A quantitative genetics model for the dynamics of phenotypic (co)variances under limited dispersal, with an application to the coevolution of socially synergistic traits

Charles Mullon*[1] and Laurent Lehmann[1]

[1] *Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland*

Abstract

The gradual transformation of heritable quantitative traits due to selection and mutation is a major topic in evolutionary biology. Here, we use a quantitative genetics approach to investigate the coevolution of multiple traits under selection, mutation, and limited dispersal. We track the dynamics of trait means and variance-covariances between pleiotropic traits that experience frequency-dependent selection. Assuming a multivariate-Normal trait distribution, we recover classical dynamics of quantitative genetics, as well as stability and evolutionary branching conditions of invasion analyses, except that due to limited dispersal, selection depends on indirect fitness effects and relatedness. Correlational selection on associations between traits *within*-individuals depends on the fitness effects of such associations *between*-individuals. These kin selection effects can be as relevant as pleiotropy for the evolutionary build-up of trait associations. We illustrate this with an example on the coevolution of two social traits whose association within-individual is costly but synergistically beneficial between-individuals. As dispersal becomes limited, associations between-traits between-individuals become increasingly important for correlational selection. Consequently, the trait distribution goes from being bimodal with a negative correlation under panmixia to unimodal with a positive correlation under limited dispersal. More broadly, our approach can help understand the evolution of intraspecific variation in social behaviour.

Keywords. Island model, G-matrix evolution, evolutionary branching, division of labour, social evolution

1 Introduction

Understanding how quantitative traits are moulded by the processes of natural selection and mutation is a major pursuit of evolutionary biology. This pursuit has led to an abundant theoretical literature that seeks to understand the roles of ecology and genetics in the gradual transformation of phenotypes. Notwithstanding this abundance, models of gradual evolution usually follow one of two broad approaches, depending on whether the focus is put on ecological or genetic processes.

One approach consists in investigating the invasion success of a rare phenotypic deviant (i.e., an invasion analysis, e.g. Eshel and Feldman, 1984, Parker and Maynard Smith, 1990, Eshel et al., 1997; also referred to as

*Electronic address: `charles.mullon@unil.ch`

"Adaptive Dynamics", e.g., Dercole and Rinaldi, 2008, for a textbook treatment) and places emphasis on ecology (or on how organisms interact with one another via effects on resources and the environment). This emphasis typically comes at the expense of genetics realism. In particular, trait dynamics inferred from invasion analyses usually assume that mutations are limiting, i.e., that mutations have weak quantitative effects and are so rare (relative to the strength of selection) that at most two alleles can segregate in the population. In this case, a sensitivity analysis of the invasion fitness of a rare mutant in a resident monomorphic population that is at its ecological equilibrium (e.g., Michod, 1979, Eshel and Motro, 1981, Eshel and Feldman, 1984, Taylor, 1989, Parker and Maynard Smith, 1990, Charlesworth, 1994) can be used to understand gradual trait evolution and the ecological transformations due to this evolution (Metz et al., 1996, Geritz et al., 1998, Rousset, 2004, Dercole and Rinaldi, 2008, Metz, 2011). The approach of invasion analysis is therefore particularly well-suited to investigate the evolution of traits under ecological feedbacks and the frequency-dependent selection that emerges due to such feedbacks (e.g., Kisdi and Geritz, 2009, Lion, 2018, and references therein). This approach has revealed that in the presence of trade-offs, gradual evolution under ecological feedbacks often leads to the emergence of polymorphism. Here, the population first evolves under directional selection to express a trait value such that any rare mutant has an advantage over the common resident (Eshel and Motro, 1981, Eshel, 1983, Taylor, 1989, Christiansen, 1991, Abrams et al., 1993b). As a result, the population subsequently splits into two lineages of distinct phenotypes, or morphs, in a process referred to as evolutionary branching (i.e., the distribution of trait values in the population goes from unimodal to bimodal, Geritz et al., 1998; see Rueffler et al., 2006, Kisdi and Geritz, 2009, for reviews).

By contrast to invasion analysis, evolutionary quantitative genetics models of gradual evolution tend to be more preoccupied with the genetic basis of traits (Roff, 1997, Lynch and Walsh, 1998). Importantly, quantitative genetics models envisage that there is substantial heritable phenotypic variation in the population (i.e., that mutations are not limiting). The continuum-of-alleles model, in particular, posits that quantitative traits are determined by a continuum of possible alleles that are produced through mutation (e.g., Kimura, 1965a, Latter, 1970, Fleming, 1979, Burger, 1986). A quantitative genetics approach generally aims to investigate the roles of selection and mutation in the gradual evolution of a phenotypic distribution of arbitrary complexity. Due to the complication of dealing with multiple variants in the population, however, analytical explorations of quantitative genetics models usually come at the expense of ecological and population structure realism. Notably, the vast majority of quantitative genetics models of traits under frequency-dependent selection, which is either implicit or emerges from ecological interactions, focuses on the evolution of mean phenotypic values in the population, assuming that heritable phenotypic variation is constant (i.e., additive genetic variances and covariances are fixed, e.g., Lande, 1976, 1981, Iwasa et al., 1991, Gomulkiewicz and Kirkpatrick, 1992, Abrams et al., 1993a, Iwasa and Pomiankowski, 1995, Day and Taylor, 1996, Tazzyman and Iwasa, 2009, Nuismer et al., 2010, Connallon, 2015).

But phenotypic variance should be especially sensitive to frequency-dependent selection (because such selection regime either favours or disfavours rare variants that differ from the most common, it increases or decreases trait variance, respectively; Slatkin, 1980, Taper and Chase, 1985, Taylor and Day, 1997, Day and Proulx, 2004, Sasaki and Dieckmann, 2011, Wakano and Iwasa, 2013, Wakano and Lehmann, 2014, Débarre et al., 2014, Débarre and Otto, 2016). In fact, recent quantitative genetics models investigating populations of individuals experiencing frequency-dependent interactions have revealed links between the dynamics of phenotypic variance and evolutionary branching (Sasaki and Dieckmann, 2011, Wakano and Iwasa, 2013, Wakano and Lehmann, 2014, Débarre et al., 2014, Débarre and Otto, 2016; thereby extending the links between the dynamics of the phenotypic mean in quantitative genetics models and directional selection in invasion analysis models, Charlesworth, 1990, Iwasa et al., 1991, Taper and Case, 1992, Abrams et al., 1993a; for reviews: Abrams, 2001, Lion, 2018). In particular, evolutionary branching occurs in a quantitative genetics model when the phe-

notypic variance is predicted to grow without bound while the phenotypic mean is held constant, under the assumption that the phenotypic distribution is Normal (this assumption allows for moment closure, i.e., to only have to consider the dynamics of the mean and variance to model the dynamics of the phenotypic distribution, Wakano and Iwasa, 2013, Wakano and Lehmann, 2014, Débarre et al., 2014, Débarre and Otto, 2016). The variance, in fact, converges to a bounded value as evolutionary branching occurs, but these dynamics cannot be captured by models that assume that the phenotypic distribution is Normal and thus unimodal (instead of a bi- or multi-modal distribution; see Sasaki and Dieckmann, 2011, for a relaxation of the unimodal assumption). In spite of this limitation, quantitative genetics approaches have been useful to investigate relevant factors for frequency-dependent selection and evolutionary branching, such as genetic drift (with fixed, Wakano and Iwasa, 2013, or fluctuating, Débarre and Otto, 2016, population size) or the interaction between multiple traits (Débarre et al., 2014).

Another factor that is particularly relevant for frequency-dependent interactions is limited dispersal. This is because limited dispersal creates genetic structure, whereby individuals that interact and compete with one another are more likely to share alleles that are identical by descent than randomly sampled individuals (Hamilton, 1964, Frank, 1998, Rousset, 2004). Using an invasion analysis, a number of models have investigated the conditions that lead to disruptive selection (usually followed by evolutionary branching) due to frequency-dependent interactions among individuals under limited dispersal (Day, 2001a, Ajar, 2003, Rousset, 2004, Mullon et al., 2016, Parvinen et al., 2018; see also Svardal et al., 2015 for evolutionary branching due to spatial and temporal heterogeneities in selection). Using a quantitative genetics approach, Wakano and Lehmann (2014) found conditions equivalent to those obtained from invasion analysis by studying the dynamics of the variance of a trait under limited dispersal. The analysis of frequency-dependent and disruptive selection under limited dispersal has helped reveal further connections between invasion analysis and fundamental branches of evolutionary theory. In particular, Ajar (2003), Wakano and Lehmann (2014) and Mullon et al. (2016) expressed disruptive selection coefficients in terms of relatedness coefficients, which are quantities central to population genetics, kin selection and social evolution theory (i.e., the evolution of traits that influence the fitness of their actor and recipient, such as helping or harming, e.g., Hamilton, 1964, Frank, 1998, Rousset, 2004, Wenseleers, 2010; see also Kisdi, 2016 for a kin selection perspective on evolutionary branching of dispersal).

In this paper, we contribute to the above-mentioned literature on the gradual evolution of quantitative traits by studying frequency-dependent selection in a continuum-of-alleles model under limited dispersal (Wakano and Lehmann, 2014; see also Lande, 1992, for a neutral model). In particular, we extend Wakano and Lehmann (2014)'s model to incorporate two factors. First, we consider the joint evolution of multiple traits (whereas Wakano and Lehmann, 2014, focuses on a single trait). This enables us to investigate how phenotypic covariances among traits within individuals are moulded by frequency-dependent selection and pleiotropic mutations (i.e., when traits share a common genetic basis so that a mutation that influences one trait tends to have a correlated effect on the other). Second, we model the joint dynamics of the phenotypic means and (co)variances (whereas Wakano and Lehmann, 2014, considered the dynamics of the variance once directional selection on phenotypic means had vanished). This allows for a more complete picture of the evolutionary dynamics of the phenotypic distribution due to the combined forces of selection and mutation. In addition, by expressing these joint dynamics of means and (co)variances in terms of relatedness coefficients, we further connect kin selection theory with classical quantitative genetics theory of multiple traits (Lande, 1979, Lande and Arnold, 1983, Phillips and Arnold, 1989, Brodie et al., 1995; in particular with theory on the evolution of the **G**-matrix of additive genetic variance-covariance, Steppan et al., 2002, Arnold et al., 2008)

The rest of this paper is organised as follows. We describe the life-cycle and population structure under consideration in section 2. Our first result, presented in section 3.1, is an integro-difference equation for the

one-generational change of a multi-variate phenotypic distribution under limited dispersal, mutation, and selection (most of our derivations can be found in the Appendices). Next, in section 3.2, we obtain a closed dynamical system for the mean vector and variance-covariance matrix of the phenotypic distribution, under the assumption that the distribution in the whole population is Normal. Further, we express this dynamical system in terms of effects on individual fitness and relatedness in section 3.3. We then highlight some equilibrium properties of our dynamical system in section 3.4. Finally, in section 3.5, we apply our framework to study the coevolution of two traits that have socially synergistic effects.

2 Model

We consider a population of haploid individuals, divided among an infinite number of groups, each of fixed size N. Each individual bears a multidimensional phenotype consisting of n quantitative traits that are genetically determined. The discrete-time life cycle of this population is as follows. (1) Adults produce a large number offspring (sufficiently for the population size to remain constant) then either survive or die, which frees up breeding spots. (2) The phenotype of each individual mutates with probability ν, causing a small (random) quantitative deviation in trait values. (3) Each offspring either remains in its natal group, or disperses to another randomly chosen group (i.e., we consider the island model of dispersal, Wright, 1931). (4) Offspring compete in each group to fill open breeding spots, if any.

3 Results

3.1 Dynamics of the phenotypic distribution

3.1.1 General recurrence

Since the constant input of mutations leads to the co-existence of many alleles, understanding evolution under our life cycle requires tracking the dynamics of the phenotypic density distribution across the global population. We denote this distribution at an arbitrary demographic time point t as $p_t(\mathbf{z})$, where $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$ is a vector variable that collects the variable z_a for each trait $a = 1, \ldots, n$. Using the theory of point processes, we show in Appendix A that at the next demographic time point $t+1$ (i.e., after one iteration of life cycle), the phenotypic distribution can be expressed as

$$p_{t+1}(\mathbf{z}) = (1-\nu)W(\mathbf{z},\phi_t)p_t(\mathbf{z}) + \nu \int_{\mathbb{R}^n} \upsilon(\mathbf{z}',\mathbf{z})W(\mathbf{z}',\phi_t)p_t(\mathbf{z}')\mathrm{d}\mathbf{z}' \tag{1}$$

(see eq. A-24 in the Appendix). The first summand of eq. (1) represents individuals that have not mutated, and the second summand, those that have ($\upsilon(\mathbf{z}',\mathbf{z})$ is the probability density function for the event that an individual mutates from $\mathbf{z}'$ to $\mathbf{z}$ given that a mutation has occurred).

The quantity $W(\mathbf{z},\phi_t)$ in eq. (1) is defined as

$$W(\mathbf{z},\phi_t) = \int_{\mathcal{S}} w_\mu(\mathbf{z},\phi_t)q_t(\mu|\mathbf{z})\mathrm{d}\mu, \tag{2}$$

where $w_\mu(\mathbf{z},\phi_t)$ is the expected number of successful offspring produced by a focal individual (possibly in-

cluding self through survival) when: (1) the focal has phenotype $\mathbf{z}$; (2) and resides in a group whose state, i.e., the phenotypes of every individuals in the group, is μ (more precisely, μ is a mathematical function – a so-called counting measure – that contains all the necessary information about the phenotypes of all the individuals of the focal group; we denote by $\mathcal{S}$ the space of all such functions; see Appendix A.1.2, eq. A-1-A-2 for details); and (3) the distribution of states among all groups in the population is ϕ_t (i.e., $\phi_t(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$ is the density distribution of groups in which individuals labelled 1 to N have phenotypes $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N$, respectively; see Appendix A.1.3). The function $q_t(\mu|\mathbf{z})$ in eq. (2) is probability density function for the event that an individual randomly picked from the collection of all carriers of the $\mathbf{z}$ phenotype in the population at time period t, resides in a group in state μ.

According to eq. (2), $W(\mathbf{z}, \phi_t)$ is the number of offspring of an individual with phenotype $\mathbf{z}$, averaged over all group states μ in which this individual can reside, when groups in the global population have phenotypic distribution ϕ_t. In fact, $W(\mathbf{z}, \phi_t)$ can be interpreted as the expected fitness of a representative member of the lineage of individuals carrying phenotype $\mathbf{z}$ at time t. To see this, consider the individual in which phenotype $\mathbf{z}$ first appears through mutation. We define its lineage as the collection of all its descendants that have phenotype $\mathbf{z}$. This consists of its direct descendants (including self through survival) bearing $\mathbf{z}$ (and thus having not mutated to any other phenotype), the descendants of its immediate descendants bearing $\mathbf{z}$, and so on. Then, because $\mathbf{z}$ will appear once through mutation (as there are an infinite number of alleles in the continuum of alleles model), the fitness of a randomly sampled member of this lineage at time t is $W(\mathbf{z}, \phi_t)$. This fitness measure can be seen as the multi-allelic version of the concept of a mutant's *lineage fitness* used previously in invasion analyses in the island model (which turns out to be the mutant's geometric growth rate when rare, Lehmann et al., 2015, Mullon et al., 2016, Lehmann et al., 2016). We will therefore refer to $W(\mathbf{z}, \phi_t)$ as the lineage fitness of type $\mathbf{z}$, keeping in mind that $W(\mathbf{z}, \phi_t)$ here applies for any frequency of $\mathbf{z}$ (rare or common) and for any resident composition (monomorphic or polymorphic).

3.1.2 Connection with previous results in well-mixed population

Our recurrence eq. (1) takes the same functional form as the classical recurrence of the phenotypic distribution in well-mixed populations (under the continuum-of-alleles model, e.g., Kimura, 1965a, eqs. 1-2; Fleming, 1979, eq. 2.4; Burger, 1986, eq. 1; Champagnat et al., 2006, eq. 4.1). In a well-mixed population, lineage fitness $W(\mathbf{z}, \phi_t)$ reduces to the individual fitness $w(\mathbf{z}, p_t(\mathbf{z}))$ of a single carrier of $\mathbf{z}$ and only depends on the population phenotypic distribution $p_t(\mathbf{z})$, which is independent of the group state μ (Taylor and Day, 1997, Champagnat et al., 2006). When, in addition, selection is frequency-independent, fitness further becomes independent of the population phenotypic distribution ϕ_t altogether (and can thus simply be written as $w(\mathbf{z})$, Kimura, 1965a, Fleming, 1979, Burger, 1986). Such differences in fitness measure between eq. (1) and previous works are due to the group structure of the population considered here. We will later highlight these differences more evidently when we connect the dynamics of eq. (1) with relatedness coefficients and individual fitness (which we generally denote by a lower case w, while lineage fitness is denoted by an upper case W).

3.2 Dynamics when the phenotypic distribution is Normal

The dynamical equation for the phenotypic distribution eq. (1) has no straightforward solution, even in the simpler case when the population is well-mixed and selection is frequency-independent (Kimura, 1965a, Fleming, 1979, Lande, 1979, Burger, 1986). Under limited dispersal, this problem is further complicated by the necessity to track the dynamics of the group distribution ϕ_t (which determines the conditional profile dis-

tribution $q_t(\mu|\mathbf{z})$ and thus lineage fitness, eq. 2). To proceed further in our analysis, we therefore make some additional assumptions.

3.2.1 Normal closure and quasi-equilibrium of local genetic associations

We first assume that we can ignore the joint effects of selection and mutation on the phenotypic distribution over one time period (i.e., that the joint effects of selection and mutation are weak, see Appendix A.3 for details). Following previous authors (e.g., see Taylor and Day, 1997, Wakano and Iwasa, 2013, Débarre et al., 2014, Wakano and Lehmann, 2014, Débarre and Otto, 2016, for the distribution of social traits), we also assume that the processes of selection and mutation are such that $p_t(\mathbf{z})$ is approximately multivariate Normal (allowing for moment closure; see Appendix C.1 for phenotypic dynamics before closure and Appendix C.2.1 for Normal closure). This is a strong assumption but it is noteworthy that it does not require that the realised distribution of phenotypes within a focal group at any given demographic time period is Normal. Rather, it means that the distribution of phenotypes within a focal group is Normal on average over many replicates of the evolutionary process. In addition, the assumption of Normality has been demonstrated to give remarkably accurate predictions for the change of mean and variance, which is our main goal, even when selection generates significant deviations from normality in well-mixed populations (Turelli and Barton, 1994).

Under the assumption of Normality, the distribution $p_t(\mathbf{z})$ is characterised by its mean vector $\bar{\mathbf{z}}_t = (\bar{z}_{1,t}, \bar{z}_{2,t}, \ldots, \bar{z}_{n,t})$, whose a-entry is,

$$\bar{z}_{a,t} = \int_{\mathbb{R}^n} z_a p_t(\mathbf{z}) \mathrm{d}\mathbf{z}, \tag{3}$$

the average value of trait a in the population at time period t; and its variance-covariance matrix $\boldsymbol{G}_t$ whose (a,b)-entry is,

$$\sigma_{ab,t} = \int_{\mathbb{R}^n} (z_a - \bar{z}_{a,t})(z_b - \bar{z}_{b,t}) p_t(\mathbf{z}) \mathrm{d}\mathbf{z}, \tag{4}$$

the (co)variance among traits a and b in the population at time period t. The dynamics of $p_t(\mathbf{z})$ can therefore be tracked through the joint dynamics of its mean vector $\bar{\mathbf{z}}_t$ and variance-covariance matrix $\boldsymbol{G}_t$.

Due to limited dispersal, however, the dynamics of $\bar{\mathbf{z}}_t$ and $\boldsymbol{G}_t$ also depend on time-dependent local genetic associations among individuals of the same group (which emerge from ϕ_t). In order to close evolutionary dynamics on $\bar{\mathbf{z}}_t$ and $\boldsymbol{G}_t$ (and avoid tracking the dynamics of ϕ_t), we further assume that selection is weak relative to dispersal so that local genetic associations reach their steady state before significant changes have occurred in the phenotypic distribution, $p_t(\mathbf{z})$, at the level of the population (see section C.2.2 for details). This "quasi-equilibrium" assumption, which is frequently used in population genetic theory (e.g., Kimura, 1965b, Nagylaki, 1993, Kirkpatrick et al., 2002, Roze and Rousset, 2005, Lehmann et al., 2007, Roze and Rousset, 2008), allows us to characterize the dynamics of $p_t(\mathbf{z})$ entirely by the coupled dynamics of its mean vector $\bar{\mathbf{z}}_t$ and variance-covariance matrix $\boldsymbol{G}_t$, which we explore in the next section.

3.2.2 Dynamics of phenotypic mean vector and variance-covariance matrix

We show in Appendices B-C that the joint change over one demographic time period of the phenotypic mean trait vector and variance-covariance matrix are respectively given by

$$\Delta\bar{\mathbf{z}}_t(\bar{\mathbf{z}}_t, \boldsymbol{G}_t) = \boldsymbol{G}_t \boldsymbol{s}(\bar{\mathbf{z}}_t) \tag{5a}$$

$$\Delta \boldsymbol{G}_t(\bar{\mathbf{z}}_t, \boldsymbol{G}_t) = \boldsymbol{M} + \boldsymbol{G}_t\left(\boldsymbol{H}(\bar{\mathbf{z}}_t) - \boldsymbol{s}(\bar{\mathbf{z}}_t)\boldsymbol{s}(\bar{\mathbf{z}}_t)^{\mathrm{T}}\right)\boldsymbol{G}_t, \tag{5b}$$

where $\boldsymbol{s}(\bar{\mathbf{z}}_t) = (s_1(\bar{\mathbf{z}}_t), \ldots, s_n(\bar{\mathbf{z}}_t))^{\mathrm{T}}$ is a $n \times 1$ is vector of directional selection coefficients (or selection gradients) i.e., $s_a(\bar{\mathbf{z}}_t)$ is the first-order, marginal, effect of an (infinitesimal) change in trait a away from the population mean $\bar{\mathbf{z}}_t$ on lineage fitness ($s_a(\bar{\mathbf{z}}_t) = \partial W(\mathbf{z}, \phi_t)/\partial z_a$). The $n \times n$ matrix $\boldsymbol{M}$ collects the effects of mutation; its (a, b)-entry,

$$(\boldsymbol{M})_{ab} = \nu \underbrace{\int_{\mathbb{R}^n} (z_a - z'_a)(z_b - z'_b)\, \upsilon(\mathbf{z}', \mathbf{z})\mathrm{d}\mathbf{z}}_{\sigma^{\mathrm{m}}_{ab}}, \tag{6}$$

is the product of the mutation probability, ν, with the (co)variance, σ^{m}_{ab}, in mutational effects on traits a and b conditional on the appearance of a mutation (which captures the pleiotropic effects of mutations on a and b: when $\sigma^{\mathrm{m}}_{ab} > 0$, mutations tend to change a and b in a similar way; and when $\sigma^{\mathrm{m}}_{ab} < 0$, in opposite ways). The $n \times n$ Hessian matrix $\boldsymbol{H}(\bar{\mathbf{z}}_t)$ collects the second-order effects of traits on lineage fitness; its (a, b)-entry $\boldsymbol{H}(\bar{\mathbf{z}}_t)_{ab} = h_{ab}(\bar{\mathbf{z}}_t)$, i.e., is the effect of a joint (infinitesimal) changes in traits a and b away from the population mean $\bar{\mathbf{z}}_t$ on lineage fitness ($h_{ab}(\bar{\mathbf{z}}_t) = \partial^2 W(\mathbf{z}, \phi_t)/(\partial z_a \partial z_b)$). Finally, the notation $\boldsymbol{s}(\bar{\mathbf{z}}_t)\boldsymbol{s}(\bar{\mathbf{z}}_t)^{\mathrm{T}}$ denotes the outer product between two column vectors, so that $\boldsymbol{s}(\bar{\mathbf{z}}_t)\boldsymbol{s}(\bar{\mathbf{z}}_t)^{\mathrm{T}}$ is $n \times n$ matrix with (a, b)-entry $s_a(\bar{\mathbf{z}}_t)s_b(\bar{\mathbf{z}}_t)$.

3.2.3 Directional, disruptive, and correlational selection coefficients

Dynamical equations (5) have the same form as the one found for the evolution of the phenotypic means and (co)variances in well-mixed populations (e.g., eqs. 1-2 of Phillips and Arnold, 1989, see also eq. 7 of Lande, 1979 and eqs. 6 and 15 of Lande and Arnold, 1983). In such previous models, the effects of selection (captured by $\boldsymbol{s}(\bar{\mathbf{z}}_t)$ and $\boldsymbol{H}(\bar{\mathbf{z}}_t)$) depend on the effects of traits on *individual* (w) rather than *lineage* (W) fitness. Nevertheless, the parallels between eq. (5) and previous works allow us to use the same vocabulary and previous general interpretations on the evolution of phenotypic means and (co)variances (Brodie et al., 1995). First, the evolution of the mean of each trait (eq. 5a) depends on the vector of directional selection coefficient (or the selection gradient), $\boldsymbol{s}(\bar{\mathbf{z}}_t)$, which points in the direction favoured by selection in multivariate phenotypic space (Lande, 1979). The effect of directional selection on the mean of each trait is constrained by the genetic variation available for each trait and how traits covary genetically with one another (Lande, 1979), and these constraints are captured by $\boldsymbol{G}_t$ in eq. (5a).

Second, the evolution of the variance-covariance matrix $\boldsymbol{G}_t$ (eq. 5b) depends on the effects of mutations ($\boldsymbol{M}$), of directional selection ($\boldsymbol{s}(\bar{\mathbf{z}}_t)\boldsymbol{s}(\bar{\mathbf{z}}_t)^{\mathrm{T}}$), and of so-called quadratic selection given by the matrix $\boldsymbol{H}(\bar{\mathbf{z}}_t)$ (Lande, 1979, Lande and Arnold, 1983, Phillips and Arnold, 1989). This matrix $\boldsymbol{H}(\bar{\mathbf{z}}_t)$ captures two important features of selection. First, the sign of its diagonal entry (a, a) indicates whether selection favours a decrease (when $h_{aa}(\bar{\mathbf{z}}_t) < 0$) or an increase (when $h_{aa}(\bar{\mathbf{z}}_t) > 0$) in the variance of trait a when this trait evolves in isolation of other traits (Phillips and Arnold, 1989), hence we refer to $h_{aa}(\bar{\mathbf{z}}_t)$ as the coefficient of disruptive selection on trait a. Second, the off-diagonal entry (a, b) tells us whether selection favours a positive (when $h_{ab}(\bar{\mathbf{z}}_t) > 0$) or negative (when $h_{ab}(\bar{\mathbf{z}}_t) < 0$) (co)variance or correlation among traits a and b. The off-diagonal entry $h_{ab}(\bar{\mathbf{z}}_t)$ is therefore referred to as the coefficient of correlational selection among a and b (Lande and Arnold, 1983, Phillips and Arnold, 1989).

