Evolution of quantitative traits: exploring the ecological, social and genetic bases of adaptive polymorphism

Laurent Lehmann* and Charles Mullon†

Abstract. Adaptive polymorphisms in quantitative traits are characterised by the coexistence of distinct genetically encoded phenotypes maintained by selection. Such intraspecific genetic and phenotypic variation makes a substantial contribution to biological diversity. Building on established results from invasion analysis, we analyse conditions under which adaptive polymorphism can emerge under long-term gradual evolution. We first note that, in general, fully characterising an adaptive polymorphism, including its number of morphs, trait values and frequencies, is a computationally hard problem. Nevertheless, standard invasion results reveal a unifying necessary condition for the gradual emergence of polymorphism: negative trait-dependent selection (NTDS), whereby population-level expression of a trait decreases selection acting on that same trait. We use NTDS as an organising principle to structure the pathways through which polymorphism can evolve. This is based on a decomposition of individual fitness that distinguishes production factors, the inputs to survival and reproduction that are endogenous to individuals, from regulating factors, the ecological, social or genetic inputs to fitness that are influenced in part by conspecifics. This decomposition allows us to clarify how NTDS arises across three broad classes of interactions – ecological, social and genetic – and to identify previously unrecognised adaptive polymorphisms in well-studied scenarios, including life-history, social preferences, and class-structured populations. More broadly, by integrating perspectives from population genetics, evolutionary ecology, adaptive dynamics, game theory and inclusive fitness theory, we aim to provide a common thread among the pathways through which selection favours intraspecific variation.

Keywords: adaptive polymorphism, adaptive dynamics, environmental feedbacks, environmental fluctuations, life-history, preference evolution

*Department of Ecology and Evolution, University of Lausanne, Switzerland.
†Department of Ecology and Evolution, University of Lausanne, Switzerland.

1 Introduction

Polymorphism–the coexistence of multiple genetically determined phenotypes within a population–is widespread in nature, from discrete colour morphs in many flowering plants and animals (e.g. Gigord et al., 2001; Sinervo and Lively, 1996), to alternative mating strategies in vertebrates (e.g. Gross, 1996; Oliveira et al., 2008), sexual dimorphism in many species (e.g. Lande, 1980; Fairbairn, 2013), and allelic diversity at metabolic loci (e.g. Gillespie, 1991; Tishkoff et al., 2001). Unlike polyphenism, where distinct phenotypes are environmentally induced by the same genotype (West-Eberhard, 2003), polymorphism necessarily involves genetic variation, maintained by the segregation of different alleles. Although non-selective processes such as mutation, recombination, and genetic drift shape genetic variation (Crow and Kimura, 1970; Lynch and Walsh, 2018), natural selection is the only evolutionary force that can maintain cohesive and clearly differentiated phenotypes over evolutionary timescales. Here, we use "adaptive polymorphism" to refer to an intraspecific polymorphism that is underlain by the segregation of different alleles whose coexistence is favoured by natural selection and by natural selection alone (thus resulting in the stable maintenance of phenotypic variation in the absence of mutation). Understanding the conditions that favour adaptive polymorphism is important to identifying the mechanisms that promote and maintain the diversity of forms and functions so characteristic of the natural world.

Different theoretical frameworks, including population genetics, evolutionary ecology, and evolutionary game theory, have proposed a variety of mechanisms that favour intraspecific adaptive polymorphism. These include temporally or spatially varying selection (e.g. Levene, 1953; Gillespie, 1991; Svardal et al., 2015; Yamamichi et al., 2023), pleiotropy and trade-offs between competitive traits due to functional constraints (e.g. de Mazancourt and Dieckmann, 2004; Ito and Sasaki, 2016), niche differentiation (e.g. Karlin and McGregor, 1972), strategic interactions producing frequency-dependent fitness effects in social contexts (e.g. Maynard Smith, 1982; McNamara and Leimar, 2020), and balancing selection via heterozygote advantage (e.g. Karlin and McGregor, 1972; Lewontin, 1974; Hartl and Clark, 2007; Sellis et al., 2011). This diversity of mechanisms can make it difficult, when analysing a specific biological scenario, to determine which intuition to apply or what general conditions, if any, govern the emergence of adaptive polymorphism.

A common feature unifying these diverse mechanisms is that some form of negative intraspecific population feedback must act to favour diversity. However, this feedback differs across theoretical frameworks. Evolutionary ecology models often invoke environmental feedbacks, mediated by density-dependent resource depletion, co-adapted parasitism, or prey- predator interactions (e.g. Geritz et al., 1998; Pásztor et al., 2016). Game-theoretic models emphasise social feedback and frequency-dependent selection, where the payoff to behavioural strategies, such as aggression, cooperation, or signalling, depends on their frequency in the population (e.g. Hofbauer and Sigmund, 1998; McNamara and Leimar, 2020). Population genetics, meanwhile, highlights genetic feedbacks, notably overdominance in diploids, where allelic fitness depends directly on allele frequencies within genotypes and thus on mating patterns (e.g. Crow and Kimura, 1970; Karlin and

McGregor, 1972; Ewens, 2004).

How do these various population feedbacks influence the emergence of adaptive polymorphism when new phenotypes regularly arise by mutation? What are the relationships between these different mechanisms? We address these questions in the context of quantitative traits evolving gradually through Darwinian adaptive dynamics, i.e. through the many successive small trait modifications. We build on the established theory of analysing the emergence of polymorphism under gradual evolution (e.g. Taylor, 1989; Christiansen, 1991; Metz et al., 1996; Eshel et al., 1997; Vincent et al., 1996; Geritz et al., 1998; Ajar, 2003; Rousset, 2004; Vincent and Brown, 2005; Otto and Day, 2007; Dercole and Rinaldi, 2008; Sasaki and Dieckmann, 2011; Brännström et al., 2013; Kisdi and Geritz, 2016; Mullon et al., 2016; Avila and Mullon, 2023), adopting evolutionary invasion analysis or long-term evolution theory ("ESS theory", see Van Cleve, 2023, for review). This provides a unifying framework to study conditions under which adaptive polymorphisms initially emerge from ecological, social, and genetic feedbacks.

Our study proceeds as follows. We begin by formalizing the concept of an adaptive polymorphism and show its useful equivalence with the Nash equilibrium of a population game. This correspondence highlights that mean fitness is maximized at an adaptive polymorphism yet it is, in general, computationally intractable to characterize this polymorphism explicitly. We therefore turn to classical invasion analysis, which identifies negative trait-dependent selection (NTDS), where mutant fitness increases when mutant and resident trait values diverge, as a key necessary condition for the gradual emergence of polymorphism. We next examine the population feedbacks that generate NTDS by decomposing individual fitness in terms of production and regulating factors, which link heritable trait values to survival and reproduction through their genetic, social and ecological effects. Building on this decomposition, we analyse how NTDS arises across three main classes of biological interactions: ecological, social, and genetic; and derive the conditions under which each can promote adaptive polymorphism. Finally, we apply these conditions to classical models of quantitative trait evolution, revealing previously overlooked polymorphisms.

2 Evolutionary model

2.1 Biological assumptions

We consider a large (ideally infinite) density-dependent regulated population. We start by assuming that individuals are haploid and reproduce asexually (this seemingly strong assumption is relaxed in section 6.1, where we consider diploidy and sexual reproduction). Individuals can belong to a finite number of different classes (e.g. be of different ages or be in different physiological state) and can be distributed in groups provided there is no isolation by distance (e.g. group or patch structure according to the infinite island model of dispersal, Wright, 1931; Rousset, 2004). The population may also experience environmental fluctuations in

space and time as long as the biotic and abiotic processes driving these fluctuations are stationary ergodic. We census this population at discrete demographic time steps between which individuals can survive, reproduce and interact.

Each individual expresses a quantitative or continuous trait; that is, a trait whose heritable value belongs to a subset $\mathcal{X} \subseteq \mathbb{R}$ of the real numbers. This trait can influence any aspect of an individual's biology, from its morphology to its physiology and behavior. We assume that the trait evolves gradually via the constant influx of rare genetic mutations of small effects.

2.2 Adaptive polymorphism

2.2.1 Background concepts

We are interested in describing what constitutes an adaptive polymorphism in the evolving trait. To that end, we introduce several necessary concepts. The first is the probability distribution p of trait values segregating in the population [p is an element of $\Delta(\mathcal{X})$, the set of probability distributions (measures) over the trait space $\mathcal{X}$]. For a finite number of morphs $p(x)$ can be thought as the frequency of the allele encoding trait value $x \in \mathcal{X}$. More generally, we do not assume that p has a density (with respect to Lebesgue measure) in order to be able to cover both the cases of a finite and infinite number of different segregating trait values. The probability distribution p is the main quantity used to characterise the state of a population in evolutionary biology, with its change in time constituting the basis of evolutionary dynamics (e.g., Wright, 1931; Hamilton, 1964; Crow and Kimura, 1970; Nagylaki, 1992; Charlesworth, 1994; Bürger, 2000, and see Bürger, 2000, chapter 4 for covering both finite and infinite number of segregating morphs). However, p alone is seldom sufficient to characterize evolutionary dynamics as these typically rely on several additional state variables (e.g., population size, environmental variables, morph distribution among individual classes, genotype frequencies for diploids, Crow and Kimura, 1970; Nagylaki, 1992; Charlesworth, 1994; Bürger, 2000). All the auxiliary state variables that are necessary to close the dynamics of the distribution p, we collect in ε. At an any demographic time step, the state of the population is therefore fully described by $\omega = (p, \varepsilon)$.

Following Ellner and Sasaki (1996), we assume that the evolutionary process in the absence of mutation is such that given any set of alleles initially segregating (i.e. given any subset of segregating morphs in $\mathcal{X}$), the population state ω converges towards an attractor that is stationary ergodic. We denote this attractor as μ, i.e. μ is the stationary distribution of ω over all states the population can be in. Thus, μ is a joint distribution of p and ε. We use μ_{p} to denote the marginal stationary ergodic distribution of p induced by the joint stationary distribution μ [μ_{p} is an element of $\Delta\Delta(\mathcal{X})$, the set of probability distributions (measures) over $\Delta(\mathcal{X})$]. Under a stationary ergodic process, an allele is either present or absent at all demographic time points; and in the absence of temporally fluctuating selection, μ_{p} simply reduces to p.

A population is polymorphic when the support $\mathcal{P}(\mu_\mathrm{p})$ of distribution μ_p contains two or more traits (or morphs). If the population is monomorphic for x, then $\mathcal{P}(\mu_\mathrm{p}) = \{x\}$. If two trait values x_1 and x_2 coexist, then $\mathcal{P}(\mu_\mathrm{p}) = \{x_1, x_2\}$. More generally, the support of the trait distribution can be finite, consisting of a finite number of traits (i.e. an oligomorphism *sensu* Sasaki and Dieckmann, 2011), or infinite consisting of either a continuum of traits or an infinite number of discrete traits. A population is polymorphic when the support $\mathcal{P}(\mu_\mathrm{p})$ contains two or more trait values (or morphs).

Let us now consider a mutant type x introduced as a single copy in a resident population with distribution μ_p induced by μ. Under our biological assumptions of section 2.1, the dynamics of the number of mutant copies descending from the initial mutant can be modelled as a multitype branching process (e.g., Harris, 1963; Mode, 1971 for constant environments and Tanny, 1981 for random environments). This branching process is assumed to satisfy two conditions: (i) the event of sure extinction of the mutant lineage occurs with probability one or the event of non-sure extinction occurs with probability one; and (ii) the process is strongly regular, biologically entailing that there is no class of individuals from which a mutant lineage starting with a single individual in that class is essentially sterile (Tanny, 1981 conditions Q and of strong regularity). It then follows from the (remarkable) classification theorem of Tanny (1981, p. 465) that the mutant lineage – started in any possible population state with a single mutant copy – goes extinct with probability one if and only if

$$\rho(x, \mu_\mathrm{p}) \leq 1. \tag{1}$$

Here, $\rho(x, \mu_\mathrm{p})$ is the long-term geometric mean of the average of the per-capita number of mutant copies produced per demographic time step in a resident population where segregating trait values are $\mathcal{P}(\mu_\mathrm{p})$ and will be called invasion fitness. We write $\rho(x, \mu_\mathrm{p})$ as depending on μ_p rather than μ because the distribution over auxiliary variables ε will not be needed here; our aim is not to compute invasion fitness explicitly but to use it to characterise polymorphism. Invasion fitness gives the long-term geometric growth ratio of the mutant lineage size when $\rho(x, \mu_\mathrm{p}) > 1$ (Tanny, 1981, p. 465) and it captures the various concepts of invasion fitness used in invasion analyses (e.g., Fisher, 1930; Eshel and Feldman, 1984; Altenberg and Feldman, 1987; Liberman, 1988; Lessard, 1990; Tuljapurkar, 1990; Metz et al., 1992; Charlesworth, 1994; Arnold et al., 1994; Rand et al., 1994; Ferrière and Gatto, 1995; Ellner and Sasaki, 1996; Metz et al., 1996; Otto and Day, 2007; Metz, 2011; Lehmann et al., 2016; Altenberg et al., 2017; Van Cleve, 2023).

2.2.2 Formalizing adaptation and biological interpretation

We now define an adaptive polymorphism as: a distribution μ_p^* such that (i) its support $\mathcal{P}(\mu_\mathrm{p}^*)$ contains two or more morphs; and (ii) this coalition of morphs is uninvadable, i.e.

$$\rho(x, \mu_\mathrm{p}^*) = 1 \qquad \text{for all } x \in \mathcal{P}(\mu_\mathrm{p}^*) \tag{2a}$$

$$\rho(x, \mu_\mathrm{p}^*) \leq 1 \qquad \text{for all } x \in \mathcal{X} \setminus \mathcal{P}(\mu_\mathrm{p}^*). \tag{2b}$$

Condition (2a) entails that each morph $x \in \mathcal{P}(\mu_{\mathrm{p}}^*)$ in the uninvadable coalition exactly replaces itself through survival and reproduction, and is therefore maintained in the population. Condition (2b), meanwhile, ensures that any type $x \in \mathcal{X} \setminus \mathcal{P}(\mu_{\mathrm{p}}^*)$ outside of the coalition goes extinct if introduced as a single copy into the population. The definition of an evolutionary stable strategy by Ellner and Sasaki (1996, p. 39) is equivalent to eq. (2), except that they impose a strong inequality in eq. (2b). The strong inequality, however, is not necessary where mutations are assumed to arrive as single (or few) copies in a large population and where a branching process is sufficient to determine their fate (eq. (1) holds).

