emptyset} - x + x(1-x)\sum_{\tau=1}^{+\infty} \frac{\Big(c_{\mathrm{f}}(1-\sigma_{\mathrm{RC}})R_\tau^{\mathbf{f}} + c_{\mathrm{m}}(1-\sigma_{\mathrm{MC}}+\sigma_{\mathrm{MB}})R_\tau^{\mathbf{m}}\Big)}{c_{\mathrm{f}}\Big(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\Big) + c_{\mathrm{m}}\big(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}}-\sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\big)} \frac{1}{\beta^\circ} \overbrace{\frac{\partial\beta_0}{\partial x_\tau}}^{\text{Eqn (S-11)}} \\ &= \widehat{x_\emptyset} - x - x\sum_{\tau=1}^{+\infty} \frac{c_{\mathrm{f}}(1-\sigma_{\mathrm{RC}})R_\tau^{\mathbf{f}} + c_{\mathrm{m}}(1-\sigma_{\mathrm{MC}}+\sigma_{\mathrm{MB}})R_\tau^{\mathbf{m}}}{c_{\mathrm{f}}\Big(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}}R_0^{\mathbf{f}}\Big) + c_{\mathrm{m}}\big(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}}-\sigma_{\mathrm{MB}})R_0^{\mathbf{m}}\big)} \Big(\alpha(1-x)\big(1-d_{\mathrm{f}}\big)\Big)^\tau \end{aligned} \tag{S-28}$$

(Lehmann 2007, 2008). So we must obtain recursions for $R_\tau^{\mathbf{f}}, R_\tau^{\mathbf{m}}$ (below).

B-10 Intra-generational consanguinities

We denote the consanguinity between (i) a juvenile female and male sharing the same patch by $Q_0^{\mathbf{fm}}$, (ii) two juvenile males sharing the same patch by $Q_0^{\mathbf{mm}}$, and (iii) two juvenile females sharing the same patch by $Q_0^{\mathbf{ff}}$, each in the same generation $\tau = 0$.

Using the standard coalescent argument for haplodioloids, we get:

$$\begin{aligned} Q_0^{\mathbf{ff}} &= \frac{1}{n}\left(\frac{Q_\bullet + 2\big(1-d_{\mathrm{m}}\big)Q_0^{\mathbf{fm}} + 1}{4}\right) + \frac{n-1}{n}\sigma_{\mathrm{RC}}\left(\frac{Q_0^{\mathbf{ff}} + 2\big(1-d_{\mathrm{m}}\big)Q_0^{\mathbf{fm}} + \big(1-d_{\mathrm{m}}\big)^2 Q_0^{\mathbf{mm}}}{4}\right), \\ Q_0^{\mathbf{fm}} &= \frac{1}{n}\left(\frac{Q_\bullet + \big(1-d_{\mathrm{m}}\big)Q_0^{\mathbf{fm}}}{2}\right) + \frac{n-1}{n}\sigma_{\mathrm{RC}}\left(\frac{Q_0^{\mathbf{ff}} + \big(1-d_{\mathrm{m}}\big)Q_0^{\mathbf{fm}}}{2}\right), \\ Q_0^{\mathbf{mm}} &= \frac{1}{n}Q_\bullet + \frac{n-1}{n}\sigma_{\mathrm{RC}}Q_0^{\mathbf{ff}} \end{aligned} \tag{S-29}$$

(Taylor 1988; Johnstone *et al.* 2012). Solving these to find the unique solution, we can determine each consanguinity coefficient.

The consanguinity of an adult female to

(i) herself is $Q_\bullet = \Big(1 + \big(1-d_{\mathrm{m}}\big)Q_0^{\mathbf{fm}}\Big)/2$,

(ii) her daughter is $Q_\bullet^{\mathbf{f}} := \Big(Q_\bullet + \big(1-d_{\mathrm{m}}\big)Q_0^{\mathbf{fm}}\Big)/2$,

(iii) her son is $Q_\bullet^{\mathbf{m}} := \left(1 + (1 - d_{\mathrm{m}}) Q_0^{\mathbf{fm}}\right)/2$,

(iv) a random juvenile female born in the same patch in the present generation is $Q_0^{\mathbf{f}} := \left(Q_\bullet^{\mathbf{f}} + (1 - d_{\mathrm{m}}) Q_0^{\mathbf{fm}}\right)/n + (1 - d_{\mathrm{f}})^2 (1 - 1/n)\left(Q_0^{\mathbf{ff}} + (1 - d_{\mathrm{m}}) Q_0^{\mathbf{fm}}\right)/2$, which turns out to equal $Q_0^{\mathbf{fm}}$, and

(v) a random juvenile male in the same patch born in the present generation is $Q_0^{\mathbf{m}} := Q_\bullet^{\mathbf{m}}/n + (1 - d_{\mathrm{f}})^2 (1 - 1/n) Q_0^{\mathbf{ff}} = Q_0^{\mathbf{mm}}$.

Also, the relatedness coefficient, from the adult female's perspective (Taylor 1988; Bulmer 1994), to

(i) her daughter is $R_\bullet^{\mathbf{f}} := Q_\bullet^{\mathbf{f}}/Q_\bullet$,

(ii) a random juvenile female on the same patch born in the present generation is $R_0^{\mathbf{f}} := Q_0^{\mathbf{f}}/Q_\bullet$,

(iii) her son is $R_\bullet^{\mathbf{m}} := Q_\bullet^{\mathbf{m}}/Q_\bullet$, and

(iv) a random juvenile male on the same patch born in the present generation is $R_0^{\mathbf{m}} := Q_0^{\mathbf{m}}/Q_\bullet$.

Proof for $\kappa = 1/n$ for $d_{\mathrm{m}} = 0$, Taylor's cancelling principle

We here show that $d_{\mathrm{m}} = 0$ gives $\kappa = 1/n$, which is a part of Taylor's (1992) cancelling principle, as shown in Eqn (11).