3.3 Selection coefficients in terms of individual fitness effects

So far, the effects of limited dispersal on evolutionary dynamics (eqs. 1 and 5) have been hidden behind the notion of lineage fitness $W(\mathbf{z},\phi_t)$ (otherwise, eqs. 1 and 5 are the same as in panmictic populations, Kimura, 1965a, Fleming, 1979, Burger, 1986, Lande, 1979, Lande and Arnold, 1983, Phillips and Arnold, 1989). In order to highlight more tangibly how selection depends on limited dispersal, we now express the selection coefficients ($\boldsymbol{s}(\bar{\mathbf{z}}_t)$ and $\boldsymbol{H}(\bar{\mathbf{z}}_t)$) in terms of the effects of traits on individual fitness and relatedness. To do so, we let $w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)$ denote the expected number of adult offspring of a focal individual, that we label as individual "i". According to this notation, the fitness of a focal individual then depends on its own phenotype $\mathbf{z}_i = (z_{i,1},z_{i,2},\ldots,z_{i,n})$, on the collection of phenotypes of its $N-1$ neighbours $\mathbf{z}_{-i} = (\mathbf{z}_1,\ldots\mathbf{z}_{i-1},\mathbf{z}_{i+1},\ldots,\mathbf{z}_N)$ (where $\mathbf{z}_j = (z_{j,1},z_{j,2},\ldots,z_{j,n})$ is the phenotype of a neighbour indexed j), and on the average trait values in the population $\bar{\mathbf{z}}_t$ (an example of such a fitness function is given later when we apply our framework, see in eq. 17). This individual fitness function is equal to the fitness function $w_\mu(\mathbf{z},\phi_t)$ that appears in eq. (2), i.e.,

$$w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t) = w_\mu(\mathbf{z},\phi_t), \tag{7}$$

when $\mathbf{z} = \mathbf{z}_i$, the state of the focal patch is $\mu = \{\mathbf{z}_i\} \cup \mathbf{z}_{-i}$, and groups other than the focal one are monomorphic for the population average $\bar{\mathbf{z}}_t$ (i.e., the distribution ϕ_t is delta peaked on $\bar{\mathbf{z}}_t$ in other groups).

3.3.1 Directional selection

We find that the directional selection gradient on a trait a can be expressed as

$$s_a(\bar{\mathbf{z}}_t) = \left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} + (N-1)r_2^\circ(\bar{\mathbf{z}}_t)\left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \tag{8}$$

(see Appendix C.1.1 to see the connection to mean trait dynamics), where the first derivative is the marginal effect of a (infinitesimal) change in trait a in a focal individual on its own fitness (the direct effect of trait a, which is the only effect that matters in a well-mixed population, i.e., $s_a(\bar{\mathbf{z}}_t) = \partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)/\partial z_{i,a}$, Phillips and Arnold, 1989). The second derivative is the marginal effect on the fitness of the focal resulting from an (infinitesimal) change in trait a in one neighbour of the focal (the indirect effect of trait a), which we arbitrarily chose to be the individual indexed as j (since the population has no class structure, a phenotypic change in any neighbour $j \neq i$ has the same effect on focal fitness; the evaluation symbol $\mathbf{z}_{-i} = \bar{\mathbf{z}}_t$ means that we evaluate it when all neighbours express the mean phenotype $\bar{\mathbf{z}}_t$, $\mathbf{z}_j = \bar{\mathbf{z}}_t$ for all $j \neq i$).

The indirect fitness effect is weighted by $r_2^\circ(\bar{\mathbf{z}}_t)$, which is the probability that, in the absence of selection and when the population phenotypic average is $\bar{\mathbf{z}}_t$, two neighbours are identical-by-descent (i.e., have a common ancestor that resided in the focal group, Rousset, 2004, and see Appendix B.2, eqs. B-41 - B-43 for details). This neutral coefficient of pairwise relatedness, $r_2^\circ(\bar{\mathbf{z}}_t)$, is also equal to the ratio of phenotypic covariance among neighbours to the total phenotypic covariance in the population in the absence of selection (see eq. B-48 in the Appendix; and for equivalent definitions of $r_2^\circ(\bar{\mathbf{z}}_t)$ when a single trait is under scrutiny, see eq. 7 of Queller, 1992a; eq. 2.13 of Frank, 1998; eq. 11 of Wakano and Lehmann, 2014).

We therefore find that the directional selection gradient effect of trait a (eq. 8) is the inclusive fitness effect of trait a (Hamilton, 1964, Rousset, 2004). Hence, in the absence of covariance among traits, the change in population mean of a trait is proportional to its inclusive fitness effect (substituting eq. 8 into 5a with the off-diagonal elements of $\boldsymbol{G}_t$ all zeros). This finding is in line with much previous modelling work on the quanti-

tative genetics in spatially- or family-structured populations (for e.g., Cheverud, 1985, Queller, 1992a,b, Frank, 1998, McGlothlin et al., 2014, Wakano and Lehmann, 2014).

3.3.2 Disruptive and correlational selection

We further find that the correlational selection coefficient on two traits a and b (or the disruptive selection coefficient when $a = b$) can be expressed as the sum of two terms,

$$h_{ab}(\bar{\mathbf{z}}_t) = h_{\mathrm{w},ab}(\bar{\mathbf{z}}_t) + h_{\mathrm{r},ab}(\bar{\mathbf{z}}_t), \tag{9a}$$

where the first term,

$$\begin{aligned} h_{\mathrm{w},ab}(\bar{\mathbf{z}}_t) = & \left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,a}\partial z_{i,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} + (N-1)r_2^{\circ}(\bar{\mathbf{z}}_t)\left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}\partial z_{j,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \\ & + (N-1)r_2^{\circ}(\bar{\mathbf{z}}_t)\left(\left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,a}\partial z_{j,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} + \left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,b}\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}}\right) \\ & + (N-1)(N-2)r_3^{\circ}(\bar{\mathbf{z}}_t)\left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}\partial z_{k,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \end{aligned} \tag{9b}$$

depends on the effects of joint changes in traits a and b away from the mean on individual fitness (see Appendix C.1.2 to see the connection with the dynamics of (co)variance). These changes can be divided according to whether they occur within- (first line) and between-individuals (second and third line). The first derivative on the first line of eq. (9b) measures the effect of a change in traits a and b in a focal individual on its own fitness (Figure 1a). In a well-mixed population, there are no other effects participating to correlational and disruptive selection (i.e., $h_{ab}(\bar{\mathbf{z}}_t) = \partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)/(\partial z_{i,a}\partial z_{i,b})$, Phillips and Arnold, 1989).

Under limited dispersal, however, indirect fitness effects are also important (as illustrated by the remaining derivatives of eq. 9b). There are three types of indirect effects that matter for correlational (and disruptive when $a = b$) selection: (i) the effect of a change in both traits a and b in a neighbour of the focal (second derivative on the first line, Figure 1b); (ii) the effect of a change in one trait in the focal and in another trait in a neighbour (the two derivatives of the second line, Figure 1c); and (iii) the effect of a change in one trait in a neighbour and in another trait in another neighbour (indexed as individual k with $i \neq k \neq j$, last derivative, Figure 1d). This latter effect is weighted by the three-way coefficient of relatedness, $r_3^{\circ}(\bar{\mathbf{z}}_t)$, which is defined as the probability that three neighbours are identical-by-descent in the absence of selection and when the population mean phenotype is $\bar{\mathbf{z}}_t$ (see eqs. B-63-B-65 for how $r_3^{\circ}(\bar{\mathbf{z}}_t)$ emerges in the calculations of second order effects; for explicit expression of $r_3^{\circ}(\bar{\mathbf{z}}_t)$, see e.g., eq. 12b of Ohtsuki, 2010).

We find that the second term of $h_{ab}(\bar{\mathbf{z}}_t)$ (eq. 9a) can be expressed as

$$h_{\mathrm{r},ab}(\bar{\mathbf{z}}_t) = (N-1)\left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \times \left.\frac{\partial r_2(\mathbf{z})}{\partial z_b}\right|_{\mathbf{z}=\bar{\mathbf{z}}_t} + (N-1)\left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \times \left.\frac{\partial r_2(\mathbf{z})}{\partial z_a}\right|_{\mathbf{z}=\bar{\mathbf{z}}_t}. \tag{9c}$$

This multiplies the indirect effect of one trait ($\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)/\partial z_{j,a}$ and $\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)/\partial z_{j,b}$), with the effect of the other trait on the probability of that two neighbours are identical-by-descent (denoted by $\partial r_2(\mathbf{z})/\partial z_a$ for the effect of trait a and $\partial r_2(\mathbf{z})/\partial z_b$ for trait b, Figure 1e, see eq. B-46 for details). This reflects that under limited dispersal, selection favours a correlation among two traits within individuals when it results in indirect fitness

benefits being preferentially directed towards relatives.

When a single traits evolves ($n = 1$) and the directional selection gradient on this trait is zero ($s_a(\bar{\mathbf{z}}_t) = 0$), the change in phenotypic variance that we obtain (eq. 9 substituted into eq. 5b) reduces to previously derived expressions from quantitative genetics models in the island model (eqs. 26 and 31 of Wakano and Lehmann, 2014). More generally, eqs. (8)-(9) are consistent with the first- and second-order effects of selection on the growth rate of a rare mutant that arises in a monomorphic, spatially-structured, population and that differs from the resident in a single (so when $a = b$; eqs. 8 & 9 of Ajar, 2003) or multiple (eqs. 12 & 13 of Mullon et al., 2016) traits. We discuss further such consistency between quantitative genetics, invasion analyses, and adaptive dynamics models in the next section, in which we study the equilibrium properties of the phenotypic distribution.

3.4 Equilibrium properties of the phenotypic distribution

Eq. (5) with eqs. (8)-(9) is a closed dynamical system that allows to track the evolution of the mean trait value and of the (co)variance between traits. In this section, we first investigate key features of the equilibrium of these phenotypic dynamics, and then discuss their connections with notions of evolutionary stability that come from invasion analyses and adaptive dynamics.

3.4.1 Equilibrium mean trait values

We denote the mean trait vector and variance-covariance matrix of the equilibrium phenotypic distribution by $\bar{\mathbf{z}}^*$ and $\boldsymbol{G}^*$, respectively. Such equilibrium simultaneously satisfies $\Delta\bar{\mathbf{z}}_t(\bar{\mathbf{z}}^*, \boldsymbol{G}^*) = \mathbf{0}$ and $\Delta\boldsymbol{G}_t(\bar{\mathbf{z}}^*, \boldsymbol{G}^*) = \mathbf{0}$ (where $\mathbf{0}$ is used to denote a n vector and $n \times n$ matrix whose entries are all zero, respectively). Rather than solving both systems of equations simultaneously, we can use the fact that in eq. (5a), the matrix $\boldsymbol{G}$ is a positive-definite matrix with real-entries (since it is a variance-covariance matrix). From standard linear algebra theory (Hines, 1980, Leimar, 2005, 2009), it then follows that the equilibrium for the phenotypic means must satisfy

$$\boldsymbol{s}(\bar{\mathbf{z}}^*) = \mathbf{0}, \tag{10}$$

i.e., all selection gradients (eq. 8) vanish at $\bar{\mathbf{z}}^*$, independently of the $\boldsymbol{G}$ matrix.

We can ask further whether a population with a mean vector that is close to an equilibrium $\bar{\mathbf{z}}^*$ will eventually converge to it as a result of selection and mutation. From the fact that $\boldsymbol{G}$ is positive-definite, it can be shown (see Leimar, 2009, for e.g.) that a necessary condition for a population to converge to $\bar{\mathbf{z}}^*$ for all possible $\boldsymbol{G}$ matrices is that the Jacobian matrix $\boldsymbol{J}(\bar{\mathbf{z}}^*)$ of the selection gradients with (a, b) entry

$$\boldsymbol{J}(\bar{\mathbf{z}}^*)_{ab} = \left.\frac{\partial s_a(\bar{\mathbf{z}})}{\partial z_b}\right|_{\bar{\mathbf{z}}=\bar{\mathbf{z}}^*} \tag{11}$$

is negative-definite at $\bar{\mathbf{z}}^*$, which means that the symmetric real part of $\boldsymbol{J}(\bar{\mathbf{z}}^*)$, $(\boldsymbol{J}(\bar{\mathbf{z}}^*) + \boldsymbol{J}(\bar{\mathbf{z}}^*)^{\mathrm{T}})/2$ has only negative eigenvalues. This type of equilibrium is referred to as (strongly) convergence stable (Leimar, 2005, 2009).

3.4.2 Equilibrium variance-covariance matrix

The dynamics of the variance-covariance matrix can then be studied at a convergence stable equilibrium $\bar{\mathbf{z}}^*$ for mean trait values (eq. 10). In this case, the equilibrium $\boldsymbol{G}^*$ for the variance-covariance matrix solves

$$\boldsymbol{M} + \boldsymbol{G}^* \boldsymbol{H}(\bar{\mathbf{z}}^*)\boldsymbol{G}^* = \mathbf{0}. \tag{12}$$

Eq. (12) has an admissible solution (i.e., such that $\boldsymbol{G}^*$ is positive-definite) if, and only if, the hessian matrix, $\boldsymbol{H}(\bar{\mathbf{z}}^*)$, is negative-definite (Bhatia, 2015). This corresponds to the case under which selection is stabilising at $\bar{\mathbf{z}}^*$. In fact, if $\boldsymbol{H}(\bar{\mathbf{z}}^*)$ is negative-definite, then the population will remain unimodally distributed around the mean vector $\bar{\mathbf{z}}^*$ and exhibit a variance-covariance matrix,

$$\boldsymbol{G}^* = \boldsymbol{M}\left[\boldsymbol{M}^{-1}\left(-\boldsymbol{H}(\bar{\mathbf{z}}^*)\right)^{-1}\right]^{1/2}, \tag{13}$$

where the operation $\mathbf{X}^{1/2}$ denotes the unique square root of $\mathbf{X}$ such that all the eigenvalues of $\mathbf{X}^{1/2}$ are positive (Bhatia, 2015; see also eq. 21c of Lande, 1980).

3.4.3 Connections with notions of stability from invasion analyses

Using a quantitative genetics approach, we have derived the conditions under which the multivariate phenotypic distribution of a dispersal limited population converges and remains at an equilibrium. Here, we highlight the connections between these conditions and notions of evolutionary stability that have emerged from invasion analyses and adaptive dynamics under limited dispersal.

Singular strategy. First, the selection gradient eq. (8) substituted into condition (10) is equivalent to the definition of evolutionarily singular strategies/phenotypes under limited dispersal (i.e., phenotypes which when expressed by the whole population, the gradient of invasion fitness is zero, e.g., Rousset, 2004; see also Geritz et al., 1998, for general definition).

Convergence stability. Second, the condition for a mean trait vector to be an attractor of directional selection (condition 11 with eq. 8) is equivalent to the condition for a multi-trait phenotype to be convergence stable in invasion analysis (Mullon et al., 2016; see also Brown and Taylor, 2010, for a graphical approach to the coevolution of two traits in a genetically structured population; and Leimar, 2009, Geritz et al., 2016, for general considerations on multi-trait in invasion analysis). It is noteworthy that in spite of this equivalence, the phenotypic dynamics envisaged by a quantitative genetics model (given by eq. 5a, see also eq. 7 of Lande, 1979, or eq. 1 of Phillips and Arnold, 1989) are quite different to the dynamics inferred from invasion analysis (which are captured by the so-called “canonical equation”, eq. 1 of Dieckmann and Law, 1996, or eq. 3 of Leimar, 2009). In a quantitative genetics model, the mean trait vector changes as a result of selection acting on a standing genetic variation, which is large enough to be captured by a statistical distribution ($\boldsymbol{G}_t$ in eq. 5a). Under the “canonical equation”, traits evolve under a trait substitution sequence, whereby selected mutants fix before other mutants arise, so that the population "jumps" from one monomorphic state to another and in principle cannot sustain polymorphism (see Fig. 1c, upper right panel of Champagnat et al., 2006, for a useful depiction of a trait substitution sequence; see Van Cleve, 2015, for a review of trait substitution sequence under limited dispersal).

Uninvadability. Third, the condition that $\boldsymbol{H}(\bar{\mathbf{z}}^*)$ with eq. (9) is negative-definite for the population to remain unimodally distributed around $\bar{\mathbf{z}}^*$ is consistent with the condition derived for $\bar{\mathbf{z}}^*$ to be locally uninvadable (i.e., that any rare mutant that arises in a population for monomorphic for $\bar{\mathbf{z}}^*$ and that causes a slight deviation from $\bar{\mathbf{z}}^*$ eventually vanishes) using evolutionary invasion analyses (Mullon et al., 2016; see also Ajar, 2003 for a single evolving trait in dispersal limited population; and Leimar, 2009, Geritz et al., 2016, for general considerations on multi-trait analyses).

Evolutionary branching. A phenotype that is convergence stable is not necessarily uninvadable (Eshel and Motro, 1981, Eshel, 1983, Taylor, 1989, Christiansen, 1991, Abrams et al., 1993b). In fact, when a singular phenotype is convergence stable but invadable, frequency-dependent disruptive selection can lead to evolutionary branching, whereby the phenotypic distribution goes from being unimodal to bimodal (Metz et al., 1996, Geritz et al., 1998). When multiple traits are coevolving, a sufficient condition for the initiation of evolutionary branching to is that the Jacobian is negative-definite and the Hessian matrix is positive-definite at the singular phenotype (note that this condition does not ensure that the resulting polyormphism is stable, Geritz et al., 2016, for further considerations). When $\boldsymbol{H}(\bar{\mathbf{z}}^*)$ is positive-definite, the dynamics of the variance-covariance matrix (eq. 5b) will diverge to infinity. This illustrates that at the onset of evolutionary branching, directional selection maintains the population mean vector at $\bar{\mathbf{z}}^*$ all the while disruptive selection favours extreme phenotypes, leading to an explosion of the variance-covariance matrix (in line with previous quantitative genetics approaches to study evolutionary branching, Wakano and Iwasa, 2013, Débarre et al., 2014, Wakano and Lehmann, 2014, Débarre and Otto, 2016).

The above connections are born out from the facts that: (1) the lineage fitness $W(\mathbf{z}, \phi_t)$ of type $\mathbf{z}$ is equivalent to its geometric growth rate when rare in a resident population held at a demographic-genetic equilibrium; and (2) that the effects of mutations are unbiased (i.e., symmetric around the mean). This suggests that adopting either invasion analyses or quantitative genetics models with normal closure do not matter for studying equilibrium properties of phenotypic evolution when dispersal is limited (as previously noted when populations are well-mixed, Day and Taylor, 1996, Lion, 2018, for reviews). This duality is implicitly endorsed by studies that locate singular strategies in subdivided populations by applying the "direct fitness method" (Frank and Taylor, 1996), which, strictly speaking, has been derived in terms of invasion analysis (Frank and Taylor, 1996, p. 33), but can be interpreted in terms of quantitative genetics (see Appendix B of Frank and Taylor, 1996 and Wenseleers, 2010). By contrast to an invasion analysis, a quantitative genetics approach may be more useful if one is interested in the evolution of the distribution under directional selection (i.e., eq. 5 with $\boldsymbol{s}(\bar{\mathbf{z}}^*) \neq \mathbf{0}$), or, as discussed in the next section, in the way that correlations among traits are shaped by selection and mutations.

3.4.4 The moulding of phenotypic correlations by selection and mutation

Invasion analyses can be used to infer on the phenotypic correlations or associations generated by disruptive selection (by studying at the eigenvector associated with the greatest eigenvalue of $\boldsymbol{H}(\bar{\mathbf{z}}^*)$, which gives the axis in phenotypic space on which selection is disruptive and along which the population becomes dimorphic, Mullon et al., 2016, Geritz et al., 2016). This approach, however, only incorporates the effect of selection and is limited to studying phenotypic correlations at the onset of evolutionary branching (inferring on the long term outcome of evolutionary branching requires studying invasion in dimorphic populations, which is typically much more involved mathematically, e.g., Geritz et al., 1998, Sasaki and Dieckmann, 2011). A quantitative genetics approach such as the one proposed here allows two further considerations on phenotypic correlations. First, it allows to incorporate the influence of pleiotropy (through the distribution of mutational input, captured by the variance-covariance matrix $\boldsymbol{M}$ in eq. 5b). Second, eq. (13) allows to study equilibrium phenotypic

correlations as a balance between mutation and stabilising selection (and not only disruptive selection). We investigate this balance in more detail in the next section, in which we apply our quantitative approach to model the evolution of traits with socially synergistic effects.

3.5 Application to the coevolution of two synergistic social traits

3.5.1 Model

We model the coevolution of two traits, labelled 1 and 2, that underlie social behaviours expressed at the adult stage. For example, one trait could be the level of energy invested into collecting food shared among the groups' offspring, and the other the energy invested into defending the group. A focal individual receives a fecundity benefit $B(\mathbf{z}_i, \mathbf{z}_{-i})$ that depends on the expression of both traits in the focal, $\mathbf{z}_i = (z_{i,1}, z_{i,2})$, as well as in its neighbours, $\mathbf{z}_{-i} = \bigcup_{j\neq i}(z_{j,1}, z_{j,2})$ (where $z_{j,a}$ is the value of trait a in individual indexed j). We assume that the focal expends a fecundity cost, $C(\mathbf{z}_i)$, that depends on it expression of both traits. The fecundity of a focal individual, $f(\mathbf{z}_i, \mathbf{z}_{-i})$, is generically written as

$$f(\mathbf{z}_i, \mathbf{z}_{-i}) = 1 + B(\mathbf{z}_i, \mathbf{z}_{-i}) - C(\mathbf{z}_i), \tag{14}$$

which we detail further in the next section.

Synergistic benefits and costs. We assume that traits have synergistic benefits among individuals, for instance because benefits rely on the coordinated expression of different behaviours among individuals. Specifically, we write the benefits as

$$B(\mathbf{z}_i, \mathbf{z}_{-i}) = \frac{b_{\mathrm{M}}}{N(N-1)} \sum_{j=1}^{N} \sum_{k=1, j\neq k}^{N} z_{j,1} z_{k,2}, \tag{15}$$

where $b_{\mathrm{M}} > 0$ tunes the strength of synergistic benefits, which increase with the extent to which different individuals have different traits (hence the product $z_{j,1} z_{k,2}$ involving different traits in different individuals, $j \neq k$). We assume that each trait is independently costly to express for a focal individual, but that additionally, costs increase when traits are positively associated within that focal individual (i.e., traits have direct multiplicative costs). A cost function that captures this situation is

$$C(\mathbf{z}_i) = \frac{c}{2}\left(z_{i,1}^2 + z_{i,2}^2\right) + c_{\mathrm{M}} z_{i,1} z_{i,2}, \tag{16}$$

where $c > 0$ tunes the independent cost of each trait, and $c_{\mathrm{M}} > 0$, the multiplicative costs of the traits. Multiplicative costs could for instance reflect a situation in which each trait underlies a task that requires a degree of specialisation and being a generalist is more costly than being a specialist.

These benefits (eq. 15) and costs (eq. 16) functions entail that it is best for a focal individual to express a negative *within*-individual association between traits, and simultaneously be in a group in which the different traits are positively associated *between*-individuals. Such a configuration is possible when the population is well-mixed (so that there are no genetic correlations – or no relatedness – among the individuals of the same group), but difficult when individuals of the same group are genetically related due to limited dispersal. As relatedness increases, phenotypic associations within and between individuals are expected to become more similar due to the co-inheritance of linked traits among individuals. In fact, we show that under neutrality, the covariance between traits between individuals is equal to the product of pairwise relatedness with total covariance (i.e., the covariance of traits a and b between individuals is equal to $r_2^{\circ}(\bar{\mathbf{z}}_t)\sigma_{ab,t}$, see Appendix B.2

and eq. B-48 for details). We therefore expect limited dispersal to be relevant to the coevolution of the two traits of our model and to the way selection associates these two traits within individuals.

Fitness. Before proceeding to the analysis, let us give the individual fitness function of a focal individual $w(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)$. For this model, we will assume that all adults die after reproduction (so that the population follows a Wright-Fisher life cycle). In this case, individual fitness is,

$$w(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t) = \underbrace{\frac{(1-m) f(\mathbf{z}_i, \mathbf{z}_{-i})}{(1-m)\sum_{i=1}^{N} f(\mathbf{z}_i, \mathbf{z}_{-i})/N + m f(\bar{\mathbf{z}}_t, \bar{\mathbf{z}}_t)}}_{w^{\mathrm{P}}(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)} + \underbrace{\frac{m f(\mathbf{z}_i, \mathbf{z}_{-i})}{f(\bar{\mathbf{z}}_t, \bar{\mathbf{z}}_t)}}_{w^{\mathrm{D}}(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)} \tag{17}$$

where $0 < m = m(\bar{\mathbf{z}}_t) \leq 1$ is the backward probability of dispersal, which here is independent from the evolving traits (fitness function eq. 17 is standard under limited dispersal, e.g., Rousset, 2004, Ohtsuki, 2010). The expected number of offspring of the focal that successfully establish in their natal group, $w^{\mathrm{P}}(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)$, is the ratio of the number of philopatric offspring of the focal to the total number of offspring that enter the competition in the focal group. The expected number of offspring of the focal that successfully settle in other groups, $w^{\mathrm{D}}(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)$, is the ratio of offspring the focal sends in a non-focal group to the expected number of offspring in such a group.

Relatedness. The final pieces of information that are necessary to apply our framework are the neutral relatedness coefficients, $r_2^{\circ}(\bar{\mathbf{z}}_t)$ and $r_3^{\circ}(\bar{\mathbf{z}}_t)$, and the effect of each trait on pairwise relatedness ($\partial r_2(\mathbf{z})/\partial z_a$). Neutral relatedness coefficients are found using standard identity-by-descent arguments (Karlin, 1968, Rousset, 2004), yielding

$$\begin{aligned} r_2^{\circ}(\bar{\mathbf{z}}_t) &= \frac{(1-m)^2}{N-(N-1)(1-m)^2} \\ r_3^{\circ}(\bar{\mathbf{z}}_t) &= \frac{(1-m)^3(1+3(N-1)r_2^{\circ}(\bar{\mathbf{z}}_t))}{N^2-(N-1)(N-2)(1-m)^3} \end{aligned} \tag{18}$$

for the Wright-Fisher life cycle (e.g., eqs. 12a & 12b of Ohtsuki, 2010). For this life-cycle, the effect of a trait on pairwise relatedness can be expressed as,

$$\begin{aligned} \left.\frac{\partial r_2(\mathbf{z})}{\partial z_a}\right|_{\mathbf{z}=\bar{\mathbf{z}}_t} = \frac{2r_2^{\circ}(\bar{\mathbf{z}}_t)}{1-m}\Bigg[& \left[1+(N-1)r_2^{\circ}(\bar{\mathbf{z}}_t)\right] \left.\frac{\partial w^{\mathrm{P}}(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)}{\partial z_{i,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t \\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \\ & + \left[2r_2^{\circ}(\bar{\mathbf{z}}_t)+(N-2)r_3^{\circ}(\bar{\mathbf{z}}_t)\right](N-1) \left.\frac{\partial w^{\mathrm{P}}(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)}{\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t \\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \Bigg] \end{aligned} \tag{19}$$

(which is the factor of $2(N-1)\partial w/\partial z_0$ in eq. 18 of Ajar, 2003 under mean field games, and the factor of $2V^2$ in eq. B.46 of Wakano and Lehmann, 2014; see also Appendix C.2.2, eq. C-19, for the expression of $\partial r_2(\mathbf{z})/\partial z_a$ for an iteroparous Moran life-cycle).