In any case, when both conditions (2a)–(2b) hold, the polymorphism can no longer be altered through natural selection on the introduction of alternative morphs by mutation. Hence, the polymorphism is an equilibrium and each morph in the coalition satisfies the verbal biological interpretation of an adaptation: "a phenotypic variant that results in the highest fitness among a specified set of variants in a given environment" (Reeve and Sherman, 1993, p. 9). This biological interpretation of adaption suggests that in an adaptive polymorphic population state, population fitness should be maximized. This turns out to be the case and to see exactly what is meant by this, consider first the case of no temporal fluctuations where we can replace μ_{p}^* by p^* in eq. (2). Then, the distribution p^* is an adaptive polymorphism if:

$$\rho(x, p^*) = 1 \qquad \text{for all } x \in \mathcal{P}(p^*) \tag{3a}$$

$$\rho(x, p^*) \leq 1 \qquad \text{for all } x \in \mathcal{X} \setminus \mathcal{P}(p^*). \tag{3b}$$

Let us now denote

$$\overline{\rho}(q, p) = \int_{\mathcal{X}} \rho(x, p) \, \mathrm{d}q(x) \tag{4}$$

as the mean invasion fitness of a mutant coalition $q \in \Delta(\mathcal{X})$ in a population where the resident coalition is $p \in \Delta(\mathcal{X})$, where $\int_{\mathcal{X}} g(x) \, \mathrm{d}q(x)$ denotes the expectation of $g(x)$ over the probability measure q and $\int_{\mathcal{X}} \mathrm{d}q(x) = 1$. That is, $\overline{\rho}(q, p)$ gives the mean fitness over the morphs that are part of a coalition q, when each of these morphs are rare in a resident population characterised by p. Existing results then allow us to say that p^* satisfies eq. (3) if and only if it satisfies

$$p^* \in \underset{p \in \Delta(\mathcal{X})}{\arg\max} \, \overline{\rho}(p, p^*) \tag{5}$$

(Anderson et al., 2013, Proposition 1). In other words, p^* is an uninvadable coalition if and only if p^* maximizes mean fitness at the population state p^*. No other coalition has greater mean fitness than p^* when the population is characterised by p^*. Thus, in an uninvadable population state, all variation decreasing mean fitness has been removed. We show in Box 1 that this is also true under fluctuating environments.

2.2.3 The challenge of characterising an adaptive polymorphism

An adaptive polymorphism may consist of a finite or infinite number of morphs. Yet previous results suggests that long-term evolution should lead to a finite number of segregating morphs, i.e. to an oligomorphism (*sensu* Sasaki and Dieckmann, 2011). In particular, Ellner and Sasaki (1996) show that when invasion fitness is an analytic function, the uninvadable polymorphism can only be discrete (a qualitative result first observed in the context of polyphenism, Sasaki and Ellner, 1995). This is line with the finding that the equilibria of the continuum replicator dynamics in the absence of mutation consist of a finite number of coexisting morphs in panmictic populations without class structure (van Veelen and Spreij, 2009; Hingu et al., 2008). More generally, Gyllenberg and Meszéna (2005) demonstrate that the coexistence of an infinite number of morphs along a continuous range of trait values is not robustly stable, as it is sensitive to parameter changes. Hence, an adaptive polymorphism (if any) should ultimately be an oligomorphism.

Ideally, one would be able to characterize such an uninvadable oligomorphism in a quantitative way for any given biological scenario; that is, one would be able to tell the number of morphs in the coalition, the traits they express, and their frequency in the population. Moreover one would like to show that such an equilibrium is evolutionarily attainable, i.e. is convergence stable (Eshel, 1983; Eshel et al., 1997; Leimar, 2009). Working towards these general goals, however, proves especially challenging, even in the absence of temporally fluctuating selection and class structure.

The challenge is revealed by the equivalence between eqs. (3) and (5), which establishes a connection between an uninvadable coalition and a Nash equilibrium of a population game (e.g. Sandholm, 2011; Anderson et al., 2013). This connection provides us with two insights (Box I for details). First, an uninvadable coalition exists (under the assumptions that invasion fitness is continuous in its arguments and $\mathcal{X}$ is a closed interval). Second, characterizing the uninvadable coalition – i.e. computing the Nash equilibrium – is likely to be a computationally hard problem. Hence, characterising the long-term equilibrium can be intractable, even leaving aside the questions of the evolutionary convergence towards adaptive polymorphism and the complication of population heterogeneity.

To circumvent these difficulties, we next focus on one relevant segment of the evolutionary dynamics and investigate the conditions that are required for adaptive polymorphism to emerge from a population initially consisting of a single morph.

2.3 Monomorphic instability through invasion analysis

2.3.1 Gradual emergence of polymorphism

The mathematical machinery for determining when a monomorphic population becomes polymorphic through gradual evolution is well-established (e.g. Taylor, 1989; Christiansen, 1991; Ellner and Sasaki, 1996; Metz et al., 1996; Eshel et al., 1997; Geritz et al., 1998; Rousset, 2004; Otto and Day, 2007; Dercole and Rinaldi, 2008; Sasaki and Dieckmann, 2011; Brännström et al., 2013; Kisdi and Geritz, 2016; Mullon et al., 2016). We summarise the established results needed for our analysis here. The starting point is the invasion fitness of a mutant with trait value x in a resident population monomorphic for trait value y. Since the resident population is monomorphic (i.e. $\mathcal{P}(\mu) = \{y\}$), we from now on use the notation $\rho(x, y)$ to represent the invasion fitness of mutant x in a y population. We assume that $\rho(x, y)$ is differentiable with respect to both of its arguments. Owing to density-dependent regulation, the invasion fitness of a single copy of a resident y must be one, $\rho(y, y) = 1$; that is, each gene copy on average replaces itself in a monomorphic population at its demographic attractor.

Whether a population monomorphic for y is evolutionarily unstable and whether such instability favours polymorphism can be examined from the following three coefficients of selection:

$$S(y) = \left.\frac{\partial \rho(x, y)}{\partial x}\right|_{x=y}, \quad H(y) = \left.\frac{\partial^2 \rho(x, y)}{\partial x^2}\right|_{x=y} \text{ and } \quad I(y) = \left.\frac{\partial^2 \rho(x, y)}{\partial x \partial y}\right|_{x=y}. \tag{6}$$

We refer to these as the selection gradient (S), the disruptive selection coefficient (H), and the selection interaction coefficient (I), respectively. The selection gradient $S(y)$ indicates the direction of selection on the trait in a population at y. When $S(y) > 0$ ($S(y) < 0$), a mutation with small positive (negative) effect on trait value has a positive probability of invading, and if it does invade, it fixes and substitutes the resident (Eshel et al., 1997; Rousset, 2004; Cai and Geritz, 2020; Priklopil and Lehmann, 2021). Therefore, $S(y)$ is sufficient to characterise gradual evolution under the constant but rare input of mutations with small effects: the population will converge either to a boundary point of the state space $\mathcal{X}$, or to a trait value y^* lying in the interior of $\mathcal{X}$ such that

$$S(y^*) = 0 \tag{7}$$

$$J(y^*) = \left.\frac{\mathrm{d}S(y)}{\mathrm{d}y}\right|_{y=y^*} = H(y^*) + I(y^*) < 0 \tag{8}$$

(e.g. Eshel, 1983; Taylor, 1989; Eshel et al., 1997; Leimar, 2009). A trait value y^* satisfying the condition in eq. (7) is known as a singular trait value, while those satisfying also the condition in eq. (8) are further said to be convergence stable. The coefficient $H(y^*)$, meanwhile, indicates whether selection is stabilizing or disruptive when all individuals in a population express a singular trait value y^*. When $H(y^*) < 0$, any change in trait value decreases invasion fitness, so that selection ousts any mutant whose value deviates

locally from y^*, i.e. selection is stabilizing for y^*. In contrast, when $H(y^*) > 0$ any trait change increases fitness so that selection favours any mutant whose value deviates locally from y^*, i.e. selection is disruptive. When a convergence stable trait value y^* is such that $H(y^*) > 0$, the population first converges to y^* and then gradually splits into two morphs (Eshel et al., 1997; Geritz et al., 1998). This represents the gradual emergence of polymorphism due to selection in a process often referred to as evolutionary branching (yet polymorphism is not necessarily systematically permanent after branching, Geritz et al., 1999).

2.3.2 Negative-trait dependent selection (NTDS) is necessary for polymorphism

It follows that two conditions must hold for a monomorphic population to evolve an adaptive polymorphism: $H(y^*) + I(y^*) < 0$ (for convergence stability) and $H(y^*) > 0$ (for disruptive selection). Together, these imply that

$$I(y^*) < 0, \tag{9}$$

i.e. the selection interaction coefficient must be negative at y^*. The selection interaction coefficient measures how selection on the trait (as in $\partial\rho(x, y)/\partial x$) changes with the trait value y expressed in the resident population (recall eq. 6), and thus quantifies trait-dependent selection. Condition (9) means that negative trait-dependent selection (NTDS) is necessary for adaptive polymorphism to emerge by gradual evolution. Biologically, this corresponds to situations where mutant invasion fitness increases when mutant and resident trait values vary in opposite directions. That is, when individuals in the population do more of something (e.g. increase y), the mutant's fitness increases when it does less (e.g. decrease x). Mutant and population trait values are thus strategic substitutes *sensu* game theory (Bulow et al., 1985).

The relevance of condition (9) is underscored by two further results. First it is also the local requirement for mutual invasibility : two types that are weakly diverged around a singularity y^* can invade one another and thus coexist in a protected polymorphism only if $I(y^*) < 0$ (Eshel et al., 1997; Geritz et al., 1998, and see Appendix B here). That is, condition (9) is necessary for negative-frequency dependent selection near y^* (Maynard Smith, 1998; Heino et al., 1998). Second, once $I(y^*) < 0$ holds, it is in principle always possible to define a trade-off or add an individual cost of trait expression such that $H(y^*) > 0$ and $J(y^*) < 0$ (Kisdi, 2006, 2015; Ito and Sasaki, 2016). Hence, $I(y^*) < 0$ often represents the limiting condition for the evolution of adaptive polymorphism, because without it the negative feedback required for coexistence cannot be.

While our focus here is on the conditions for instability of monomorphism, the same principles can be extended to study the instability of polymorphism. Examining invasion fitness and selection coefficients, this time in polymorphic resident populations, allows one to determine whether a given oligomorphism is convergence stable and uninvadable (e.g. Geritz et al., 1998; Ito and Dieckmann, 2007; Sasaki and Dieckmann, 2011). In such analyses, NTDS remains the key diagnostic quantity, capturing the negative population feedback on which adaptive polymorphism ultimately rests (Appendix C for details). In what follows, we use $I(y^*)$ as a unifying measure to delineate the bases of adaptive polymorphism.

3 Fitness factors: production, regulation, and their interactions

To specify the negative population trait feedback embodied by $I(y^*) < 0$ and understand better the biological processes via which it can be generated, we need to be more explicit about what factors influence individual survival and reproduction – "fitness factors" for short – and thus contribute to invasion fitness. From its definition, invasion fitness $\rho(x, y)$ can be seen as the average of the individual fitness of mutants. This average, which may be arithmetic, geometric, or a combination of the two depending on the biological situation, needs to be taken over all the relevant contexts in which the gene that code for x may find itself; that is, all types of genetic (e.g. genomic background), demographic (e.g. population size), ecological (e.g. prey density), and environmental states (e.g. temperature, see examples below and Box 2). Invasion fitness may thus be influenced by an untold number of factors. Yet, these can often be usefully grouped into different categories, as we next outline.

3.1 Limiting, regulating, and production factors

Fitness factors that have always been considered fundamental are those limiting survival and reproduction, i.e. limiting factors (Grovers, 1997). These fall into two types. On the one hand, limiting factors may be exogenous, such as temperature, sunlight, or humidity. On the other hand, limiting factors can be endogenous, meaning that their availability depends on conspecifics, conspecifics' abundance, and the traits they express. Examples include available space or nutrients. These endogenous limiting factors are special cases of a broader category usually referred to as *regulating factors* (e.g. May, 1973; Armstrong and McGehee, 1980; Grovers, 1997; Pásztor et al., 2016). Regulating factors are defined here as inputs to individual survival and reproduction that depend directly or indirectly on the traits expressed by *other* individuals in the population. Regulating factors are typically thought to have negative effects on individual fitness, for instance through competition for limited resources (Grovers, 1997). However, regulating factors may also enhance fitness, as occurs when mutualistic interactions improve resource availability. In this paper, regulating factors represent the first of two broad categories of fitness factors we explicitly consider.

The second category, which we refer to as *production factors*, comprises inputs to individual fitness that depend, at least partially, on the trait x of the focal individual. Unlike regulating factors, which depend solely on others, production factors reflect intrinsic characteristics of the focal individual, such as baseline fecundity in the absence of competition, germination probability, energy absorption rates, or behavioral predisposition. As these examples suggest, production factors are typically associated with improved survival and/or reproduction. However, production outputs of one individual often affect the regulating factors experienced by others, creating indirect interactions among individuals. These interactions between regulating and production factors may combine in complex ways to shape overall individual reproduction and survival. To disentangle these effects and explore their evolutionary implications, we next introduce a model explicitly

linking trait x to its invasion fitness through a structured hierarchy of regulating and production factors.

3.2 Invasion fitness in terms of regulating and production factors

The first production factor we consider is the phenotype, defined broadly here as any physiological, morphological, or behavioural phenotype expressed by an individual (e.g. nutrient uptake rate, coat colour, effort into alloparental care). The phenotype expressed by an individual carrying trait x may depend not only on x but also on its direct or indirect interactions with others in the population. To capture this idea, we write the n_{z}-dimensional phenotype $\mathbf{z}$ of an individual as a function $\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x, y)) \in \mathbb{R}^{n_{\mathrm{z}}}$ that depends on: (i) the heritable trait value x it carries (first argument); and (ii) a set of r_{z} regulating factors, which can be formalised as a function $\mathbf{e}^{\mathrm{z}}(x, y) \in \mathbb{R}^{r_{\mathrm{z}}}$ of mutant x and resident y trait values

For example, in a model of social interactions where individuals can respond to each other's actions (e.g. Akçay and Van Cleve, 2012), the behaviour of a mutant carrier depends on the trait of its social partners (section 5.2 for examples). These interaction partners may be residents (under random interactions), or possibly other mutants (under preferential interactions among relatives or limited dispersal; section 6.2 for examples). Similarly, in diploid populations, this formulation allows for gene interactions within individuals on the phenotype, where mutant carriers are heterozygotes under random mating but may be homozygotes (carrying two copies of x) under inbreeding (section 6.1 for more details on diploids). More generally, $\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x, y))$ allows us to capture biotic and abiotic regulating factors that determine the expression of quantitative traits, including environmental effects, non-additive genetic effects, gene-by-environment effects, and indirect genetic effects (Lynch and Walsh, 2018).