First, plugging $d_{\mathrm{m}} = 0$ into Eqn (S-29), we get:

$$Q_0^{\mathbf{ff}} = \frac{1}{n}\left(\frac{Q_\bullet + 2Q_0^{\mathbf{fm}} + 1}{4}\right) + \frac{n-1}{n}\sigma_{\mathrm{RC}}\left(\frac{Q_0^{\mathbf{ff}} + 2Q_0^{\mathbf{fm}} + Q_0^{\mathbf{mm}}}{4}\right), \tag{S-30}$$

$$Q_0^{\mathbf{fm}} = \frac{1}{n}\left(\frac{Q_\bullet + Q_0^{\mathbf{fm}}}{2}\right) + \frac{n-1}{n}\sigma_{\mathrm{RC}}\left(\frac{Q_0^{\mathbf{ff}} + Q_0^{\mathbf{fm}}}{2}\right), \tag{S-31}$$

$$Q_0^{\mathbf{mm}} = \frac{1}{n}Q_\bullet + \frac{n-1}{n}\sigma_{\mathrm{RC}} Q_0^{\mathbf{ff}}, \tag{S-32}$$

with:

$$\begin{aligned} Q_\bullet &= \frac{1 + Q_0^{\mathbf{fm}}}{2}, \\ Q_0^{\mathbf{f}} &= Q_0^{\mathbf{fm}}, \\ Q_0^{\mathbf{m}} &= Q_0^{\mathbf{mm}}, \\ Q_\bullet^{\mathbf{f}} &= \frac{Q_\bullet + Q_0^{\mathbf{fm}}}{2}, \\ Q_\bullet^{\mathbf{m}} &= \frac{1 + Q_0^{\mathbf{fm}}}{2}. \end{aligned} \tag{S-33}$$

Second, computing $Q_0^{\mathbf{ff}} - \left(Q_0^{\mathbf{fm}} + Q_0^{\mathbf{mm}}\right)/2$ from Eqns (S-30) to (S-32) (separating LHS and RHS, while

using Eqn (S-33)) yields:

$$Q_0^{\mathbf{ff}} - \frac{Q_0^{\mathbf{fm}} + Q_0^{\mathbf{mm}}}{2} = -\sigma_{\mathrm{RC}} \frac{n-1}{n} \left(Q_0^{\mathbf{ff}} - \frac{Q_0^{\mathbf{fm}} + Q_0^{\mathbf{mm}}}{2} \right), \tag{S-34}$$

thus implying

$$Q_0^{\mathbf{ff}} = \frac{Q_0^{\mathbf{fm}} + Q_0^{\mathbf{mm}}}{2} \tag{S-35}$$

at equilibrium.

Third, let us repeat the definition of κ (for any d_{m}):

$$\kappa = \frac{c_{\mathrm{f}} R_0^{\mathbf{f}} (1 - \sigma_{\mathrm{RC}}) + c_{\mathrm{m}} R_0^{\mathbf{m}} \left(1 - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}})\right)}{c_{\mathrm{f}} \left(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}} R_0^{\mathbf{f}} \right) + c_{\mathrm{m}} \left(R_\bullet^{\mathbf{m}} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}}) R_0^{\mathbf{m}} \right)}, \tag{S-36}$$

into which we substitute $d_{\mathrm{m}} = 0$ to get:

$$\begin{aligned} \kappa \Big|_{d_{\mathrm{m}}=0} &= \frac{c_{\mathrm{f}} R_0^{\mathbf{f}} (1 - \sigma_{\mathrm{RC}}) + c_{\mathrm{m}} R_0^{\mathbf{m}} (1 - \sigma_{\mathrm{RC}})}{c_{\mathrm{f}} \left(R_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}} R_0^{\mathbf{f}} \right) + c_{\mathrm{m}} \left(R_\bullet^{\mathbf{m}} - \sigma_{\mathrm{RC}} R_0^{\mathbf{m}} \right)} \\ &= \frac{c_{\mathrm{f}} Q_0^{\mathbf{f}} (1 - \sigma_{\mathrm{RC}}) + c_{\mathrm{m}} Q_0^{\mathbf{m}} (1 - \sigma_{\mathrm{RC}})}{c_{\mathrm{f}} \left(Q_\bullet^{\mathbf{f}} - \sigma_{\mathrm{RC}} Q_0^{\mathbf{f}} \right) + c_{\mathrm{m}} \left(Q_\bullet^{\mathbf{m}} - \sigma_{\mathrm{RC}} Q_0^{\mathbf{m}} \right)} \end{aligned} \tag{S-37}$$

(where we used $\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}} \equiv \sigma_{\mathrm{RC}}$ when $d_{\mathrm{m}} = 0$). If we, for now, define $\rho = \left(c_{\mathrm{f}} Q_0^{\mathbf{mm}} + c_{\mathrm{m}} Q_0^{\mathbf{ff}} \right) / \left(c_{\mathrm{f}} Q_\bullet^{\mathbf{f}} + c_{\mathrm{m}} Q_\bullet^{\mathbf{m}} \right)$, then:

$$\kappa \Big|_{d_{\mathrm{m}}=0} = \frac{\rho (1 - \sigma_{\mathrm{RC}})}{1 - \rho \sigma_{\mathrm{RC}}}. \tag{S-38}$$

$c_{\mathrm{f}}\times$Eqn (S-31) plus $c_{\mathrm{m}}\times$Eqn (S-32) gives:

$$c_{\mathrm{f}} Q_0^{\mathbf{ff}} + c_{\mathrm{m}} Q_0^{\mathbf{mm}} = \frac{1}{n} \left(c_{\mathrm{f}} Q_\bullet^{\mathbf{f}} + c_{\mathrm{m}} Q_\bullet^{\mathbf{m}} \right) + \frac{n-1}{n} \sigma_{\mathrm{RC}} \underbrace{\left(c_{\mathrm{f}} Q_0^{\mathbf{ff}} + c_{\mathrm{m}} Q_0^{\mathbf{mm}} \right)}_{=(\star)} \tag{S-39}$$

(where $(\star)$ is equal to $c_{\mathrm{f}} Q_0^{\mathbf{ff}} + c_{\mathrm{m}} Q_0^{\mathbf{mm}}$ after algebraic manipulations), which, using ρ, implies:

$$\rho = \frac{1}{n} + \frac{n-1}{n} \sigma_{\mathrm{RC}} \rho, \tag{S-40}$$

which by solving for $1/n$ gives:

$$\frac{1}{n} = \frac{\rho (1 - \sigma_{\mathrm{RC}})}{1 - \rho \sigma_{\mathrm{RC}}}, \tag{S-41}$$

which is equal to Eqn (S-38), therefore $\kappa = 1/n$ when $d_{\mathrm{m}} = 0$. This gives the result for $d_{\mathrm{m}} = 0$ of the main text (including Eqn 11).