3.5.2 Analysis

We now proceed to analyse the evolution of both social traits using the framework given in section 3.2. We first focus on the equilibrium properties of the phenotypic distribution.

Convergence of the mean trait values. Substituting eqs. (17) & (18) into eq. (8), we obtain that the selection gradient vector is

$$\boldsymbol{s}(\bar{\mathbf{z}}_t) = \frac{K}{f(\bar{\mathbf{z}}_t, \bar{\mathbf{z}}_t)} \begin{pmatrix} -c\bar{z}_{1,t} + (-c_{\mathrm{M}} + b_{\mathrm{M}}/N)\bar{z}_{2,t} \\ -c\bar{z}_{2,t} + (-c_{\mathrm{M}} + b_{\mathrm{M}}/N)\bar{z}_{1,t} \end{pmatrix}, \tag{20}$$

where

$$K = N r_2^{\circ}(\bar{\mathbf{z}}_t) \frac{1-(1-m)^2}{(1-m)^2} > 0, \tag{21}$$

is always positive. Solving eq. (20) for zero then yields the unique singular strategy

$$\bar{\mathbf{z}}^* = (\bar{z}_1^*, \bar{z}_2^*) = (0,0). \tag{22}$$

In order to establish whether the population phenotypic distribution will converge to have mean $\bar{\mathbf{z}}^* = (0,0)$, we substitute eq. (20) into the symmetric part of the Jacobian matrix eq. (11), which we evaluate at the equilibrium $\bar{\mathbf{z}}^* = (0,0)$. It is straightforward to show that the two eigenvalues of the resulting matrix are given by

$$K\left\{-c - c_{\mathrm{M}} + b_{\mathrm{M}}/N, -c + c_{\mathrm{M}} - b_{\mathrm{M}}/N\right\}. \tag{23}$$

Both are negative provided

$$-c < -c_{\mathrm{M}} + b_{\mathrm{M}}/N < c, \tag{24}$$

i.e., when the difference between the multiplicative costs, c_{M}, and direct synergistic benefits, b_{M}/N, are small compared to the independent cost, c. Then, $(\bar{z}_1^*, \bar{z}_2^*) = (0,0)$ is convergence stable.

Stabilization of the distribution around the equilibrium mean. Assuming eq. (24) holds true, whether or not the population distribution stabilises around the equilibrium trait values (eq. 22) depends on the Hessian matrix, $\boldsymbol{H}(\bar{\mathbf{z}}^*)$. Let us start with analysing the diagonal elements of $\boldsymbol{H}(\bar{\mathbf{z}}^*)$, which reveal whether selection on each trait is independently stabilising or disruptive. Substituting eqs. (17)–(19) into eq. (9) for traits 1 and 2 (i.e., $a = b = 1$ and $a = b = 2$), and evaluating it at equilibrium eq. (22), we obtain that the diagonal entries of $\boldsymbol{H}(\bar{\mathbf{z}}^*)$ are

$$h_{11}(\bar{\mathbf{z}}^*) = h_{22}(\bar{\mathbf{z}}^*) = -Kc. \tag{25}$$

Since $K > 0$, the diagonal entries of $\boldsymbol{H}(\bar{\mathbf{z}}^*)$ are always negative (i.e., $h_{11}(\bar{\mathbf{z}}^*) < 0$ and $h_{22}(\bar{\mathbf{z}}^*) < 0$), which means that selection on each trait is stabilising when they evolve independently from one another. Whether selection is stabilising when both traits co-evolve also depends on the correlational coefficient of selection, $h_{12}(\bar{\mathbf{z}}^*)$. In particular, stabilising selection requires that

$$h_{12}(\bar{\mathbf{z}}^*)^2 < h_{11}(\bar{\mathbf{z}}^*) h_{22}(\bar{\mathbf{z}}^*), \tag{26}$$

i.e., that the correlational selection coefficient is weak relative to the strength of stabilising selection on both independent traits (because a 2×2 symmetric matrix Hessian matrix is negative-definite if and only if its diagonal entries are both negative and the off-diagonal satisfies eq. 26, e.g., Horn and Johnson, 1994). If eq. (26) does not hold (and $h_{11}(\bar{\mathbf{z}}^*) < 0$ and $h_{22}(\bar{\mathbf{z}}^*) < 0$), then selection is disruptive due to correlational selection (regardless of its sign). The correlational coefficient of selection is derived by first substituting eq. (17) into eq. (9) with $a = 1$ and $b = 2$, and second, evaluating the result at equilibrium eq. (22). This yields

$$h_{12}(\bar{\mathbf{z}}^*) = K\left(-c_{\mathrm{M}} + \alpha b_{\mathrm{M}}/N\right), \tag{27}$$

where

$$0 \leq \alpha = \frac{N r_3^{\circ}(\bar{\mathbf{z}}^*) m^2}{[1+3(N-1) r_2^{\circ}(\bar{\mathbf{z}}^*)][1-(1-m)^2)]} < 1, \tag{28}$$

is a function that decreases as dispersal and group size increases (i.e., α decreases as relatedness coefficients decrease, see Figure 2a). Eq. (27) reveals that as α (and relatedness) increases, the within-individual association favoured by selection goes from negative (Figure 2b, grey region) to positive (Figure 2b, black region). This is because as relatedness increases, the fitness effects of joint changes in traits 1 and 2 between individuals, which in this example are positive, play an increasingly important role in correlational selection (second and third line of eq. 9b, Figure 1c-d).

Substituting eqs. (25) and (27) into eq. (26), we find that selection is stabilising $(0,0)$ when

$$-c < -c_{\mathrm{M}} + \alpha b_{\mathrm{M}}/N < c, \tag{29}$$

which further reveals that high relatedness tends to favour stabilising selection (note that we assume that eq. 24 holds here; Figure 2b, dark grey and black regions). Conversely, low relatedness tends to favour disruptive selection (when eq. 24 holds but eq. 29 does not). This latter finding is in line with previous results on the evolution of single traits that have found that evolutionary branching is inhibited by limited dispersal (e.g., Day, 2001b, Ajar, 2003, Wakano and Lehmann, 2014, Parvinen et al., 2017). In such models and ours, limited dispersal inhibits evolutionary branching because it creates genetic correlations among competing individuals, so that a mutant cannot be as different to the common type as in well-mixed population. As a result, frequency-dependent disruptive selection is weaker under limited dispersal.

Effect of selection on the equilibrium phenotypic correlation. Putting our stability analyses together (especially eqs. 22, 24, 27, and 29) and validating them using individual-based simulations (see Appendix D for details), we find that there are three possible outcomes for the phenotypic distribution once it has converged to be unimodal around (0,0) due to selection: (1) when relatedness is low, selection tends to be disruptive and correlational selection negative, which leads to evolutionary branching and a stable dimorphism such that individuals are specialised in either trait 1 or 2 (Figure 3a); (2) over a critical level of relatedness, selection becomes stabilising but correlational selection remains negative, which prevents evolutionary branching and thus specialisation, but still results in a negative association among traits within individuals (Figure 3b); (3) over another threshold of relatedness, correlational selection becomes positive, so that the traits become positively associated within individuals (Figure 3c).

Effect of pleiotropy on the equilibrium phenotypic correlation. So far, our analysis has focused on the effects of selection on the stability of jointly evolving traits (an analysis that could have equally been performed using invasion analysis; see Mullon et al., 2016, for such an approach to the joint evolution of multiple traits under limited dispersal). But selection is not the only process that is relevant for the way phenotypic distributions are shaped. As highlighted by the present quantitative genetic approach, the equilibrium variance-covariance matrix of the phenotypic distributions also depends on the patterns of mutation (captured by matrix $\boldsymbol{M}$ in eq. 13). In particular, we expect pleiotropy (when mutations have correlated effects among traits) to influence the correlations among traits within individuals at an evolutionary equilibrium.

In order to investigate the joint effects of pleiotropy and correlational selection, let us assume that the variance of mutational effect on both traits is the same ($\sigma_{11}^{\mathrm{m}} = \sigma_{22}^{\mathrm{m}} = \sigma_{\mathrm{m}}$), in which case the variance-covariance matrix

of mutation effects can be written as

$$\boldsymbol{M} = \nu\sigma_{\mathrm{m}}\begin{pmatrix} 1 & \rho_{\mathrm{m}} \\ \rho_{\mathrm{m}} & 1 \end{pmatrix}, \tag{30}$$

where $\rho_{\mathrm{m}} = \sigma^{\mathrm{m}}_{12}/\sigma_{\mathrm{m}}$ is the correlation of the effect of mutations on traits 1 and 2. The parameter $-1 < \rho_{\mathrm{m}} < 1$ thus captures the degree of pleiotropy between both traits (when it is zero, both traits change independently due to mutation, when it positive, they tend to change in similar ways, and when it is negative, in opposite ways).

Substituting eqs. (25), (27) and (30) into eq. (13), we find that the correlation ρ^*_{12} among traits 1 and 2 at equilibrium is

$$-1 < \rho^*_{12} = \frac{\sigma^*_{12}}{\sqrt{\sigma^*_{11}\sigma^*_{22}}} = \frac{\rho_{\mathrm{m}} + \rho^*_{\mathrm{s}}}{1 + \rho_{\mathrm{m}}\rho^*_{\mathrm{s}} + \sqrt{(1-\rho^2_{\mathrm{m}})(1-\rho^{*2}_{\mathrm{s}})}} < 1, \tag{31}$$

where

$$-1 < \rho^*_{\mathrm{s}} = \frac{h_{12}(\bar{\mathbf{z}}^*)}{\sqrt{h_{11}(\bar{\mathbf{z}}^*)h_{22}(\bar{\mathbf{z}}^*)}} = \frac{-c_{\mathrm{M}} + \alpha b_{\mathrm{M}}/N}{c} < 1, \tag{32}$$

is the strength of correlational selection, relative to the strength of stabilising selection on each independent trait (which must be bounded between -1 and 1 because eq. 26 holds when selection is stabilising overall). Eq. (31) shows that at equilibrium, the sign of the correlation among between two traits reflects the balance, $\rho_{\mathrm{m}} + \rho^*_{\mathrm{s}}$, between the degree of pleiotropy, ρ_{m}, and the relative strength of correlational selection ρ^*_{s} (see Figure 4a for more details on how ρ^*_{12} depends on ρ_{m} and ρ^*_{s} according to eq. 31; note that eq. 31 holds whenever the variance of mutational effect on both traits is the same, $\sigma^{\mathrm{m}}_{11} = \sigma^{\mathrm{m}}_{22}$, and the coefficients of disruptive selection on independent traits are equal, i.e., $h_{11}(\bar{\mathbf{z}}^*) = h_{22}(\bar{\mathbf{z}}^*)$).

We additionally checked that our model captured pleiotropy correctly by comparing the phenotypic correlation among the two traits at equilibrium predicted by our model (eq. 31) and that observed in simulations for different levels of pleiotropy. We found that model predictions and observations from simulations also matched well in the presence of pleiotropy (Figure 4b).

Dynamics of the distribution. We further tested the accuracy of our dynamical model by comparing individual-based simulations with numerical recursions of eqs. (5). This showed that eqs. (5) tend to under evaluate phenotypic variance (Figure 5). This is probably due to genetic drift, which our model ignores and which depletes phenotypic variance (as shown in well-mixed populations, e.g., Wakano and Iwasa, 2013, Débarre and Otto, 2016), and/or the presence of phenotypic skew, which is ignored under our assumption that the phenotypic distribution in the population is Normal, but which can influence the dynamics of phenotypic variance (Appendix C, eq. C-14). Nonetheless, we observed overall a good fit between the predicted and observed dynamics of the phenotypic distribution, at least qualitatively (Figure 5). This suggests that the assumption of normality yields accurate predictions for the change of mean and variance when dispersal is limited (like in well-mixed populations, Turelli and Barton, 1994).

4 Discussion

In this paper, we have modelled the evolution of the distribution of genetically determined, quantitative, phenotypic traits under limited dispersal, selection and mutation. By doing so, we have generalised two classical quantitative genetics models to include limited dispersal, first for the general recurrence formula eq. (1) of the phenotypic distribution under the continuum of alleles model (Kimura, 1965a, Fleming, 1979, Lande,

1979, Burger, 1986); and second for the closed dynamical system eq. (5) of the vector of means and matrix of variance-covariance when the distribution is Normal (Lande, 1979, Lande and Arnold, 1983, Phillips and Arnold, 1989). In both cases, genetic structure due to limited dispersal leads to the replacement of individual (or personal) fitness in classical quantitative genetics equations by lineage fitness, which is the fitness of a typical bearer of a given phenotype (randomly sampled from the lineage of all members of that phenotype). This fitness depends on the phenotypes expressed in the whole population and how they are distributed among groups (eq. 2).

The lineage fitness concept was introduced to study the invasion of rare mutants in otherwise monomorphic populations and thus characterise convergent stable and uninvadable phenotypes under limited dispersal (Lehmann et al., 2015, Mullon et al., 2016, Lehmann et al., 2016, see Akçay and Van Cleve, 2016 for discussion). Comparing such evolutionary invasion analyses results with our own results highlighted the exact correspondence between concepts of evolutionary stability and quantitative genetics models: the vector of means evolves to convergence stable phenotypic values (eqs. 10-11) and the distribution remains unimodal around such values when they are locally uninvadable or may become bimodal when they are invadable (eq. 12; in agreement with previous results in well-mixed populations, see Charlesworth, 1990, Iwasa et al., 1991, Taper and Case, 1992, Abrams et al., 1993a, Abrams, 2001, Lion, 2018, for the dynamics of the mean, and Sasaki and Dieckmann, 2011, Wakano and Iwasa, 2013, Débarre et al., 2014, Débarre and Otto, 2016, for the dynamics of the variance; under limited dispersal, see Cheverud, 1985, Queller, 1992a,b, Frank, 1998, McGlothlin et al., 2014 for the dynamics of the means, and Wakano and Lehmann, 2014 for the dynamics of the variance of a single trait under limited dispersal around a singular strategy). While it may be felt that this correspondence is intuitive and must hold, it was only obtained after surprisingly lengthy and complicated calculations (see Appendix).

By contrast to evolutionary invasion analyses, the quantitative genetics framework we employ here allows to specify the phenotypic distribution at a mutation-selection balance (by using eq. 13). In particular, it allows to study the effects of selection and mutation on the phenotypic associations that emerge among traits at equilibrium (eq. 31). Our analyses of such associations suggest that kin selection due to limited dispersal can mould phenotypic associations as much as pleiotropic mutations do (eq. 31 and Fig. 4). By expressing correlational selection on traits in terms of their direct and indirect fitness effects, we gained insights into the influence of kin selection on phenotypic associations (eq. 9). These kin selection effects are equivalent with those previously inferred from invasion analyses under limited dispersal (Mullon et al., 2016). Motivated by our present explicit formula for the variance-covariance matrix (eq. 13) and our example (section 3.5), we complement here the discussion found in Mullon et al. (2016) on the implications of kin selection for the evolution of within individual phenotypic associations. As indicated by the decomposition of correlational selection eq. (9a), there are two ways kin selection influences such associations.

The first is through the fitness effects that traits have when co-expressed among relatives (eq. 9b, Figure 1b-d). We found that under limited dispersal, selection favours an association among two traits within individuals, when such an association between individuals has indirect fitness benefits. Due to such kin selection effects, different levels of dispersal can lead to different evolutionary outcomes for phenotypic associations, as highlighted by our example on the coevolution of two traits whose association within-individual is costly but beneficial between-individuals due to synergy (Figure 3). In this example, it is noteworthy that relatedness had a substantial influence on the shape of the phenotypic distribution without altering the mean of this distribution (Figure 3, eq. 22 and 23). Hence, these effects of genetic structure on phenotypic evolution would have gone unnoticed from the study of the dynamics of the mean only (which is the focus of the vast majority of study of quantitative genetics in family-structured populations, e.g., Cheverud, 1985, Queller, 1992a,b, Frank, 1998, McGlothlin et al., 2014), or from the analysis of the directional selection gradient vector only (as done

in the majority of evolutionary analyses to synergistic social traits, e.g., Gandon, 1999, Perrin and Mazalov, 2000, Reuter and Keller, 2001, Lehmann and Perrin, 2002, Rousset and Gandon, 2002, Gardner and West, 2004, Leturque and Rousset, 2004, Hochberg et al., 2008, Brown and Taylor, 2010, Kuijper and Johnstone, 2017). Overall, this example highlights that when traits have synergistic indirect effects (Figure 1c-d), relatedness is relevant for the way natural selection moulds phenotypic associations among traits within individuals.

A relevant pair of traits likely to be influenced by such kin selection effects is costly helping and punishment, which have synergistic indirect benefits when expressed by different individuals (e.g., Raihani et al., 2012, and references therein). According to our results, kin selection should favour a positive association among helping and punishment, which interestingly, has been observed in humans (Fehr and Gächter, 2000). Another pair of traits whose evolution is likely to be influenced by their joint expression in different individuals is the production and exploitation of a public good, such as the secretion and use of siderophores by microorganisms (West et al., 2006). Under limited diffusion of siderophores and limited bacterial dispersal (Nadell et al., 2009), we thus expect kin selection effects to be ecologically relevant for how secretion and use of siderophores are associated, and more generally for patterns of variation within microbial populations (Cordero and Polz, 2014).

The second way kin selection influences phenotypic associations is via the combination of the indirect effect of one trait with the effect of the other on the tendency to interact with relatives (eq. 9c, Figure 1e). Specifically, selection favours an association among two traits when it results in fitness benefits being preferentially directed towards relatives. For example, if trait a has positive indirect fitness effects (e.g., altruistic helping) and trait b decreases the tendency to interact with relatives (e.g., dispersal), then selection favours a negative correlation between traits a and b (e.g., Koella, 2000, Purcell et al., 2012, Mullon et al., 2018). We refer readers interested in this effect to Mullon et al. (2016), in which it is discussed at greater length, in particular in the context of dispersal syndromes (Edelaar and Bolnick, 2012, Ronce and Clobert, 2012).

Our present results further provide insight into the effects of limited dispersal on how selection influences the **G** matrix of additive genetic variances-covariances (Steppan et al., 2002, Arnold et al., 2008). Previous theoretical works have studied how linkage disequilibrium, pleiotropy and epistasis influence **G** under selection (Lande, 1980, 1984, Turelli, 1985, Turelli and Barton, 1990, Jones et al., 2014), but the effects of limited dispersal on **G** have either been assessed in the absence of selection (Lande, 1992), or has assumed that selection was frequency-independent (Jones et al., 2004, Guillaume and Whitlock, 2007, Guillaume, 2011, Björklund and Gustafsson, 2015). Here, we have shown that kin selection effects can be important for the way phenotypic associations are selected (i.e., for correlational selection, eq. 9), which can in turn lead to genetic correlations through genetic integration (Roff and Fairbairn, 2012). Of course, our model ignores many relevant features for quantitative genetics, such as environmental effects, genetic dominance, genetic linkage, sexual reproduction, or genetic drift at the population level (which could for instance be included in our model by applying the method of Wakano and Iwasa, 2013 or Débarre and Otto, 2016). Incorporating such features into our framework is likely to make the analysis of selection more complicated, but it would allow to study important questions on the genetic basis of variation, such as how genetic architecture and its evolution influence trait associations (e.g., Saltz et al., 2017).

One other significant limitation to our present approach is that it assumes that the phenotypic distribution is Normal. This assumption is likely to be violated under frequency-dependent selection, which may lead to skewed and complicated distributions. In particular, the Normal assumption precludes investigating what happens to the phenotypic distribution once evolutionary branching has occurred (like adaptive dynamics models based on the invasion analyses of monomorphic populations). To relax this assumption would entail tracking the dynamics of higher moments of the phenotypic distribution. One possible way to retain some

mathematical tractability would be to use the oligomorphic approximation proposed by Sasaki and Dieckmann (2011). This approximation decomposes a multimodal trait distribution into a sum of unimodal distributions, each corresponding to a morph. Applying Sasaki and Dieckmann (2011)'s approach, which was developed for a large and well-mixed population, to a dispersal limited one, would be an interesting avenue of future research, as well as including class-structure (e.g., age- or sex-structure).

To conclude, we have derived a framework to study gradual evolution of multiple traits under frequency-dependent selection and pleiotropic mutations in the island model of (limited) dispersal. This framework is especially relevant to study how associations within individuals between social traits emerge in response to mutation and selection under limited dispersal. Our results could therefore help understand patterns of intra-specific variation in social behaviour (such as behavioural syndromes, Dall et al., 2004, Dingemanse et al., 2012; social niche specialisation, Bergmüller and Taborsky, 2007, Montiglio et al., 2013; or social division of labour, Boesch, 2002, Wright et al., 2014), which are increasingly thought to be ecologically significant (Bolnick et al., 2011, Wolf and Weissing, 2012, Sih et al., 2012, Canestrelli et al., 2016, Chaturvedi et al., 2017). More broadly, by connecting different branches of theoretical evolutionary biology, from invasion analysis to adaptive dynamics to quantitative genetics, the present framework further bolsters that natural selection cannot be divorced from kin selection under limited dispersal, whatever modelling approach is taken (Hamilton, 1964, Frank, 1998, Rousset, 2004, van Baalen M, 2013, Lehmann et al., 2016).

References

Abrams. 2001. Modelling the adaptive dynamics of traits involved in inter- and intraspecific interactions: An assessment of three methods. *Ecology Letters* 4:166–175.

Abrams, P. A., Y. Harada, and H. Matsuda. 1993a. On the Relationship Between Quantitative Genetic and Ess Models. *Evolution (N. Y)*. 47:982–985.

Abrams, P. A., H. Matsuda, and Y. Harada. 1993b. Evolutionarily unstable fitness maxima and stable fitness minima of continuous traits. *Evol. Ecol.* 7:465–487.

Ajar, É. 2003. Analysis of disruptive selection in subdivided populations. *BMC Evol. Biol.* 3:22.

Akçay, E. and J. Van Cleve. 2016. There is no fitness but fitness, and the lineage is its bearer. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20150085.

Arnold, S. J., R. Burger, P. A. Hohenlohe, B. C. Ajie, and A. G. Jones. 2008. Understanding the evolution and stability of the G-matrix. *Evolution* 62:2451–2461.

Bergmüller, R. and M. Taborsky. 2007. Adaptive behavioural syndromes due to strategic niche specialization. *BMC Ecology* 7:12–7.

Bhatia, R. 2015. Positive Definite Matrices. Princeton University Press, Princeton, NJ, USA.

Björklund, M. and L. Gustafsson. 2015. The stability of the G-matrix: The role of spatial heterogeneity. *Evolution* 69:1953–1958.

Boesch, C. 2002. Cooperative hunting roles among taï chimpanzees. *Human nature (Hawthorne, N.Y.)* 13:27–46.

Bolnick, D. I., P. Amarasekare, M. S. Araújo, R. Burger, J. M. Levine, M. Novak, V. H. W. Rudolf, S. J. Schreiber, M. C. Urban, and D. A. Vasseur. 2011. Why intraspecific trait variation matters in community ecology. *Trends in Ecology and Evolution* 26:183–192.

Brodie, E. D., A. J. Moore, and F. J. Janzen. 1995. Visualizing and quantifying natural selection. *Trends Ecol. Evol.* 10:313–318.

Brown, S. P. and P. D. Taylor. 2010. Joint evolution of multiple social traits: a kin selection analysis. *Proc. R. Soc. B Biol. Sci.* 277:415–422.

Burger, R. 1986. On the maintenance of genetic variation: global analysis of Kimura's continuum-of-alleles model. *J. Math. Biol.* 24:341–351.

Canestrelli, D., R. Bisconti, and C. Carere. 2016. Bolder Takes All? The Behavioral Dimension of Biogeography. *Trends in Ecology and Evolution* 31:35–43.

Champagnat, N., R. Ferrière, and S. Méléard. 2006. Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. *Theoretical Population Biology* 69:297–321.

Champagnat, N. and A. Lambert. 2007. Evolution of discrete populations and the canonical diffusion of adaptive dynamics. *Theor. Popul. Biol.* 17:102–155.

Charlesworth, B. 1990. Optimization Models, Quantitative Genetics, and Mutation. *Evolution (N. Y).* 44:520–538.

Charlesworth, B. 1994. Evolution in Age-Structured Populations. Cambridge University Press, Cambridge, 2th edn.

Chaturvedi, A., D. Croll, M. C. Fischer, F. Guillaume, S. Karrenberg, B. Kerr, G. Rolshausen, and J. Stapley. 2017. Can Evolution Supply What Ecology Demands? *Trends in Ecology and Evolution* 32:187–197.

Chesson, P. L. 1981. Models for spatially distributed populations: The effect of within-patch variability. *Theor. Popul. Biol.* 19:288–325.

Chesson, P. L. 1984. Persistence of a Markovian population in a patchy environment. *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete* 66:97–107.

Cheverud, J. M. 1985. A Quantitative Genetic Model of Altruistic Selection. *Behav. Ecol. Sociobiol.* 16:239–243.

Christiansen, F. B. 1991. On Conditions for Evolutionary Stability for a Continuously Varying Character. *Am. Nat* 138:37–50.

Connallon, T. 2015. The geography of sex-specific selection, local adaptation, and sexual dimorphism. *Evolution* 69:2333–2344.

Cordero, O. X. and M. F. Polz. 2014. Explaining microbial genomic diversity in light of evolutionary ecology. *Nature Publishing Group* 12:263–273.

Daley, D. J. and D. Vere-Jones. 2003. An Introduction to the Theory of Point Processes, vol. II of *Probability and Its Applications*. Springer New York.

Dall, S. R. X., A. I. Houston, and J. M. McNamara. 2004. The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecology Letters* 7:734–739.

Day, T. 2001a. Population structure inhibits evolutionary diversification under competition for resources. *Genetica* 112:71–86.

Day, T. 2001b. Population structure inhibits evolutionary diversification under competition for resources. *Genetica* 112-113:71–86.

Day, T. and S. R. Proulx. 2004. A general theory for the evolutionary dynamics of virulence. *Am. Nat* 163:E40–E63.

Day, T. and P. D. Taylor. 1996. Evolutionarily stable versus fitness maximizing life histories under frequency-dependent selections. *Proceedings of the Royal Society B: Biological Sciences* 263:333–338.

Débarre, F., S. L. Nuismer, and M. Doebeli. 2014. Multidimensional (co)evolutionary stability. *Am. Nat.* 184:158–171.

Débarre, F. and S. P. Otto. 2016. Evolutionary dynamics of a quantitative trait in a finite asexual population. *Theor. Popul. Biol.* 108:75–88.

Dercole, F. and S. Rinaldi. 2008. Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications. Princeton University Press, Princeton, NJ.

Dieckmann, U. and R. Law. 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* 34:579–612.

Dingemanse, N. J., N. A. Dochtermann, and S. Nakagawa. 2012. Defining behavioural syndromes and the role of 'syndrome deviation' in understanding their evolution. *Behav. Ecol. Sociobiol.* 66:1543–1548.

Edelaar, P. and D. I. Bolnick. 2012. Non-random gene flow: An underappreciated force in evolution and ecology. *Trends Ecol. Evol.* 27:659–665.

Eshel, I. 1983. Evolutionary and Continuous Stability. *Journal of Theoretical Biology* 103:99–111.