The second production factor we consider explicitly is termed payoff, an umbrella term for all the material and informational inputs that are acquired by the individual and that serve for its survival or reproduction (e.g. number of mates, prey caught, quality of breeding spot, calories, knowledge, Maynard Smith, 1982; Stephens and Krebs, 1987; Houston and McNamara, 1999). We conceive the n_{π}-dimensional payoff of an individual as a function $\boldsymbol{\pi}(\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x, y)), \mathbf{e}^{\pi}(x, y)) \in \mathbb{R}^{n_{\pi}}$ that depends on: (i) the phenotype $\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x, y))$ it expresses (first argument); and (ii) a number r_{π} of regulating factors, given by a function $\mathbf{e}^{\pi}(x, y) \in \mathbb{R}^{r_{\pi}}$ of mutant and resident trait values. In evolutionary games for instance, such regulating factors would include the behavioural phenotype of others that the individual interacts with, who may also be mutant carriers where interactions occurs in the family or in dispersal-limited groups, or whose behaviour depend on the individual's own trait under social responsiveness.

Finally, the translation of payoff into actual survival and reproduction may itself depend on a further regulating factors, that we collect in $\mathbf{e}^{\mathrm{w}}(x, y) \in \mathbb{R}^{r_{\mathrm{w}}}$. These additional r_{w} regulating factors include every biotic and abiotic factors outside the individual that may influence its survival and reproduction and that are affected by conspecifics (holding the phenotype and payoff of the individual fixed). Examples include population

density, the payoff of competitors, or prey density. In a well-mixed population, such regulating factors would typically depend on the resident trait y, but under limited dispersal, they may also depend on the mutant x (e.g. Rousset and Ronce, 2004; Mullon and Lehmann, 2018). Putting it all together, we obtain that the invasion fitness of the mutant x in a resident population characterised by trait y can be written as

$$\rho(x,y) = W\big(\boldsymbol{\pi}(\mathbf{z}(x,\mathbf{e}^{\mathrm{z}}(x,y)),\mathbf{e}^{\pi}(x,y)),\mathbf{e}^{\mathrm{w}}(x,y)\big), \tag{10}$$

where the function W has as its first argument the payoff $\boldsymbol{\pi}$ and second argument the regulating factors $\mathbf{e}^{\mathrm{w}}(x,y)$ (see Fig. 1 for an illustration).

3.3 Intensive versus extensive regulating factors and their trait dependence

To connect eq. (10) to the literature, it is useful to recognise that a regulating factor can be one of two types. The first type a regulating factor can be is an extensive variable, which is a variable indicating the size of a system under focus (Kondepuli and Prigogine, 2015), such as population size, number of competitors, or total resource availability. In evolutionary ecology, such variables are often treated as state variables whose equilibrium values are determined by the demographic structure or dynamics of the resident population (e.g. Diekmann et al., 2003; Pásztor et al., 2016). We illustrate this notion in Box 2, where regulating factors are determined by resident equilibrium equations (see also Diekmann et al., 2003, eq. 2.1; Kisdi and Geritz, 2016, eq. 8).

Alternatively, a regulating factor can be an intensive variable, which is a variable that is not indicative of the size of the system, i.e. variables that do not scale with system size (Kondepuli and Prigogine, 2015). Instead, these reflect per capita quantities, such as average survival, average fecundity, or local competitive pressure. This perspective is particularly common in population and quantitative genetics, or evolutionary game theory models with constant population size, where fitness is defined in terms of relative success (e.g., ratio of fecundities). This is the case for instance in the classical Wright-Fisher and Moran models of population genetics, where an individual's fitness is expressed directly as a function of the survival probability and/or fecundity of others (e.g. Maynard Smith, 1982; Gillespie, 2004; Hartl and Clark, 2007; Ewens, 2004; Lynch and Walsh, 2018).

The expression of invasion fitness in eq. (10) in terms of regulating factors shows that mutant and resident traits can interact in many different ways, and hence that $I(y)$ in eq. (6) is generally non-zero. Because fitness results from a hierarchy of nested interactions (trait $\to$ phenotype $\to$ payoff $\to$ fitness), x and y can influence $\rho(x,y)$ through multiple biological pathways. The structure of eq. (10) allows us to identify three main pathways leading to NTDS: when interactions are (i) ecologically, (ii) socially, or (iii) genetically mediated, which we explore in turn below. For each, we characterise $I(y)$ and illustrate how NTDS may arise through examples. Although these examples are based on scenarios that have been studied previously, all

the polymorphisms we report here are new. In cases where selection leads to evolutionary branching, we also analyse the stability of the resulting oligomorphism for completeness in the Appendix (following the analysis detailed in Appendix C).

4 Ecologically mediated interactions

4.1 Ecological setup

Many interactions among individuals are indirect, occurring through environmental change, e.g., through resource consumption or habitat modification. In such cases, trait dependence arises via ecological feedbacks, where changes in population trait values alter environmental conditions, which in turn impact individual reproduction and survival. To formalise this, we set $\boldsymbol{\pi}(\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x,y)), \mathbf{e}^{\pi}(x,y)) = \mathbf{z}(x)$ and $\mathbf{e}^{\mathrm{w}}(x,y) = \mathbf{e}(y)$ in eq. (10) so that an individual's payoff depends only on its own trait x, while the environment is determined solely by the resident trait y. Invasion fitness then simplifies to $\rho(x,y) = W(\mathbf{z}(x), \mathbf{e}(y))$. We refer to "ecologically mediated interactions" as cases where fitness takes this form, with interactions mediated through a regulating factor $\mathbf{e}(y)$ exogenous to the individual. These interactions can nevertheless involve complex ecological feedback loops, potentially spanning multiple environmental state variables (e.g., multitrophic consumer-resource models). In the ecological literature, the production factors $\mathbf{z}(x)$ are sometimes referred to as environmentally dependent parameters, while regulating factors $\mathbf{e}(y)$ are called competition parameters (Chesson, 2000a, p. 212-213).

For simplicity, we further assume that

$$\rho(x,y) = W(\underbrace{z_1(x), z_2(x)}_{=\mathbf{z}(x)}, \underbrace{e_1(y), e_2(y)}_{=\mathbf{e}(y)}), \tag{11}$$

so that the function W depends on only four arguments, which are sufficient to capture standard pathways to trait-dependent selection under ecological interactions. The first two arguments, $z_1(x)$ and $z_2(x)$, are production factors, assumed to never decrease fitness such as fecundity or survival ($\partial W/\partial z_1 \geq 0$ and $\partial W/\partial z_2 \geq 0$). The second two, $e_1(y)$ and $e_2(y)$, are regulating factors, which are assumed to only feedback negatively on fitness ($\partial W/\partial e_1 \leq 0$ and $\partial W/\partial e_2 \leq 0$). As discussed in section 3.3, the regulating factors can be conceptualized as intensive variables, in which case one can set $e_1(y) = z_1(y)$ and $e_2(y) = z_2(y)$, or as extensive variables, which would then satisfy an equilibrium system of equations as detailed in Box 2.

Substituting eq. (11) into eq. (7), shows that any singularity y^* must satisfy

$$\frac{\partial W}{\partial z_1}\frac{\mathrm{d}z_1(x)}{\mathrm{d}x}\bigg|_{x=y=y^*} + \frac{\partial W}{\partial z_2}\frac{\mathrm{d}z_2(x)}{\mathrm{d}x}\bigg|_{x=y=y^*} = 0. \tag{12}$$

(here and throughout, we drop the argument of functions with multiple arguments when the notation becomes too cumbersome and there is no risk of confusion). Eq. (12) shows that for an evolutionary equilibrium y^* to exist, there must be a trade-off between the fitness returns obtained via the two production factors. Specifically, since fitness increases with both production factors ($\partial W/\partial z_1 \geq 0$ and $\partial W/\partial z_2 \geq 0$), an evolving trait cannot increase both factors indefinitely, otherwise selection would continuously favour larger trait values. Instead, the trait must increase one production factor while reducing the other, with these opposing effects on fitness balanced at y^*.

Because invasion fitness in a monomorphic population must equal one, i.e. $W(z_1(y), z_2(y), e_1(y), e_2(y)) = 1$, we also have

$$\frac{\partial W}{\partial e_1}\frac{\mathrm{d}e_1(y)}{\mathrm{d}y}\bigg|_{y=y^*} + \frac{\partial W}{\partial e_2}\frac{\mathrm{d}e_2(y)}{\mathrm{d}y}\bigg|_{y=y^*} = 0 \tag{13}$$

(using eq. 12). Since $\partial W/\partial e_1 \leq 0$ and $\partial W/\partial e_2 \leq 0$, eq. (13) indicates that $\mathrm{d}e_1(y)/\,\mathrm{d}y$ and $\mathrm{d}e_2(y)/\,\mathrm{d}y$ must have opposite signs at $y = y^*$. In other words, a trait increase in the population must be associated with opposite effects on regulating variables 1 and 2. This result expresses a fundamental ecological or demographic constraint: due to density-dependent regulation, an increase in competition for one resource must be offset by a decrease in competition for another at the singularity.

4.2 Interaction coefficient

We now turn to specifying the interaction coefficient $I(y^*)$ under the above setup. Substituting eq. (11) into eq. (6) and using eqs. (12)–(13), we obtain

$$I(y^*) = L(y^*)K(y^*), \tag{14}$$

where $L(y^*) = (\mathrm{d}z_1(y)/\,\mathrm{d}y)(\mathrm{d}e_1(y)/\,\mathrm{d}y)|_{y=y^*}$ is the product of the marginal changes in the first production and regulating factors; and

$$K(y^*) = \left[\left(\frac{\partial}{\partial z_1} - M_\mathrm{z}(y^*)\frac{\partial}{\partial z_2}\right)\left(\frac{\partial}{\partial e_1} - M_\mathrm{e}(y^*)\frac{\partial}{\partial e_2}\right)W(\mathbf{z},\mathbf{e})\right]_{x=y=y^*}. \tag{15}$$

Here, $M_\mathrm{z}(y^*)$ and $M_\mathrm{e}(y^*)$ are marginal rates of factor substitution (Jehle and Reny, 2011, p. 128) (see Appendix D.1 for calculations). Specifically, $M_\mathrm{z}(y^*) = (\partial W/\partial z_1)/(\partial W/\partial z_2)|_{x=y=y^*} \geq 0$ is the rate at which production factor z_2 can be substituted for z_1 without changing fitness: it measures the change in the number of units of z_2 required to compensate for a one-unit increase in z_1 in terms of fitness effects. Similarly, $M_\mathrm{e}(y^*) = (\partial W/\partial e_1)/(\partial W/\partial e_2)|_{x=y=y^*} \geq 0$ is the rate at which limiting factor e_2 can be substituted for e_1 without changing fitness.

Before interpreting $K(y^*)$, let us remark that for most ecologically relevant models, $L(y^*)$ will be nonzero ($L(y^*) \neq 0$) since in the presence of a trade-off, neither a production factor nor a regulating factor will be

pushed to a maximum by evolution (i.e. $\mathrm{d}z_1(y)/\,\mathrm{d}y \neq 0$ and $\mathrm{d}e_1(y)/\,\mathrm{d}y \neq 0$ at $y = y^*$). In fact, for any given model one should be able to define fitness factors such that $L(y^*) > 0$, since when a trait increases the production factor for one individual (e.g., a trait increases its fecundity), this is likely to increase a regulating factor when expressed by others (e.g., greater density of offspring before density-dependent competition for recruitment). In particular, when regulating factors are taken as intensive variables (such that $e_1(y) = z_1(y)$, $e_2(y) = z_2(y)$), this leads to $L(y^*) = (\mathrm{d}z_1(y^*)/\,\mathrm{d}y^*)^2 \geq 0$. Either way, condition (9) simplifies to

$$K(y^*) < 0, \tag{16}$$

which therefore provides the condition for NTDS.

Eq. (15) can then be interpreted as follows. The first operator, $(\partial/\partial z_1 - M_{\mathrm{z}}(y^*)\,\partial/\partial z_2)$, can be seen as measuring the net effect of a marginal change in the mutant trait x via both production factors, i.e., taking into account the fact that near y^*, an increase in one production factor necessarily entails a reduction in the other (recall eq. 12). This trade-off is captured by the coefficient $-M_{\mathrm{z}}(y^*)$. The second operator, $(\partial/\partial e_1 - M_{\mathrm{e}}(y^*)\,\partial/\partial e_2)$, measures the net effect of a marginal change in the resident trait y via the regulating factors, i.e., taking into account the constraint that at resident demographic equilibrium, an increase in one regulating factor is necessarily associated with a reduction in the other (recall eq. 13). This constraint is captured by the coefficient $-M_{\mathrm{e}}(y^*)$. When applied successively to $W(\mathbf{z}, \mathbf{e})$, the two operators thus quantify the effect of joint changes in mutant and resident traits on mutant fitness, as expected.

The structure of eq. (15) helps understand the origin of NTDS from ecologically mediated interactions. It shows that NTDS arises when the trait induced trade-off in production factors (first operator) interacts negatively with the trait induced constraint on ecological feedback (second operator). NTDS thus requires that once these feedbacks have been taken into account, net trait-dependent selection occurring via each production factor are decoupled. This can be illustrated with a local adaptation scenario (e.g. Levene, 1953; Gillespie, 1991; Kisdi and Geritz, 1999; Svardal et al., 2015). Suppose z_1 and z_2 are fecundity in two habitats (say a wet and a dry patch), while e_1 and e_2 are the corresponding local densities. Suppose also that a mutant increasing its trait value increases its fecundity under wet conditions, which by trade-off, necessarily means decreased fecundity in the dry patch. If at the same time, the population trait varies in the opposite direction, then density increases in the dry patch and decreases in the wet one. When these two changes exactly compensate one another – so that once feedbacks are accounted for, the fitness gains by the mutant in the wet patch is exactly offset by the loss in the dry one – then $K(y^*) = 0$. When, by contrast, these two changes reinforce each other, so that the fitness gain in the wet patch is greater than the loss in the dry one, then $K(y^*) < 0$, thus enabling NTDS.

A direct consequence of eq. (14) is that polymorphism cannot emerge when invasion fitness depends on a single production or a single regulating factor, since in this case, $K(y^*) = 0$ and the necessary condition for polymorphism can never be met. Specifically, when only one production (regulating) factor, say z_2 (e_2)

affects fitness while the other z_1 (e_1) does not, it follows that $M_{\mathrm{z}}(y^*) = 0$ ($M_{\mathrm{e}}(y^*) = 0$) and $\partial^2 W/(\partial z_1 \partial e_1) = \partial^2 W/(\partial z_1 \partial e_2) = 0$ [$\partial^2 W/(\partial z_1 \partial e_1) = \partial^2 W/(\partial z_2 \partial e_1) = 0$]. For regulating factors, this restates a classical result in theoretical ecology: at least two independent regulating factors are needed to sustain interspecific coexistence (e.g. Levin, 1970; Haigh and Maynard Smith, 1972; Meszéna et al., 2006; Pásztor et al., 2016). Eq.(15) additionally shows that an analogous requirement holds for production factors in the intraspecific case : polymorphism requires at least two production factors.