B-11 Trans-generational consanguinities

We focus on haplodiploid genetics, and designate $Q_\tau^{\mathbf{f}}$ (or $Q_\tau^{\mathbf{m}}$) for the probability of consanguinity between (i) a juvenile female (or male) randomly sampled right after its birth in the present generation in a patch and (ii) a random adult female in the same patch in the τ-th generation ($\tau \geq 0$). To develop the recursion between $\mathbf{q}_{\tau+1} = \left(Q_{\tau+1}^{\mathbf{f}}, Q_{\tau+1}^{\mathbf{m}}\right)^\top$ and $\mathbf{q}_\tau = \left(Q_\tau^{\mathbf{f}}, Q_\tau^{\mathbf{m}}\right)^\top$ (with $^\top$ for transpose), we sample three classes of individuals living in the same patch but potentially in different time epochs: (i) adult females in the $\tau+1$-st generation (representatively named $\mathcal{A}_{\tau+1}$), (ii) juvenile females or males born from the adult females living in the generation $\tau = 1$ (named $\mathcal{J}_1^{\mathbf{f}}$ for female and $\mathcal{J}_1^{\mathbf{m}}$ for male), and (iii) juvenile females or males born from the adult females in the current generation $\tau = 0$ (named $\mathcal{J}_0^{\mathbf{f}}$ for female and $\mathcal{J}_0^{\mathbf{m}}$ for male). The consanguinity between (i) and (ii) is by definition $\mathbf{q}_\tau$ and that between (i) and (iii) is $\mathbf{q}_{\tau+1}$. The initial condition ($\tau = 0$) reads $Q_0^{\mathbf{f}} = Q_0^{\mathbf{fm}}$ (which is the consanguinity of a random juvenile female born in $\tau = 0$ and a random adult reproducing in the same generation) and $Q_0^{\mathbf{m}} = Q_0^{\mathbf{mm}}$ (which is the consanguinity of a random juvenile male born in $\tau = 0$ and a random adult reproducing in the same generation). We depict a conceptual illustration in SI Fig 2, in which actors' and recipients' perspectives (or inclusive-fitness and neighbor-modulated fitness approaches) do agree, as in Taylor *et al.* (2007).

Recursion for $Q_{\tau+1}^{\mathbf{f}}$

First consider the probability that $\mathcal{A}_{\tau+1}$ (who is a random adult female sampled from a patch in the $\tau+1$-st generation) and $\mathcal{J}_0^{\mathbf{f}}$ (who is a random juvenile female born from an adult female in the same patch in the present generation) share an allele of identity-by-descent (IBD). For this to occur, it entails that $\mathcal{J}_0^{\mathbf{f}}$'s mother be of philopatric origin ($1 - d_{\mathrm{f}}$); given this, $\mathcal{J}_0^{\mathbf{f}}$'s allele derives maternally with a probability of 1/2, in which case IBD probability is $Q_\tau^{\mathbf{f}}$; otherwise, it derives from a father (1/2) born in the same patch ($1 - d_{\mathrm{m}}$), in which case IBD probability is $Q_\tau^{\mathbf{m}}$. Hence,

$$Q_{\tau+1}^{\mathbf{f}} = \frac{1}{2}\left(1 - d_{\mathrm{f}}\right) Q_\tau^{\mathbf{f}} + \frac{1}{2}\left(1 - d_{\mathrm{f}}\right)\left(1 - d_{\mathrm{m}}\right) Q_\tau^{\mathbf{m}}. \tag{S-42}$$

Recursion for $Q_{\tau+1}^{\mathbf{m}}$

Similarly consider the probability that $\mathcal{A}_{\tau+1}$ (a random adult female sampled from a patch in the $\tau+1$-st generation) and $\mathcal{J}_0^{\mathbf{m}}$ (a random juvenile male born to the adult females in the same patch in the present generation) share an allele of IBD. Given that $\mathcal{J}_0^{\mathbf{m}}$'s mother be of philopatric origin (which occurs with a probability of $1 - d_{\mathrm{f}}$), $\mathcal{J}_0^{\mathbf{m}}$'s allele derives certainly maternally (probability 1), in which case IBD probability is $Q_\tau^{\mathbf{f}}$, yielding:

$$Q_{\tau+1}^{\mathbf{m}} = \left(1 - d_{\mathrm{f}}\right) Q_\tau^{\mathbf{f}}. \tag{S-43}$$

Vector form of the recursions

In a vector form, Eqn (S-42) and Eqn (S-43), divided by $Q_\bullet$ read:

$$\begin{aligned}\begin{pmatrix} R^{\mathbf{f}}_{\tau+1} \\ R^{\mathbf{m}}_{\tau+1} \end{pmatrix} &= \underbrace{\begin{pmatrix} 1-d_{\mathrm{f}} & 0 \\ 0 & 1-d_{\mathrm{f}} \end{pmatrix}}_{\text{female dispersal}} \underbrace{\begin{pmatrix} \pi_{\mathrm{ff}} & \pi_{\mathrm{fm}} \\ \pi_{\mathrm{mf}} & \pi_{\mathrm{mm}} \end{pmatrix}}_{\text{inheritance}} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1-d_{\mathrm{m}} \end{pmatrix}}_{\text{male dispersal}} \begin{pmatrix} R^{\mathbf{f}}_{\tau} \\ R^{\mathbf{m}}_{\tau} \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} \frac{1-d_{\mathrm{f}}}{2} & \frac{(1-d_{\mathrm{f}})(1-d_{\mathrm{m}})}{2} \\ 1-d_{\mathrm{f}} & 0 \end{pmatrix}}_{=\mathbf{P}} \begin{pmatrix} R^{\mathbf{f}}_{\tau} \\ R^{\mathbf{m}}_{\tau} \end{pmatrix},\end{aligned} \tag{S-44}$$

where π_{XY} represents the probability that a juvenile of sex X derives its gene from an adult of sex Y. πs are determined by ploidy levels: πs are all $1/2$ for haploids and diploids, while $\pi_{\mathrm{ff}} = \pi_{\mathrm{fm}} = 1/2, \pi_{\mathrm{mf}} = 1$ and $\pi_{\mathrm{mm}} = 0$ for haplodiploids.

In Eqn (S-44), the first matrix determines the decay of genetic relatedness due to female emigration, in which mated females disperse and therefore males' gametes may emigrate together (the bottom-right element should not be 1); the second describes the genetic inheritance, or movement of genes between sexes due to mating; the third represents the male dispersal, by which juvenile females' gametes do not disperse.