Eshel, I. and M. W. Feldman. 1984. Initial Increase of New Mutants and Some Continuity Properties of ESS in Two-Locus Systems. *Am. Nat* 124:631.

Eshel, I. and U. Motro. 1981. Kin selection and strong evolutionary stability of mutual help. *Theoretical Population Biology* 19:420–433.

Eshel, I., U. Motro, and E. Sansone. 1997. Continuous Stability and Evolutionary Convergence. *Journal of Theoretical Biology* 185:333–343.

Fehr, E. and S. Gächter. 2000. Cooperation and punishment in public goods experiments. *The American Economic Review* 90:980–994.

Fleming, W. H. 1979. Equilibrium Distributions of Continuous Polygenic Traits. *SIAM Journal on Applied Mathematics* 36:148–168.

Frank, S. A. 1998. Foundations of Social Evolution. Monogr. Behav. Ecol. Princeton University Press, Princeton.

Frank, S. A. and P. D. Taylor. 1996. How to make a kin selection model, vol. 180. Journal of Theoretical Biology.

Gandon, S. 1999. Kin competition, the cost of inbreeding and the evolution of dispersal. *J. Theor. Biol.* 200:345–364.

Gardner, A. and S. A. West. 2004. Cooperation and Punishment, Especially in Humans. *Am. Nat* 164:753–764.

Geritz, S., É. Kisdi, G. Meszéna, and J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* 12:35–57.

Geritz, S. A. H., J. a. J. Metz, and C. Rueffler. 2016. Mutual invadability near evolutionarily singular strategies for multivariate traits, with special reference to the strongly convergence stable case. *J. Math. Biol.* 72:1081–1099.

Gomulkiewicz, R. and M. Kirkpatrick. 1992. Quantitative Genetics and the Evolution of Reaction Norms. *Evolution* 46:390.

Grafen, A. 1985. A geometric view of relatedness. In Dawkins, R. and M. Ridley (eds.), *Oxford Surveys in Evolutionary Biology*, pp. 28–90. Oxford University Press, Oxford.

Guillaume, F. 2011. Migration-induced phenotypic divergence: the migration-selection balance of correlated traits. *Evolution* 65:1723–1738.

Guillaume, F. and M. C. Whitlock. 2007. Effects of migration on the genetic covariance matrix. *Evolution* 61:2398–2409.

Hamilton, W. D. 1964. The genetical evolution of social behaviour. I. *J. Theor. Biol.* 7:1–16.

Harris, T. E. 1963. The Theory of Branching Processes, vol. 6 of *Umn*. Courier Corporation, Mineola.

Hines, W. G. S. 1980. Three characterizations of population strategy stability. *Journal of applied probability* 17:333–340.

Hochberg, M. E., D. J. Rankin, and M. Taborsky. 2008. The coevolution of cooperation and dispersal in social groups and its implications for the emergence of multicellularity. *BMC Evol. Biol.* 8:238.

Horn, R. A. and C. R. Johnson. 1994. Topics in Matrix Analysis. Cambridge University Press.

Iwasa, Y. and A. Pomiankowski. 1995. Continual Change in Mate Preferences. *Nature* 377:420–422.

Iwasa, Y., A. Pomiankowski, and S. Nee. 1991. The evolution of costly mate preferences II. The 'handicap' principle. *Evolution* 45:1431–1442.

Jones, A. G., S. J. Arnold, and R. Burger. 2004. Evolution and Stability of the G-Matrix on a Landscape with a Moving Optimum. *Evolution (N. Y)*. 58:1639.

Jones, A. G., R. Bürger, and S. J. Arnold. 2014. Epistasis and natural selection shape the mutational architecture of complex traits. *Nat. Commun.* 5:3709.

Karlin, S. 1968. Equilibrium Behaviour of Population Genetic Models with Non-Random Mating. *J. Appl. Probab.* 5:231–313.

Kimura, M. 1965a. A stochastic model concerning the maintenance of genetic variability in quantitative character. *Proceedings of the National Academy of Sciences of the United States of America* 54:731–736.

Kimura, M. 1965b. Attainment of quasi-linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* 52:875–890.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. *Genetics* 161:1727–1750.

Kisdi, E. 2016. Dispersal polymorphism in stable habitats. *Journal of Theoretical Biology* 392:69–82.

Kisdi, E. and S. A. H. Geritz. 2009. Adaptive dynamics: a framework to model evolution in the ecological theatre. *J. Math. Biol.* 61:165–169.

Koella, J. C. 2000. The spatial spread of altruism versus the evolutionary response of egoists. *Proceedings of the Royal Society of London Series B-Biological Sciences* 267:1979–1985.

Kuijper, B. and R. A. Johnstone. 2017. How Sex-Biased Dispersal Affects Sexual Conflict over Care. *Am. Nat* 189:501–514.

Lande, R. 1976. Natural-Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution (N. Y).* 30:314–334.

Lande, R. 1979. Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. *Evolution* 33:402.

Lande, R. 1980. The Genetic Covariance between Characters Maintained by Pleiotropic Mutations. *Genetics* 94:203–215.

Lande, R. 1981. Models of Speciation by Sexual Selection on Polygenic Traits. *Proc. Natl. Acad. Sci. U. S. A.* 78:3721–3725.

Lande, R. 1984. The genetic correlation between characters maintained by selection, linkage and inbreeding. *Genet. Res.* 44:309–320.

Lande, R. 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution* 46:381–389.

Lande, R. and S. J. Arnold. 1983. The Measurement of Selection on Correlated Characters. *Evolution* 37:1210–1226.

Latter, B. 1970. Selection in Finite Populations with Multiple Alleles .2. Centripetal Selection, Mutation, and Isoallelic Variation. *Genetics* 66:165–&.

Lehmann, L., I. Alger, and J. Weibull. 2015. Does evolution lead to maximizing behavior? *Evolution* 69:1858–1873.

Lehmann, L., C. Mullon, E. Akçay, and J. Van Cleve. 2016. Invasion fitness, inclusive fitness, and reproductive numbers in heterogeneous populations. *Evolution* 70:1689–1702.

Lehmann, L. and N. Perrin. 2002. Altruism, dispersal and phenotype-matching kin recognition. *American Naturalist* 159:451–468.

Lehmann, L. and F. Rousset. 2014. The genetical theory of social behaviour. *Philosophical transactions of the Royal Society B* 369:1–18.

Lehmann, L., F. Rousset, D. Roze, and L. Keller. 2007. Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. *The American Naturalist* 170:21–36.

Leimar, O. 2005. The evolution of phenotypic polymorphism: randomized strategies versus evolutionary branching. *Am. Nat.* 165:669–681.

Leimar, O. 2009. Multidimensional convergence stability. *Evolutionary Ecology Research* 11:191–208.

Leturque, H. and F. Rousset. 2004. Intersexual competition as an explanation for sex-ratio and dispersal biases in polygynous species. *Evolution (N. Y).* 58:2398–2408.

Lion, S. 2018. Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a Unifying Perspective. *Am. Nat* 191:21–44.

Lynch, M. and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer, Massachusetts.

McGlothlin, J. W., J. B. Wolf, E. D. Brodie, and A. J. Moore. 2014. Quantitative genetic versions of Hamilton's rule with empirical applications. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130358–20130358.

Metz, J., S. Geritz, G. Meszéna, and F. Jacobs. 1996. Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction. In van Strien, S. J. and S. M. V. Lunel (eds.), *Stochastic and Spatial Structures of Dynamical Systems*. North-Holland, Amsterdam.

Metz, J. a. J. 2011. Thoughts on the geometry of meso-evolution: collecting mathematical elements for a post-modern synthesis. In Chalub, F. A. C. C. and J. Rodrigues (eds.), *The Mathematics of Darwin's Legacy*, pp. 193–231. Birkhäuser, Basel.

Meyn, S. and R. L. Tweedie. 2009. Markov Chains and Stochastic Stability. Cambridge University Press, Cambridge, 2th edn.

Michod, R. E. 1979. Evolution of Life Histories in Response to Age-Specific Mortality Factors. *Am. Nat* 113:531–550.

Montiglio, P. O., C. Ferrari, and D. Reale. 2013. Social niche specialization under constraints: personality, social interactions and environmental heterogeneity. *Philos. Trans. R. Soc. B Biol. Sci.* 368:20120343–20120343.

Morale, D., V. Capasso, and K. Oelschläger. 2005. An interacting particle system modelling aggregation behavior: from individuals to populations. *J. Math. Biol.* 50:49–66.

Mullon, C., L. Keller, and L. Lehmann. 2016. Evolutionary stability of jointly evolving traits in subdivided populations. *Am. Nat.* 188:175–195.

Mullon, C., L. Keller, and L. Lehmann. 2018. Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nature Ecology & Evolution* 2:132–140.

Nadell, C. D., J. B. Xavier, and K. R. Foster. 2009. The sociobiology of biofilms. *FEMS Microbiology Reviews* 33:206–224.

Nagylaki, T. 1993. The evolution of multilocus systems under weak selection. *Genetics* 134:627–647.

Nuismer, S. L., R. Gomulkiewicz, and B. J. Ridenhour. 2010. When Is Correlation Coevolution? *Am. Nat* 175:525–537.

Ohtsuki, H. 2010. Evolutionary games in Wright's island model: kin selection meets evolutionary game theory. *Evolution* 64:3344–3353.

Parker, G. A. and J. Maynard Smith. 1990. Optimality theory in evolutionary biology. *Science* 349:27–33.

Parvinen, K., H. Ohtsuki, and J. Y. Wakano. 2017. The effect of fecundity derivatives on the condition of evolutionary branching in spatial models. *Journal of Theoretical Biology* 416:129–143.

Parvinen, K., H. Ohtsuki, and J. Y. Wakano. 2018. Spatial heterogeneity and evolution of fecundity-affecting traits. *Journal of Theoretical Biology* 454:190–204.

Perrin, N. and V. Mazalov. 2000. Local Competition, Inbreeding, and the Evolution of Sex-Biased Dispersal. *Am. Nat* 155:116–127.

Phillips, P. C. and S. J. Arnold. 1989. Visualizing Multivariate Selection. *Evolution* 43:1209.

Price, G. R. 1970. Selection and covariance. *Nature* 227:520–521.

Purcell, J., A. Brelsford, and L. Avilés. 2012. Co-evolution between sociality and dispersal: The role of synergistic cooperative benefits. *Philos. Trans. R. Soc. B Biol. Sci.* 312:44–54.

Queller, D. C. 1992a. A General-Model for Kin Selection. *Evolution (N. Y).* 46:376–380.

Queller, D. C. 1992b. Quantitative genetics, inclusive fitness, and group selection. *Am. Nat* 139:540.

Raihani, N. J., A. Thornton, and R. Bshary. 2012. Punishment and cooperation in nature. *Trends in Ecology and Evolution* 27:288–295.

Reuter, M. and L. Keller. 2001. Sex ratio conflict and worker production in eusocial Hymenoptera. *American Naturalist* 158:166–177.

Roff, D. 1997. Evolutionary Quantitative Genetics. Evolutionary biology : Genetics. Springer US.

Roff, D. A. and D. J. Fairbairn. 2012. A test of the hypothesis that correlational selection generates genetic correlations. *Evolution* 66:2953–2960.

Ronce, O. and J. Clobert. 2012. Dispersal syndromes. In *Dispersal Ecology and Evolution*, pp. 119–138. Oxford University Press.

Rousset, F. 2002. Inbreeding and relatedness coefficients: what do they measure? *Heredity (Edinb).* 88:371–380.

Rousset, F. 2004. Genetic Structure and Selection in Subdivided Populations. Princeton University Press, Princeton, NJ.

Rousset, F. and S. Gandon. 2002. Evolution of the distribution of dispersal distance under distance-dependent cost of dispersal. *J. Evol. Biol.* 15:515–523.

Roze, D. and F. Rousset. 2005. Inbreeding depression and the evolution of dispersal rates: a multilocus model. *Am. Nat.* 166:708–721.

Roze, D. and F. Rousset. 2008. Multilocus models in the infinite island model of population structure. *Theoretical Population Biology* 73:529–542.

Rueffler, C., T. J. M. Van Dooren, O. Leimar, and P. A. Abrams. 2006. Disruptive selection and then what? *Trends in Ecology and Evolution* 21:238–245.

Saltz, J. B., F. C. Hessel, and M. W. Kelly. 2017. Trait Correlations in the Genomics Era. *Trends in Ecology and Evolution* 32:1–12.

Sasaki, A. and U. Dieckmann. 2011. Oligomorphic dynamics for analyzing the quantitative genetics of adaptive speciation. *J. Math. Biol.* 63:601–635.

Sih, A., J. Cote, M. Evans, S. Fogarty, and J. Pruitt. 2012. Ecological implications of behavioural syndromes. *Ecology Letters* 15:278–289.

Slatkin, M. 1980. Ecological Character Displacement. *Ecology* 61:163–177.

Steppan, S. J., P. C. Phillips, and D. Houle. 2002. Comparative quantitative genetics: Evolution of the G matrix. *Trends Ecol. Evol.* 17:320–327.

Svardal, H., C. Rueffler, and J. Hermisson. 2015. A general condition for adaptive genetic polymorphism in temporally and spatially heterogeneous environments. *Theoretical Population Biology* 99:76–97.

Taper, M. L. and T. J. Case. 1992. Models of character displacement adn the theoretical robustness of taxon cycles. *Evolution* 46:317–333.

Taper, M. L. and T. J. Chase. 1985. Quantitative Genetic Models for the Coevolution of Character Displacement. *Ecology* 66:355–371.

Taylor, P. and T. Day. 1997. Evolutionary stability under the replicator and the gradient dynamics. *Evol. Ecol.* 11:579–590.

Taylor, P. D. 1989. Evolutionary Stability in One-Parameter Models Under Weak Selection. *Theoretical Population Biology* 36:125–143.

Tazzyman, S. J. and Y. Iwasa. 2009. Sexual Selection Can Increase the Effect of Random Genetic Drift-a Quantitative Genetic Model of Polymorphism in Oophaga Pumilio, the Strawberry Poison-Dart Frog. *Evolution* 64:1719–1728.

Turelli, M. 1985. Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* 111:165–195.

Turelli, M. and N. H. Barton. 1990. Dynamics of polygenic characters under selection. *Theor. Popul. Biol.* 38:1–57.

Turelli, M. and N. H. Barton. 1994. Genetic and Statistical-Analyses of Strong Selection on Polygenic Traits - What, Me Normal. *Genetics* 138:913–941.

van Baalen M. 2013. The unit of adaptation, the emergence of individuality, and the loss of evolutionary sovereignty. In Huneman, P. and F. Bouchard (eds.), *From Groups to Individuals. Evolution and Emerging Individuality*, pp. 117–140. MIT Press.

Van Cleve, J. 2015. Social evolution and genetic interactions in the short and long term. *Theor. Popul. Biol.* 103:2–26.

Wakano, J. Y. and Y. Iwasa. 2013. Evolutionary branching in a finite population: deterministic branching vs. stochastic branching. *Genetics* 193:229–241.

Wakano, J. Y. and L. Lehmann. 2014. Evolutionary branching in deme-structured populations. *J. Theor. Biol.* 351:83–95.

Wenseleers, T. 2010. Social evolution theory: a review of methods and approaches. In Szekely, T., A. Moore, and J. Komdeur (eds.), *Soc. Behav. Genes, Ecol. Evol.*, pp. 132–158. Cambridge University Press, Cambridge.

West, S. A., A. S. Griffin, A. Gardner, and S. P. Diggle. 2006. Social evolution theory for microorganisms. *Nature Reviews Microbiology* 4:597–607.

Wolf, M. and F. J. Weissing. 2012. Animal personalities: consequences for ecology and evolution. *Trends in Ecology and Evolution* 27:452–461.

Wolfram Research, Inc. 2016. Mathematica. Wolfram Research, Inc., Champaign, Illinois.

Wright, C. M., C. T. Holbrook, and J. N. Pruitt. 2014. Animal personality aligns task specialization and task proficiency in a spider society. *Proc. Natl. Acad. Sci. U. S. A.* 111:9533–9537.

Wright, S. 1931. Evolution in Mendelian Populations. *Genetics* 16:290.

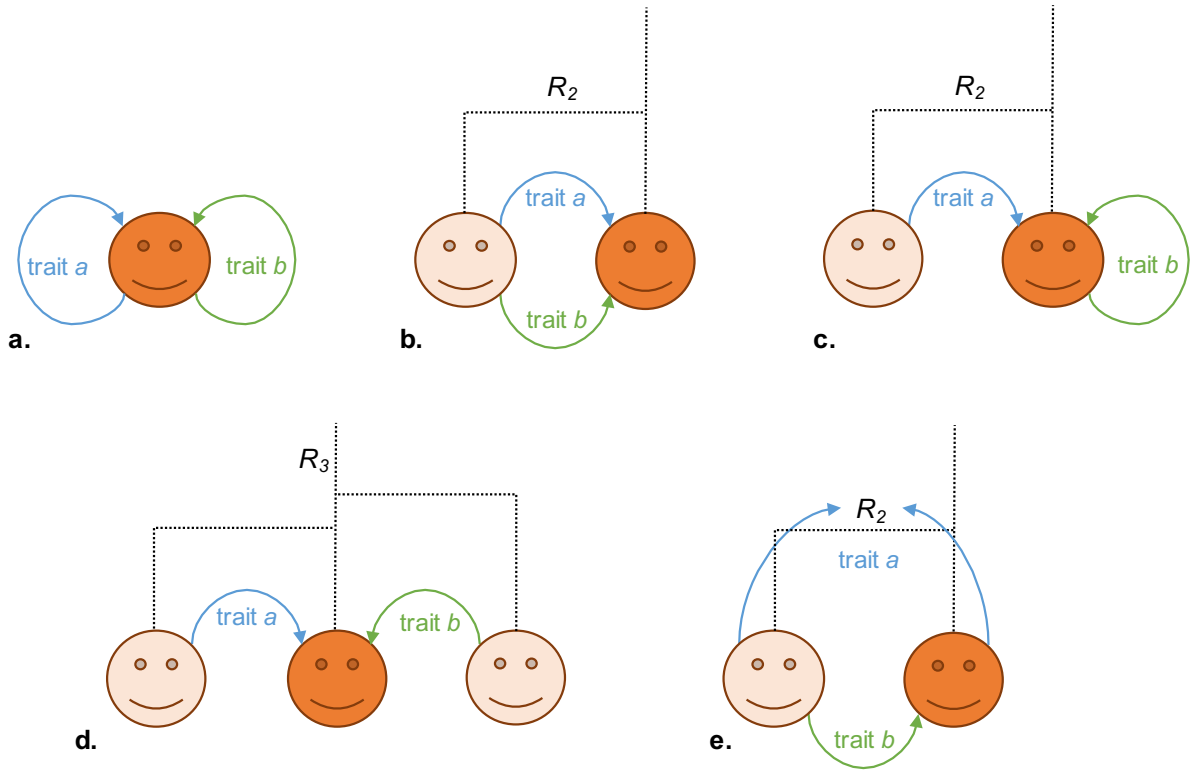

Figure 1: Within- and between-individual trait effects relevant for correlational selection. As revealed by eq. (9), there are five types of fitness effects due to perturbations in two traits a and b that are relevant for correlational selection when dispersal is limited: **a.** effect of a joint changes in a and b within the focal individual (first term of eq. 9b); **b.** effect of joint changes in a and b within neighbours of the focal (second term of eq. 9b, weighted by the probability that a neighbour and the focal are identitcal-by-descent in the absence of selection, $r_2^\circ(\bar{\mathbf{z}}_t)$); **c.** effect of joint changes in a and b between the focal (here b) and its neighbours (here, a; second line of eq. 9b, weighted by $r_2^\circ(\bar{\mathbf{z}}_t)$); **d.** effect of joint changes in a and b between neighbours of the focal (third line of eq. 9b, weighted by the probability that two neighbours and the focal are identitcal-by-descent in the absence of selection, $r_3^\circ(\bar{\mathbf{z}}_t)$); **e.** the effect of the indirect effect of one trait (here b) combined with the effect of the other (here a) on pairwise identity-by-descent, which reflects the tendency of relatives to receive the effects of b (eq. 9c).

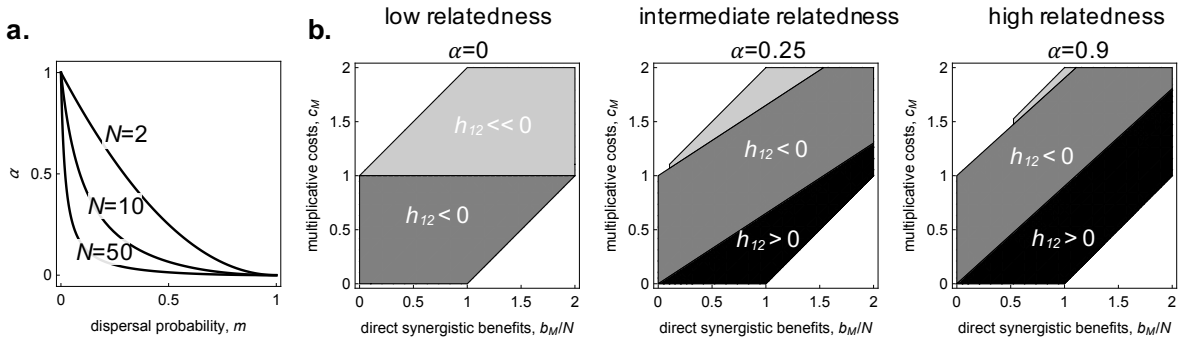

Figure 2: Correlational selection on synergistic social traits. a. Factor α to synergistic benefits in coefficient correlational selection (i.e., the greater α is, the more correlational selection depends on the synergistic effects of both traits, see eq. 27; Figure computed by substituting eq. 18 into eq. 28). **b.** Parameter combinations for which correlational selection at the equilibrium $(0,0)$ is: (1) strongly negative (and causes selection to be disruptive, $h_{12}(\bar{\mathbf{z}}^*) \ll 0$, i.e., eq. 26 does not hold, light gray regions); (2) negative (and selection is stabilizing, $h_{12}(\bar{\mathbf{z}}^*) < 0$, i.e., eq. 26 holds, dark gray regions); and (3) positive (and selection is stabilizing, $h_{12}(\bar{\mathbf{z}}^*) > 0$, black regions). White regions correspond to parameter combinations under which the equilibrium is not evolutionary convergent (i.e., eq. 24 does not hold).

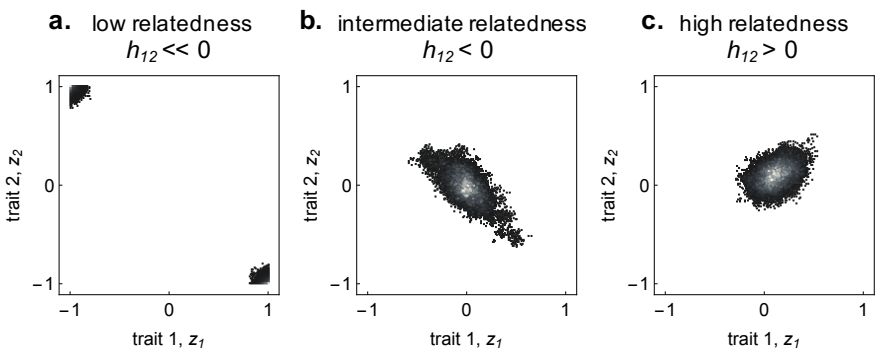

Figure 3: The effect of correlational selection on the phenotypic distribution due to relatedness and synergy. Density distribution of a simulated population, initially monomorphic for $(0,0)$, sampled every 500 generations over 20'000 generations after 30'000 generations of evolution (population composed of 1000 groups of size $N = 10$, other parameters: $c = .1$, $b_M/N = .1$, $\nu = 0.01$, $\sigma^m_{11} = \sigma^m_{22} = 0.02$, $\sigma^m_{12} = 0$; see Appendix D for details on simulation procedure). **a**. Strong negative association with social polymorphism (with $m = 0.8$ and $c_M = 0.005$). **b**. Negative association ($m = 0.5$ and $c_M = .1$). **c**. Positive association ($m = 0.05$ and $c_M = 0.005$).

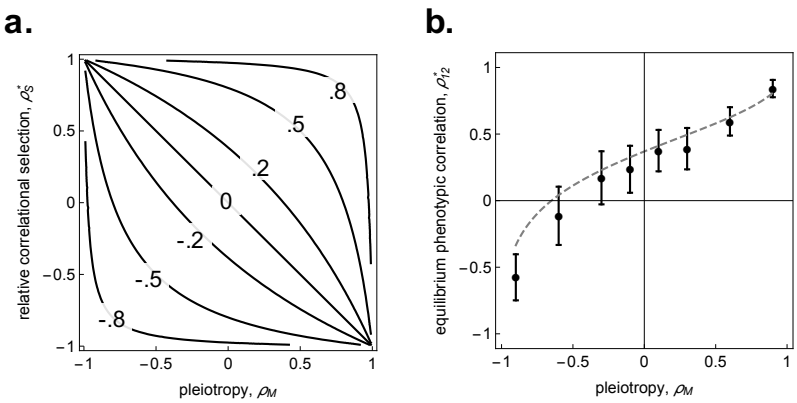

Figure 4: The effect of pleiotropy on phenotypic correlation at equilibrium. **a**. Predicted phenotypic correlation among traits 1 and 2 according to pleiotropy, ρ_{m}, and the relative strength of selection, ρ_{s}^* (from eq. 31). **b**. Predicted phenotypic correlation among traits 1 and 2 (dashed grey curve, from eq. 31), and corresponding observations from individual based simulations of a population initially monomorphic for $(0,0)$ divided among 1000 groups of size $N = 10$ (black dots, averaged correlation over 10'000 generations after a burnin 40'000 generations of evolution, error bars indicate standard deviation; other parameters: $c = .1$, $b_{\mathrm{M}}/N = .1$, $c_{\mathrm{M}} = 0.005$, $m = 0.05$, $v = 0.01$, $\sigma_{11}^{\mathrm{m}} = \sigma_{22}^{\mathrm{m}} = 0.02$; see Appendix D for details).