In addition to aiding interpretation, eq. (15) is also methodologically useful. Importantly, eq. (15) does not rely on an explicitly specification of how traits influence fitness factors (i.e. how x and y influence $\mathbf{z}$ and $\mathbf{e}$). This generality makes it particularly valuable as it provides a direct means to identify whether polymorphism can be promoted in a given biological scenario. In the next sections, we apply this approach to two different scenarios, each revealing previously undescribed polymorphisms.

4.3 Life-history evolution and limited dispersal

We first consider a classic scenario of life-history evolution involving a trade-off between fecundity and survival, two fundamental production factors (e.g. Charnov, 1997; Case, 2000). Although such trade-offs have been studied extensively, it remains unresolved whether polymorphism can emerge when dispersal is limited and generations overlap. To address this, we consider a subdivided iteroparous population with the following life-cycle events: (a) Each adult individual occupies a site where it produces a large number of juveniles and then either survives or dies. (b) Each juvenile either migrates out of the natal site with probability m or remains philopatric with probability $1 - m$. (c) Density-dependent competition occurs locally among juveniles for the breeding spots vacated by adult mortality, so that the adult population is regulated back to a constant level. With $z_1(x)$ denoting the survival probability of an adult expressing trait x and $z_2(x)$ its fecundity, invasion fitness for this model is

$$\rho(x, y) = z_1(x) + (1 - z_1(x))\frac{(1-m)z_2(x)}{(1-m)z_2(x) + me_2(y)} + (1 - e_1(y))\frac{mz_2(x)}{e_2(y)}, \tag{17}$$

where the regulating factors are $e_1(y) = z_1(y)$ and $e_2(y) = z_2(y)$, i.e. the resident's survival and fecundity (intensive variables). With complete dispersal $m = 1$, eq. (17) reduces to the standard reproductive-effort model in panmictic populations, for which polymorphism cannot emerge (Charnov, 1997; Case, 2000). More generally, eq. (17) coincides with the model of Pen (2000) for one individual per patch and to the model of Ronce and Promislow (2010) in the absence of senescence; however, those studies focused on directional selection and did not analyse the emergence of polymorphism.

Under model (17), the singular reproductive effort y^* involves a positive allocation to reproduction, and this allocation increases as dispersal becomes more limited (Pen, 2000, eq. 8; see also eq. A-14 in Appendix here). Limited dispersal favours reduced survival (and hence greater reproductive effort) because surviving adults

occupy breeding sites that would otherwise become available to their philopatric offspring. What remains to be determined is whether selection is negatively trait-dependent at this singularity, thereby opening the door to the emergence of polymorphism.

Substituting eq. (17) into eq. (15) and using eq. (12), we obtain

$$K(y^*) = -\frac{(1-m)m^2}{(2-m)^2[1-z_1(y^*)]} \leq 0. \tag{18}$$

This shows that when dispersal is complete ($m = 1$), $K(y^*) = 0$ and there is no NTDS, whereas under limited dispersal ($m < 1$), $K(y^*) < 0$. To understand why, it is useful to consider that the mutant allocates more to reproduction (and less to survival), while the resident allocates more to survival (and less to reproduction). The mutant's change increases the number of juveniles it produces, whereas the resident's change reduces the number of breeding sites available. When dispersal is complete, these two effects compensate exactly: the increase in fecundity in the mutant is offset by the smaller number of breeding spots left open by the resident. When dispersal is limited, however, this symmetry breaks down because decreased mutant survival increases the number of breeding spots available to mutants more than it does for residents. As a result, mutant invasion fitness increases when its trait diverges from that of the resident. This NTDS should enable the coexistence of different life-history strategies, given an appropriate trade-off between survival and reproduction.

We investigated this specifically with a trade off induced by $z_1(x) = x^\alpha$ and $z_2(x) = (1-x)^\beta$ with parameters $\alpha < 1$ and $\beta < 1$. Our analysis in Appendix D.2 shows that, for a range of parameter values, polymorphism in resource allocation can then be favoured by gradual evolution as long as $m < 1$. Interestingly, the end-point of evolution consists of one morph investing entirely in survival (becoming effectively immortal), while the other morphs can consist of any mixed allocation to survival and fecundity. We conducted individual-based simulations that confirm these findings. As illustrated in Fig. 2, the gradual emergence of a quasi immortal morph (mutations preclude converging to exactly $x = 1$ for the immortal morph), coexisting with a second type that invests in both survival and reproduction. The population thus eventually shows a marked dimorphism in life-history strategies (see Appendix D.2 for more details), a dimorphism impossible under well-mixed conditions (Charnov, 1997; Case, 2000).

4.4 Temporal heterogeneity and limited dispersal

The second scenario we examine is temporal heterogeneity, where production factors vary between generations, which is a classical source of diversity in both evolutionary ecology and population genetics (e.g. Chesson and Warner, 1981; Gillespie, 1991; Turelli et al., 2001). Gradual evolution under temporal heterogeneity and limited dispersal has previously been analysed in models assuming infinitely large local populations, effectively excluding kin competition (Svardal et al., 2015). Here we consider the complementary case of populations with limited dispersal and one individual per site, where kin competition arises under hard

selection.

Consider again a population subject to limited dispersal, and that experiences two environmental states, say a wet and a dry environment, occurring independently in each generation with probabilities q and $1-q$, respectively. At each demographic time step, each adult individual occupies a site where it produces a large number of juveniles, with fecundity $z_i(x)$ depending on both its trait x and the environmental state $i \in \{1,2\}$. Adults then either survive with probability s or die, freeing up breeding spots. Each juvenile either disperses with probability m or remains philopatric. Density-dependent competition among juveniles occurs locally for the breeding spots vacated by adult mortality. Invasion fitness for this scenario then takes the form of a geometric mean across environmental states, i.e.

$$W(z_1(x), z_2(x), e_1(y), e_2(y)) = \left(s + (1-s)\left[\frac{(1-m)z_1(x)}{(1-m)z_1(x) + me_1(y)} + \frac{mz_1(x)}{e_1(y)}\right]\right)^q \times \left(s + (1-s)\left[\frac{(1-m)z_2(x)}{(1-m)z_2(x) + me_2(y)} + \frac{mz_2(x)}{e_2(y)}\right]\right)^{1-q}, \quad (19)$$

where $e_1(y) = z_1(y)$ and $e_2(y) = z_2(y)$.

Substituting eq. (19) into eq. (15) and eq. (12) yields

$$K(y^*) = -\left[s(2-m)^2 - (1-m)^2\right] \frac{q\, m^2(1-s)}{(1-q)z_1(y^*)^2}. \quad (20)$$

Whether $K(y^*) < 0$ depends on survival s and dispersal m. Consider first the case of complete dispersal ($m = 1$). When generations do not overlap ($s = 0$), $K(y^*) = 0$ and polymorphism is impossible, whereas when generations overlap ($s > 0$), $K(y^*) < 0$ and polymorphism becomes possible. To understand this, consider a mutant and a resident whose traits change in opposite directions from y^*, so that the mutant performs better in one environment and the resident in the other. When $s = 0$, the fitness loss experienced by the mutant in one environment is exactly offset by the gain in the other. When generations overlap ($s > 0$), this balance changes because surviving individuals make the environment they perform best in more crowded when it recurs. As a result, the mutant experiences weaker competition in the environment where it performs better, thus enabling NTDS.

Eq. (20) shows that this effect weakens as dispersal decreases, with $s > (1-m)^2/(2-m)^2$ required for NTDS to occur. Survival and limited dispersal thus have opposing effects on NTDS. Overlapping generations promote NTDS because surviving individuals make their own environment more crowded when it recurs, which favours a mutant performing better in the other environment than the resident. Limited dispersal, by contrast, means offspring remain near their parents and mostly compete with others sharing the same trait. This removes the advantage of a mutant performing well in a different environment from the resident and erodes the asymmetry that drives negative trait-dependent selection.

Using $z_i(x) = \exp(-(x-o_i)^2/(2\sigma^2))$, it can readily be shown that with fitness eq. (19) there exists a range

of parameter values leading to evolutionary branching, and that this range becomes smaller as dispersal becomes limited (see Appendix D.3). This differs from the results of Svardal et al. (2015), where dispersal affected diversification only under spatial heterogeneity (i.e. their eq. 15), and had no influence under purely temporal heterogeneity with effectively infinite local population size (and hence no kin competition). In our model, by contrast, limited dispersal generates kin competition, which weakens NTDS and inhibits branching, consistent with previous results showing that kin competition tends to oppose diversification under constant environments (Day, 2001; Ajar, 2003). The role of kin selection in polymorphism is explored further in section 6.2.

5 Socially mediated interactions

We now turn to another important class of trait-dependent interactions: those mediated by direct social interactions influencing payoffs. Classic examples include contests for mates or territories, public-good investments, cooperative breeding, and various forms of aggressive or cooperative encounters (e.g. Maynard Smith, 1982; Vincent and Brown, 2005; McNamara and Leimar, 2020). Here, an individual's payoff explicitly depends on the traits expressed by others in the population, creating a social feedback loop that may lead to stable polymorphisms.

5.1 Selection and interaction coefficients at the payoff level

To capture socially mediated interactions in a simple way covering multiple scenarios, we assume that the payoff to an individual is one-dimensional, given by $\boldsymbol{\pi}(\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x, y)), \mathbf{e}^{\pi}(x, y)) = \pi(x, y)$, and depends explicitly on both mutant and resident traits. Additionally, we assume that the regulating variable $\mathbf{e}^{\mathrm{w}}(x, y) = e(y)$ is also one dimensional and depends solely on the resident population. Under these assumptions, invasion fitness (10) takes the form

$$\rho(x, y) = W(\pi(x, y), e(y)), \tag{21}$$

for some mapping W. We assume that invasion fitness is monotonically increasing with payoff (i.e. W increases with its first argument). Eq. (21) represents the standard fitness formulation that is implicitly used in evolutionary game theory, encompassing both pairwise interactions and cases where individuals interact collectively ("playing against the field"; e.g. Maynard Smith, 1982; Vincent and Brown, 2005; McNamara and Leimar, 2020). This formulation captures the essence of socially mediated interactions, where fitness depends explicitly on payoffs (determined by the traits of interacting individuals) while environmental factors influence selection only indirectly through their effect on population-wide competition or regulation.

Substituting eq. (21) into eq. (6), we show in Appendix E.1 that the invasion analysis can be fully worked

out from the payoff function. In particular, any singular strategy y^* must be such that

$$\left.\frac{\partial \pi(x,y)}{\partial x}\right|_{x=y=y^*} = 0, \tag{22}$$

and the necessary condition for protected polymorphism (eq. 9) reduces to

$$I(y^*) \propto I_\pi(y^*) = \left.\frac{\partial^2 \pi(x,y)}{\partial x \partial y}\right|_{x=y=y^*} < 0. \tag{23}$$

When condition (23) holds, this indicates that the mutant's payoff increases when the mutant and the resident population change trait values in opposite ways. Since fitness is monotonically increasing with payoff, these payoff effects directly generate NTDS. We explore further how such selection can emerge owing to social interactions by making more specific assumptions concerning payoff in the next section.

5.2 Action based feedback

To investigate how a social feedback can promote polymorphism, we consider a class of models in which an individual's payoff depends explicitly on actions taken during interactions. Specifically, we assume that the payoff can be expressed as:

$$\pi(x,y) = \Pi(a(x,y), a(y,x)), \tag{24}$$

for a mapping Π, where $a(x,y)$ and $a(y,x)$, are the actions of a mutant and a resident individual, respectively (formally actions here are equivalent to phenotypes z but for ease of interpretation in the social context, we'll refer to actions and use the variable a in this section). These actions may depend not only on an individual's own trait but also on the trait of its interaction partner.

This formulation covers two broad classes of interactions. The first is pairwise interactions, where $a(x,y)$ and $a(y,x)$ describe the actions of two individuals in a direct encounter. The dependence of an individual's action on its own trait and that of its partner allows us to model cases where actions arise as equilibrium outcomes of a dynamic behavioural process, such as learning, information exchange, or negotiation (as used in models of evolving response rules, motivations, and preference evolution, e.g. McNamara et al., 1999; André and Day, 2007; Heifetz et al., 2007a; Akçay and Van Cleve, 2012; Alger and Lehmann, 2023). This also captures simpler pairwise interactions where one's action depends only on own trait, $a(x,y) = a(x)$, as is very commonly studied in evolutionary game theory (Maynard Smith, 1982; McNamara and Leimar, 2020). In any case, the resident's trait y in $a(x,y)$ corresponds to the regulating factor $\mathbf{e}^{\mathrm{z}}(x,y) = y$ acting at the level of the phenotype (i.e. the action), while the resident's action $a(y,x)$ correspond to the regulating factor $\mathbf{e}^{\pi}(x,y) = a(y,x)$ at the level of payoff (recall eq. 10).

The second class of interactions covered by eq. (24) corresponds to "playing against the field" multiplayer interactions (Maynard Smith, 1982). In this case, actions depend only on the focal individual's own trait, such

that $a(x,y) = a(x)$ and $a(y,x) = a(y)$. This extends beyond pairwise interactions, as multiple individuals expressing action $a(y)$ can simultaneously influence the focal individual's payoff. Here, there is no regulating factor at phenotype level, and the regulating factor at the payoff level simplifies to $\mathbf{e}^{\pi}(x,y) = a(y)$.

Substituting eq. (24) into eq. (22) reveals that any singularity y^* must be such that

$$\left.\frac{\partial \Pi(a,b)}{\partial a}\right|_{\substack{x=y=y^* \\ a=b=a(y^*,y^*)}} + \lambda(y^*) \left.\frac{\partial \Pi(a,b)}{\partial b}\right|_{\substack{x=y=y^* \\ a=b=a(y^*,y^*)}} = 0, \tag{25}$$

where

$$\lambda(y^*) = \frac{\partial a(y,x)}{\partial x} \Big/ \left.\frac{\partial a(x,y)}{\partial x}\right|_{x=y=y^*} \tag{26}$$

is a response coefficient, which measures how much the partner to a focal individual changes her action in response to a change in the action of the focal (a change that is due to a trait change in the focal). When $\lambda(y^*) > 0$, an increase in one's action due to a trait change leads to a corresponding increase in the partner's action. Conversely, when $\lambda(y^*) < 0$, the response is negative, meaning that a trait-driven increase in one's action is associated with a decrease in the partner's action. Eq. (25) expresses that at equilibrium, selection balances two opposing effects on payoff: a direct effect due to changes in the focal individual's own action (first term) and an indirect effect due to the response of the partner's action (second term), weighted by the response coefficient $\lambda(y^*)$.