Now let $\mathbf{V} := \alpha(1-d_{\mathrm{f}})(1-x)\mathbf{P}$; then its spectral radius is lower than unity in modulus and therefore $(\mathbf{I}-\mathbf{V})^{-1}\mathbf{V} = \sum_{\tau=1}^{\infty} \mathbf{V}^{\tau}$ does exist (with $\mathbf{I}$ the identity). Then, Eqn (S-28) becomes:

$$\begin{aligned} H(x) &= \widehat{x_{\emptyset}} - x - x\left(u_{\mathrm{f}}, u_{\mathrm{m}}\right)^{\top} \sum_{\tau=1}^{+\infty} \frac{\left(\alpha(1-d_{\mathrm{f}})(1-x)\right)^{\tau}}{c_{\mathrm{f}}\left(R^{\mathbf{f}}_{\bullet} - \sigma_{\mathrm{RC}} R^{\mathbf{m}}_{\bullet}\right) + c_{\mathrm{m}}\left(R^{\mathbf{m}}_{\bullet} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}}) R^{\mathbf{m}}_{\bullet}\right)} \begin{pmatrix} R^{\mathbf{f}}_{\tau} \\ R^{\mathbf{m}}_{\tau} \end{pmatrix} \\ &= \widehat{x_{\emptyset}} - x - x\left(u_{\mathrm{f}}, u_{\mathrm{m}}\right)^{\top} \frac{1}{c_{\mathrm{f}}\left(R^{\mathbf{f}}_{\bullet} - \sigma_{\mathrm{RC}} R^{\mathbf{m}}_{\bullet}\right) + c_{\mathrm{m}}\left(R^{\mathbf{m}}_{\bullet} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}}) R^{\mathbf{m}}_{\bullet}\right)} \sum_{\tau=1}^{+\infty} \mathbf{V}^{\tau} \begin{pmatrix} R^{\mathbf{f}}_{0} \\ R^{\mathbf{m}}_{0} \end{pmatrix} \\ &= \widehat{x_{\emptyset}} - x - x\left(u_{\mathrm{f}}, u_{\mathrm{m}}\right)^{\top} \frac{1}{c_{\mathrm{f}}\left(R^{\mathbf{f}}_{\bullet} - \sigma_{\mathrm{RC}} R^{\mathbf{m}}_{\bullet}\right) + c_{\mathrm{m}}\left(R^{\mathbf{m}}_{\bullet} - (\sigma_{\mathrm{MC}} - \sigma_{\mathrm{MB}}) R^{\mathbf{m}}_{\bullet}\right)} (\mathbf{I}-\mathbf{V})^{-1}\mathbf{V} \begin{pmatrix} R^{\mathbf{f}}_{0} \\ R^{\mathbf{m}}_{0} \end{pmatrix}. \end{aligned} \tag{S-45}$$

As any nonsingular bidimensional square matrix has a formula for its inverse matrix (one may want to use *Mathematica*® notebook file we submitted), we can find that the trans-generational subcomponent, $g_{\mathtt{tr}}(x)$, is given by

$$-\frac{}{2} \bullet \frac{c_{\mathrm{f}}(1-\sigma_{\mathrm{RC}})\left(\left(1+(1-d_{\mathrm{m}})\right)R^{\mathbf{f}}_{0} + (1-d_{\mathrm{m}})R^{\mathbf{m}}_{0}\right) + c_{\mathrm{m}}(1-\sigma_{\mathrm{m}})\left(2R^{\mathbf{f}}_{0} + (1-d_{\mathrm{m}})R^{\mathbf{m}}_{0}\right)}{\left(1 - \frac{}{2}\left(1+(1-d_{\mathrm{m}})\right)\right)} \tag{S-46}$$

for haplodiploids, where $\sigma_{\mathrm{m}} := \left(1-d_{\mathrm{m}}\right)^2 - \left(1-d_{\mathrm{m}}\right)\left(1-\left(1-d_{\mathrm{f}}\right)^2\right)$ and $:= \sigma_{\mathrm{RC}}\alpha(1-x)$.

C Robustness under $d_{\mathrm{m}} = 0$

C-1 Different functional form: additive

Intra-generational model

Suppose the fecundity of individuals in the present generation is given by:

$$\beta_0 = \frac{K}{1 + A\Big(\left(1-d_{\mathrm{f}}\right)x_0 + d_{\mathrm{f}}x\Big)} \tag{S-47}$$

where A represents the strength of LRE, and we therefore have:

$$\frac{\partial \beta_0}{\partial x_0} = -\beta^{\circ}\frac{\left(1-d_{\mathrm{f}}\right)A}{1+Ax}. \tag{S-48}$$

Replacing α in the main text with $A = \alpha \,/\, (1-\alpha)$ gives the corresponding analyses; therefore, the present extension is analogous to the analysis given in the main text.

Trans-generational model

Suppose that β obeys the following recursion:

$$\beta_{\tau-1} = K - A\Big(\left(1-d_{\mathrm{f}}\right)\beta_{\tau}x_{\tau} + d_{\mathrm{f}}\bar{\beta}x\Big), \tag{S-49}$$

with:

$$\begin{aligned} \beta^{\circ} = \bar{\beta} &= \frac{K}{(1+Ax)}, \\ \frac{\partial \beta_{\tau-1}}{\partial \beta_{\tau}} &= -A\left(1-d_{\mathrm{f}}\right)x, \\ \frac{\partial \beta_{\tau-1}}{\partial x_{\tau}} &= -A\left(1-d_{\mathrm{f}}\right)\beta^{\circ}, \end{aligned} \tag{S-50}$$

which constitutes a negative niche-construction (fecundities of the present and previous generations are negatively correlated), or aggressiveness towards offspring (of any sex) from males. In this case, a straightforward calculation similar to the analyses performed above yields:

$$g_{\mathtt{tr}}(x) = -(1-\sigma_{\mathrm{RC}})\Big(c_{\mathrm{f}}Q_0^{\mathbf{f}} + c_{\mathrm{m}}Q_0^{\mathbf{m}}\Big)\frac{A\sigma_{\mathrm{RC}}}{1+A\sigma_{\mathrm{RC}}x}. \tag{S-51}$$

Rewriting $A = \alpha \,/\, (1 - \alpha\sigma_{\mathrm{RC}})$, we have:

$$g_{\mathtt{tr}}(x) = -(1-\sigma_{\mathrm{RC}})\left(c_{\mathrm{f}}Q_0^{\mathbf{f}} + c_{\mathrm{m}}Q_0^{\mathbf{m}}\right)\frac{\alpha\sigma_{\mathrm{RC}}}{1-\alpha\sigma_{\mathrm{RC}}+\alpha\sigma_{\mathrm{RC}}x}, \tag{S-52}$$

which is equivalent to the one obtained in the original, trans-generational model, with a difference in feasibility conditions: the original α varies from 0 to 1 while A may vary from 0 to $1 \,/\, (1-\sigma_{\mathrm{RC}}) > 1$, suggesting that the upper bound for the strength of LRE depends on the dispersal rate. Therefore, our results are robust against this simple recursion for β_τ, and the original approach presented in the main text may be suitable for experimental confirmation.