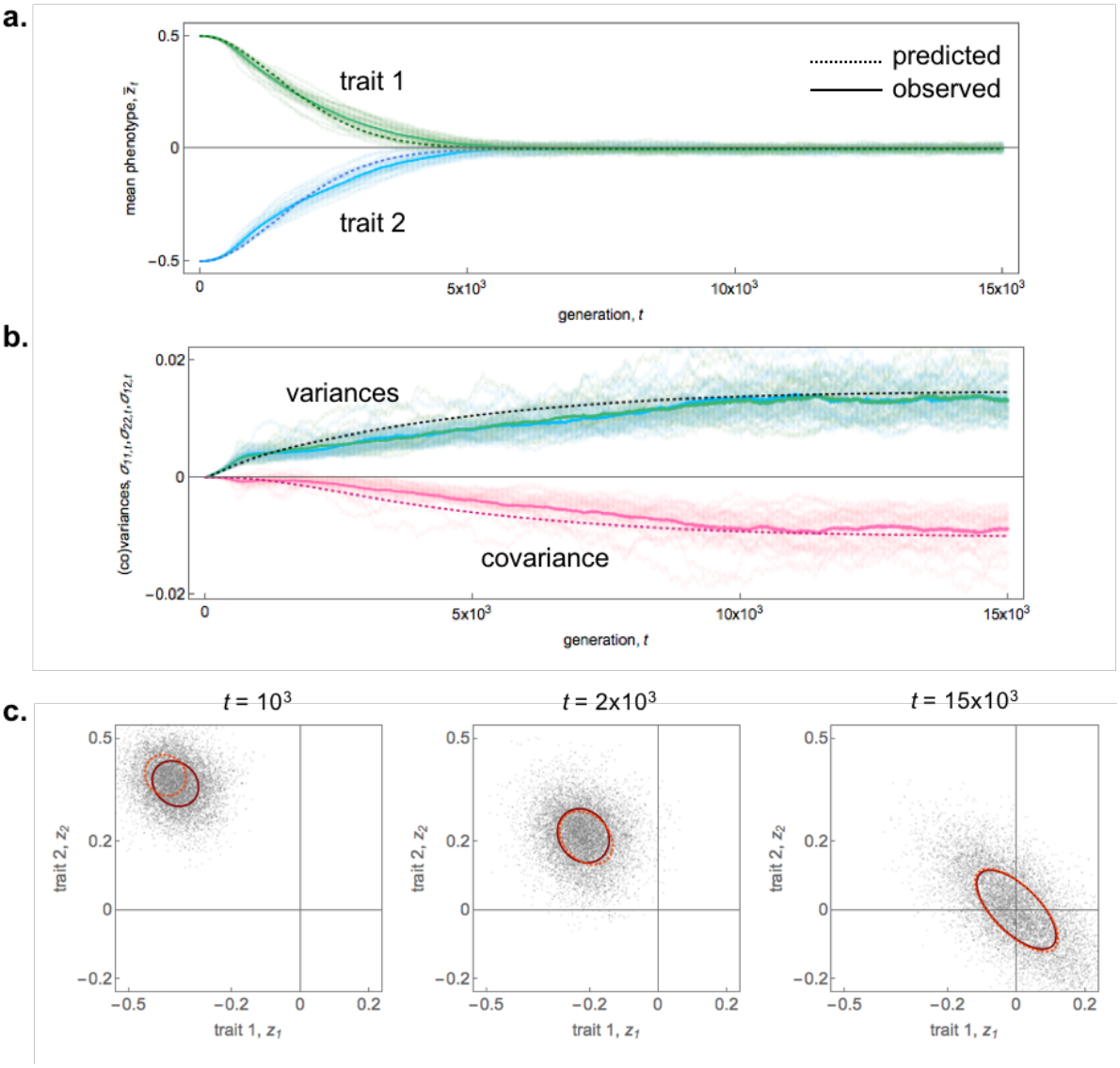

Figure 5: Observed and predicted evolution of the phenotypic distribution. The observed (full lines, from individual based simulations) and predicted (dashed lines, from eq. 5) evolution of the traits' means (**a**. trait 1 in blue and 2 in green), variances (**b**. trait 1 in blue and 2 in green) and covariance (**b**. pink) for 30 replicates (each replicate in lighter shade, average over replicates in darker shade, initial population monomorphic with $z_1 = -0.5$ and $z_2 = 0.5$, distributed over 5000 groups of size $N = 10$, other parameters: $c = .1$, $b_M/N = .1$, $c_M = .1$, $m = 0.5$, $\nu = 0.01$, $\sigma^m_{11} = \sigma^m_{22} = 0.02$, $\sigma^m_{12} = 0$; see Appendix D for details). **c**. Snapshot of the population (10'000 individuals randomly sampled across 30 replicates shown by grey points) and variance-covariance ellipses given by the (right) eigenvectors of the **G** matrix (observed in full lines and predicted in dashed), at generations: 1000 (left panel); 2000 (middle panel); and 15000 (right panel).

Appendix

The main goal of this (lengthy) appendix is to derive the recurrence equations for the mean and variance to the second order of selection intensity (eqs. 5 with eqs. 8 and 9 of the main text). To that end, we first derive in Appendix A the dynamics of the phenotypic distribution at the group level and then at the individual level to reach eq. (1), which depends on lineage fitness. Second, in Appendix B, we approximate lineage fitness to the second order in order to allow the consideration of selection on phenotypic (co)variances. Our derivations rely on the notion of counting measures, which allow for a more abstract representation of Markov chains (Daley and Vere-Jones, 2003) than usually encountered in evolutionary biology (but see Morale et al., 2005, Champagnat et al., 2006, Champagnat and Lambert, 2007). We would like to point out that we resorted to this formalism after trying to carry out our derivations using the Price equation (Price, 1970, as in Wakano and Lehmann, 2014), but this turned out to be more complex than expected. In appendix C, we use the so-obtained approximation to derive a second order approximation for the dynamics of the mean vector and variance-covariance matrix, and then show that these dynamics can be expressed as eqs. (5) with eqs. (8) and (9) under the assumption that the distribution is Normal. Finally, in appendix D, we describe the simulation procedure used to generate Figures 3-5. A detailed table of contents of the appendix is as follows.

Contents of appendix

A Phenotypic distribution dynamics **2**

A.1 Phenotypic distribution dynamics at the group level . . . 2

A.1.1 Markov chain at the group level . . . 2

A.1.2 Markov chain in terms of counting measures . . . 2

A.1.3 State dynamics . . . 3

A.2 Phenotypic distribution dynamics at the individual level . . . 4

A.3 Weak mutation and selection . . . 7

B Second-order approximation of lineage fitness **7**

B.1 Conditional individual fitness approximation . . . 8

B.1.1 First-order effects on individual fitness . . . 8

B.1.2 Second-order effects on individual fitness . . . 10

B.1.3 Effect of joint trait changes in the same individual, $\beta_{ab,t}^{\text{same}}$ . . . 11

B.1.4 Effect of joint trait changes in neighbours, $\beta_{ab,t}^{\text{neigh}}$ . . . 11

B.1.5 Effect of joint trait changes in two individuals from different groups, $\beta_{ab,t}^{\text{cross isle}}$ . . . 13

B.1.6 Second order approximation of individual fitness . . . 13

B.2 Phenotypic covariance among neighbours . . . 14

B.2.1 Conditional phenotypic distribution in neighbors . . . 14

B.2.2 Relatedness, identity-by-descent & neutral phenotypic covariance among neighbours. . . 16

B.2.3 Phenotypic covariance among neighbours under selection . . . 16

B.3 Second-order approximation of lineage fitness . . . 16
B.3.1 Average individual fitness effect of one trait change . . . 17
B.3.2 Average individual fitness effects of a joint change in two traits . . . 17
B.3.3 Polynomial approximation to lineage fitness . . . 19

C Dynamics of phenotypic mean and (co)variance **20**
C.1 General recursions for the mean and (co)variance . . . 20
C.1.1 Dynamics of mean phenotype . . . 20
C.1.2 Dynamics of the phenotypic (co)variance . . . 21
C.2 Moment closure . . . 22
C.2.1 Normal closure . . . 23
C.2.2 Quasi-equilibrium . . . 23

D Individual-based simulations **24**

A Phenotypic distribution dynamics

A.1 Phenotypic distribution dynamics at the group level

A.1.1 Markov chain at the group level

The set of phenotypic values $\{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$, where $\mathbf{z}_i = \{z_{i,1}, ..., z_{i,n}\} \in \mathbb{R}^n$ is the phenotype of individual indexed $i \in \{1, \ldots, N\}$ residing in a focal group at a given generation (where $z_{i,a}$ is the value of a in individual i), gives the phenotypic state in such a group. We call this the *state* of the group. We are interested in describing how the state changes from one demographic time period to the next (i.e., over one iteration of the life-cycle, section 2).

The state in each group changes stochastically from one time period to the next, and we assume it does so according to a discrete time Markov chain on a continuous state space (since we are considering the evolution of continuous traits, see Meyn and Tweedie, 2009, for a formulation of Markov chain on general state spaces). Because groups affect each other through dispersal, the transition kernel for the Markov chain on the focal group depends on the state of all the other groups. But since there is an infinite number of groups and there is no isolation-by-distance (i.e., all groups are connected to one another through dispersal), the infinite set of interacting Markov chains (one for each group) can be described as a single Markov chain, whose transition kernel is a function of the expected value of the process (Chesson, 1981, 1984). In other words, we can ignore the stochasticity stemming from groups other than the focal one, and the phenotypic dynamics in the focal group captures the phenotypic dynamics at the population level.

A.1.2 Markov chain in terms of counting measures

Because we are not interested in knowing which precise individual bears which phenotype, the space Ω of group states that we are considering is the space of all unordered sets with N vectors in $\mathbb{R}^n$ (i.e., the order of

elements in $\{\mathbf{z}_1, \ldots, \mathbf{z}_N\}$ does not matter). From this, we can represent the state of the focal group by a function that is called a *counting measure* and that will be mathematically convenient for our analysis. The counting measure $\mu : \wp(\mathbb{R}^n) \to \mathbb{Z}_+$, which takes elements from the power set $\wp(\mathbb{R}^n)$ (the set of all measurable subsets of $\mathbb{R}^n$, including the empty set and $\mathbb{R}^n$ itself) and sends them to non-negative integers $\mathbb{Z}_+$, counts the number of individuals within the focal group with given phenotypes. More precisely, applied to a subset $E \subseteq \wp(\mathbb{R}^n)$, $\mu(E)$ returns the number of individuals in the focal group with phenotypes that belong to E according to the following definition

$$\mu(E) = \sum_{i=1}^{N} \delta_{\mathbf{z}_i}(E), \tag{A-1}$$

where δ is the dirac measure,

$$\delta_{\mathbf{z}_i}(E) = \begin{cases} 1, & \text{if } \mathbf{z}_i \in E; \\ 0, & \text{otherwise,} \end{cases} \tag{A-2}$$

(p. 51 of Harris, 1963 and p. 3 of Daley and Vere-Jones, 2003). Applied to a single phenotypic value $\mathbf{z} = \{z_1, ..., z_n\} \in \mathbb{R}^n$, where z_a is the value of trait a, $\mu(\mathbf{z})$ returns the number of individuals with phenotype $\mathbf{z}$, and applied to the whole set of phenotypic values, $\mu(\mathbb{R}^n) = N$, it gives the size of the group. The application of μ to a subset $E \subseteq \wp(\mathbb{R}^n)$ can be equivalently written as,

$$\mu(E) = \int_E \mu(\mathrm{d}\mathbf{z}) = \int_E \mu(\mathbf{z})\mathrm{d}\mathbf{z}. \tag{A-3}$$

A counting measure uniquely determines the state of a group, i.e., for each unordered set of N vectors in $\mathbb{R}^n$, there exists a unique counting measure (p. 52 of Harris, 1963 and p. 7 of Daley and Vere-Jones, 2003). We can therefore study the dynamics of the state of the focal group by studying the dynamics of the equivalent counting measure (Daley and Vere-Jones, 2003, p. 13-15). Thus, if S_t denotes the (random) counting measure of a focal group at time period t, we study the Markov chain $\{S_t\}$ on the space of all finite counting measures, which we write as $\mathcal{S}$. The equivalence between state and counting measure allows us to use the terms “state" and “counting measure" interchangeably. This type of construction, which are analysed using the theory of point processes (e.g., Daley and Vere-Jones, 2003, Harris, 1963, p. 51 for rigorous approaches), has already been used to study phenotypic evolution, in particular in populations that are well-mixed and when time is continuous (e.g., Champagnat et al., 2006, Champagnat and Lambert, 2007, but see Morale et al., 2005, for populations in explicit space).

A.1.3 State dynamics

In order to describe the stochastic dynamics of the state of a focal group, we let

$$\phi_t(T) = \Pr[S_t \in T], \tag{A-4}$$

denote the probability that the focal group is in a state that belongs to a subset $T \subseteq \mathcal{S}$ at time period t (equivalent to eq. 3.3 of Harris, 1963, p. 55). We write the Markov kernel transition function as

$$P_t(\mu, \phi_t; T) = \Pr[S_{t+1} \in T | S_t = \mu, \phi_t], \tag{A-5}$$

which is the probability that a group will be in a state that belongs to a subset $T \subseteq \mathcal{S}$ at time period $t+1$, given that it was in state μ at time period t and that the population distribution of states is ϕ_t (eq. 6.1 of Harris, 1963, p. 60). We are therefore considering a non-homogeneous Markov chain, i.e., the transition kernel of the focal

group depends on the distribution of states, ϕ_t. This representation is feasible owing to our assumptions that i) there are is infinite number of groups; ii) none are isolated by distance; and iii) the initial conditions for all groups are exchangeable random variables (Chesson, 1981, 1984).

We can then write the probability that the focal group is in a state that belongs to $T \subseteq \mathcal{S}$ at time period $t+1$ as

$$\phi_{t+1}(T) = \int_{\mathcal{S}} P_t(\mu, \phi_t; T)\phi_t(\mathrm{d}\mu) = \int_{\mathcal{S}} \int_{T} P_t(\mu, \phi_t; \mathrm{d}\mu')\phi_t(\mathrm{d}\mu) = \int_{\mathcal{S}} \int_{T} P_t(\mu, \phi_t; \mu')\phi_t(\mu)\mathrm{d}\mu'\mathrm{d}\mu. \tag{A-6}$$

Because one iteration of the life cycle of our population encompasses many events, like selection, mutation, and dispersal, the transition kernel for our model is difficult to characterise (studies like Champagnat et al., 2006, are capable of constructing explicit transition kernels by considering time steps small enough so that only one event can occur per step). We therefore seek to model the evolutionary process in a more practical way. We will achieve this by focusing on the distribution of phenotypes at the individual level rather than on the distribution of group states. In so doing, we will repeatedly make use of the conditional expectation of arbitrary functions g that depend on the state of the group at time period $t+1$,

$$\mathrm{E}[g(T)|S_t = \mu] = \int_{T} g(\mu')P_t(\mu, \phi_t; \mu')\mathrm{d}\mu', \tag{A-7}$$

given that the group state at time period t is μ (see theorem 9.1 of Harris, 1963, p.65 for proof that such expectation exists and is well-defined).

A.2 Phenotypic distribution dynamics at the individual level

Having formalised the phenotypic state dynamics at the group level (eq. A-6), our aim is now to derive the dynamical equations for the phenotypic distribution at the individual level. We start by observing that the probability density function (p.d.f) that an individual randomly sampled in the focal group carries phenotype $\mathbf{z}$ at time period t is

$$p_t(\mathbf{z}) = \int_{\mathcal{S}} \frac{\mu(\mathbf{z})}{N}\phi_t(\mu)\mathrm{d}\mu, \tag{A-8}$$

where $\mu(\mathbf{z})/N$ is the frequency of individuals with phenotype $\mathbf{z}$ within a group in state μ. Now, since there is an infinite number of groups, $p_t(\mathbf{z})$ is also the density function of individuals with phenotype $\mathbf{z}$ in the whole population. We now derive a recurrence equation for $p_t(\mathbf{z})$ to capture the dynamics of the phenotypic distribution in the whole population.

Using eq. (A-6), the individual phenotypic distribution at time period $t+1$ can be written as

$$p_{t+1}(\mathbf{z}) = \int_{\mathcal{S}} \frac{\mu(\mathbf{z})}{N}\phi_{t+1}(\mu)\mathrm{d}\mu = \frac{1}{N}\int_{\mathcal{S}} \int_{\mathcal{S}} \mu'(\mathbf{z})P_t(\mu, \mu')\phi_t(\mu)\mathrm{d}\mu'\mathrm{d}\mu = \frac{1}{N}\int_{\mathcal{S}} \lambda_\mu(\mathbf{z}, \phi_t)\phi_t(\mu)\mathrm{d}\mu, \tag{A-9}$$

where

$$\lambda_\mu(\mathbf{z}, \phi_t) = \int_{\mathcal{S}} \mu'(\mathbf{z})P_t(\mu, \mu')\mathrm{d}\mu', \tag{A-10}$$

is the expected number of individuals with phenotype $\mathbf{z}$ in the focal group in the descendant time period (after one iteration of the life-cycle), given that it was in state μ in the parental time period and that the state of the population is ϕ_t.

We can decompose the expected number of individuals after one iteration of the life-cycle with phenotype $\mathbf{z}$

as

$$\lambda_\mu(\mathbf{z},\phi_t) = \lambda_\mu^{\mathrm{P}}(\mathbf{z},\phi_t) + \lambda_\mu^{\mathrm{I}}(\mathbf{z},\phi_t), \tag{A-11}$$

where $\lambda_\mu^{\mathrm{P}}(\mathbf{z},\phi_t)$ is the expected number of philopatric individuals with phenotype $\mathbf{z}$ (i.e., surviving adults or successful natal offspring with phenotype $\mathbf{z}$) and $\lambda_\mu^{\mathrm{I}}(\mathbf{z},\phi_t)$ is the expected number of immigrant offspring with phenotype $\mathbf{z}$, given that the group was in state μ in the parental time period.

The number of philopatric individuals with phenotype $\mathbf{z}$ can be expressed in terms of phenotypes in the parental time period by integrating over all possible parental phenotypes $\mathbf{z}'$ that contributed to $\mathbf{z}$ in a group in state μ,

$$\lambda_\mu^{\mathrm{P}}(\mathbf{z},\phi_t) = \int_{\mathbb{R}^n} \mu(\mathbf{z}')\, w_\mu^{\mathrm{P}}(\mathbf{z}',\phi_t)\, u(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}', \tag{A-12}$$

where $w_\mu^{\mathrm{P}}(\mathbf{z}',\phi_t)$ is the expected number of offspring (including itself if it survives) produced by a *single* individual bearing $\mathbf{z}'$, given that it resides in a state μ group, and $u(\mathbf{z}',\mathbf{z})$, which captures the process of mutation, is the p.d.f. for the event that the offspring produced by an individual with phenotype $\mathbf{z}'$ has phenotype $\mathbf{z}$ (recall that according to the life-cycle, offspring and adults mutate alike). Likewise, we can write the expected number of immigrant offspring as

$$\lambda_\mu^{\mathrm{I}}(\mathbf{z},\phi_t) = \int_{\mathbb{R}^n} \int_{\mathcal{S}} \mu'(\mathbf{z}')\, w_{\mu,\mu'}^{\mathrm{D}}(\mathbf{z}',\phi_t)\, u(\mathbf{z}',\mathbf{z})\phi_t(\mu')\mathrm{d}\mu'\mathrm{d}\mathbf{z}', \tag{A-13}$$

where $w_{\mu,\mu'}^{\mathrm{D}}(\mathbf{z}',\phi_t)$ is the expected number of successful emigrant offspring of a *single* individual with phenotype $\mathbf{z}'$, given that it resides in a group in state $\mu' \in \mathcal{S}$, and that the colonized group (i.e., the group the offspring lands in) was in state μ in the parental generation.

Substituting eqs. (A-12) and (A-13) into eq. (A-11), which is in turn substituted into eq. (A-9), the phenotypic distribution reads as

$$p_{t+1}(\mathbf{z}) = \frac{1}{N} \int_{\mathcal{S}} \int_{\mathbb{R}^n} \left(\mu(\mathbf{z}')\, w_\mu^{\mathrm{P}}(\mathbf{z}',\phi_t) + \int_{\mathcal{S}} \mu'(\mathbf{z}')\, w_{\mu,\mu'}^{\mathrm{D}}(\mathbf{z}',\phi_t)\, u(\mathbf{z}',\mathbf{z})\phi_t(\mu')\mathrm{d}\mu' \right) \mathrm{d}\mathbf{z}'\phi_t(\mu)\mathrm{d}\mu. \tag{A-14}$$

By exchanging dummy variables, we can re-arrange the second summand of the right hand side of eq. (A-14) to

$$\int_{\mathcal{S}} \int_{\mathcal{S}} \mu(\mathbf{z}')\, w_{\mu,\mu'}^{\mathrm{D}}(\mathbf{z}',\phi_t)\, u(\mathbf{z}',\mathbf{z})\phi_t(\mu')\phi_t(\mu)\mathrm{d}\mu'\mathrm{d}\mu = \int_{\mathcal{S}} \int_{\mathcal{S}} \mu(\mathbf{z}')\, w_{\mu',\mu}^{\mathrm{D}}(\mathbf{z}',\phi_t)\, u(\mathbf{z}',\mathbf{z})\phi_t(\mu)\phi_t(\mu')\mathrm{d}\mu\mathrm{d}\mu', \tag{A-15}$$

and therefore re-write eq. (A-14) as

$$\begin{aligned} p_{t+1}(\mathbf{z}) &= \frac{1}{N} \int_{\mathcal{S}} \int_{\mathbb{R}^n} \mu(\mathbf{z}') \left(w_\mu^{\mathrm{P}}(\mathbf{z}',\phi_t) + \int_{\mathcal{S}} \mu'(\mathbf{z}')\, w_{\mu',\mu}^{\mathrm{D}}(\mathbf{z}',\phi_t)\phi_t(\mu')\mathrm{d}\mu' \right) u(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}'\phi_t(\mu)\mathrm{d}\mu \\ &= \frac{1}{N} \int_{\mathcal{S}} \int_{\mathbb{R}^n} \mu(\mathbf{z}') \left(w_\mu^{\mathrm{P}}(\mathbf{z}',\phi_t) + w_\mu^{\mathrm{D}}(\mathbf{z}',\phi_t) \right) u(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}'\phi_t(\mu)\mathrm{d}\mu, \end{aligned} \tag{A-16}$$

where

$$w_\mu^{\mathrm{D}}(\mathbf{z}',\phi_t) = \int_{\mathcal{S}} w_{\mu',\mu}^{\mathrm{D}}(\mathbf{z}',\phi_t)\phi_t(\mu')\mathrm{d}\mu' \tag{A-17}$$

is the expected number of successful dispersing offspring that establish into a representative group produced

by an individual with phenotype $\mathbf{z}'$, given that it resides in a group in state μ. So, with

$$w_\mu(\mathbf{z}',\phi_t) = w_\mu^{\mathrm{P}}(\mathbf{z}',\phi_t) + w_\mu^{\mathrm{D}}(\mathbf{z}',\phi_t), \tag{A-18}$$

as the expected number of successful offspring of an individual with phenotype $\mathbf{z}'$ (including itself if it survives), given that this individual resides in a group in state μ (i.e., $w_\mu(\mathbf{z}',\phi_t)$ is the *conditional individual fitness* of this focal individual), the phenotypic distribution in time period $t+1$ is,

$$p_{t+1}(\mathbf{z}) = \frac{1}{N}\int_{\mathcal{S}}\int_{\mathbb{R}^n} \mu(\mathbf{z}')w_\mu(\mathbf{z}',\phi_t)u(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}'\phi_t(\mu)\mathrm{d}\mu. \tag{A-19}$$

We can further define

$$W(\mathbf{z}',\phi_t) = \int_{\mathcal{S}} w_\mu(\mathbf{z}',\phi_t)\frac{\mu(\mathbf{z}')}{N}\frac{\phi_t(\mu)}{p_t(\mathbf{z}')}\mathrm{d}\mu = \int_{\mathcal{S}} w_\mu(\mathbf{z}',\phi_t)q_t(\mu|\mathbf{z}')\mathrm{d}\mu, \tag{A-20}$$

as the average individual fitness of a typical bearer of phenotype $\mathbf{z}'$, where

$$q_t(\mu|\mathbf{z}') = \frac{\mu(\mathbf{z}')}{N}\frac{\phi_t(\mu)}{p_t(\mathbf{z}')} \tag{A-21}$$

is the p.d.f. for the event that an individuals resides in a group in state μ given that this individual bears phenotype $\mathbf{z}'$. In other words, $q_t(\mu|\mathbf{z}')$ is the experienced-profile phenotypic distribution of an individual randomly sampled from the collection of all individuals with phenotype $\mathbf{z}'$ in the population (the "$\mathbf{z}'$ lineage"). Hence, $W(\mathbf{z}',\phi_t)$ is the marginal fitness of a typical bearer of phenotype $\mathbf{z}'$ and a multi-allelic version of *lineage fitness* (Lehmann et al., 2015, Mullon et al., 2016, Lehmann et al., 2016). We will therefore refer to $W(\mathbf{z}',\phi_t)$ as the lineage fitness of phenotype $\mathbf{z}'$.

Substituting eq. (A-20) into eq. (A-19), we obtain that the individual phenotypic density distribution is

$$p_{t+1}(\mathbf{z}) = \int_{\mathbb{R}^n} W(\mathbf{z}',\phi_t)u(\mathbf{z}',\mathbf{z})p_t(\mathbf{z}')\mathrm{d}\mathbf{z}', \tag{A-22}$$

which combines the forces of mutation and selection on phenotypic change.

When the probability of a mutation is independent from parental phenotype, we can rewrite the p.d.f. for the event that the offspring of an individual with phenotype $\mathbf{z}'$ has phenotype $\mathbf{z}$ as

$$u(\mathbf{z}',\mathbf{z}) = (1-\nu)\delta(\mathbf{z}'-\mathbf{z}) + \nu\upsilon(\mathbf{z}',\mathbf{z}), \tag{A-23}$$

where ν is the probability that an offspring has a mutant phenotype (i.e., $1-u(\mathbf{z},\mathbf{z}) = \nu$ for all $\mathbf{z}$), $\delta(\mathbf{z}'-\mathbf{z})$ is the Dirac delta function, and $\upsilon(\mathbf{z}',\mathbf{z})$ is the conditional probability of mutating from $\mathbf{z}'$ to $\mathbf{z}$ given that a mutation has occurred. So the first term of eq. (A-23) captures the event of no mutation, in which case the offspring has the same phenotype than its parent, and the second term captures the event of a mutation. Substituting eq. (A-23) into eq. (A-22), we obtain that the individual phenotypic density distribution is

$$p_{t+1}(\mathbf{z}) = W(\mathbf{z},\phi_t)p_t(\mathbf{z}) + \nu\left(\int_{\mathbb{R}^n} W(\mathbf{z}',\phi_t)p_t(\mathbf{z}')\upsilon(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}' - W(\mathbf{z},\phi_t)p_t(\mathbf{z})\right), \tag{A-24}$$

which is eq. (1) in the main text. Eq. (A-24) takes the same functional form as the classical equation for a continuum of alleles under selection and mutation in well-mixed populations (e.g., eq. 2.4 of Fleming, 1979, or

eq. 4.1 of Champagnat et al., 2006). The difference is that in well-mixed populations, $W(\mathbf{z},\phi_t)$ is the individual fitness of a single carrier of $\mathbf{z}$, while $W(\mathbf{z},\phi_t)$ here is the expected individual fitness of a randomly selected carrier of $\mathbf{z}$ (eq. A-20), which depends on the experienced-profile phenotype distribution of carriers of $\mathbf{z}$, and thus on the structure of the population.