The balance captured by eq. (25) helps to understand the evolution of many well-known social interactions. In particular, when payoff $\Pi(a,b)$ is decreasing in its first argument and increasing in an accelerating manner in its second argument, individual actions can be regarded as levels of cooperation. If actions increase with trait values, and if $a(x,y) = a(x)$ and $a(y,x) = a(y)$ so that individuals cannot respond to their partner's action, then $\lambda(y^*) = 0$. In this case, no singularity exists, and eq. (25) correctly predicts the loss of unconditional cooperation in a well-mixed population. By contrast, when cooperation is conditional, so that actions depend on both players traits enabling individuals who express cooperative behaviour to receive cooperative responses in return, then conditional cooperation via reciprocity can evolve when $\lambda(y^*) > 0$. This occurs for instance where traits determine response slopes such that equilibrium actions are given by $a(x,y) = \alpha + xa(y,x)$ and $a(y,x) = \alpha + ya(x,y)$ for some $\alpha > 0$, e.g. McNamara et al., 1999; Taylor and Day, 2004). More broadly, the coefficient $\lambda(y^*)$ plays a central role across many scenarios of social evolution (e.g. Akçay and Van Cleve, 2012; Akçay and Cleve, 2014; Lehmann, 2022; Alger and Lehmann, 2023).

Using eq. (25) and eq. (26) on substituting eq. (24) into eq. (23), we find that $I_{\pi}(y) = L_{\pi}(y^*)[K_{\pi}(y^*)+K_{\lambda}(y^*)]$, where $L_{\pi}(y^*) = \left.(\partial a(x,y)/\partial x)^2\right|_{x=y=y^*} \geq 0$, such that NTDS selection occurs when

$$K_{\pi}(y^*) + K_{\lambda}(y^*) < 0, \tag{27}$$

in which

$$K_\pi(y^*) = \left[\left(1+\lambda(y^*)^2\right)\frac{\partial^2\Pi(a,b)}{\partial a\partial b} + \lambda(y^*)\left(\frac{\partial^2\Pi(a,b)}{\partial a^2} + \frac{\partial^2\Pi(a,b)}{\partial b^2}\right)\right]_{\substack{a=a(y^*,y^*)\\ b=a(y^*,y^*)}} \tag{28a}$$

and

$$K_\lambda(y^*) = \underbrace{\frac{\partial^2 a(y,x)}{\partial x\partial y}\Big/\left(\frac{\partial a(x,y)}{\partial x}\right)^2\Bigg|_{x=y=y^*}}_{=A(y^*)} \times \left(\frac{\partial\Pi(a,b)}{\partial a} + \frac{\partial\Pi(a,b)}{\partial b}\right)_{\substack{a=a(y^*,y^*)\\ b=a(y^*,y^*)}}. \tag{28b}$$

The first term, $K_\pi(y^*)$, describes trait-dependent selection via direct payoff effects, which considers how selection on the mutant changes as the population trait is varied, while keeping the response coefficient $\lambda(y^*)$ unchanged. This is relevant in cases where the effects of actions on payoff are antagonistic, meaning that the actions of interacting individuals interfere in a way that reduces payoff, i.e. $\partial^2\Pi(a,b)/(\partial a\partial b) < 0$. When this holds, a mutant obtains greater payoffs when the actions of interacting individuals change oppositely to the mutant.

The second term, $K_\lambda(y^*)$, describes trait-dependent selection via response effects, i.e. through changes in $\lambda(y)$. This is governed by $A(y^*)$, which measures how one's action changes in response to a similar trait change in others. When $A(y^*) < 0$ changing traits similarly reduces the response coefficient, whereas when $A(y^*) > 0$ it increases it. For example, if $\partial\Pi/\partial a + \partial\Pi/\partial b > 0$ as in cooperative interactions and $A(y^*) < 0$ (for instance, due to interference in cooperative investments) an increase in population trait leads to weaker cooperation overall, favouring NTDS. We next illustrate how eq. (27) helps understand two more specific examples of social interactions generating NTDS.

5.3 Examples

5.3.1 Contests over resources

Consider a scenario where payoffs are determined by contest competition among N individuals for access to a resource of value v. Suppose the payoff function is

$$\Pi(a(x), b(y)) = \frac{a(x)}{a(x) + (N-1)a(y)}v - C(a(x)), \tag{29}$$

where the action $a(x)$ measures the fighting or competitive effort of an individual with trait x and $C(a(x))$ is the cost associated with such effort. The probability of winning the contest depends on an individual's relative effort, given by the ratio $a(x)/[a(x)+(N-1)a(y)]$, which is standard of conflict and contest competition (e.g. Hirshleifer, 1987; Skaperdas, 1996; Cant, 2012). This is a scenario of "playing against the field", where actions depend only on one's own trait and not on the trait of interacting partners so that the response coefficient (eq. 26) is zero $\lambda(y) = 0$, which implies $A(y^*) = 0$ and as a result, $K_\lambda(y^*) = 0$ here.

Whether trait-dependent selection is negative therefore depends solely on whether actions are strategic substitutes at the payoff level (from eq. 28), which using eq. (29) yields

$$K_\pi(y^*) = -\frac{(N-1)(N-2)va'(y^*)^2}{N^3 a(y^*)^2} \leq 0, \tag{30}$$

where the prime denotes the derivative. This equation shows that as long as $N > 2$ (meaning that at least three individuals compete for a resource), NTDS occurs. When $N = 2$, payoffs depend only on the direct outcome of the contest between two individuals, so the marginal effects of mutant and resident trait changes combine additively and NTDS cannot occur. This contrasts with the classic Hawk-Dove game, where pairwise interactions (i.e. $N = 2$) can generate NTDS. In the Hawk-Dove game, payoffs decrease for aggressive individuals when interacting with other aggressive individuals, meaning that NTDS arises from interaction effects on costs (this would correspond here to writing the cost function as $C(x, y)$ with $\partial^2 C(x, y)/(\partial x \partial y) > 0$, where the cost of aggression increases as the opponent's aggression increases). Without such interaction effects on costs, competition must involve multiple individuals for NTDS to emerge ($N > 2$). Nevertheless, the ease with which NTDS arises in our model exemplifies how conflicts in general provide a natural setting for polymorphism, since if one individual increases its competitive effort, this is likely be be mitigated if others in the population change similarly. In fact, evolutionary diversification has been found to occur through gradual evolution in multiple conflict models (McNamara and Leimar, 2020).

To see polymorphism emerge under fitness model (29), one can use $a(x) = e^{\omega x}$ and $C(a(x)) = x^2$ with $\omega > 0$ as a parameter. For $N > 2$, there exists a range of parameter values leading to evolutionary branching with these functions. Over the long term, gradual evolution results in a convergence-stable and uninvadable dimorphism, with one morph investing heavily in competition while the other adopts a more moderate investment strategy. Further details on the analysis are provided in Appendix E.2.1, and Fig. 2 shows the emergence of this previously unreported polymorphism in individual-based simulations. This model with fixed number of contestants can readily be extended to a more realistic setting with random number of contestants N (independent of traits). In this case, NTDS requires that the expectation over the (random) number of contests yields $\mathrm{E}[(N-1)(N-2)/N^3] > 0$.

5.3.2 Evolution of preferences

The previous example had no behavioural response ($\lambda(y^*) = 0$); we now consider a case where individuals adjust their behaviour in response to their partner's traits ($\lambda(y^*) \neq 0$). We assume individuals interact in a pairwise manner and aim to maximise a utility function that evolves under a situation of complete information (see Alger, 2023 for discussion of information settings). This follows models of preference or motivation evolution, where selection acts not directly on the actions individuals take but on the underlying behavioral mechanism guiding their decision-making (e.g. Heifetz et al., 2007b; Akçay and Van Cleve, 2012; Alger and Lehmann, 2023; see Alger, 2023 for a review).

Building on Heifetz et al. (2007a,b), we assume that each individual's utility consists of their own material payoff and a self-esteem bias, which distorts their perception of their own actions. Specifically, the utility functions for a mutant with action a and trait x, and a resident with action b and trait y, are given by

$$\begin{aligned} u_x(a,b) &= \Pi(a,b) + xa \\ u_y(b,a) &= \Pi(b,a) + yb, \end{aligned} \tag{31}$$

respectively. Here, the self-esteem bias is modulated by the evolving trait (x in mutants, y in residents) and determines whether individuals overestimate or underestimate the benefits of their own actions. When $x = 0$, a mutant individual simply seeks to maximise its material payoff. But when $x > 0$, the mutant overestimates the benefits of its own action, acting as if its actions are more rewarding than they actually are. Conversely, when $x < 0$, the mutant underestimates its returns and behaves more cautiously, even when a greater action value would be materially beneficial. In this way, the evolving trait x governs an individual's "disposition" (Heifetz et al., 2007a,b), shaping whether they are overconfident optimists ($x > 0$) or underconfident pessimists ($x < 0$) in how they perceive and respond to interactions.

To complete the model, we assume that material payoffs follow a linear-quadratic structure, given by

$$\Pi(a,b) = (\alpha - \gamma a - \beta b)a, \tag{32}$$

for parameters $\alpha > 0$, $\gamma > 0$, and $\beta \geq \gamma$. Eq. (32) is consistent with a model where actions constitute the effort or energy allocated to the extraction of a biotic resource that pairs of individuals exploit (e.g. pairs of individuals exploit different resource patches), and where the resource has logistic growth and a linear functional response to extraction. In this case, parameters of eq. (32) can be interpreted as follows. The parameter α is proportional to the replenishment rate of the resource and γa to the rate of resource extraction by an individual with action a. Assuming fast resource dynamics, the equilibrium amount of resource is then proportional to $(\alpha - \gamma a - \gamma b)$, and so the amount of resources individual with action a extracts per unit time is $(\alpha - \gamma a - \gamma b)a$. Finally, we add to this $ab(\gamma - \beta)$, which can be seen as a cost stemming from interference competition: when $\beta > \gamma$, joint efforts lead to diminishing returns, similarly to interference competition where competition over shared resources reduces overall resource uptake rate (e.g., Begon and Townsend, 2021, chapter 5).

By maximising their utility given the action of their partner, each individual plays a Nash equilibrium strategy. The necessary first-order conditions for a Nash equilibrium require that neither individual can improve their utility by unilaterally changing their action, meaning $\partial u_x(a,b)/\partial a = 0$ and $\partial u_y(b,a)/\partial b = 0$. Using eqs. (31)

and (32), these conditions can be expressed as the system:

$$a = \frac{\alpha + x}{2\gamma} - b\frac{\beta}{2\gamma} \qquad (33)$$
$$b = \frac{\alpha + y}{2\gamma} - a\frac{\beta}{2\gamma}.$$

Solving these for $a = a(x, y)$ and $b = a(y, x)$ yields

$$a(x, y) = \frac{2\gamma(\alpha + x) - \beta(\alpha + y)}{4\gamma^2 - \beta^2}. \qquad (34)$$

This solution represents a Nash equilibrium action because it is unique and satisfies the second-order condition for stability: $\partial^2 u_x(a, b)/\partial a^2 = \partial^2 \Pi(a, b)/\partial a^2 = -2\gamma \leq 0$ and $\partial^2 u_y(b, a)/\partial b^2 = \partial^2 \Pi(b, a)/\partial b^2 = -2\gamma \leq 0$. Eq. (34) shows that an individual's action can increase with their own optimism (x) but decreases with that of their partner (y), provided $4\gamma^2 > \beta^2$. This is the case when $\gamma > 0$ where the joint increase in action then reduce individual resource extraction. To avoid these deleterious effects, it is best to reduce one's own action when faced with more optimistic individuals (who increase their action as a result). This relationship is captured by the negative response coefficient $\lambda(y^*) = -\beta/(2\gamma)$ (found by plugging eq. 34 in eq. 26), indicating that an increase in the partner's action leads to a decrease in one's own.

A key feature of eq. (33) is that an individual's action is a linear response to their partner's. This is formally identical to models for the evolution of response rules, where individuals adjust their behaviour based on social interactions (e.g., McNamara et al., 1999; Taylor and Day, 2004; see Alger, 2023 for a discussion on the connection between preference evolution and response rule evolution models). This linearity means that, here and in these aforementioned models, $A(y) = 0$ (eq. 28b). As a result, there can be no trait-dependent selection via a feedback on behaviour rules (i.e. $K_\lambda(y^*) = 0$).

Therefore, NTDS arises when $K_\pi(y^*) < 0$ in this model. Substituting eqs. (32) and (34) into eq. (28a), we obtain

$$K_\pi(y^*) = -\frac{\beta^3}{4\gamma^2}, \qquad (35)$$

which is negative because when both a focal individual and population member increase their optimism bias, payoff is reduced in the presence of interference competition (where $\beta > \gamma > 0$ according to our assumptions). It is then straightforward to show that evolutionary branching in this self-esteem trait can happen as long as $\beta > \sqrt{2}\gamma$. When it occurs, evolutionary branching leads to ever more "optimists" (with increasingly positive trait values) and ever more "pessimists" (with increasingly negative trait values) as both strategy fuel one another. This open ended evolution of exaggerated self-perception comes to a halt where there is a cost to overconfidence (Appendix E.2.2 for analysis and Fig. 2 for individual-based simulations).

The original study by Heifetz et al. (2007a,b) did not analyse the scope for polymorphism in self-esteem biases, and in fact considered parameter values generating evolutionarily stable singularities only. Heifetz

et al. (2007a,b) also considered an alternative model of preference, in which the evolving trait x measures an individual's other-regard. In this formulation, mutant and resident utilities take the form $u_x(a,b) = \Pi(a,b) + x\Pi(b,a)$ and $u_y(b,a) = \Pi(b,a) + y\Pi(a,b)$. We analyse this model in Appendix E.2.3 using eq. (32) for payoff. We show that for this model NTDS arises through payoff effects ($K_\pi(y^*) < 0$) as well as via action response effects ($K_\lambda(y^*) < 0$). Thus, polymorphism can emerge in this model too.

6 Genetically mediated interactions

Up until now, we have assumed that individuals are haploid and that the population is well mixed. We relax here both of these assumptions in turn, and investigate their implications for adaptive polymorphism. We will see that these implications depend on the nature of interactions between genes, respectively within and between individuals (hence "genetically mediated interactions").