C-2 Different functional form: multiplicative

Intra-generational model

Here we assume that LRE is generated by a multiplicative functional form given by:

$$\beta_0 = (1-x_0)^{\alpha}K^{\alpha}\beta_0^{\alpha} \tag{S-53}$$

(for $0 \leq \alpha < 1$), which satisfies:

$$\begin{aligned} \beta^{\circ} &= (1-x)^{\frac{\alpha}{1-\alpha}}K^{\frac{\alpha}{1-\alpha}}, \\ \frac{\partial\beta_0}{\partial x_0} &= -\frac{\beta^{\circ}}{1-x}\bullet\frac{\alpha}{1-\alpha} \end{aligned} \tag{S-54}$$

at neutrality. Therefore,

$$g_{\mathtt{in}}(x) = -\left(c_{\mathrm{f}}Q_0^{\mathbf{f}} + c_{\mathrm{m}}Q_0^{\mathbf{m}}\right)(1-\sigma_{\mathrm{RC}})\frac{1}{1-x}\bullet\frac{\alpha}{1-\alpha}, \tag{S-55}$$

telling us that the effect of LRE is independent of the dispersal rate. Hamilton's rule is thus given by:

$$x_{\emptyset}^{\mathrm{LMC}} - x - \frac{1}{n}\bullet x\bullet\frac{\alpha}{1-\alpha} > 0, \tag{S-56}$$

which determines the cESS as:

$$\widehat{x_{\mathtt{in}}} = \frac{x_{\emptyset}^{\mathrm{LMC}}}{1+\dfrac{\alpha}{(1-\alpha)n}}. \tag{S-57}$$

This expression immediately tells us that $\alpha \to 1$ leads to $\widehat{x_{\mathtt{in}}} \to 0$.

Trans-generational model

Similarly, consider:

$$\begin{aligned}\beta_\tau &= B(\beta_{\tau+1}, x_{\tau+1}) \\ &= \Big(\big(1-d_{\mathrm{f}}\big)(1-x_{\tau+1})\beta_{\tau+1} + d_{\mathrm{f}}(1-x)\overline{\beta}\Big)^{\alpha} K^{\alpha},\end{aligned} \tag{S-58}$$

which, at equilibrium, should satisfy:

$$\beta^{\circ} = (1-x)^{\frac{\alpha}{1-\alpha}} K^{\frac{\alpha}{1-\alpha}}. \tag{S-59}$$

Partial differentiation gives:

$$\begin{aligned}\frac{\partial \beta_0}{\partial x_1} &= -\alpha \frac{1-d_{\mathrm{f}}}{1-x}\beta^{\circ}, \\ \frac{\partial \beta_\tau}{\partial \beta_{\tau+1}} &= \alpha\big(1-d_{\mathrm{f}}\big).\end{aligned} \tag{S-60}$$

Hence,

$$g_{\mathtt{tr}}(x) = -\sum_{\tau=1}^{\infty}\Big(c_{\mathrm{f}}Q_0^{\mathbf{f}} + c_{\mathrm{m}}Q_0^{\mathbf{m}}\Big)(1-\sigma_{\mathrm{RC}}) \bullet \frac{(\alpha\sigma_{\mathrm{RC}})^{\tau}}{1-x}, \tag{S-61}$$

which, for $0 \leq \alpha < 1$, converges to:

$$-\Big(c_{\mathrm{f}}Q_0^{\mathbf{f}} + c_{\mathrm{m}}Q_0^{\mathbf{m}}\Big)(1-\sigma_{\mathrm{RC}})\frac{1}{1-x} \bullet \frac{\alpha\sigma_{\mathrm{RC}}}{1-\alpha\sigma_{\mathrm{RC}}}. \tag{S-62}$$

Therefore, Hamilton's rule is given by:

$$x_{\emptyset}^{\mathrm{LMC}} - x - \frac{1}{n} \bullet \frac{\alpha\sigma_{\mathrm{RC}}}{1-\alpha\sigma_{\mathrm{RC}}}x > 0, \tag{S-63}$$

supplying the cESS, of:

$$\widehat{x_{\mathtt{tr}}} = \frac{x_{\emptyset}^{\mathrm{LMC}}}{1 + \frac{1}{n} \bullet \frac{\alpha\sigma_{\mathrm{RC}}}{1-\alpha\sigma_{\mathrm{RC}}}}. \tag{S-64}$$

Overall, we can observe that the effects of LRE on $\widehat{x}$ are much more pronounced compared to the additive one (SI Fig 3), which is because with this multiplicative formula $x = 1$ gives $\beta = 0$: producing females is a prerequisite for producing offspring. Though much simpler in any associated expressions, the model may overestimate the LRE effect. For this reason, we presented the additive formula for LRE in the main text.

C-3 LRE from males: additive

We here demonstrate a minimal analysis for LRE supplied by juvenile males.

Intra-generational model

Suppose that the fecundity of the females in the focal patch (with its average phenotype x_0) is given by:

$$\beta_0 = K + \alpha x_0 \beta_0 = \frac{K}{1 - \alpha x_0}, \tag{S-65}$$

which means that:

$$\begin{aligned} \beta^\circ &= \frac{K}{1 - \alpha x}, \\ \frac{\partial \beta_0}{\partial x_0} &= \beta^\circ \frac{\alpha}{1 - \alpha x} \end{aligned} \tag{S-66}$$

at neutrality. The preceding logic shows that Hamilton's rule is given by:

$$x_\emptyset^{\mathrm{LMC}} - x + \frac{1}{n} \bullet \frac{\alpha x}{1 - \alpha x}(1 - x) > 0, \tag{S-67}$$

which gives:

$$\widehat{x_{\mathtt{in}}} = \frac{x_\emptyset^{\mathrm{LMC}} + \frac{1}{n} \bullet \frac{\alpha \widehat{x_{\mathtt{in}}}}{1 - \alpha \widehat{x_{\mathtt{in}}}}}{1 + \frac{1}{n} \bullet \frac{\alpha \widehat{x_{\mathtt{in}}}}{1 - \alpha \widehat{x_{\mathtt{in}}}}}, \tag{S-68}$$

from which, by the intermediate value theorem, we can find a unique $\widehat{x_{\mathtt{in}}}$ lying between 0 and 1.

Trans-generational model

We use the function of the form:

$$\beta_0 = K + \alpha \Big((1 - d_{\mathrm{f}}) x_1 \beta_1 + d_{\mathrm{f}} x \overline{\beta} \Big), \tag{S-69}$$

which has the equilibrium solution given by:

$$\beta^\circ = \overline{\beta} = \frac{K}{1 - \alpha x} \tag{S-70}$$

at neutrality, and satisfies:

$$\begin{aligned} \frac{\partial \beta_0}{\partial x_1} &= \alpha \beta^\circ (1 - d_{\mathrm{f}}), \\ \frac{\partial \beta_0}{\partial \beta_1} &= \alpha x (1 - d_{\mathrm{f}}), \end{aligned} \tag{S-71}$$

once again at neutrality. Therefore we can obtain Hamilton's rule of:

$$x_{\emptyset}^{\mathrm{LMC}} - x + \frac{1}{n} \bullet \frac{\alpha x \sigma_{\mathrm{RC}}}{1 - \alpha x \sigma_{\mathrm{RC}}}(1 - x) > 0, \tag{S-72}$$

which gives the cESS $\widehat{x_{\mathtt{tr}}}$ as $\widehat{x_{\mathtt{in}}}|_{\alpha \to \alpha\sigma_{\mathrm{RC}}}$ in Eqn (S-68).