A.3 Weak mutation and selection

We proceed further by assuming that selection is weak (see for instance eq. 2.7 of Fleming, 1979) and that mutations are sufficiently rare so that we can ignore the joint effects of mutation and selection on the phenotypic distribution over one time period ($\nu W(\mathbf{z},\phi_t)$ is small). In this case, we can write eq. (A-24) as

$$p_{t+1}(\mathbf{z}) = W(\mathbf{z},\phi_t)p_t(\mathbf{z}) + \left(\nu \int_{\mathbb{R}^n} W^\circ(\mathbf{z}',\phi_t)p_t(\mathbf{z}')\upsilon(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}' - \nu W^\circ(\mathbf{z},\phi_t)p_t(\mathbf{z})\right), \tag{A-25}$$

where the superscript $^\circ$ is used to denote quantities that are measured in the absence of selection throughout (i.e., under neutrality). The quantity $W^\circ(\mathbf{z},\phi_t)$ is therefore the lineage fitness of phenotype $\mathbf{z}$ in the absence of selection. To evaluate $W^\circ(\mathbf{z},\phi_t)$, note that since the total population size remains constant, conditional individual fitness in the absence of selection is one, i.e., $w^\circ_\mu(\mathbf{z}',\phi_t) = 1$. Using eq. (A-8), the lineage fitness of phenotype $\mathbf{z}$ in the absence of selection is then also one

$$W^\circ(\mathbf{z},\phi_t) = \int_{\mathcal{S}} \frac{\mu(\mathbf{z})}{N}\frac{\phi_t(\mu)}{p_t(\mathbf{z})}\mathrm{d}\mu = 1. \tag{A-26}$$

So, the phenotypic distribution at time period $t+1$ reads as

$$p_{t+1}(\mathbf{z}) = W(\mathbf{z},\phi_t)p_t(\mathbf{z}) + \nu\left(\int_{\mathbb{R}^n} p_t(\mathbf{z}')\upsilon(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}' - p_t(\mathbf{z})\right), \tag{A-27}$$

which is eq. (1) of the main text, and where the first term captures the effects of selection only, and the next term, the effects of mutation only. Again, eq. (A-27) takes the same functional form as classical recurrence for the phenotypic distribution in well-mixed populations when selection is weak (under the continuum-of-alleles model, e.g., eq. 1 of Burger, 1986; for fluctuating population size, see eq. 4.9 of Champagnat et al., 2006, which is in line with A-25), but with lineage fitness $W(\mathbf{z},\phi_t)$ instead of individual fitness.

B Second-order approximation of lineage fitness

In this appendix, we develop a second order approximation to lineage fitness $W(\mathbf{z},\phi_t)$ (eq. A-20). This will subsequently allow us to derive a second-order approximation to the phenotypic dynamics of $p_t(\mathbf{z})$ (by substituting our approximation for $W(\mathbf{z},\phi_t)$ into eq. A-27).

When selection is weak, we can approximate lineage fitness $W(\mathbf{z},\phi_t)$ as a polynomial by way of a Taylor expansion around average trait values, i.e., around

$$\bar{z}_{a,t} = \int_{\mathbb{R}^n} z_a p_t(\mathbf{z})\mathrm{d}\mathbf{z}, \tag{B-1}$$

where z_a refers to the entry a of the vector $\mathbf{z} \in \mathbb{R}^n$; i.e., the value of trait a. Such Taylor expansion yields a

second-order polynomial approximation for lineage fitness that we can write generically as

$$W(\mathbf{z},\phi_t) = W(\bar{\mathbf{z}},\phi_t) + \sum_{a=1}^{n}(z_a - \bar{z}_{a,t})s_{a,t}(\bar{\mathbf{z}}_t) + \frac{1}{2}\sum_{a=1}^{n}\sum_{b=1}^{n}(z_a - \bar{z}_{a,t})(z_b - \bar{z}_{b,t})h_{ab,t}(\bar{\mathbf{z}}_t), \tag{B-2}$$

where $W(\bar{\mathbf{z}},\phi_t)$ is the lineage fitness of an individual bearing the average phenotype $\bar{\mathbf{z}}_t = (\bar{z}_{1,t}, \bar{z}_{2,t}, \ldots, \bar{z}_{n,t})^{\mathrm{T}}$, $s_{a,t}(\bar{\mathbf{z}}_t)$ is the first-order effect on lineage fitness of a change in trait a away from the population mean, and $h_{ab,t}(\bar{\mathbf{z}}_t)$ is the second-order effect of changes in traits a and b away from their population means. Hence, $s_{a,t}(\bar{\mathbf{z}}_t)$ captures the directional coefficient of selection on trait a and $h_{ab,t}(\bar{\mathbf{z}}_t)$ the correlational (and disruptive when $a = b$) coefficient of selection on traits a and b (Phillips and Arnold, 1989). We will now evaluate, $W(\bar{\mathbf{z}},\phi_t)$, $s_{a,t}(\bar{\mathbf{z}}_t)$, and $h_{ab,t}(\bar{\mathbf{z}}_t)$ coefficients in several steps.

B.1 Conditional individual fitness approximation

Since lineage fitness depends on conditional individual fitness, $w_\mu(\mathbf{z},\phi_t)$ (eq. A-20), we first derive a second-order polynomial approximation for conditional individual fitness, $w_\mu(\mathbf{z},\phi_t)$. To obtain this second-order polynomial, we start by rewriting $w_\mu(\mathbf{z},\phi_t)$ as a function $w : \mathbb{R}^n \times \mathbb{R}^{(N-1)\times n} \times \mathbb{R}^{N\times n} \to \mathbb{R}_+$,

$$w_\mu(\mathbf{z},\phi_t) = w(\mathbf{z}, \mathrm{n}_\mu(\mathbf{z}), \mathrm{e}(\phi_t)), \tag{B-3}$$

that depends explicitly on the phenotype $\mathbf{z} \in \mathbb{R}^n$ of the focal individual (the individual whose fitness is under scrutiny), on the set of phenotypes $\mathrm{n}_\mu(\mathbf{z}) \in \mathbb{R}^{(N-1)\times n}$ of the $N-1$ neighbours of the focal individual, and on the set $\mathrm{e}(\phi_t) \in \mathbb{R}^{N\times n}$ of N phenotypes from a representative (or average) group other than the one the focal resides in (which depends on ϕ_t).

We approximate conditional individual fitness, $w(\mathbf{z}, \mathrm{n}_\mu(\mathbf{z}), \mathrm{e}(\phi_t))$, by Taylor expanding it around the average phenotype $\bar{\mathbf{z}}_t$, and write it as

$$w(\mathbf{z}, \mathrm{n}_\mu(\mathbf{z}), \mathrm{e}(\phi_t)) = 1 + \sum_{a=1}^{n}\underbrace{\sum_{\hat{z}\in\mathcal{Z}}(\hat{z}_a - \bar{z}_{a,t})\frac{\partial w}{\partial \hat{z}_a}\bigg|_{\bar{\mathbf{z}}_t}}_{\beta_{a,t}} + \frac{1}{2}\sum_{a=1}^{n}\sum_{b=1}^{n}\underbrace{\sum_{\hat{z}\in\mathcal{Z}}\sum_{\hat{z}\in\mathcal{Z}}(\hat{z}_a - \bar{z}_{a,t})(\hat{z}_a - \bar{z}_{b,t})\frac{\partial^2 w}{\partial \hat{z}_a \partial \hat{z}_a}\bigg|_{\bar{\mathbf{z}}_t}}_{\beta_{ab,t}}, \tag{B-4}$$

where 1 is individual fitness in a monomorphic population (i.e., when all individuals have phenotype $\bar{\mathbf{z}}_t$) and the set $\mathcal{Z} = \{\mathbf{z}\} \cup \mathrm{n}_\mu(\mathbf{z}) \cup \mathrm{e}(\phi_t)$ collects the phenotypes of the focal, its neighbours, and those in other groups. The term $\beta_{a,t}$ thus measures the total first-order effect of a change in trait a on focal fitness: it sums the first-order effects of changing trait a ($\partial w/\partial \hat{z}_a|_{\bar{\mathbf{z}}_t}$, where $.|_{\bar{\mathbf{z}}_t}$ means that the derivative is evaluated where all individuals have mean phenotype $\bar{\mathbf{z}}_t$) across all individuals that belong to $\mathcal{Z}$. Similarly, $\beta_{ab,t}$ measures the total second-order effects of changes in traits a and b on focal fitness (summing the second-order effects of changes in traits a and b within, and between, individuals that are in $\mathcal{Z}$).

B.1.1 First-order effects on individual fitness

We first develop the first order effect $\beta_{a,t}$ of varying trait a, and to obtain more compact expressions, we will henceforth use the notation

$$\xi_t(z_a) = z_a - \bar{z}_{a,t}, \tag{B-5}$$

to denote the difference between a value z_a and the average trait value a. Also, we will use the symbols

$$\mathbf{z}_{\mathrm{n}} = \{z_{\mathrm{n},a}, ..., z_{\mathrm{n},n}\}, \mathbf{z}_{\mathrm{m}} = \{z_{\mathrm{m},a}, ..., z_{\mathrm{m},n}\} \in \mathrm{n}_{\mu}(\mathbf{z}), \tag{B-6}$$

to denote the phenotype of two distinct individuals from the focal group (and distinct from the focal individual), and

$$\mathbf{z}_{\mathrm{nn}} = \{z_{\mathrm{nn},a}, ..., z_{\mathrm{nn},n}\}, \mathbf{z}_{\mathrm{mm}} = \{z_{\mathrm{mm},a}, ..., z_{\mathrm{mm},n}\} \in \mathrm{e}(\phi_t), \tag{B-7}$$

to denote the phenotype of two distinct individuals a different group different to the focal. With these notations in hand, the first-order effect on conditional fitness can be decomposed into effect from individuals living in the same and different groups than the focal, and written as

$$\beta_{a,t} = \xi_t(z_a)\frac{\partial w}{\partial z_a} + \sum_{\mathbf{z}_{\mathrm{n}} \in \mathrm{n}_{\mu}(\mathbf{z})} \xi_t(z_{\mathrm{n},a})\frac{\partial w}{\partial z_{\mathrm{n},a}} + \sum_{\mathbf{z}_{\mathrm{nn}} \in \mathrm{e}(\phi_t)} \xi_t(z_{\mathrm{nn},a})\frac{\partial w}{\partial z_{\mathrm{nn},a}}, \tag{B-8}$$

where we have dropped the evaluation sign $.|_{\bar{\mathbf{z}}_t}$ for derivatives with the understanding that here and hereafter all derivatives are evaluated at $\bar{\mathbf{z}}_t$.

Now, we note that under our assumptions, the fitness function $w(\mathbf{z}, \mathrm{n}_{\mu}(\mathbf{z}), \mathrm{e}(\phi_t))$ is insensitive to permutations of elements within the sets $\mathrm{n}_{\mu}(\mathbf{z})$ and $\mathrm{e}(\phi_t)$ (i.e., it does not matter to individual fitness which precise neighbour or individual from another group expresses which phenotype). The derivatives in eq. (B-8) can therefore be taken out of their sums,

$$\beta_{a,t} = \xi_t(z_a)\frac{\partial w}{\partial z_a} + \left(\sum_{\mathbf{z}_{\mathrm{n}} \in \mathrm{n}_{\mu}(\mathbf{z})} \xi_t(z_{\mathrm{n},a})\right)\frac{\partial w}{\partial z_{\mathrm{n},a}} + \left(\sum_{\mathbf{z}_{\mathrm{nn}} \in \mathrm{e}(\phi_t)} \xi_t(z_{\mathrm{nn},a})\right)\frac{\partial w}{\partial z_{\mathrm{nn},a}}. \tag{B-9}$$

In order to make the sums more convenient for later calculations, we seek to express them in terms of the counting measure $\mu(\mathbf{z})$ (eq. A-1), which counts the number of individuals with phenotype $\mathbf{z}$ in a group in state μ. This is easily achieved from definitions for the second sum in eq. (B-9), i.e., the summed deviations in a representative group other than the focal, which is given by

$$\sum_{\mathbf{z}_{\mathrm{nn}} \in \mathrm{e}(\phi_t)} \xi_t(z_{\mathrm{nn},a}) = \int_{\mathbb{R}^n} \int_{\mathcal{S}} \xi_t(z_{\mathrm{nn},a})\mu(\mathbf{z})\phi_t(\mu)\mathrm{d}\mu\mathrm{d}\mathbf{z}_{\mathrm{nn}}. \tag{B-10}$$

The first sum in eq. (B-9) needs a bit more work because the sum excludes the focal individual, when $\mu(\mathbf{z})$ counts all individual in the group (and thus includes the focal individual). We define the conditional counting measure,

$$\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}}) = \delta(\mathbf{z}_{\mathrm{n}} - \mathbf{z})(\mu(\mathbf{z}) - 1) + (1 - \delta(\mathbf{z}_{\mathrm{n}} - \mathbf{z})))\mu(\mathbf{z}_{\mathrm{n}}), \tag{B-11}$$

where $\delta(.)$ is the Dirac-Delta function. According to eq. (B-11), $\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}})$ counts the number of neighbours of the focal that have phenotype $\mathbf{z}_{\mathrm{n}}$ given that the focal individual has phenotype $\mathbf{z}$. To see this, consider that when integrated over $\mathbf{z}_{\mathrm{n}} \in \mathbb{R}^n$, the first term of eq. (B-11), $\delta(\mathbf{z}_{\mathrm{n}} - \mathbf{z})(\mu(\mathbf{z}) - 1)$, counts the number of neighbours of the focal that have the same phenotype as the focal, while the second term, $(1 - \delta(\mathbf{z}_{\mathrm{n}} - \mathbf{z})))\mu(\mathbf{z}_{\mathrm{n}})$, counts those with a different phenotype. With eq. (B-11), we can write the first sum of eq. (B-9) as

$$\sum_{\mathbf{z}_{\mathrm{n}} \in \mathrm{n}_{\mu}(\mathbf{z})} \xi_t(z_{\mathrm{n},a}) = \xi_t(z_a)(\mu(\mathbf{z}) - 1) + \sum_{\substack{\mathbf{z}_{\mathrm{n}} \in \mathrm{n}_{\mu}(\mathbf{z}) \\ \mathbf{z}_{\mathrm{n}} \neq \mathbf{z}}} \xi_t(z_{\mathrm{n},a})\mu(\mathbf{z}_{\mathrm{n}}) = \int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}})\mathrm{d}\mathbf{z}_{\mathrm{n}}. \tag{B-12}$$

Substituting eqs. (B-10) & (B-12) into eq. (B-9) then gives

$$\beta_{a,t} = \xi_t(z_a)\frac{\partial w}{\partial z_a} + \left(\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_\mathrm{n})\mathrm{d}\mathbf{z}_\mathrm{n}\right)\frac{\partial w}{\partial z_{\mathrm{n},a}} + \left(\int_{\mathbb{R}^n}\int_{\mathcal{S}} \xi_t(z_{\mathrm{nn},a})\mu(\mathbf{z})\phi_t(\mu)\mathrm{d}\mu\mathrm{d}\mathbf{z}_\mathrm{nn}\right)\frac{\partial w}{\partial z_{\mathrm{nn},a}}. \tag{B-13}$$

for the first order effects of trait a.

Next, note that eq. (A-8) and the definition of the mean eq. (B-1) entail that the third summand of eq. (B-13) is zero:

$$\int_{\mathbb{R}^n}\int_{\mathcal{S}} \xi_t(z_{\mathrm{nn},a})\mu(\mathbf{z})\phi_t(\mu)\mathrm{d}\mu\mathrm{d}\mathbf{z} = N\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{nn},a})p_t(\mathbf{z}_\mathrm{nn})\mathrm{d}\mathbf{z}_\mathrm{nn} = 0. \tag{B-14}$$

So we obtain that the first-order effect of trait a on focal fitness (eq. B-13) is

$$\beta_{a,t} = \xi_t(z_a)\frac{\partial w}{\partial z_a} + \left(\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\frac{\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_\mathrm{n})}{N-1}\mathrm{d}\mathbf{z}_\mathrm{n}\right)(N-1)\frac{\partial w}{\partial z_{\mathrm{n},a}}, \tag{B-15}$$

which only depends on the effect of a change in trait value in the focal individual and a representative neighbour.

B.1.2 Second-order effects on individual fitness

We now tackle the second-order effect on focal fitness of traits a and b, $\beta_{ab,t}$ (see eq. B-4). Calculations follow the same argument as for the first-order effect. Using notation eqs. (B-5)-(B-7), we first decompose $\beta_{ab,t}$ as

$$\beta_{ab,t} = \beta_{ab,t}^{\text{same}} + \beta_{ab,t}^{\text{neigh}} + \beta_{ab,t}^{\text{cross isle}}, \tag{B-16}$$

where

$$\beta_{ab,t}^{\text{same}} = \xi_t(z_a)\xi_t(z_b)\frac{\partial^2 w}{\partial z_a \partial z_b} + \sum_{\mathbf{z}_\mathrm{n}\in \mathrm{n}_\mu(\mathbf{z})} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{n},b})\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{n},b}} + \sum_{\mathbf{z}_\mathrm{nn}\in \mathrm{e}(\phi_t)} \xi_t(z_{\mathrm{nn},a})\xi_t(z_{\mathrm{nn},b})\frac{\partial^2 w}{\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nn},b}}, \tag{B-17}$$

collects the second-order effects of a and b when changed in the same individual, while

$$\begin{aligned}\beta_{ab,t}^{\text{neigh}} = &\sum_{\mathbf{z}_\mathrm{n}\in \mathrm{n}_\mu(\mathbf{z})} \xi_t(z_a)\xi_t(z_{\mathrm{n},b})\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{n},b}} + \sum_{\mathbf{z}_\mathrm{n}\in \mathrm{n}_\mu(\mathbf{z})} \xi_t(z_b)\xi_t(z_{\mathrm{n},a})\frac{\partial^2 w}{\partial z_b \partial z_{\mathrm{n},a}} \\ &+ \sum\sum_{\mathbf{z}_\mathrm{n},\mathbf{z}_\mathrm{m}\in \mathrm{n}_\mu(\mathbf{z})} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{m},b})\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}} + \sum\sum_{\mathbf{z}_\mathrm{nn},\mathbf{z}_\mathrm{nm}\in \mathrm{e}(\phi_t)} \xi_t(z_{\mathrm{nn},a})\xi_t(z_{\mathrm{nm},b})\frac{\partial^2 w}{\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nm},b}},\end{aligned} \tag{B-18}$$

collects the changes in two neighbours, and

$$\begin{aligned}\beta_{ab,t}^{\text{cross isle}} = &\sum_{\mathbf{z}_\mathrm{nn}\in \mathrm{e}(\phi_t)} \xi_t(z_a)\xi_t(z_{\mathrm{nn},b})\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{nn},b}} + \sum_{\mathbf{z}_\mathrm{nn}\in \mathrm{e}(\phi_t)} \xi_t(z_b)\xi_t(z_{\mathrm{nn},a})\frac{\partial^2 w}{\partial z_b \partial z_{\mathrm{nn},a}} \\ &+ \sum\sum_{\substack{\mathbf{z}_\mathrm{n}\in \mathrm{n}_\mu(\mathbf{z}) \\ \mathbf{z}_\mathrm{nn}\in \mathrm{e}(\phi_t)}} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{nn},b})\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{nn},b}} + \sum\sum_{\substack{\mathbf{z}_\mathrm{n}\in \mathrm{n}_\mu(\mathbf{z}) \\ \mathbf{z}_\mathrm{nn}\in \mathrm{e}(\phi_t)}} \xi_t(z_{\mathrm{n},b})\xi_t(z_{\mathrm{nn},a})\frac{\partial^2 w}{\partial z_{\mathrm{n},b}\partial z_{\mathrm{nn},a}},\end{aligned} \tag{B-19}$$

in one individual from the focal group, and one in a representative group other than the focal. We proceed to specify each of these terms in terms of counting measures.

B.1.3 Effect of joint trait changes in the same individual, $\beta^{\text{same}}_{ab,t}$

Using the counting measures eqs. (A-1) and (B-11), the $\beta^{\text{same}}_{ab,t}$ term of eq. (B-17) can be expressed as

$$\begin{aligned}\beta^{\text{same}}_{ab,t} = \xi_t(z_a)\xi_t(z_b)\frac{\partial^2 w}{\partial z_a \partial z_b} + \left(\int_{\mathbb{R}^n} \xi_t(z_{\text{n},a})\xi_t(z_{\text{n},b})\mu_{\text{n},\mathbf{z}}(\mathbf{z}_\text{n})\mathrm{d}\mathbf{z}_\text{n}\right)\frac{\partial^2 w}{\partial z_{\text{n},a}\partial z_{\text{n},b}} \\ + \left(\int_{\mathbb{R}^n}\int_{\mathcal{S}} \xi_t(z_{\text{nn},a})\xi_t(z_{\text{nn},b})\mu(\mathbf{z}_\text{nn})\phi_t(\mu)\mathrm{d}\mathbf{z}_\text{nn}\mathrm{d}\mu\right)\frac{\partial^2 w}{\partial z_{\text{nn},a}\partial z_{\text{nn},b}},\end{aligned} \tag{B-20}$$

where the first term consists of the effect of joint changes in the focal individual, the second in a neighbour of the focal and the third in an individual from another than the focal. Because the population size remains constant, we have

$$\frac{\partial^2 w}{\partial z_a \partial z_b} + (N-1)\frac{\partial^2 w}{\partial z_{\text{n},a}\partial z_{\text{n},b}} + N\frac{\partial^2 w}{\partial z_{\text{nn},a}\partial z_{\text{nn},b}} = 0 \tag{B-21}$$

(see eq. B.13 of Wakano and Lehmann, 2014). In addition, using eq. (A-8), the third term of eq. (B-20) can be expressed as

$$\int_{\mathbb{R}^n}\int_{\mathcal{S}} \xi_t(z_{\text{nn},a})\xi_t(z_{\text{nn},b})\mu(\mathbf{z}_\text{nn})\phi_t(\mu)\mathrm{d}\mu\mathrm{d}\mathbf{z}_\text{nn} = N\int_{\mathbb{R}^n}\int_{\mathcal{S}} \xi_t(z_{\text{nn},a})\xi_t(z_{\text{nn},b})p_t(\mathbf{z}_{\text{nn},a})\mathrm{d}\mathbf{z}_\text{nn} = N\sigma_{ab,t}, \tag{B-22}$$

where

$$\sigma_{ab,t} = \int_{\mathbb{R}^n} (z_a - \bar{z}_{a,t})(z_b - \bar{z}_{b,t})p_t(\mathbf{z})\mathrm{d}\mathbf{z}, \tag{B-23}$$

is the (co)variance between traits a and b in the population at time period t. Thus, substituting eqs. (B-21) and (B-22) into eq. (B-20), we find that the effects of changing traits a and b in the same individual can be written as

$$\beta^{\text{same}}_{ab,t} = (\xi_t(z_a)\xi_t(z_b) - \sigma_{ab,t})\frac{\partial^2 w}{\partial z_a \partial z_b} + \left(\int_{\mathbb{R}^n} \xi_t(z_{\text{n},a})\xi_t(z_{\text{n},b})\frac{\mu_{\text{n},\mathbf{z}}(\mathbf{z}_\text{n})}{N-1}\mathrm{d}\mathbf{z}_\text{n} - \sigma_{ab,t}\right)(N-1)\frac{\partial^2 w}{\partial z_{\text{n},a}\partial z_{\text{n},b}}, \tag{B-24}$$

which only depends on effects from the focal group.

B.1.4 Effect of joint trait changes in neighbours, $\beta^{\text{neigh}}_{ab,t}$

The first two terms of $\beta^{\text{neigh}}_{ab,t}$ (eq. B-18) correspond to the effect on focal fitness of changing one trait in the focal individual, and another one in one of its neighbours. These can be expressed using the counting measures eqs. (A-1) and (B-11) as

$$\begin{aligned}\sum_{\mathbf{z}_\text{n}\in \text{n}_\mu(\mathbf{z})} \xi_t(z_a)\xi_t(z_{\text{n},b})\frac{\partial^2 w}{\partial z_a \partial z_{\text{n},b}} &= \left(\int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_{\text{n},b})\mu_{\text{n},\mathbf{z}}(\mathbf{z}_\text{n})\mathrm{d}\mathbf{z}_\text{n}\right)\frac{\partial^2 w}{\partial z_a \partial z_{\text{n},b}} \\ \sum_{\mathbf{z}_\text{n}\in \text{n}_\mu(\mathbf{z})} \xi_t(z_b)\xi_t(z_{\text{n},a})\frac{\partial^2 w}{\partial z_b \partial z_{\text{n},a}} &= \left(\int_{\mathbb{R}^n} \xi_t(z_b)\xi_t(z_{\text{n},a})\mu_{\text{n},\mathbf{z}}(\mathbf{z}_\text{n})\mathrm{d}\mathbf{z}_\text{n}\right)\frac{\partial^2 w}{\partial z_b \partial z_{\text{n},a}}.\end{aligned} \tag{B-25}$$

The third term of $\beta^{\text{neigh}}_{ab,t}$ (eq. B-18),

$$\sum_{\mathbf{z}_\text{n},\mathbf{z}_\text{m}\in \text{n}_\mu(\mathbf{z})}\sum \xi_t(z_{\text{n},a})\xi_t(z_{\text{m},b})\frac{\partial^2 w}{\partial z_{\text{n},a}\partial z_{\text{m},b}} \tag{B-26}$$

collects the effects on focal fitness of changing two traits in two neighbours to the focal. In order to express it in terms of the counting measure $\mu(\mathbf{z})$, we introduce another conditional counting measure,

$$\begin{aligned}\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}},\mathbf{z}_{\mathrm{m}}) &= \delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z})\,\delta(\mathbf{z}_{\mathrm{m}}-\mathbf{z}_{\mathrm{n}})\,(\mu(\mathbf{z})-1)(\mu(\mathbf{z})-2)\\ &+\delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z})\,(1-\delta(\mathbf{z}_{\mathrm{m}}-\mathbf{z}))\,(\mu(\mathbf{z})-1)\mu(\mathbf{z}_{\mathrm{m}})+\delta(\mathbf{z}_{\mathrm{m}}-\mathbf{z})\,(1-\delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z}))\,(\mu(\mathbf{z})-1)\mu(\mathbf{z}_{\mathrm{n}})\\ &+\delta(\mathbf{z}_{\mathrm{m}}-\mathbf{z}_{\mathrm{n}})\,(1-\delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z}))\,\mu(\mathbf{z}_{\mathrm{n}})\,(\mu(\mathbf{z}_{\mathrm{n}})-1)\\ &+(1-\delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z}))\,(1-\delta(\mathbf{z}_{\mathrm{m}}-\mathbf{z}))\,(1-\delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z}_{\mathrm{m}}))\,\mu(\mathbf{z}_{\mathrm{n}})\mu(\mathbf{z}_{\mathrm{m}})\end{aligned} \tag{B-27}$$

which counts the number of unordered pairs of neighbours that have phenotypes $\mathbf{z}_{\mathrm{n}}$ and $\mathbf{z}_{\mathrm{m}}$, given that the focal individual has phenotype $\mathbf{z}$. Integrated over $\mathbb{R}^n$, the first line of eq. (B-27) counts the number of pairs with focal phenotype $\mathbf{z}$; the second line, the number of pairs in which only one neighbour has the focal phenotype $\mathbf{z}$; the third line, the number of pairs in which both neighbours have the same phenotype that is different to the the focal; and the fourth line, the number of pairs in which neighbours have phenotypes that are different to one another, and to the focal. Using eq. (B-27), we can re-write the third term of $\beta^{\mathrm{neigh}}_{ab,t}$ (eq. B-26) as

$$\sum_{\mathbf{z}_{\mathrm{n}},\mathbf{z}_{\mathrm{m}}\in \mathrm{n}_{\mu}(\mathbf{z})}\sum \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{m},b})\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}} = \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}},\mathbf{z}_{\mathrm{m}})\mathrm{d}\mathbf{z}_{\mathrm{n}}\mathrm{d}\mathbf{z}_{\mathrm{m}}\right)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}}. \tag{B-28}$$

The fourth and final term of $\beta^{\mathrm{neigh}}_{ab,t}$ (see eq. B-18), which is the effect of changing two traits in two different individuals that reside in a representative group other than the focal, can be expressed using eq. (B-11) as

$$\sum_{\mathbf{z}_{\mathrm{nn}},\mathbf{z}_{\mathrm{nm}}\in \mathrm{e}(\phi_t)}\sum \xi_t(z_{\mathrm{nn},a})\xi_t(z_{\mathrm{nm},b})\frac{\partial^2 w}{\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nm},b}} = N(N-1)\varphi_{ab,t}\frac{\partial^2 w}{\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nm},b}}, \tag{B-29}$$

where

$$\varphi_{ab,t} = \int_{\mathcal{S}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\xi_t(z_{\mathrm{nn},a})\xi_t(z_{\mathrm{nm},b})\frac{\mu_{\mathrm{n},\mathbf{z}_{\mathrm{nn}}}(\mathbf{z}_{\mathrm{nm}})\mu(\mathbf{z}_{\mathrm{nn}})}{N(N-1)}\phi_t(\mu)\mathrm{d}\mathbf{z}_{\mathrm{nm}}\mathrm{d}\mathbf{z}_{\mathrm{nn}}\mathrm{d}\mu \tag{B-30}$$

is the covariance among traits a and b between individuals within groups (as $\mu_{\mathrm{n},\mathbf{z}_{\mathrm{nn}}}(\mathbf{z}_{\mathrm{nm}})\mu(\mathbf{z}_{\mathrm{nn}})/[N(N-1)]$ gives the frequency of pairs of individuals in the same group in state μ that have phenotypes $\mathbf{z}_{\mathrm{nn}}$ and $\mathbf{z}_{\mathrm{nm}}$).