6.1 Diploidy and heterozygote advantage

We consider now that individuals are diploid and that an individual's phenotype depends on the two homologous alleles it carries. To model this, we denote as $z_i(x,y)$ the phenotype i of an individual carrying a mutant allele with heritable trait value x and and a resident allele with heritable trait value y. As such an individual homozygote for the resident allelic value y expresses phenotype $z_i(y,y)$. We assume that mating is random so that the mutant allelic form x is only ever present in heterozygotic state. We also assume that the genotype-to-phenotype map is: (i) homogeneous (by which we mean $z_i(x,y) = z_i(y,x)$) and (ii) such that variation in allelic value always leads to variation in phenotype, i.e. that phenotypes always show some degree of heritability (which entails $\partial z_i(x,y)/\partial x|_{x=y} = \partial z_i(x,y)/\partial y|_{x=y} \neq 0$, where the first equality unfolds from assumption (i)). Where the phenotype of heterozygotes is a convex combination of the phenotypes of homozygotes, assumptions (i) and (ii) entail additive gene action (i.e. entail that $z_i(x,y) = z_i(x,x)/2 + z_i(y,y)/2$); our forthcoming results apply more broadly (though the assumption that $\partial z_i(x,y)/\partial x|_{x=y} = \partial z_i(x,y)/\partial y|_{x=y}$ amounts to local additivity).

6.1.1 No effect of diploidy on the conditions for the emergence of polymorphism of a single phenotype

Let us first consider the case where invasion fitness depends on a single phenotype such that we can write

$$\rho(x,y) = W(z(x,y), z(y,y)), \tag{36}$$

where the function W gives the invasion fitness of a x/y heterozygote in a population of y homozygotes. Substituting eq. (36) into eqs. (6)-(7) and using our assumptions concerning $z(x, y)$, we show in Appendix F.1 that the population converges to a singular allelic value y^*, which generates phenotype $z^* = z(y^*, y^*)$, and subsequently undergoes evolutionary branching when the following holds:

$$\begin{aligned} &\left.\frac{\partial W(z,\bar{z})}{\partial z}\right|_{z=\bar{z}=z^*} = 0 \\ &\left.\frac{\partial^2 W(z,\bar{z})}{\partial z^2}\right|_{z=\bar{z}=z^*} + \left.\frac{\partial^2 W(z,\bar{z})}{\partial z \partial \bar{z}}\right|_{z=\bar{z}=z^*} < 0 \\ &\left.\frac{\partial^2 W(z,\bar{z})}{\partial z^2}\right|_{z=\bar{z}=z^*} > 0. \end{aligned} \tag{37}$$

In these expressions, the derivatives are taken with respect to phenotype (where $\bar{z} = z(y, y)$ denotes resident phenotype) and not genetic values. These conditions are equivalent to those derived under haploidy (mirroring eq. 7 and conditions discussed thereafter). This shows that diploidy does not affect the conditions for polymorphism, at least when underlying a single phenotype. In particular, the emergence of polymorphism requires that the cross derivative of the fitness function with respect to mutant and resident phenotypes is negative:

$$\left.\frac{\partial^2 W(z,\bar{z})}{\partial z \partial \bar{z}}\right|_{z=\bar{z}=z^*} < 0, \tag{38}$$

indicating that directional selection depends negatively on the phenotype expressed in the resident population (as in eq. 9 at the level of the trait value). Consequently, all our results in sections (4)–(5), which are based on a invasion fitness function of the form $\rho(x, y)$, still apply unchanged to diploids, with x now standing for the phenotypic value of a mutant individual (z) and y for the phenotypic value in the resident population ($\bar{z}$).

Although the conditions favoring the emergence of polymorphism are left unchanged by diploidy under the above assumptions, diploidy does affect the condition for mutual invasibility (in fact, unlike the haploid case, it is possible to have mutual invasibility without NTDS owing to genetic constraints under diploidy, Appendix eq. A-44 for details). Additionally, evolution after the onset of evolutionary branching will likely differ between haploid and diploid populations owing to heterozygotes with intermediate phenotypes in the latter. In a haploid population, evolutionary divergence leads to the coexistence of two morphs. In sexual diploids, by contrast, mating between homozygotes for these diverged alleles creates a third, intermediate heterozygote. Evolution in this case leads one of the three resulting types – either one of the two homozygotes or the heterozygote – to be maladapted (which type evolves to be maladapted depends on model parameters; Kisdi and Geritz, 1999). As a result of the genetic constraint of additive effects, a proportion of the phenotypic polymorphism in the population is therefore non-adaptive. This in turn creates selection on the genetic architecture of the trait to reduce such genetic load, e.g. through the evolution of dominance (Van Dooren, 1999). More generally, genetic constraints that lead to a genetic load favour the evolution modifiers of recombination, epistasis, or gene duplication, that should eventually lead to a phenotypic polymorphism

mirroring that seen in haploid asexuals, where such genetic constraints are absent.

6.1.2 Heterozygote advantage when there are multiple phenotypes

Next, we consider invasion fitness when it depends on multiple phenotypes that are influenced by the evolving locus. Specifically, we assume that the biological scenario is such that invasion fitness can be written as

$$\rho(x,y) = W(\pi(x,y), e(y)), \tag{39}$$

where

$$\pi(x,y) = \Pi(z_1(x,y), z_2(x,y)) \tag{40}$$

is the payoff to an heterozygote showing phenotypes $z_1(x,y)$ and $z_2(x,y)$, and $e(y)$ is a regulating variable determined by the resident population (where individuals are homozygotes for the resident allele y). This form of invasion fitness, which is functionally equivalent to eq. (21), could for instance capture a situation where the fecundity of an individual depends on two of its phenotypes $z_1(x,y)$ and $z_2(x,y)$ (e.g. expression levels of two genes). Whatever the specific scenario, $z_1(x,y)$ and $z_2(x,y)$ are genetically correlated because they both depend on allelic values x and y. We assume that this dependency is given by

$$z_i(x,y) = \frac{g_i(x) + g_i(y)}{2}, \tag{41}$$

where $g_i(h)$ are the effects of an allele with genetic value $h \in \{x,y\}$ on phenotype i. In other words, we assume additive genetic effects on the phenotype between homologous loci (which will allow us to connect to a number of existing results). These effects however may translate into non-additive effects on fitness.

Substituting eqs. (40)–(41) into eq. (22) shows that any singularity y^* must satisfy

$$\frac{\partial \Pi(z_1,z_2)}{\partial z_1} \frac{\mathrm{d}g_1(x)}{\mathrm{d}x}\bigg|_{\substack{x=y^* \\ z_1=z_1(y^*,y^*) \\ z_2=z_2(y^*,y^*)}} + \frac{\partial \Pi(z_1,z_2)}{\partial z_2} \frac{\mathrm{d}g_2(x)}{\mathrm{d}x}\bigg|_{\substack{x=y^* \\ z_1=z_1(y^*,y^*) \\ z_2=z_2(y^*,y^*)}} = 0. \tag{42}$$

This says that at y^*, the effects on fitness of varying allelic value via phenotypes 1 and 2 must cancel each other out. Consider for example the case where fitness increases with both phenotypes (i.e. $\partial\Pi/\partial z_1 > 0$ and $\partial\Pi/\partial z_2 > 0$). Eq. (42) then tells us that for a singular trait value y^* to exist, phenotypes 1 and 2 need to change in opposite directions with x, namely $g_1'(y^*)$ and $g_2'(y^*)$ have opposite signs, such that a change in x has antagonistic effects on fitness via phenotypes 1 and 2 : antagonistic effects which are balanced at y^*.

If we substitute eqs. (40)–(41) into eq. (23) and use eq. (42), we readily obtain that the interaction coefficient can be expressed as $I_\pi(y) = L_\pi(y^*)K_\pi(y^*)/2$ where $L_\pi(y^*) = g_2'(y^*)^2 \geq 0$, such that the sign of trait-

dependent selection is governed by

$$K_\pi(y^*) = \frac{1}{4}\left[M(y^*)^2\frac{\partial^2\Pi(z_1,z_2)}{\partial z_1^2} + \frac{\partial^2\Pi(z_1,z_2)}{\partial z_2^2} - 2M(y^*)\frac{\partial^2\Pi(z_1,z_2)}{\partial z_1\partial z_2}\right]_{\substack{y=y^* \\ z_1=z_1(y,y) \\ z_2=z_2(y,y)}}. \tag{43}$$

Here, $M(y^*) = (\partial\Pi/\partial z_2)/(\partial\Pi/\partial z_1)|_{y=y^*}$ gives the rate at which phenotype 1 can be substituted for phenotype 2 without changing payoff (i.e. the change in the number of units of phenotype 1 that correspond to the effect on payoff of a change in one unit of phenotype 2). Eq. (43) shows that NTDS tends to occur when the payoff function decelerates with phenotypes ($\partial^2\Pi/\partial z_i^2 < 0$ for $i \in \{1,2\}$) and phenotypes have complementary effects on payoff ($\partial^2\Pi/\partial z_1\partial z_2 > 0$). This is because such situations lead to heterozygote advantage, i.e. where heterozygotes for divergent allelic effects close to the singular allelic value y^* have greater fitness that either homozygotes. This occurs because heterozygotes benefit from expressing an intermediate value for both phenotypes. In particular, when alleles contribute additively to two phenotypes (eq. 41), and payoff decelerates with both phenotypes in a concave way, heterozygotes can have greater fitness than either homozygotes (see Fig. 3 panel A for a graphical argument). When the payoff exhibit complementary effects, this also favours a situation such that heterozygotes have a greater payoff (see Fig. 3 panel B for a graphical argument). Heterozygote advantage thus arises because the contribution of an allele to fitness is context-dependent here, in the sense that its fitness effect varies across different phenotypes.

This mechanism is in fact similar to the one described in the seminal SAS-CFF population genetic model ("Stochastic Additive Scale - Concave Fitness Function", Gillespie, 1978, 1991). That model assumes that alleles at an homologous locus contribute additively to physiological activity (which is a phenotype in our model, eq. 41) in a fashion that varies stochastically with environmental conditions (say phenotype 1 in environment 1 and phenotype 2 in environment 2); and that fitness increases in a decelerating – concave – manner with physiological activity. To see the connection with our approach here, consider that eq. (43) applies to this SAS-CFF scenario with Π taken as the geometric mean fecundity of the mutant and the regulating variable e as the geometric mean fecundity of a resident individual under a Wright-Fisher process (as assumed in Gillespie, 1978, 1991 see Appendix F.2 for details on how eq. 43 applies to the SAS-CFF scenario and how evolutionary branching can occur in this scenario).

While eq. (43) applies to a broader set of situations, its connection to the SAS-CFF model provides a couple of insights. Firstly, it reinforces the idea that heterozygote advantage should arise from context-dependent allelic effect that are summed over different contexts. In the original SAS-CFF model (Gillespie, 1978), these contexts are temporal (over changing environments), but the same principle applies across multiple phenotype or biological functions. Secondly, since the SAS-CFF model is closely linked to the storage effect of ecological coexistence (Schreiber, 2020), this connection extends to our framework and can be phrased as follows. When conditions vary, diploidy allows alleles that are poorly adapted to current conditions to persist by buffering them in heterozygotes. This buffering effect can prevent the loss of alleles that may be beneficial under different environmental or functional conditions. In turn, the adaptive polymorphism that

evolves under the above model leads to the divergence of two morphs, each showing a better fit for these different environmental or functional conditions, showing heterozygote advantage in total.

6.2 Negative trait-dependent kin selection

We now consider the case where owing to limited dispersal, individuals interacting with one another show genetic associations generated by identity-by-descent, i.e., individuals are related (Hamilton, 1964; Michod, 1982 and see Rousset, 2004 for a review). This generalises the kin competition effects seen in sections 4.3-4.4. When relatedness is high, this means that a globally rare mutant is nevertheless likely to interact with other mutants locally. Relatedness should therefore modulate trait-dependent selection depending on the fitness effects of mutant-mutant interactions. To examine this, we analyse $I(y^*)$ in the infinite island model of dispersal with two individuals per island (or patch or deme). Although the conditions for evolutionary branching in this model have been characterised previously (e.g. Ajar, 2003), $I(y^*)$ has not yet been analysed.

For simplicity and to focus on genetic interactions among individuals, we assume individuals are haploid and reproduce clonally. Invasion fitness in this case can be written as

$$\rho(x,y) = r(x,y)w(x,x,y) + [1 - r(x,y)]w(x,y,y), \tag{44}$$

where $r(x,y)$ is relatedness, that is, the probability that a rare mutant x in a resident y population has another mutant as patch neighbor (conversely, it has a resident as patch neighbor with probability $1-r(x,y)$); $w(x,x,y)$ is the individual fitness of a focal mutant x (first argument of w) when its patch neighbor is a mutant x (second argument), and the rest of the population in other patches consists of residents y (third argument); similarly, $w(x,y,y)$ is the fitness of the mutant when its patch neighbor is a resident y (second argument).

Plugging eq. (44) into eq. (7), we first recover the classical result that the singular trait value x^* must be such that the inclusive fitness effect, i.e. Hamilton's rule in marginal form, vanishes:

$$\underbrace{\left.\frac{\partial w(x,y,y)}{\partial x}\right|_{x=y=y^*}}_{-c} + r(y^*,y^*)\underbrace{\left.\frac{\partial w(x,y,y^*)}{\partial y}\right|_{x=y=y^*}}_{b} = 0 \tag{45}$$

(e.g., Rousset, 2004). Therefore at x^*, the direct fitness effects of the trait (often conceived as a cost $-c$) must balance the relatedness-weighted indirect effects (i.e. the marginal effect a trait change in a neighbour has on an individual's fitness; typically conceived as a benefit b), where relatedness here is evaluated in a population monomorphic for x^*. The equilibrium x^* can thus be seen as the result of a trade-off between different components of inclusive fitness: that is, between the effect an individual has on its own fitness and the effect it has on the fitness of relatives.

Substituting eq. (44) into eq. (6), we obtain that the interaction coefficient can be decomposed as

$$I(y^*) = I_{\mathrm{w}}(y^*) + I_{\mathrm{q}}(y^*), \tag{46}$$

where

$$\begin{aligned} I_{\mathrm{w}}(y^*) &= r(y^*, y^*) \left.\frac{\partial^2 w(y^*, x, y)}{\partial x \partial y}\right|_{x=y=y^*} + [1 - r(y^*, y^*)] \left.\frac{\partial^2 w(x, y, y^*)}{\partial x \partial y}\right|_{x=y=y^*} + \left.\frac{\partial^2 w(x, y^*, y)}{\partial x \partial y}\right|_{x=y=y^*} \\ I_{\mathrm{q}}(y^*) &= \left.\frac{\partial w(y^*, y, y^*)}{\partial y}\right|_{x=y=y^*} \left(\left.\frac{\partial r(y^*, y)}{\partial y}\right|_{x=y=y^*} - \left.\frac{\partial r(x, y^*)}{\partial x}\right|_{x=y=y^*} \right), \end{aligned} \tag{47}$$

each capturing a pathway to negative trait-dependent selection.