C-4 LRE from males: multiplicative

Intra-generational model

Suppose the fecundity function of the form:

$$\beta_0 = x_0^{\alpha} K^{\alpha} \beta_0^{\alpha} = (x_0 K)^{\frac{\alpha}{1-\alpha}}, \tag{S-73}$$

with:

$$\frac{\partial \beta_0}{\partial x_0} = \beta^{\circ} \frac{\alpha}{1 - \alpha} \bullet \frac{1}{x}, \tag{S-74}$$

implying that Hamilton's rule reads:

$$x_{\emptyset}^{\mathrm{LMC}} - x + \frac{1}{n} \bullet \frac{\alpha}{1 - \alpha}(1 - x) > 0, \tag{S-75}$$

which gives:

$$\widehat{x_{\mathtt{in}}} = \frac{x_{\emptyset}^{\mathrm{LMC}} + \frac{1}{n}\frac{\alpha}{1-\alpha}}{1 + \frac{1}{n}\frac{\alpha}{1-\alpha}} \tag{S-76}$$

Trans-generational model

The same logic supplies Hamilton's rule as:

$$x_{\emptyset}^{\mathrm{LMC}} - x + \frac{1}{n} \bullet \frac{\alpha \sigma_{\mathrm{RC}}}{1 - \alpha \sigma_{\mathrm{RC}}}(1 - x) > 0, \tag{S-77}$$

with the resulting cESS of:

$$\widehat{x_{\mathtt{tr}}} = \frac{x_{\emptyset}^{\mathrm{LMC}} + \frac{1}{n}\frac{\alpha\sigma_{\mathrm{RC}}}{1-\alpha\sigma_{\mathrm{RC}}}}{1 + \frac{1}{n}\frac{\alpha\sigma_{\mathrm{RC}}}{1-\alpha\sigma_{\mathrm{RC}}}}. \tag{S-78}$$

We here remark that the trans-generational LRE from juvenile males is unlikely, because males do not disperse and therefore have no access to post-dispersal, founding females. One may want to apply an alternative model to specific organisms with a specific lifecycle.

Combining additivity and multiplicativity: Combinational LRE

We here consider

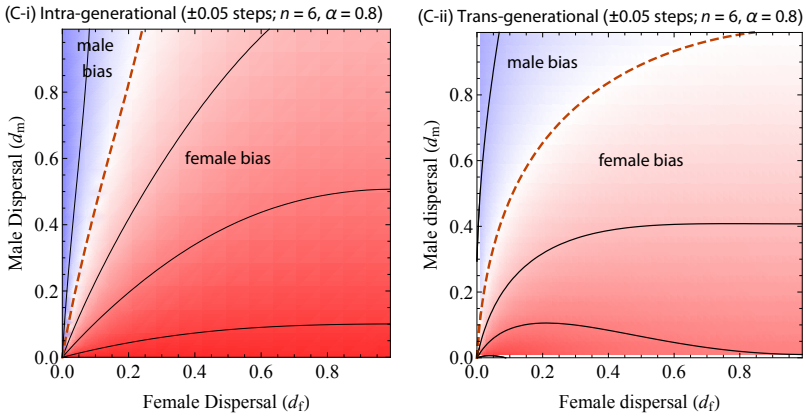
$$\beta_0 = K + \alpha(1 - x_0)^{\theta}\beta_0, \tag{S-79}$$

or

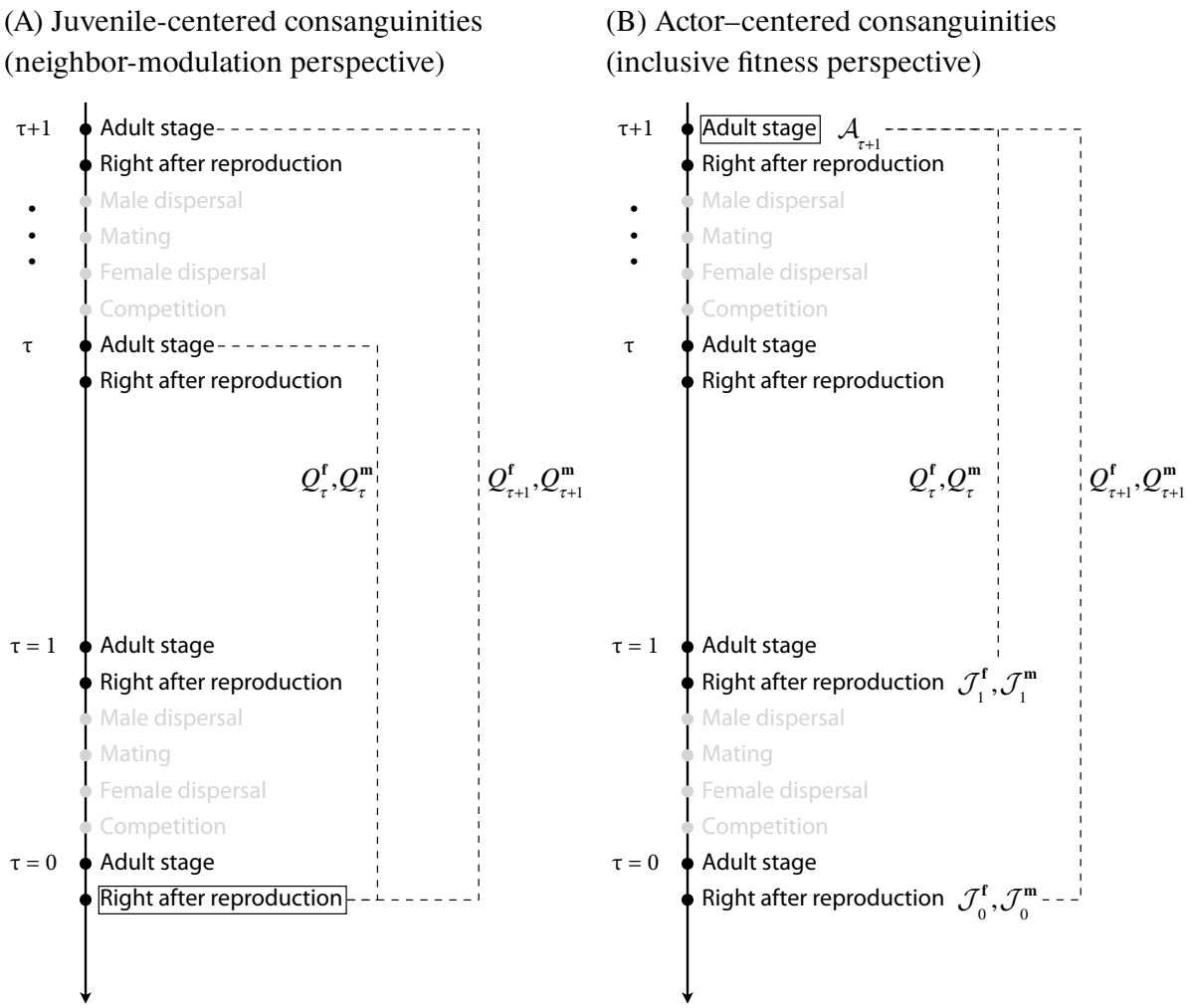
$$\beta_0 = \frac{K}{1 - \alpha(1 - x_0)^{\theta}}. \tag{S-80}$$

for $0 < \theta < 1$, and the resulting cESS does not have analytical formula. The numerical solutions are given in SI Fig 6.