Note that the covariance $\varphi_{ab,t}$ already weighs on a second-order fitness effect ($\partial^2 w/(\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nm},b})$ in eq. B-29). In our aim of expressing evolutionary dynamics to the second-order effects of selection, it is therefore sufficient to estimate $\varphi_{ab,t}$ under neutrality. In section B.2.1 below, we show that the within-group phenotypic covariance among traits a and b under neutrality satisfies

$$\varphi^{\circ}_{ab,t} = r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}, \tag{B-31}$$

where $r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)$ is the probability that two individuals sampled without replacement from a group at time t have a common ancestor who resided in that same group in the absence of selection (see section B.2.1 for more details, eq. B-48).

Substituting eq. (B-31) into eq. (B-29), we find that the effect of changing two traits in two different individuals that reside in a representative group other than the focal then is

$$\sum_{\mathbf{z}_{\mathrm{nn}},\mathbf{z}_{\mathrm{nm}}\in \mathrm{e}(\phi_t)}\sum \xi_t(z_{\mathrm{nn},a})\xi_t(z_{\mathrm{nm},b})\frac{\partial^2 w}{\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nm},b}} = N(N-1)r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\frac{\partial^2 w}{\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nm},b}}. \tag{B-32}$$

Further, since the total population size is constant, the following holds,

$$(N-1)\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{n},b}} + (N-1)\frac{\partial^2 w}{\partial z_b \partial z_{\mathrm{n},a}} + (N-1)(N-2)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}} + N(N-1)\frac{\partial^2 w}{\partial z_{\mathrm{nn},a}\partial z_{\mathrm{nm},b}} = 0 \tag{B-33}$$

(see eq. B.14 of Wakano and Lehmann, 2014), so that eq. (B-32) is equal to

$$-r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\left((N-1)\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{n},b}} + (N-1)\frac{\partial^2 w}{\partial z_b \partial z_{\mathrm{n},a}} + (N-1)(N-2)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}}\right). \tag{B-34}$$

Finally, we add eqs. (B-25), (B-28) and (B-34) in order to obtain,

$$\begin{aligned}\beta^{\mathrm{neigh}}_{ab,t} = &\left(\int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_{\mathrm{n},b})\frac{\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}})}{N-1}\mathrm{d}\mathbf{z}_{\mathrm{n}} - r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\right)(N-1)\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{n},b}} \\ &+\left(\int_{\mathbb{R}^n} \xi_t(z_b)\xi_t(z_{\mathrm{n},a})\frac{\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}})}{N-1}\mathrm{d}\mathbf{z}_{\mathrm{n}} - r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\right)(N-1)\frac{\partial^2 w}{\partial z_b \partial z_{\mathrm{n},a}} \\ &+\left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{m},b})\frac{\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_{\mathrm{n}},\mathbf{z}_{\mathrm{m}})}{(N-1)(N-2)}\mathrm{d}\mathbf{z}_{\mathrm{n}}\mathrm{d}\mathbf{z}_{\mathrm{m}} - r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\right)(N-1)(N-2)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}},\end{aligned} \tag{B-35}$$

for the effects on focal fitness of changing traits a and b in neighbours.

B.1.5 Effect of joint trait changes in two individuals from different groups, $\beta^{\text{cross isle}}_{ab,t}$

The final term necessary for our approximation of individual fitness is $\beta^{\text{cross isle}}_{ab,t}$ (eq. B-19), which is the effect of changing traits a and b in two individuals from different groups. Let us first consider the first summand of eq. (B-19), which measures the effect of a change in trait a of the focal and a trait b in an individual from a group other than the focal. It is straightforward to show that this term is zero:

$$\begin{aligned}\sum_{\mathbf{z}_{\mathrm{nn}}\in \mathrm{e}(\phi_t)} \xi_t(z_a)\xi_t(z_{\mathrm{nn},b})\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{nn},b}} &= \int_{\mathbb{R}^n}\int_{\mathcal{S}} \xi_t(z_a)\xi_t(z_{\mathrm{nn},b})\mu(\mathbf{z}_{\mathrm{nn}})\phi_t(\mu)\mathrm{d}\mu\mathrm{d}\mathbf{z}_{\mathrm{nn}} \\ &= N\xi_t(z_a)\int_{\mathbb{R}^n}\int_{\mathcal{S}} \xi_t(z_{\mathrm{nn},b})\frac{\mu(\mathbf{z}_{\mathrm{nn}})}{N}\phi_t(\mu)\mathrm{d}\mu\mathrm{d}\mathbf{z}_{\mathrm{nn}} \\ &= N\xi_t(z_a)\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{nn},b})p_t(\mathbf{z}_{\mathrm{nn}})\mathrm{d}\mathbf{z}_{\mathrm{nn}} = 0\end{aligned} \tag{B-36}$$

where we have used eqs. (A-8) and (B-1). In fact, by a similar argument, all the other summands of $\beta^{\text{cross isle}}_{ab,t}$ can be shown to be zero, so that

$$\beta^{\text{cross isle}}_{ab,t} = 0, \tag{B-37}$$

i.e., there is no effect of changing traits a and b in two individuals from different group on individual fitness.

B.1.6 Second order approximation of individual fitness

To summarise the above calculations (sections B.1.1-B.1.5), we have derived a second-order approximation for conditional individual fitness as

$$w_\mu(\mathbf{z},\phi_t) = w(\mathbf{z},\mathrm{n}_\mu(\mathbf{z}),\mathrm{e}(\phi_t)) = 1 + \sum_{a=1}^{n}\beta_{a,t} + \frac{1}{2}\sum_{a=1}^{n}\sum_{b=1}^{n}\left(\beta^{\text{same}}_{ab,t} + \beta^{\text{neigh}}_{ab,t}\right), \tag{B-38}$$

where first-order effects $\beta_{a,t}$ are given by eq. (B-15), and the second-order effects $\beta_{ab,t}^{\text{same}}$ and $\beta_{ab,t}^{\text{neigh}}$ are respectively given by eqs. (B-24) and (B-35).

B.2 Phenotypic covariance among neighbours

Here, we evaluate the phenotypic covariance $\varphi_{ab,t}$ to the first-order of selection, which will be useful later, and connect it to notions of identity-by-descent and definitions of relatedness.

B.2.1 Conditional phenotypic distribution in neighbors

In order to gain greater insight into the phenotypic covariance among neighbours, we rewrite eq. (B-30) – using eq. (A-21) and changing dummy variables – as

$$\varphi_{ab,t} = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_{\text{n},b}) q_{2,t}(\mathbf{z}_\text{n}|\mathbf{z}) p_t(\mathbf{z}) \mathrm{d}\mathbf{z}_\text{n} \mathrm{d}\mathbf{z}, \tag{B-39}$$

where,

$$q_{2,t}(\mathbf{z}_\text{n}|\mathbf{z}) = \int_{\mathcal{S}} \frac{\mu_{\text{n},\mathbf{z}}(\mathbf{z}_\text{n})}{N-1} q_t(\mu|\mathbf{z}_\text{n}) \mathrm{d}\mu \tag{B-40}$$

is by definition the conditional p.d.f. for the event of sampling an individual with phenotype $\mathbf{z}_\text{n}$ within a group, given that within this group, a focal individual with phenotype $\mathbf{z}$ has already been sampled (and removed). Substituting eq. (B-11) into eq. (B-40), we decompose this conditional p.d.f. in the following form,

$$q_{2,t}(\mathbf{z}_\text{n}|\mathbf{z}) = r_{2,t}(\mathbf{z})\delta(\mathbf{z}_\text{n} - \mathbf{z}) + k_t(\mathbf{z}, \mathbf{z}_\text{n}) p_t(\mathbf{z}_\text{n}), \tag{B-41}$$

where

$$r_{2,t}(\mathbf{z}) = \int_{\mathcal{S}} \frac{\mu(\mathbf{z}) - 1}{N-1} q_t(\mu|\mathbf{z}) \mathrm{d}\mu \tag{B-42}$$

is the p.d.f for the event of sampling an individual with the same phenotype $\mathbf{z}$ to the focal among its neighbours at time t. Hence, the first summand of eq. (B-41) captures the event that the sampled neighbour has the same phenotype to the focal individual ($\mathbf{z}_\text{n} = \mathbf{z}$). Note that with a continuum-of-alleles in the infinite island model of dispersal, the only way for the sampled neighbour to have the same phenotype as the focal is that these two individuals have a common ancestor that resided in that same group (i.e., that they are identical-by-descent; recall that here, we ignore the influence of mutations, which are taken into account in the second term of eq. A-27). Note also that $r_{2,t}(\mathbf{z})$ may depends on the whole phenotypic distribution ϕ_t, but this dependence is captured by the time index in $r_{2,t}(\mathbf{z})$.

The second summand of eq. (B-41) then captures the event of sampling an individual with phenotype $\mathbf{z}_\text{n}$ among the neighbours of an individual with phenotype $\mathbf{z}$, given that they have different phenotypes ($\mathbf{z}_\text{n} \neq \mathbf{z}$). We have written arbitrarily this p.d.f. as $k_t(\mathbf{z}, \mathbf{z}_\text{n}) p_t(\mathbf{z}_\text{n})$, where $p_t(\mathbf{z}_\text{n})$ is the marginal probability of sampling an individual with phenotype $\mathbf{z}_\text{n}$ from the global population. The function $k_t(\mathbf{z}, \mathbf{z}_\text{n})$ can therefore be viewed as the multiplicative effect of having already sampled an individual with phenotype $\mathbf{z}$ in the group on this marginal probability.

We seek to express $k_t(\mathbf{z}, \mathbf{z}_\text{n})$ in terms of $r_{2,t}(\mathbf{z})$ to the first-order of selection (as it will later be necessary to have an expression for $q_{2,t}(\mathbf{z}_\text{n}|\mathbf{z})$ to the first order of selection). To do so, we first Taylor expand both $r_{2,t}(\mathbf{z})$ and

$k_t(\mathbf{z},\mathbf{z}_\mathrm{n})$ in eq. (B-41) to the first-order around $\bar{\mathbf{z}}_t$, and obtain

$$\begin{aligned} q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) = &\left(r^\circ_{2,t}(\bar{\mathbf{z}}_t) + \sum_{a=1}^{n} \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a} \xi_t(z_a) \right) \delta\left(\mathbf{z}_\mathrm{n} - \mathbf{z}\right) \\ &+ \left(k_t(\bar{\mathbf{z}}_t,\bar{\mathbf{z}}_t) + \sum_{a=1}^{n} \frac{\partial k_t(\mathbf{z},\mathbf{z}_\mathrm{n})}{\partial z_a} \xi_t(z_a) + \sum_{a=1}^{n} \frac{\partial k_t(\mathbf{z},\mathbf{z}_\mathrm{n})}{\partial z_{\mathrm{n},a}} \xi_t(z_{\mathrm{n},a}) \right) p_t(\mathbf{z}_\mathrm{n}), \end{aligned} \tag{B-43}$$

where $r^\circ_{2,t}(\bar{\mathbf{z}}_t)$ is the neutral probability of the event of sampling an individual with the same phenotype $\mathbf{z}$ as the focal in the focal's group. Under neutrality, this is independent of the phenotype $\mathbf{z}$. It depends only on the mean phenotype, $\bar{\mathbf{z}}_t$, in the population and is equivalent to the probability that distinct individuals individuals sampled from the same group are identical-by-descent at time t.

Second, we use three identities that must hold for $q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})$ to be a valid p.d.f.:

$$\int_{\mathbb{R}^n} q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) p_t(\mathbf{z}) \mathrm{d}\mathbf{z}_\mathrm{n} = p_t(\mathbf{z}). \tag{B-44a}$$

$$\int_{\mathbb{R}^n} q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) p_t(\mathbf{z}) \mathrm{d}\mathbf{z} = p_t(\mathbf{z}_\mathrm{n}), \tag{B-44b}$$

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) p_t(\mathbf{z}) \mathrm{d}\mathbf{z}_\mathrm{n} \mathrm{d}\mathbf{z} = 1, \tag{B-44c}$$

which all stem from the observation that $q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) p_t(\mathbf{z})$ is the (unconditional) p.d.f. for the event of sampling one individual with phenotype $\mathbf{z}$ and another with $\mathbf{z}_\mathrm{n}$ without replacement from a group. Substituting eq. (B-43) into eq. (B-44c), we obtain

$$r^\circ_{2,t}(\bar{\mathbf{z}}_t) + k_t(\bar{\mathbf{z}}_t,\bar{\mathbf{z}}_t) = 1 \Rightarrow k_t(\bar{\mathbf{z}}_t,\bar{\mathbf{z}}_t) = 1 - r^\circ_{2,t}(\bar{\mathbf{z}}_t). \tag{B-45a}$$

Then, substituting eq. (B-45a) into eq. (B-43), which is in turn substituted into eqs. (B-45b) and (B-45c), these yield,

$$p_t(\mathbf{z}) + \sum_{a=1}^{n} \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a} \xi_t(z_a) p_t(\mathbf{z}) + \sum_{a=1}^{n} \frac{\partial k_t(\mathbf{z},\mathbf{z}_\mathrm{n})}{\partial z_a} \xi_t(z_a) p_t(\mathbf{z}) = p_t(\mathbf{z}) \Rightarrow \frac{\partial k_t(\mathbf{z},\mathbf{z}_\mathrm{n})}{\partial z_a} = \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a} \tag{B-45b}$$

and

$$p_t(\mathbf{z}_\mathrm{n}) + \sum_{a=1}^{n} \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a} \xi_t(z_{\mathrm{n},a}) p_t(\mathbf{z}_\mathrm{n}) + \sum_{a=1}^{n} \frac{\partial k_t(\mathbf{z},\mathbf{z}_\mathrm{n})}{\partial z_{\mathrm{n},a}} \xi_t(z_{\mathrm{n},a}) p_t(\mathbf{z}_\mathrm{n}) = p_t(\mathbf{z}_\mathrm{n}) \Rightarrow \frac{\partial k_t(\mathbf{z},\mathbf{z}_\mathrm{n})}{\partial z_{\mathrm{n},a}} = \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a}, \tag{B-45c}$$

respectively.

Finally, substituting eqs. (B-45) into eq. (B-44c), we obtain to the first order of selection,

$$q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) = q^\circ_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) + \left(\sum_{a=1}^{n} \xi_t(z_a) \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a} \right) \delta\left(\mathbf{z}_\mathrm{n} - \mathbf{z}\right) - \left(\sum_{a=1}^{n} \left(\xi_t(z_a) + \xi_t(z_{\mathrm{n},a})\right) \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a} \right) p_t(\mathbf{z}_\mathrm{n}), \tag{B-46}$$

where

$$q^\circ_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z}) = r^\circ_{2,t}(\bar{\mathbf{z}}_t) \delta\left(\mathbf{z}_\mathrm{n} - \mathbf{z}\right) + \left(1 - r^\circ_{2,t}(\bar{\mathbf{z}}_t)\right) p_t(\mathbf{z}_\mathrm{n}), \tag{B-47}$$

is the neutral (conditional) p.d.f. for the event of sampling an individual with phenotype $\mathbf{z}_\mathrm{n}$ within the neighbourhood of an individual with phenotype $\mathbf{z}$.

B.2.2 Relatedness, identity-by-descent & neutral phenotypic covariance among neighbours.

Eq. (B-47) has the same functional form as the (conditional) expected frequency of an allele in an individual, given the frequency of that allele in a focal individual when there is a finite number of alleles in the population (e.g., eq. 21 of Rousset, 2002, eq. 1.14 of Lehmann and Rousset, 2014; except that eq. (B-47) depends explicitly on time). Such expression underpins the "regression definition of relatedness" (e.g., Grafen, 1985, eq. 2.13 of Frank, 1998), whereby relatedness is defined as the regression of neighbour phenotype on focal phenotype; i.e., relatedness is defined as the ratio of phenotypic covariance among neighbours, $\varphi^{\circ}_{ab,t}$, to the total variance in the population, $\sigma_{ab,t}$. Substituting eq. (B-47) into eq. (B-39) to compute this ratio, we indeed obtain

$$\frac{\varphi^{\circ}_{ab,t}}{\sigma_{ab,t}} = r^{\circ}_{2,t}(\bar{\mathbf{z}}_t). \tag{B-48}$$

We will therefore refer to $r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)$ interchangeably as the (pairwise) coefficient of relatedness or probability that two neighbours are identical-by-descent (under neutrality).

B.2.3 Phenotypic covariance among neighbours under selection

According to eq. (B-43), the derivative $\partial r_{2,t}(\mathbf{z})/\partial z_a$ can be interpreted as the effect of a (infinitesimal) change in trait a away from population mean in a focal lineage on the probability that two individuals randomly sampled from a group both belong to this lineage. As such, it is reasonable to view $\partial r_{2,t}(\mathbf{z})/\partial z_a$ as the effect of trait a on relatedness at time t. Further, this view is consistent with the regression definition of relatedness. To see this, consider the regression coefficient, $\varphi_{ab,t}/\sigma_{ab,t}$, to the first order of selection (substituting eq. B-46 into eq. B-39),

$$\frac{\varphi_{ab,t}}{\sigma_{ab,t}} = r^{\circ}_{2,t}(\bar{\mathbf{z}}_t) + \sum_{c=1}^{n} \frac{\kappa_{abc,t}}{\sigma_{ab,t}} \frac{\partial r_{2,t}(\mathbf{z})}{\partial z_c}. \tag{B-49}$$

where

$$\kappa_{abc,t} = \int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_b)\xi_t(z_c)p_t(\mathbf{z})\mathrm{d}\mathbf{z} \tag{B-50}$$

is the skew between traits a, b and c within individuals at time period t. Eq. (B-49) shows that $\partial r_{2,t}(\mathbf{z})/\partial z_c$ is also the first order effect of an (infinitesimal) change in trait c on the regression coefficient of an individuals phenotype on its neighbour's phenotype.

B.3 Second-order approximation of lineage fitness

We now use our approximation for individual fitness (eq. B-38) to compute the components, $W(\bar{\mathbf{z}},\phi_t)$, $s_{a,t}(\bar{\mathbf{z}}_t)$ and $h_{ab,t}(\bar{\mathbf{z}}_t)$ of the Taylor expansion of lineage fitness (eq. B-2). To so we first substitute individual fitness eq. (B-38) into the definition of lineage fitness, eq. (A-20). This yields,

$$W(\mathbf{z},\phi_t) = 1 + \sum_{a=1}^{n} \underbrace{\int_{\mathcal{S}} \beta_{a,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu}_{\text{average individual fitness effect of } a} + \frac{1}{2}\sum_{a=1}^{n}\sum_{b=1}^{n} \underbrace{\int_{\mathcal{S}} \left(\beta^{\text{same}}_{ab,t} + \beta^{\text{neigh}}_{ab,t}\right) q_t(\mu|\mathbf{z})\mathrm{d}\mu}_{\text{average individual fitness effect of } a \text{ and } b}, \tag{B-51}$$

where the first underbraced term corresponds to the effect of a change in one trait (trait a) on individual fitness, averaged over the lineage of carriers of that change; similarly, the second term corresponds to the average effect of a joint change in two traits (a and b) on individual fitness. We evaluate these average effects below

sequentially.

B.3.1 Average individual fitness effect of one trait change

Substituting eq. (B-15) into eq. (B-51) and using the definition of $q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})$, eq. (B-40), we obtain that the average individual fitness effect of a can be written as

$$\begin{aligned}\int_{\mathcal{S}} \beta_{a,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu &= \xi_t(z_a)\frac{\partial w}{\partial z_a} + \left(\int_{\mathcal{S}}\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\frac{\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_\mathrm{n})}{N-1} q_t(\mu|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n}\mathrm{d}\mu\right)(N-1)\frac{\partial w}{\partial z_{\mathrm{n},a}} \\ &= \xi_t(z_a)\frac{\partial w}{\partial z_a} + \left(\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n}\right)(N-1)\frac{\partial w}{\partial z_{\mathrm{n},a}}.\end{aligned} \tag{B-52}$$

The term $\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})$ in eq. (B-52) is the expected deviation from the population mean in a neighbour of an individual with phenotype $\mathbf{z}$. Since this deviation weighs on a first-order effect in eq. (B-52) (i.e., it weighs on $\partial w/\partial z_{\mathrm{n},a}$), and we seek a second-order approximation for lineage fitness, we need to evaluate this deviation to the first order of selection. To that end, it is necessary to use the first-order evaluation of $q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})$ given by eq. (B-46). Doing so, we find that the expected deviation from the population mean in neighbours of an individual with phenotype $\mathbf{z}$ to the first-order of selection is

$$\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n} = \xi_t(z_a) r^{\circ}_{2,t}(\bar{\mathbf{z}}_t) + \sum_{b=1}^{n}\left(\xi_t(z_a)\xi_t(z_b) - \sigma_{ab,t}\right)\frac{\partial r_{2,t}(\mathbf{z})}{\partial z_b}. \tag{B-53}$$

Substituting eq. (B-53) into eq. (B-52), we therefore obtain

$$\int_{\mathcal{S}} \beta_{a,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu = \xi_t(z_a)\left(\frac{\partial w}{\partial z_a} + (N-1) r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\frac{\partial w}{\partial z_{\mathrm{n},a}}\right) + \sum_{b=1}^{n}\left(\xi_t(z_a)\xi_t(z_b) - \sigma_{ab,t}\right)\frac{\partial r_{2,t}(\mathbf{z})}{\partial z_b}(N-1)\frac{\partial w}{\partial z_{\mathrm{n},a}}, \tag{B-54}$$

for the average individual fitness effect of a.

B.3.2 Average individual fitness effects of a joint change in two traits

We now turn to the average individual fitness effects of a joint change in two traits, which can be decomposed between two effects (see eq. B-51), those of changing two traits in the same individual

$$\int_{\mathcal{S}} \beta^{\mathrm{same}}_{ab,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu, \tag{B-55}$$

and those in two neighbours

$$\int_{\mathcal{S}} \beta^{\mathrm{neigh}}_{ab,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu. \tag{B-56}$$

We specify both of these terms below.

Average individual fitness effects of a joint change in two traits in the same individual. Substituting eq. (B-24) into (B-55) and using eq. (B-40) gives

$$\int_{\mathcal{S}} \beta^{\text{same}}_{ab,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu = (\xi_t(z_a)\xi_t(z_b) - \sigma_{ab,t})\frac{\partial^2 w}{\partial z_a \partial z_b} + \left(\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{n},b}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n} - \sigma_{ab,t}\right)(N-1)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{n},b}}, \tag{B-57}$$

where the term $\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{n},b}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n}$ measures the within-individual covariance between traits a and b within the neighbours of an individual with phenotype $\mathbf{z}$. Note that since this covariance term already weighs on a second-order effect of selection in eq. (B-57) ($\partial^2 w/(\partial z_{\mathrm{n},a}\partial z_{\mathrm{n},b})$), it is sufficient to evaluate it in the absence of selection, which using eq. (B-47) gives

$$\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{n},b}) q^{\circ}_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n} = \xi_t(z_a)\xi_t(z_b) r^{\circ}_{2,t}(\bar{\mathbf{z}}_t) + \sigma_{ab,t}(1 - r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)). \tag{B-58}$$

Substituting eq. (B-58) into eq. (B-57) then gives

$$\int_{\mathcal{S}} \beta^{\text{same}}_{ab,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu = (\xi_t(z_a)\xi_t(z_b) - \sigma_{ab,t})\left(\frac{\partial^2 w}{\partial z_a \partial z_b} + (N-1) r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{n},b}}\right), \tag{B-59}$$

for the effect of changing traits a and b in the same individual on lineage fitness.

Average individual fitness effects of a joint change in two traits among neighbours. Substituting eq. (B-35) into eq. (B-56), and using eq. (B-39), the effect on lineage fitness of changing traits a and b in two neighbours is given by

$$\begin{aligned}\int_{\mathcal{S}} \beta^{\text{neigh}}_{ab,t} q_t(\mu|\mathbf{z})\mathrm{d}\mu = &\left(\int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_{\mathrm{n},b}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n} - r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\right)(N-1)\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{n},b}} \\ &+ \left(\int_{\mathbb{R}^n} \xi_t(z_b)\xi_t(z_{\mathrm{n},a}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n} - r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\right)(N-1)\frac{\partial^2 w}{\partial z_b \partial z_{\mathrm{n},a}} \\ &+ \left(\rho_{ab,t} - r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\sigma_{ab,t}\right)(N-1)(N-2)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}},\end{aligned} \tag{B-60}$$

where

$$\rho_{ab,t} = \int_{\mathcal{S}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{m},b})\frac{\mu_{\mathrm{n},\mathbf{z}}(\mathbf{z}_\mathrm{n},\mathbf{z}_\mathrm{m})}{(N-1)(N-2)} q_t(\mu|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n}\mathrm{d}\mathbf{z}_\mathrm{m}\mathrm{d}\mu \tag{B-61}$$

measures the covariance between traits a and b among the neighbours of an individual with phenotype $\mathbf{z}$. The moments of the phenotypic distribution that appear in eq. (B-60), $\int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_{\mathrm{n},b}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n}$, $\int_{\mathbb{R}^n} \xi_t(z_b)\xi_t(z_{\mathrm{n},a}) q_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n}$, and $\rho_{ab,t}$, all weigh on second-order effects in eq. (B-60) ($\partial^2 w/(\partial z_a \partial z_{\mathrm{n},b})$, $\partial^2 w/(\partial z_b \partial z_{\mathrm{n},a})$, $\partial^2 w/(\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b})$, respectively). We can therefore evaluate these moments under neutrality.