The first term, $I_{\mathrm{w}}(y^*)$, captures how invasion fitness changes when the mutant and resident traits are varied simultaneously in the same direction, while keeping relatedness constant, hence relatedness $r(y^*, y^*)$ being evaluated in a population monomorphic for x^*. This term shows NTDS is weakened in populations showing higher relatedness when joint trait changes in a focal and its patch neighbour have negative fitness effects (i.e. where $\partial^2 w(x, y, y^*)/\partial x \partial y < 0$; note that $\partial^2 w(x, y^*, y)/\partial x \partial y$, in contrast, consists in the fitness effect of joint changes in the focal and individuals from *other* patches which are always non-relatives). Such a fitness effect could arise for various reasons. One relevant instance would be when the trait mediates a social interaction among patch neighbours that shows complementary or negative synergistic effects (as given by eq. 23, e.g. snowdrift games, Doebeli et al., 2004; Wakano and Lehmann, 2014, or trait based competition for limiting resources leading to character displacement, Ajar, 2003; Schmid et al., 2024). In that case, high relatedness makes polymorphism more difficult simply because it makes it more difficult for social partners to show differentiated phenotypes, and therefore enjoy the inherent benefit that being different yields under this type of interaction. Kin selection thus opposes NTDS here and this effect explains why several studies have found that limited dispersal tends to make evolutionary branching more difficult (e.g. Day, 2001; Ajar, 2003; Wakano and Lehmann, 2014; Mullon and Lehmann, 2019; Schmid et al., 2024).

In contrast, the second term of eq. (46), $I_{\mathrm{q}}(y^*)$, reveals how kin selection can contribute to NTDS through relatedness mediated effects. Suppose the trait under consideration has positive indirect fitness effects, i.e. the effect of a neighbour varying its trait has positive fitness effects: $\partial w(y^*, y, y^*)/\partial y > 0$. Then NTDS tends to arise via relatedness mediated effects when

$$\left.\frac{\partial r(y^*, y)}{\partial y}\right|_{x=y=y^*} - \left.\frac{\partial r(x, y^*)}{\partial x}\right|_{x=y=y^*} < 0. \tag{48}$$

The two derivatives capture two different feedbacks on mutant invasion fitness owing to joint changes in mutant and resident traits. The first term measures how the resident trait alters relatedness among mutants. If this derivative is negative (e.g. because higher resident trait values increase dispersal into mutant patches), then a joint increase in trait values reduces mutant fitness, thereby generating NTDS, since the indirect

benefits of helping by mutants are directed more towards unrelated residents. The second term of eq. (48) measures how the mutant trait itself alters relatedness among mutants. If this derivative is positive (e.g. because higher mutant trait values increase philopatry), then again a joint increase with resident trait reduces mutant fitness: residents receive more of the indirect benefits they produce as mutants interact more with each other. Condition (48) shows that relatedness can either promote or oppose NTDS, depending on how trait expression feeds back on inclusive fitness. This relies on traits affecting both the fitness of neighbours and the probability that those neighbours are relatives. An example of such a process leading to the evolution of a social–dispersal syndrome is given in Appendix F.3. More generally, any trait that modifies relatedness – such as those affecting budding dispersal, dispersal success, establishment probability, generational overlap, or group extinction – may generate negative trait-dependent kin selection.

6.3 The role of modifiers of genetic context in NTDS

We saw in the previous section how group structure and the heterogeneity it entails allow for a form of NTDS that is relatedness-mediated. There, it arose when trait expression increased individual fitness in one context (e.g. when being in a group with individuals expressing a similar trait due to identity-by-descent) and the population trait moving in the same direction reduces the probability of individuals being in that context (e.g. due to dispersal decreasing identity-by-descent). This mechanism rests on the existence of heterogeneity in the population: mutant individuals do not interact randomly but instead experience different fitness effects depending on the local genetic composition of their group or patch. More generally, population heterogeneity and the existence of different contexts in which individuals can find themselves open new pathways for NTDS: it can arise whenever a trait influences both the contexts an individual experiences and fitness within those contexts.

Box 3 formalises this idea by considering populations structured into discrete classes to capture heterogeneity either within or between generations. These classes may correspond to different genetic backgrounds, phenotypic states, environmental conditions, or stages in an individual's life cycle. The general form of the interaction coefficient in such populations (eqs. 3.E–3.F) shows that trait-dependent selection is influenced not only by how population trait expression affects fitness within each class but also how it modifies the class structure. In particular, NTDS emerges when an increase in the resident trait value reduces either; (i) the probability that mutants occur in the class in which their fitness is increased owing to a mutant trait increase; or (ii) the reproductive value of individuals produced by mutants in that class (see Box 3 for further details).

The mechanism we have described here should be especially relevant to the maintenance of polymorphism in traits that influence how individuals experience environmental or demographic variation, such as sex ratio, habitat choice, bet-hedging mechanisms, aging, or life-history switches. Consider for instance a trait influencing habitat choice and reproductive success in that habitat, such that mutants expressing greater trait value preferentially settle in a specific habitat that confers them higher fitness (Ravigné et al., 2009). However,

if the resident population also expresses higher trait values and preferentially settles in the same habitat, this increases local competition and thus reduces the per capita benefit of the trait, leading to NTDS. More broadly, these considerations show how population heterogeneity should favor diversity by allowing different morphs to specialise in different contexts, favouring the coexistence of different trait values that associate their reproductive success with the conditions where they perform best [see also Avila and Mullon, 2023 for similar considerations of context-dependence for the coefficient $H(y^*)$].

7 Discussion

Polymorphism is widespread in nature. Numerous models from population genetics, evolutionary ecology, and game theory have proposed mechanisms by which polymorphisms can be maintained by natural selection. Here, we have aimed to organise these mechanisms by the common thread of negative trait-dependent selection (NTDS), which helps to clarify when adaptive polymorphism can be expected under the gradual evolution of quantitative traits.

7.1 Monomorphic instability, the role of NTDS, and its underlying mechanisms

We started by noting that characterizing the adaptive polymorphism favoured by long-term evolution is, in general, a computationally hard problem (section 2.2.3). This difficulty perhaps explains why the bulk of theoretical models about quantitative trait evolution focus on direction selection or on identifying singular trait values in monomorphic populations. Indeed, the phenotypic selection gradient is the central tool to understand adaptation, from evolutionary ecology (Vincent and Brown, 2005; Pásztor et al., 2016), through kin selection (Frank, 1998; Rousset, 2004), to quantitative genetics (Lande, 1980; Arnold, 2023). Nevertheless, necessary conditions for the emergence of polymorphism by gradual evolution can be derived from the evolutionary instability of convergence stable singularities (section 2.3). This requires the selection interaction coefficient to be negative, which, biologically, means that selection on a trait is reduced when the population expresses a similar trait value, or equivalently, invasion fitness of a mutant increases when the mutant and resident trait values diverge. This negative trait-dependent selection (NTDS) is a necessary condition for adaptive polymorphism (eq. 9). As such, NTDS provides a general computable condition to identify candidate polymorphism under gradual evolution, independent of specific biological mechanisms.

To clarify the biological origins of NTDS, we structured invasion fitness around two classes of fitness factors (section 3). Production factors are inputs to an individual's survival and reproduction that depend directly on its own trait (e.g. baseline fecundity, germination rate). Regulating factors, meanwhile, are inputs to survival and reproduction that depend on the traits expressed by conspecifics (e.g. local density, other's behaviour). These two types of factors can act at different biological levels: from phenotypes, to payoffs from interactions, to invasion fitness (Fig. 1).

When regulating factors at the level of invasion fitness depend solely on the resident trait (e.g. eq. 11), our model captures ecologically mediated interactions. We showed that NTDS under such interactions requires trait-specific responses to production factors, and a decoupling between production and regulating factors across contexts (eq. 15). This decoupling enables niche partitioning, with different trait values adapting to different environments associated with different regulating variables. Focusing on NTDS allowed us to uncover a previously undescribed life-history polymorphism under a standard fecundity–survival trade-off (e.g. Charnov, 1993; Case, 2000). Limited dispersal, by generating local kin competition, enables the coexistence of a morph with low survival and high fecundity alongside a quasi-immortal morph with low fecundity. NTDS arises because higher fecundity raises local offspring density, intensifying kin competition and reducing the fitness of the more fecund morph. This favours allocation towards survival as an alternative strategy.

Our general analysis emphasises that regulating factors can depend on both resident and mutant traits (section 3.3), unlike earlier related formalisations of invasion fitness (e.g. Metz, 2011; Pásztor et al., 2016; Lion, 2018). This generalisation is useful for capturing several biologically relevant scenarios: (1) behavioural responsiveness, where resident individuals adjust their actions in response to mutant traits; (2) diploid inheritance, where mutant alleles co-occur with resident alleles in heterozygotes; and (3) mutant–mutant interactions, as occur under limited dispersal, family interactions, or assortative mating. In such cases, NTDS can arise through interaction effects on regulating factors, without requiring direct effects on production factors. We illustrated these broader mechanisms through examples involving socially and genetically mediated interactions.

Socially mediated interactions involve direct behavioural encounters between individuals (e.g. Maynard Smith, 1982; Akçay and Van Cleve, 2012; McNamara and Leimar, 2020). When fitness increases with the payoff obtained from such interactions, and regulating factors at the fitness level depend only on the resident population (due to e.g. density-dependent regulation), the interaction coefficient can be evaluated directly at the level of payoffs (eq. 23). However, the regulating variables that determine these payoffs may themselves depend on mutant traits. This occurs, for instance, when individuals modulate their behaviour in response to their partner's traits, i.e. under social responsiveness. Social responsiveness introduces additional pathways to NTDS, which can act independently of direct trait effects on payoffs (eq. 28). This led us to revisit models of the evolution of traits that bias how individuals evaluate the consequences of their own actions in pairwise interactions under complete information (section 5.3.2). These dispositions influence not only an individual's own behaviour but also, via social responsiveness, the behaviour of their partners. When interactions have antagonistic effects on payoff (as per eq. 32), we found that selection can favour a dimorphism between "optimists," who overvalue the payoff benefits of their actions, and "pessimists," who undervalue them.

Finally, we considered how genetically mediated interactions can influence NTDS. One relevant instance of this are interactions between mutant and resident alleles within diploid individuals. Under random mating, conditions for the emergence of polymorphism depend solely on heterozygote fitness, making them equivalent to conditions identified in haploid or asexual systems (eq. 37). This implies that diploidy, in itself, does not

change the ecological or social conditions under which gradual evolution leads to polymorphism. Therefore, if one's interest lies primarily in understanding adaptive polymorphism, explicit modelling of genetic architecture is unnecessary: a form of phenotypic gambit for adaptive trait variation. Nevertheless, diploidy can generate NTDS at the level of alleles when there are pleiotropic constraints over multiple production factors (e.g. survival and fecundity). In particular, such constraints can lead to overdominance, or heterozygote advantage, a classical mechanism maintaining genetic variation. Our analysis clarified that, under gradual evolution, complementary allelic effects within heterozygotes (rather than simply intermediate optimal phenotypes) are required for overdominance to drive diversification (eq. 43).

Another important type of genetic interaction arises when limited dispersal generates identity-by-descent between interacting individuals so that kin selection occurs (e.g. Hamilton, 1964; Michod, 1982; Rousset, 2004). Here, we revealed a form of trait-dependent kin selection, which when negative, can contribute to diversification. Specifically, our analysis demonstrates that NTDS can emerge when a trait simultaneously increases the fitness of neighbouring individuals and local relatedness (eq. 48). Such traits effectively alter their own genetic context, creating conditions in which indirect fitness benefits are disproportionately high when the mutant and resident are different. More broadly, when population structure creates different contexts in which the trait can be expressed and influence fitness (e.g. sex, age or environmental structure), NTDS is increased if the trait expressed by an individual increases its reproductive success in a particular context and trait expression at the population level reduces the representation of individuals in this context (Box 3). This pathway, which appears to be underappreciated in the context of NTDS, should be particularly relevant for traits influencing habitat choice, senescence, and sex determination, as these traits naturally interface with context-dependent selection. In this sense, population heterogeneity creates distinct niches, allowing genetic effects to specialise within particular contexts, a process through which diversity begets diversity.

7.2 Oligomorphic stability and open questions

Once selection favours polymorphism, an essential question is where the evolutionary dynamics will lead the population in the long-run. Comparatively few results are available to characterize these long-term outcomes of adaptive polymorphism in quantitative traits. Nevertheless, a useful insight is that such a polymorphism should be oligomorphic so that only a discrete number of morphs stably coexist in the uninvadable coalition (e.g. Ellner and Sasaki, 1996; Gyllenberg and Meszéna, 2005). For most of our illustrative examples, we analyzed this oligomorphism in the Appendix (Appendix D.1 and Appendix E.1). But in general, analytically characterizing this oligomorphism, especially when environments fluctuate through time, remains extremely challenging. A relevant research avenue would therefore be developing approaches or techniques to better infer and characterize the convergence stable and uninvadable coalition.

Across the different examples analysed in the Appendix, we found that at the convergence-stable and un-

invadable coalition, trait-dependent selection often vanishes completely (see Appendix E.2.1 and Appendix E.2.2). This pattern has been established previously in scenarios where regulating factors depend solely on resident trait values and where the number of coexisting morphs matches the number of independent regulating variables (Kisdi and Geritz, 2016, Proposition 6). In such cases, the resulting polymorphism is said to be saturated (Kisdi and Geritz, 2016). Our analysis suggests that saturation – in the broader sense of all trait-dependent selection vanishing – can occur under more general conditions. Future work could therefore aim to establish sufficient conditions under which adaptive evolutionary endpoints results in vanishing trait-dependent selection, going beyond ecologically mediated interactions. Such knowledge would be particularly valuable because it would allow evolutionary end-points to be fully characterized by the selection gradient alone (since local uninvadability coincides precisely with convergence stability under vanishing trait-dependent selection). Given that the most biologically explicit representation of the selection gradient is the generalized marginal inclusive fitness effect for polymorphic populations (which is capable of capturing both class and group structure; Priklopil and Lehmann, 2021, eq. 101-102), this would further solidify the central role that Hamilton's marginal rule plays in understanding the nature and process of adaptation of quantitative traits.

7.3 Conclusions

By decomposing invasion fitness into production and regulating factors and focusing on negative trait-dependent selection (NTDS) we obtained necessary conditions for adaptive polymorphism applicable across diverse biological systems. This allows for (i) a consistent criterion to identify the presence or absence of NTDS; (ii) a systematic way of categorising regulating factors at the phenotype, payoff, and fitness levels, clarifying their roles in driving NTDS; and (iii) a synthesis of ecological, social, and genetic scenarios that are conducive to NTDS and thus adaptive polymorphism. These illustrative scenarios, while not exhaustive, highlight the diverse origins of polymorphism. They reinforce the notion that given the hierarchical organisation of biological systems and the ubiquity of competition, NTDS may be common rather than exceptional. If NTDS is indeed so prevalent, it may hold the key to understanding why intra-specific trait diversity is so pervasive, continuously providing the fuel for evolutionary change.