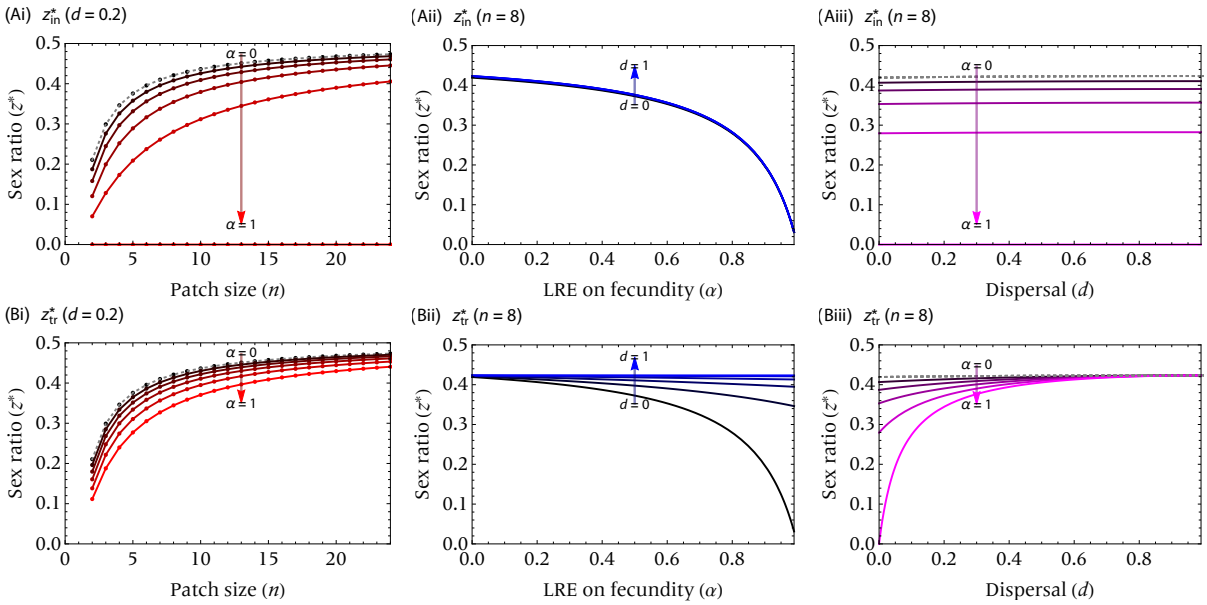
Inserting $d_{\mathrm{m}} = 0$ or $d_{\mathrm{m}} = 1$ into the above equation and then rescaling the whole equation (**??**) leads to the expression given in the main text.



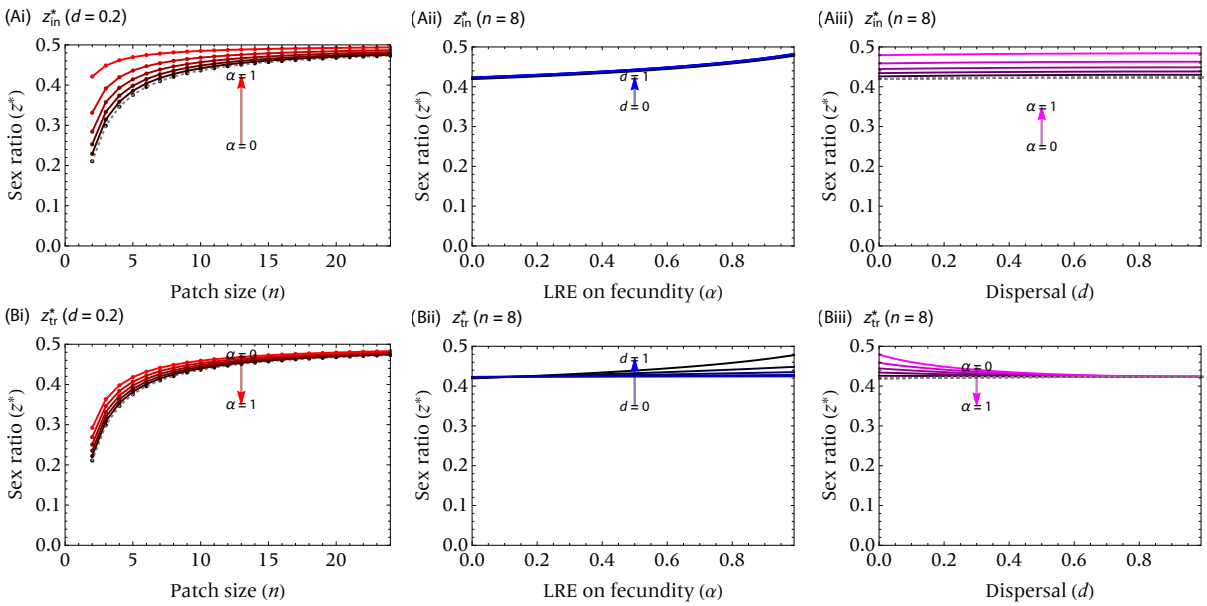
SI Figure 1: Evolutionary outcomes plotted against sex-dependent dispersal rates. Orange dotted contour: Fisherian sex ratio ($x = .5$), and the others are ± 0.05 steps deviation from Fisherian.