First, using eq. (B-47), we have

$$\int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_{\mathrm{n},b}) q^{\circ}_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n} = \int_{\mathbb{R}^n} \xi_t(z_b)\xi_t(z_{\mathrm{n},a}) q^{\circ}_{2,t}(\mathbf{z}_\mathrm{n}|\mathbf{z})\mathrm{d}\mathbf{z}_\mathrm{n} = \xi_t(z_a)\xi_t(z_b) r^{\circ}_{2,t}(\bar{\mathbf{z}}_t), \tag{B-62}$$

for the covariance among one trait of the focal and the other trait of a neighbour.

Second, we rewrite the covariance between traits a and b among the neighbours of an individual with phenotype $\mathbf{z}$ (eq. B-61) under neutrality as

$$\rho^{\circ}_{ab,t} = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \xi_t(z_{\mathrm{n},a})\xi_t(z_{\mathrm{m},b})q^{\circ}_{3,t}(\mathbf{z}_{\mathrm{n}},\mathbf{z}_{\mathrm{m}}|\mathbf{z})\mathrm{d}\mathbf{z}_{\mathrm{m}}\mathrm{d}\mathbf{z}_{\mathrm{n}} \tag{B-63}$$

where $q^{\circ}_{3,t}(\mathbf{z}_{\mathrm{n}},\mathbf{z}_{\mathrm{m}}|\mathbf{z})$ is the conditional p.d.f. for the event of sampling two individuals without replacement, one with phenotype $\mathbf{z}_{\mathrm{n}}$ and another with phenotype $\mathbf{z}_{\mathrm{m}}$ in a group, given that one with phenotype $\mathbf{z}$ has already been sampled in that group (under neutrality). Using a coalescence argument (see eq. B.10 of Wakano and Lehmann, 2014 for e.g.), we can rewrite this p.d.f. as

$$\begin{aligned} q^{\circ}_{3,t}(\mathbf{z}_{\mathrm{n}},\mathbf{z}_{\mathrm{m}}|\mathbf{z}) = {} & \delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z})\,\delta(\mathbf{z}_{\mathrm{m}}-\mathbf{z})\,r^{\circ}_{3,t}(\bar{\mathbf{z}}_t) \\ & + \delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z})\,(r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)-r^{\circ}_{3,t}(\bar{\mathbf{z}}_t))\,p_t(\mathbf{z}_{\mathrm{m}}) + \delta(\mathbf{z}_{\mathrm{m}}-\mathbf{z})\,(r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)-r^{\circ}_{3,t}(\bar{\mathbf{z}}_t))\,p_t(\mathbf{z}_{\mathrm{n}}) \\ & + \delta(\mathbf{z}_{\mathrm{n}}-\mathbf{z}_{\mathrm{m}})\,(r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)-r^{\circ}_{3,t}(\bar{\mathbf{z}}_t))\,p_t(\mathbf{z}_{\mathrm{n}}) \\ & + (1-3r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)+2r^{\circ}_{3,t}(\bar{\mathbf{z}}_t))\,p_t(\mathbf{z}_{\mathrm{n}})\,p_t(\mathbf{z}_{\mathrm{m}}), \end{aligned} \tag{B-64}$$

where $r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)$ is the probability of sampling three individuals without replacement from a group are identical-by-descent in the absence of selection at time t (i.e., three-way relatedness). Each summand of eq. (B-64) capture a different possible relationship among the sampled ($\mathbf{z}_{\mathrm{n}}$ and $\mathbf{z}_{\mathrm{m}}$) and focal ($\mathbf{z}$) phenotype. The first line on the right hand side of eq. (B-64) captures the event of sampling two individuals with the same phenotype as the focal ($\mathbf{z} = \mathbf{z}_{\mathrm{n}} = \mathbf{z}_{\mathrm{m}}$, which occurs with probability $r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)$); the next two lines, the event of sampling two individuals with the same phenotype (the three summands respectively capture the cases $\mathbf{z} = \mathbf{z}_{\mathrm{n}}$, $\mathbf{z} = \mathbf{z}_{\mathrm{m}}$ and $\mathbf{z}_{\mathrm{n}} = \mathbf{z}_{\mathrm{m}}$, that each occur with probability $r^{\circ}_{2,t}(\bar{\mathbf{z}}_t) - r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)$); and the last line, the event of sampling two individuals with different phenotypes ($\mathbf{z}_{\mathrm{n}} \neq \mathbf{z}_{\mathrm{m}}$) that are also different to the phenotype of the focal ($\mathbf{z} \neq \mathbf{z}_{\mathrm{n}}$ and $\mathbf{z} \neq \mathbf{z}_{\mathrm{m}}$). Substituting eq. (B-64) into eq. (B-63) gives us

$$\rho^{\circ}_{ab,t} = \xi_t(z_a)\xi_t(z_b)r^{\circ}_{3,t}(\bar{\mathbf{z}}_t) + \sigma_{ab,t}(r^{\circ}_{2,t}(\bar{\mathbf{z}}_t) - r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)), \tag{B-65}$$

for the covariance between traits among the neighbours of an individual with phenotype $\mathbf{z}$.

Finally, substituting eqs. (B-62) and (B-65) into eq. (B-60), we obtain that the effect on lineage of changing traits a and b in two neighbours can be written as

$$\begin{aligned} \int_{\mathcal{S}} \beta^{\mathrm{neigh}}_{ab,t}\, q_t(\mu|\mathbf{z}')\mathrm{d}\mu = {} & (\xi_t(z_a)\xi_t(z_b) - \sigma_{ab,t}) \\ & \times (N-1)\left(r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)\left(\frac{\partial^2 w}{\partial z_a \partial z_{\mathrm{n},b}} + \frac{\partial^2 w}{\partial z_b \partial z_{\mathrm{n},a}}\right) + (N-2)r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)\frac{\partial^2 w}{\partial z_{\mathrm{n},a}\partial z_{\mathrm{m},b}}\right). \end{aligned} \tag{B-66}$$

B.3.3 Polynomial approximation to lineage fitness

To summarise, we have derived a second-order polynomial approximation to lineage fitness (substituting eqs. B-54, B-59 and B-66 into eq. B-51). In order to express this approximation in a way that is more transparently connected with the results of the main text, we make a couple of changes to our notation so far. First, we make a change of variable express the derivatives of individual fitness $w(\mathbf{z}, \mathrm{n}_\mu(\mathbf{z}), \mathrm{e}(\phi_t))$ with respect to z_a, $z_{\mathrm{n},a}$, and $z_{\mathrm{nn},a}$ in terms of the derivatives of individual fitness $w(\mathbf{z}_i, \mathbf{z}_{-i}, \bar{\mathbf{z}}_t)$ (see eq. 7 of the main) with respects to the arguments $z_{i,a}$, $z_{j,a}$, and $z_{k,a}$, respectively. Second, we add evaluation signs of all derivatives to the population mean $\bar{\mathbf{z}}_t$. Our second-order polynomial approximation to lineage fitness can then be re-arranged to

read as

$$W(\mathbf{z},\phi_t) = W(\bar{\mathbf{z}},\phi_t) + \sum_{a=1}^{n} \xi_t(z_a) s_{a,t}(\bar{\mathbf{z}}_t) + \frac{1}{2}\sum_{a=1}^{n}\sum_{b=1}^{n} \xi_t(z_a)\xi_t(z_b) h_{ab,t}(\bar{\mathbf{z}}_t), \tag{B-67}$$

where

$$W(\bar{\mathbf{z}},\phi_t) = 1 - \frac{1}{2}\sum_{a=1}^{n}\sum_{b=1}^{n} \sigma_{ab,t} h_{ab,t}(\bar{\mathbf{z}}_t), \tag{B-68}$$

is the lineage fitness of the average phenotype $\bar{\mathbf{z}}_t$,

$$s_{a,t}(\bar{\mathbf{z}}_t) = \left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} + (N-1) r^{\circ}_{2,t}(\bar{\mathbf{z}}_t) \left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}}, \tag{B-69}$$

is the first-order effect of trait a on lineage fitness and

$$h_{ab,t}(\bar{\mathbf{z}}_t) = h_{\mathrm{w},ab,t}(\bar{\mathbf{z}}_t) + h_{\mathrm{r},ab,t}(\bar{\mathbf{z}}_t), \tag{B-70}$$

with

$$\begin{aligned}
h_{\mathrm{w},ab,t}(\bar{\mathbf{z}}_t) = {} & \left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,a}\partial z_{i,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} + r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)(N-1) \left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}\partial z_{j,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \\
& + r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)(N-1)\left(\left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,a}\partial z_{j,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} + \left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,b}\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} \right) \\
& + r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)(N-1)(N-2) \left.\frac{\partial^2 w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}\partial z_{k,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}}, \\
h_{\mathrm{r},ab,t}(\bar{\mathbf{z}}_t) = {} & \left.\frac{\partial r_{2,t}(\mathbf{z})}{\partial z_a}\right|_{\mathbf{z}=\bar{\mathbf{z}}_t} (N-1) \left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,b}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}} + \left.\frac{\partial r_{2,t}(\mathbf{z})}{\partial z_b}\right|_{\mathbf{z}=\bar{\mathbf{z}}_t} (N-1) \left.\frac{\partial w(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}},
\end{aligned} \tag{B-71}$$

is the second-order effect of traits a and b on lineage fitness.

C Dynamics of phenotypic mean and (co)variance

With our approximation for lineage fitness, we are now in a position to derive the dynamics of the mean of each trait $a \in \{1\ldots,n\}$ (eq. B-1) and of the (co)variances between all traits $a \in \{1\ldots,n\}$ and $b \in \{1\ldots,n\}$ (eq. B-23). We first derive these dynamics under the general case, and then under the assumption that the phenotypic distribution is Normal to derive eqs. (5) of the main text.

C.1 General recursions for the mean and (co)variance

C.1.1 Dynamics of mean phenotype

Substituting eq. (A-24) into the definition of the mean, eq. (B-1), we obtain that the mean value of trait a at time period $t+1$ is

$$\bar{z}_{a,t+1} = \int_{\mathbb{R}^n} z_a p_{t+1}(\mathbf{z})\mathrm{d}\mathbf{z} = \int_{\mathbb{R}^n} z_a W(\mathbf{z},\phi_t) p_t(\mathbf{z})\mathrm{d}\mathbf{z} + \nu\left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} z_a p_t(\mathbf{z}')\upsilon(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}'\mathrm{d}\mathbf{z} - \bar{z}_{a,t} \right). \tag{C-1}$$

Since the effects of mutation are assumed to be unbiased, we have

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} z_a p_t(\mathbf{z}')\upsilon(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}'\mathrm{d}\mathbf{z} = \bar{z}_{a,t}, \tag{C-2}$$

so that the bracketed term of eq. (C-1) vanishes, and we are left with

$$\bar{z}_{a,t+1} = \int_{\mathbb{R}^n} z_a W(\mathbf{z},\phi_t)p_t(\mathbf{z})\mathrm{d}\mathbf{z}. \tag{C-3}$$

The one-generational change in the mean of trait a then is

$$\Delta\bar{z}_{a,t} = \bar{z}_{a,t+1} - \bar{z}_{a,t} = \int_{\mathbb{R}^n} \xi_t(z_a)W(\mathbf{z},\phi_t)p_t(\mathbf{z})\mathrm{d}\mathbf{z}, \tag{C-4}$$

which corresponds to the first term of Price's equation, i.e., the change in average trait value in a population is equal to the covariance between trait and fitness (e.q. 4 of Price, 1970).

Substituting the second order approximation for lineage fitness, eq. (B-67), into eq. (C-4), we obtain that the change in the mean of trait a up to the second-order effects of selection is,

$$\Delta\bar{z}_{a,t} = \sum_{b=1}^{n}\sigma_{ab,t}s_{b,t}(\bar{\mathbf{z}}_t) + \frac{1}{2}\sum_{b=1}^{n}\sum_{c=1}^{n}\kappa_{abc,t}h_{bc,t}(\bar{\mathbf{z}}_t), \tag{C-5}$$

which depends on the skew,

$$\kappa_{abc,t} = \int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_b)\xi_t(z_c)p_t(\mathbf{z})\mathrm{d}\mathbf{z}, \tag{C-6}$$

between traits a, b and c in the population at time period t (in line with previously derived expressions for the change in the mean of single trait – which depends on the third-order centred moment of the phenotypic distribution; e.g., eq. 8a of Wakano and Iwasa, 2013; eq. A20 b of Débarre and Otto, 2016 in well-mixed populations; eq. 17 of Wakano and Lehmann, 2014 in the island model).

C.1.2 Dynamics of the phenotypic (co)variance

From definition (eq. B-23), the (co)variance within individuals between two traits a and b at time period $t+1$ can be written as

$$\begin{aligned}\sigma_{ab,t+1} &= \int_{\mathbb{R}^n}(z_a - \bar{z}_{a,t+1})(z_b - \bar{z}_{b,t+1})p_{t+1}(\mathbf{z})\mathrm{d}\mathbf{z} \\ &= \int_{\mathbb{R}^n}\xi_t(z_a)\xi_t(z_b)p_{t+1}(\mathbf{z})\mathrm{d}\mathbf{z} - \Delta\bar{z}_{a,t}\Delta\bar{z}_{b,t}.\end{aligned} \tag{C-7}$$

Substituting eq. (A-24) into the above, we obtain

$$\sigma_{ab,t+1} = \int_{\mathbb{R}^n}\xi_t(z_a)\xi_t(z_b)W(\mathbf{z},\phi_t)p_t(\mathbf{z})\mathrm{d}\mathbf{z} + \nu\left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\xi_t(z_a)\xi_t(z_b)p_t(\mathbf{z}')\upsilon(\mathbf{z}',\mathbf{z})\mathrm{d}\mathbf{z}'\mathrm{d}\mathbf{z} - \sigma_{ab,t}\right) - \Delta\bar{z}_{a,t}\Delta\bar{z}_{b,t}. \tag{C-8}$$

The bracketed term on the right hand side of eq. (C-8), which captures the effects of mutations (from parents with phenotype $\mathbf{z}'$ to descendant with phenotype $\mathbf{z}$), can be simplified by first writing out the product of

deviations in terms of parental phenotype as

$$\xi_t(z_a)\xi_t(z_b) = \xi_t(z'_a)\xi_t(z'_b) + (z_a - z'_a)(z_b - z'_b) + \xi_t(z_a)(z_b - z'_b) + (z_a - z'_a)\xi_t(z_b), \tag{C-9}$$

and second, by noting that since mutations are assumed to be unbiased, the covariance between parental phenotype and mutation effect is zero:

$$\int_{\mathbb{R}^n} (\xi_t(z_a)(z_b - z'_b) + (z_a - z'_a)\xi_t(z_b))\, \upsilon(\mathbf{z}', \mathbf{z})\mathrm{d}\mathbf{z} = 0. \tag{C-10}$$

Using eqs. (C-9)-(C-10), the effect of mutations in eq. (C-8) can then be written as

$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_b) p_t(\mathbf{z}')\upsilon(\mathbf{z}', \mathbf{z})\mathrm{d}\mathbf{z}'\mathrm{d}\mathbf{z} &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (\xi_t(z'_a)\xi_t(z'_b) + (z_a - z'_a)(z_b - z'_b))\, p_t(\mathbf{z}')\upsilon(\mathbf{z}', \mathbf{z})\mathrm{d}\mathbf{z}'\mathrm{d}\mathbf{z} \\ &= \sigma_{ab,t} + \sigma^{\mathrm{m}}_{ab},\end{aligned} \tag{C-11}$$

where we have defined

$$\sigma^{\mathrm{m}}_{ab} = \int_{\mathbb{R}^n} (z_a - z'_a)(z_b - z'_b)\upsilon(\mathbf{z}', \mathbf{z})\mathrm{d}\mathbf{z}, \tag{C-12}$$

as the (co)variance in mutational effects on traits a and b. Substituting eq. (C-11) into eq. (C-8), the (co)variance between two traits a and b at time period $t+1$ then is

$$\sigma_{ab,t+1} = \int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_b) W(\mathbf{z}, \phi_t) p_t(\mathbf{z})\mathrm{d}\mathbf{z} + \nu\sigma^{\mathrm{m}}_{ab} - \Delta\bar{z}_{a,t}\Delta\bar{z}_{b,t}. \tag{C-13}$$

Next, we substitute the approximation for lineage fitness, eqs. (B-67), and the change in mean we have derived above, eq. (C-5), into eq. (C-13), we obtain after some re-arrangements that the one-generational change in phenotypic (co)variance between traits a and b to the second-order effects of selection is

$$\Delta\sigma_{ab,t} = \nu\sigma^{\mathrm{m}}_{ab} + (W(\bar{\mathbf{z}}, \phi_t) - 1)\sigma_{ab,t} + \sum_{c=1}^{n} \kappa_{abc,t} s_{c,t}(\bar{\mathbf{z}}_t) + \frac{1}{2}\sum_{c=1}^{n}\sum_{d=1}^{n} \sigma_{abcd,t} h_{cd,t}(\bar{\mathbf{z}}_t) - \sum_{c=1}^{n}\sum_{d=1}^{n} \sigma_{ac,t}\sigma_{bd,t} s_{c,t}(\bar{\mathbf{z}}_t) s_{d,t}(\bar{\mathbf{z}}_t), \tag{C-14}$$

which depends on

$$\sigma_{abcd,t} = \int_{\mathbb{R}^n} \xi_t(z_a)\xi_t(z_b)\xi_t(z_c)\xi_t(z_d) p_t(\mathbf{z})\mathrm{d}\mathbf{z}, \tag{C-15}$$

i.e., the fourth central moment of the phenotypic distribution between traits a, b, c and d at time period t (in line with previously derived expressions for the change in the mean of single trait; e.g., eq. 8b of Wakano and Iwasa, 2013; eq. A24 b of Débarre and Otto, 2016 in well-mixed populations; eqs. B1-B8 of Wakano and Lehmann, 2014 in the island model).

C.2 Moment closure

Here, we derive the closed dynamical eqs. (5) for the means and variance-covariances found in the main text from the open dynamical eqs. (C-5) and (C-14) above. We achieve this closure in two steps.

C.2.1 Normal closure

First, we close the dynamical system on the means and covariances only (i.e., get rid of the dependency on skew, $\kappa_{abc,t}$, and fourth central moments, $\sigma_{abcd,t}$). To do so, we follow a number of authors that have explored the continuum of alleles model (e.g., Kimura, 1965a, Taylor and Day, 1997, Wakano and Iwasa, 2013, Débarre et al., 2014, Wakano and Lehmann, 2014, Débarre and Otto, 2016) and assume that the phenotypic distribution across the whole population is Normal. Under this assumption, the skew in the phenotypic distribution is zero,

$$\kappa_{abc,t} = 0, \tag{C-16}$$

and the fourth central moments can be expressed in terms of the (co)-variances,

$$\sigma_{abcd,t} = \sigma_{ab,t}\sigma_{cd,t} + \sigma_{ac,t}\sigma_{bd,t} + \sigma_{ad,t}\sigma_{bc,t}. \tag{C-17}$$

Substituting eqs. (C-16) and (C-17) into eqs (C-5) and (C-14), we obtain that the one-generational changes in means and covariances are respectively given by

$$\begin{aligned} \Delta\bar{z}_{a,t} &= \sum_{b=1}^{n} \sigma_{ab,t} s_{b,t}(\bar{\mathbf{z}}_t) \\ \Delta\sigma_{ab,t} &= \nu\sigma^{\mathrm{m}}_{ab} - \sum_{c=1}^{n}\sum_{d=1}^{n} \sigma_{ac,t}\sigma_{bd,t} s_{c,t}(\bar{\mathbf{z}}_t) s_{d,t}(\bar{\mathbf{z}}_t) + \frac{1}{2}\sum_{c=1}^{n}\sum_{d=1}^{n} \left(\sigma_{ac,t}\sigma_{bd,t} + \sigma_{ad,t}\sigma_{bc,t}\right) h_{cd,t}(\bar{\mathbf{z}}_t). \end{aligned} \tag{C-18}$$

In vector and matrix form, eq. (C-18) corresponds to eq. (5) of the main text, except that in eq. (C-18), the selection coefficients depend on time t. This is because the relatedness coefficients that appear in the selection coefficients ($r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)$, $r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)$, and $\partial r_{2,t}(\mathbf{z})/\partial z_a$ in eqs. B-69-B-71), which capture local (or group-level) moments of the phenotypic distribution, ϕ_t, still depend on time. We get rid off of this dependency and finally achieve closure in the next section.

C.2.2 Quasi-equilibrium

Our second step to close the dynamical system eq. (C-18) is to assume that dispersal is strong enough (relative to the intensity of selection) so that phenotypic associations between individuals within groups reach their steady-state values before any significant changes has occurred in the phenotypic distribution, $p_t(\mathbf{z})$, at the level of the population. This *quasi-equilibrium* assumption, which is frequently used in population genetic and social evolution theory (e.g., Kimura, 1965b, Nagylaki, 1993, Kirkpatrick et al., 2002, Roze and Rousset, 2005, Lehmann et al., 2007, Roze and Rousset, 2008) is in line with our assumption that selection is weak. It entails that we can evaluate $r^{\circ}_{2,t}(\bar{\mathbf{z}}_t)$, $r^{\circ}_{3,t}(\bar{\mathbf{z}}_t)$, and $\partial r_{2,t}(\mathbf{z})/\partial z_a$ in eqs. (B-69)-(B-71) at their quasi-equilibrium, i.e., we take the limits $\lim_{\tau\to\infty} r^{\circ}_{2,\tau}(\bar{\mathbf{z}}_t) = r^{\circ}_{2}(\bar{\mathbf{z}}_t)$, $\lim_{\tau\to\infty} r^{\circ}_{3,\tau}(\bar{\mathbf{z}}_t) = r^{\circ}_{3}(\bar{\mathbf{z}}_t)$, and $\lim_{\tau\to\infty} \partial r_{2,\tau}(\mathbf{z})/\partial z_a|_{\mathbf{z}=\bar{\mathbf{z}}_t} = \partial r_2(\mathbf{z})/\partial z_a|_{\mathbf{z}=\bar{\mathbf{z}}_t}$, while holding $p_t(\mathbf{z})$ constant. We thus denote by $r^{\circ}_{3}(\bar{\mathbf{z}}_t)$,$r^{\circ}_{3}(\bar{\mathbf{z}}_t)$, and $\partial r_2(\mathbf{z})/\partial z_a|_{\mathbf{z}=\bar{\mathbf{z}}_t}$, the steady-state values of neutral pairwise relatedness, neutral three-way relatedness, and the first-order perturbation of pairwise relatedness, respectively. Substituting these steady-state into the selection coefficients eqs. (B-69)-(B-71) (now independent of time so written as $s_a(\bar{\mathbf{z}}_t)$ and $h_{ab}(\bar{\mathbf{z}}_t)$), which are in turn substituted into eq. (C-18), we finally obtain the closed dynamical eqs. (5) of the main text.

Computing relatedness coefficients. Computing equilibrium relatedness coefficients under neutrality (i.e., $r^{\circ}_{2}(\bar{\mathbf{z}}_t)$, $r^{\circ}_{3}(\bar{\mathbf{z}}_t)$) is standard in population genetics and uses straightforward identity-by-descent arguments (e.g., Karlin, 1968, Rousset, 2004). Calculating the first-order effect of selection on pairwise relatedness,

$\partial r_2(\mathbf{z})/\partial z_a$, however, is more complicated. Under the quasi-equilibrium assumption, a perturbation on genetic associations between individuals within groups will generally depend on first-order perturbations of individual fitness and neutral relatedness coefficients (see Roze and Rousset, 2008 for a general treatment, in particular their eq. 67). So far, the first-order effect of selection on pairwise relatedness, $\partial r_2(\mathbf{z})/\partial z_a$, has been explicitly derived for two standard life-cycles, the semelparous Wright-Fisher life-cycle (in which all adults die after reproduction; see eq. 18 of Ajar, 2003 and eq. 28 of Wakano and Lehmann, 2014) and the iteroparous birth-death Moran life-cycle (in which a single adult dies after reproduction in each patch; see eq. 14 of Mullon et al., 2016). As noted in eq. (15) of Mullon et al. (2018), the effect of selection on relatedness can be written in both cases as

$$\begin{aligned}\left.\frac{\partial r_2(\mathbf{z})}{\partial z_a}\right|_{\mathbf{z}=\bar{\mathbf{z}}_t} = \kappa \frac{r_2^{\circ}(\bar{\mathbf{z}}_t)}{1-m}\Bigg[&\left[1+(N-1)r_2^{\circ}(\bar{\mathbf{z}}_t)\right]\left.\frac{\partial w^{\mathrm{P}}(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{i,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}}\\ &+\left[2r_2^{\circ}(\bar{\mathbf{z}}_t)+(N-2)r_3^{\circ}(\bar{\mathbf{z}}_t)\right](N-1)\left.\frac{\partial w^{\mathrm{P}}(\mathbf{z}_i,\mathbf{z}_{-i},\bar{\mathbf{z}}_t)}{\partial z_{j,a}}\right|_{\substack{\mathbf{z}_i=\bar{\mathbf{z}}_t\\ \mathbf{z}_{-i}=\bar{\mathbf{z}}_t}}\Bigg]\end{aligned} \tag{C-19}$$

where $\kappa = 2$ for the Wright-fisher and $\kappa = 1$ for the Moran life cycle (watch out for an unfortunate typo in eq. 15 of Mullon et al., 2018, which has "$\kappa = N$" under the Moran life cycle).

D Individual-based simulations

We performed individual based simulations for a population composed of N_{d} groups, each populated by N haploid individuals, using Mathematica 11.0.1.0 (Wolfram Research, Inc., 2016). Starting with a population monomorphic, we track the evolution of the multidimensional phenotypic distribution under the constant influx of mutations. Each individual $i \in \{1,\ldots,N_{\mathrm{d}}N\}$ is characterised by a vector of traits (z_1, z_2). At the beginning of a generation, we calculate the fecundity f_i of each individual according to its traits and those of its neighbours (using eq. 14). Then, we form the next generation of adults by sampling N individuals in each group with replacement, where each individual from the parental generation is weighted according to whether they belong to the group on which the breeding spot is filled or not. If an individual belongs to the same group in which a breeding spot is filled, then its weight is $f_i(1-m)$, where m is the backward dispersal probability. If it belongs to another group, then its weight is $f_i m/(N_{\mathrm{d}}-1)$. Once an individual is chosen to fill the breeding spot, it mutates with probability ν. If the offspring does not mutate, then it has the same phenotypic values as its parent. If it mutates, then we add to parental values a small perturbation that is sampled from a multivariate normal distribution with mean $(0,0)$ and variance-covariance matrix $\begin{pmatrix}\sigma_{11}^{\mathrm{m}} & \sigma_{12}^{\mathrm{m}}\\ \sigma_{12}^{\mathrm{m}} & \sigma_{22}^{\mathrm{m}}\end{pmatrix}$. The resulting phenotypic values are truncated to remain between -1 and 1. We repeat the procedure for a fixed number of generations (see Figures for details on parameter values used).