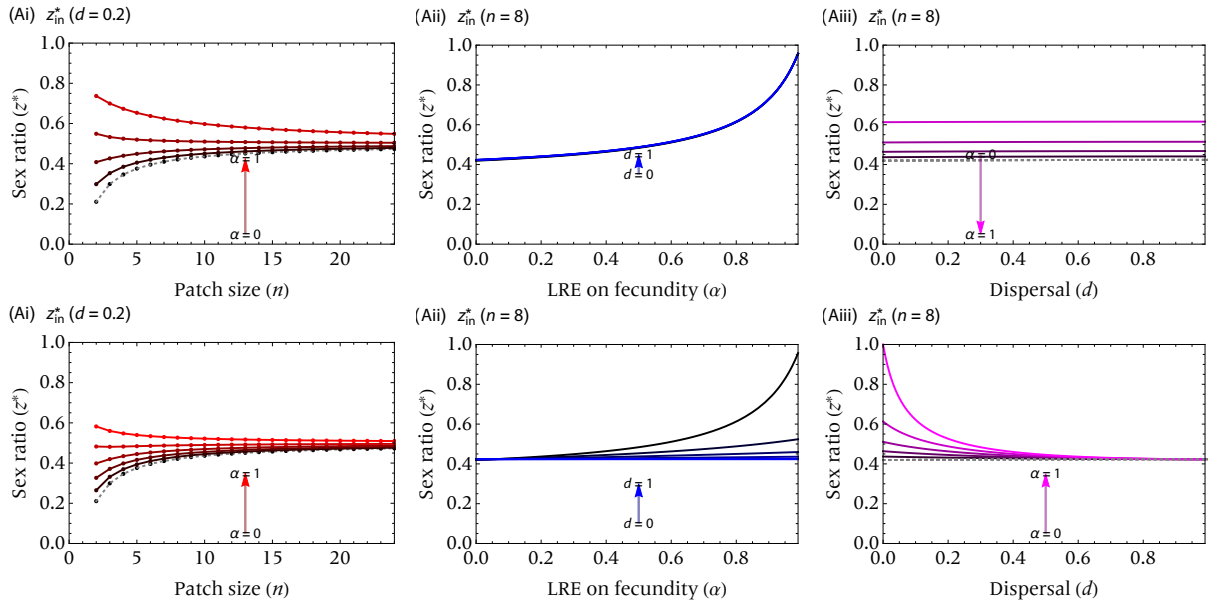
SI Figure 2: Schematic illustration to develop recursive equations for the trans-generational consanguinity



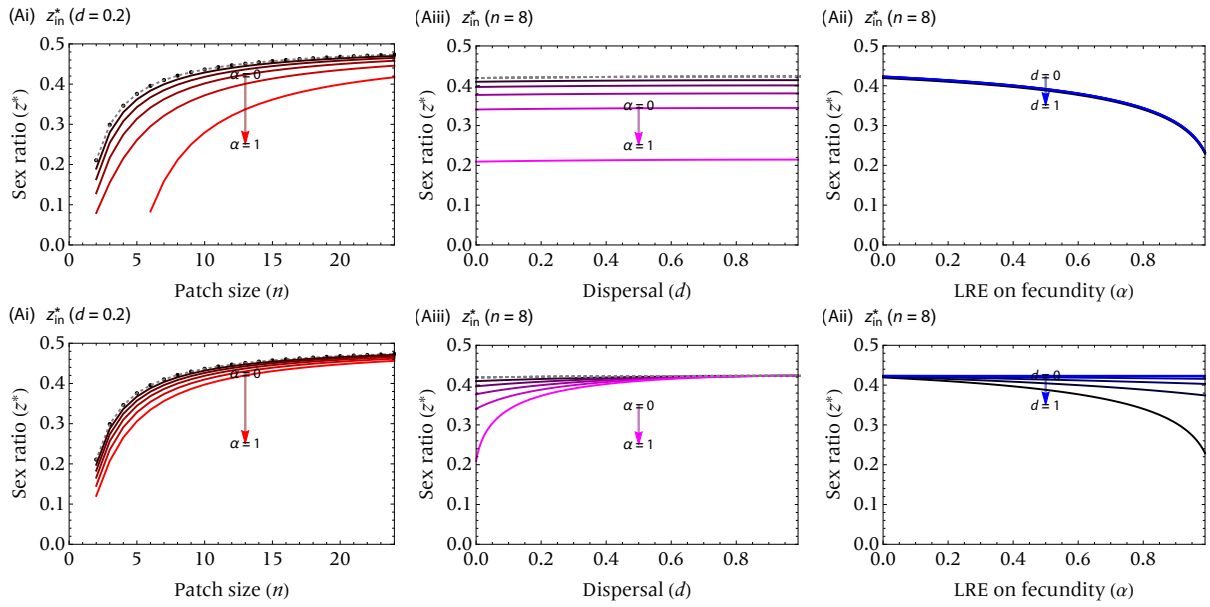
SI Figure 3: Multiplicative female-LRE.



SI Figure 4: Additive male-LRE.



SI Figure 5: Multiplicative male-LRE. Note the difference from the previous graphic figures in the ordinate range (0 to 1).



SI Figure 6: Combinational female-LRE. $\theta = \alpha = 0.5$. Other parameters as shown.

Literatures cited in Appendices

Bulmer, M. G. (1994). *Evolutionary theoretical ecology*. Sinauer Associates, Sunderland, MA.

Caswell, H. (2001). *Matrix population models*. Wiley Online Library.

Gardner, A., Arce, A., & Alpedrinha, J. (2009). Budding dispersal and the sex ratio. *Journal of Evolutionary Biology*, **22**.5, pp. 1036–1045. DOI: 10.1111/j.1420-9101.2009.01719.x.

Johnstone, R. A., Cant, M. A., & Field, J. (2012). Sex-biased dispersal, haplodiploidy and the evolution of helping in social insects. *Proceedings of the Royal Society B: Biological Sciences*, **279**.1729, pp. 787–793. DOI: 10.1098/rspb.2011.1257.

Lehmann, L. (2007). The evolution of trans-generational altruism: kin selection meets niche construction. *Journal of Evolutionary Biology*, **20**.1, pp. 181–189. DOI: 10.1111/j.1420-9101.2006.01202.x.

— (2008). The adaptive dynamics of niche constructing traits in spatially subdivided populations: evolving posthumous extended phenotypes. *Evolution*, **62**.3, pp. 549–566. DOI: 10.1111/j.1558-5646.2007.00291.x.

Taylor, P. D. (1988). Inclusive fitness models with two sexes. *Theoretical Population Biology*, **34**.2, pp. 145–168. DOI: 10.1016/0040-5809(88)90039-1.

— (1990). Allele-frequency change in a class-structured population. *The American Naturalist*, **135**, pp. 95–106. DOI: 10.1086/285034.

— (1992). Altruism in viscous populations—an inclusive fitness model. *Evolutionary Ecology*, **6**.4, pp. 352–356. DOI: 10.1007/bf02270971.

Taylor, P. D., Wild, G., & Gardner, A. (2007). Direct fitness or inclusive fitness: how shall we model kin selection? *Journal of Evolutionary Biology*, **20**.1, pp. 301–309. DOI: 10.1111/j.1420-9101.2006.01196.x.