Acknowledgments We thank Claus Rueffler for many useful discussions, providing references, and spotting an oddity within a previous version of one model. We gratefully acknowledge two anonymous reviewers who provided extremely useful and very extensive comments that fundamentally improved the manuscript, the results, and the presentation.

Box 1. Uninvadable coalition, mean fitness, Nash equilibrium, and computational complexity

The correspondence between eq. (3) and eq. (5) extends to the case of between generation environmental fluctuations (i.e., when the uninvadable coalition is given by eq. (2)). To see this, let $\mathcal{M} = \Delta(\Delta(\mathcal{X}))$ denote the set of probability measures over the set $\Delta(\mathcal{X})$ of population distributions; that is, element $v \in \mathcal{M}$ is a random probability measure. Then, the mean fitness of measure $v \in \mathcal{M}$ in a population with measure $\mu_{\mathrm{p}} \in \mathcal{M}$ is given by

$$\overline{\rho}(v, \mu_{\mathrm{p}}) = \int_{\Delta(\mathcal{X})} \int_{\mathcal{X}} \rho(x, \mu_{\mathrm{p}}) \, \mathrm{d}q(x) \, \mathrm{d}v(q). \tag{1.A}$$

With this, an uninvadable distribution μ_{p}^* satisfies eq. (2) if and only if it satisfies

$$\mu_{\mathrm{p}}^* \in \arg\max_{\mu_{\mathrm{p}} \in \mathcal{M}} \overline{\rho}(\mu_{\mathrm{p}}, \mu_{\mathrm{p}}^*). \tag{1.B}$$

In other words, μ_{p}^* is an uninvadable distribution if and only if it maximizes mean fitness in a population with distribution μ_{p}^*. This correspondence between eq. (2) and eq. (1.B) can by understood by applying the standard relationship between pure and mixed strategy Nash equilibria from game theory [Mas-Colell et al., 1995, p. 250 and Narahari, 2014, p 99 for discrete strategy (trait) space, and for metric spaces, see Anderson et al. (2013), Proposition 1]. Namely, eq. (2) says that all morphs (pure strategies) in the support of μ_{p} are best responses (Nash equilibria) to the population distribution μ_{p} (mixed strategy). To see that this is equivalent to the distribution μ_{p} being a best response to itself in terms of mean fitness consider the contradiction that μ_{p} contains in its support a morph that is not a best response. Then, the average population fitness could be improved by decreasing the frequency of this morph and concomitantly increasing the frequency of any morph that is a best response. This then contradicts the assumptions that the population distribution is a best response. Conversely, if all morphs in the support of μ_{p} are best responses, then the population distribution μ_{p} must be a best response to itself.

The correspondence between uninvadability and Nash equilibrium allows us to lift two results from game theory to the problem of evaluating long-term evolutionary outcomes. To present these, it is sufficient to focus on the case of no environmental fluctuations, when the uninvadable coalition is characterized by eq. (5). First, when $\mathcal{X}$ is compact and the payoff function $\rho(x, p)$ is continuous in p, then a Nash equilibrium is guaranteed to exist (see Appendix A for a proof). Second, finding such a Nash equilibrium is computationally hard (e.g. Papadimitriou, 2007; Conitzer and Sandholm, 2008; Papadimitriou, 2014; Narahari, 2014, Theorem 2.1). Just finding the mixed Nash equilibrium in a two-player symmetric finite game belongs to the PPAD complexity class (e.g. Papadimitriou, 2007; Conitzer and Sandholm, 2008; Papadimitriou, 2014; Narahari, 2014, Theorem 2.1). There is strong evidence that problems in this class cannot be solved with a polynomial algorithm, i.e., PPAD$\neq$P (Papadimitriou, 2007, p.39). While our population game does not necessarily reduce to a two-player symmetric game in mixed strategies (it can under certain cases), it is an infinite game and thus computational issues are likely to be harder. This does not mean that the uninvadable coalition can never be computed. If the invasion fitness has a special structure, it can be solved in certain cases. For instance, in "succinct games" (Papadimitriou, 2007), the analysis becomes more tractable, as seen in congestion games. In biological terms, this corresponds to models where competition for resources depends on some aggregate effect, such as foraging models with an ideal free distribution (e.g. Křivan et al., 2008).

The conclusion is twofold. First, one should be aware of the computational complexity of evaluating a uninvadable coalitions and results from algorithmic game theory may be useful to evolutionary biology. Second, identifying conditions necessary for the emergence of adaptive polymorphism without computing the long term equilibrium is a valuable first-step to understanding polymorphism.

Box 2. Intensive versus extensive variables and explicit ecological models

When regulating factors are extensive variables they represent dynamical state variables that have reached some equilibrium (or more generally some attractor). For instance, when the regulating factors $e_1(y)$ and $e_2(y)$ in invasion fitness eq. (11) are extensive variables, they can be written to satisfy the following system of equations:

$$\begin{aligned} R_1(z_1(y), z_2(y), e_1(y), e_2(y)) &= 0 \\ R_2(z_1(y), z_2(y), e_1(y), e_2(y)) &= 0. \end{aligned} \tag{2.A}$$

Here, R_1 and R_2 are mappings that characterise the equilibrium of the regulating variables, which are assumed to be locally asymptotically stable (see also Kisdi and Geritz, 2016, eq. 8). Since individuals interact with their environment via their production factor, the equilibrium values of both regulating factors $e_1(y)$ and $e_2(y)$ depend on both production factors $z_1(y)$ and $z_2(y)$. This could describe how competition for food resources $e_1(y)$ and breeding sites $e_2(y)$ is impacted by individuals' efforts in foraging $z_1(y)$ or reproduction $z_2(y)$.
An explicit example of an ecological map of the form eq. (2.A) and how extensive variables can be expressed in terms of intensive variables can be obtained as follows. Let us consider a population of iteroparous individuals that survive from one time step to the next with probability $z_1(x)$ and produce offspring according to a Beverton-Holt model of reproduction $z_2(x)/[1 + \gamma e_1(y) + \eta e_2(y)]$, where $z_2(x)$ is the baseline fecundity in the absence of density-dependence, $e_1(y)$ is the density of the focal species, $e_2(y)$ is the density of some other species, and γ and η are parameter tuning the importance of density-dependence. Invasion fitness of a mutant in the focal species then takes the form

$$\rho(x, y) = z_1(x) + \frac{z_2(x)}{1 + \gamma e_1(y) + \eta e_2(y)}, \tag{2.B}$$

which is a Leslie-Gower model (Leslie and Gower, 1958, 1960); the discrete analogue of the continuous time Lotka-Voltera model (Leslie and Gower, 1960). Since both $e_1(y)$ and $e_2(y)$ are extensive variables, we need maps describing population dynamics at equilibrium in these two species: $e_1(y) = \rho(y, y)e_1(y)$ and $e_2(y) = w_2(y)e_2(y)$ where $w_2(y)$ is the fitness of an individual from the second species, which may depend on all quantities $z_1(y), z_2(y), e_1(y)$ and $e_2(y)$. Hence, in order to arrive at a system of the form of eq. (2.A), we can use

$$R_1(z_1(y), z_2(y), e_1(y), e_2(y)) = \rho(y, y) - 1 = z_1(y) + \frac{z_2(y)}{1 + \gamma e_1(y) + \eta e_2(y)} - 1 = 0 \tag{2.C}$$

for the focal species, while for the second species we can use

$$R_2(z_1(y), z_2(y), e_1(y), e_2(y)) = w_2(y) - 1. \tag{2.D}$$

These allow us to solve for the population dynamic equilibrium of the focal species regardless of the functional form of eq. (2.D). Indeed, by solving eq. (2.C) for $e_2(y)$ and by substituting into eq. (2.B) produces

$$\rho(x, y) = z_1(x) + (1 - z_1(y))\frac{z_2(x)}{z_2(y)}. \tag{2.E}$$

Thus, invasion fitness can be expressed in terms of mutant and resident production factors only, thus in terms of intensive variables. Interestingly, invasion fitness (2.E) can be seen as abiding to a lottery model for iteroparous reproduction with the following life-cycle (as in eq. 17 with $m = 1$): (a) a large fixed number of adults produce a large number of juveniles with fecundity e_1 and then each adult either survives or dies independently (with probabilities z_1 and $1-z_1$, respectively); (b) density-dependent competition occurs among juveniles for the breeding spots vacated by the death of adults so that the population is regulated back to the fixed adult size. Such life-cycles are standard in population genetics. Assuming further $z_1(x) = 0$ gives the Wright-Fisher model (Ewens, 2004; Hartl and Clark, 2007). These connections illustrate that standard models of population genetics can implicitly incorporate ecological feedbacks, even when these are not made explicit (as is well established, e.g. Crow and Kimura, 1970; Nagylaki, 1992). In such models, ecological feedbacks do not modify the selection pressures being studied, which is why they can be omitted without loss of generality.

Box 3. Class structured populations

We consider here heterogeneous populations where individuals can belong to a finite number of classes, where $\mathcal{C}$ denotes the set of possible classes (e.g., genetic background, condition, age, environmental state). Assuming that this heterogeneity occurs either only within-generations (e.g. stage) or only between-generations – in which case heterogeneity is independently and identically distributed (e.g. temporal fluctuations in climate) – invasion fitness is given by

$$\rho(x,y) = \begin{cases} \sum_{c'\in\mathcal{C}}\sum_{c\in\mathcal{C}} w_{c'c}(x,y)q_c(x,y) & \text{if only within generation heterogeneity} \quad (3.A) \\ \prod_{c\in\mathcal{C}} w_c(x,y)^{q_c(y)} & \text{if only between generation heterogeneity,} \quad (3.B) \end{cases}$$

where as before mutant individuals express trait value x while the resident population has trait y. In eq. (3.A), $w_{c'c}(x,y)$ is the expected number of offspring of class c' (possibly including the surviving self) produced by a mutant in class c, and $q_c(x,y)$ is the probability that a mutant is in class c (see Appendix G.1 or Lehmann et al., 2016, Box 3). In eq. (3.B) $w_c(x,y)$ is the expected number of offspring produced by a mutant in class or environmental state c and $q_c(y)$ is the probability of occurrence of environmental state $c \in \mathcal{C}$, which depends only on resident traits (as in standard invasion analysis, e.g., Ferrière and Gatto, 1995). Note that owing to between-generation heterogeneity, there is only type of offspring so $w_c(x,y)$ depends only on the environmental state c in which the individual reproduces. We show in Appendix G.3 how these class-specific quantities can be expressed in terms of regulation and production factors such that invasion fitness in both cases of heterogeneity can be conceptualised as a mapping from fitness factors to geometric growth ratio (i.e. as in eq. 10).

Using eqs. (3.A)-(3.B), the selection gradient can be expressed as

$$S(y) = \begin{cases} \sum_{c'\in\mathcal{C}}\sum_{c\in\mathcal{C}} v_{c'}(y)\left.\frac{\partial w_{c'c}(x,y)}{\partial x}\right|_{x=y} q_c(y,y) & \text{if only within generation heterogeneity} \quad (3.C) \\ \sum_{c\in\mathcal{C}} \left.\frac{\partial \log(w_c(x,y))}{\partial x}\right|_{x=y} q_c(y) & \text{if only between generation heterogeneity,} \quad (3.D) \end{cases}$$

where $v_c(y)$ is the reproductive value of an individual in class c in a (monomorphic) resident population y, satisfying $v_c(y) = \sum_{c'\in\mathcal{C}} v_{c'}(y)w_{c'c}(y,y)$ while $q_{c'}(y,y) = \sum_{c\in\mathcal{C}} w_{c'c}(y,y)q_c(y,y)$ (eq. 3.C is a standard result, e.g., Taylor and Frank, 1996; Rousset, 2004; Lehmann et al., 2016 and eq. 3.D is derived in Appendix G.2). The singular trait value y^* thus reflects a trade-off of fitness effects among different classes, which are appropriately weighted according to how often an individual is in class c (via $q_c(y,y)$ or $q_c(y)$) and offspring reproductive value under within generation heterogeneity (via $v_{c'}(y)$ in eq. 3.C).

The interaction coefficient, meanwhile, can be written as,

$$\begin{aligned} I(y^*) = \sum_{c'\in\mathcal{C}}\sum_{c\in\mathcal{C}} v_{c'}(y^*)q_c(y^*,y^*)\Bigg[&\frac{\partial^2 w_{c'c}(x,y)}{\partial x\partial y} \\ &+ \frac{\partial w_{c'c}(x,y)}{\partial x}\left(\frac{\partial \log(q_c(x,y))}{\partial y} - \frac{\partial \log(q_c(x,y))}{\partial x} + \frac{\partial \log(v_{c'}(y))}{\partial y}\right)\Bigg]_{x=y=y^*} \end{aligned} \quad (3.E)$$

for within generation heterogeneity, and

$$I(y^*) = \sum_{c\in\mathcal{C}} q_c(y^*)\left[\frac{\partial^2 \log(w_c(x,y))}{\partial x\partial y} + \frac{\partial \log(w_c(x,y))}{\partial x}\frac{\partial \log(q_c(y))}{\partial y}\right]_{x=y=y^*} \quad (3.F)$$

for between generation heterogeneity (Appendix G for derivation). The first terms between square brackets in eqs. (3.E) and (3.F) show that NTDS can simply be due to negative interaction effects on class specific fitness components (i.e. via $\partial^2 w_{c'c}(x,y)/(\partial x\partial y)$ and $\partial^2 \log(w_c(x,y))/(\partial x\partial y)$). The rest of eqs. (3.E) and (3.F) highlight how population heterogeneity opens new pathways for NTDS. These can be understood by assuming that mutant trait expression increases mutant fitness in a given class ($\partial w_{c'c}(x,y)/\partial x > 0$ or $\partial \log(w_c(x,y)/\partial x > 0$). A joint increase in the resident trait value then tends to have negative effects on mutant invasion when such an increase reduces: (i) the probability that mutants belong to the class in which their fitness is enhanced (via $\partial \log(q_c(x,y))\partial y < 0$ in eq. 3.E or $\partial \log(q_c(y))\partial y < 0$ in eq. 3.F – in eq. 3.E, the term $[-\partial \log(q_c(x,y))/\partial x]$ emerges owing to a scaling of invasion fitness, Appendix G.1 for details); and/or (ii) the reproductive value of offspring produced in that class under within generation heterogeneity (via $\partial \log(v_{c'}(y))\partial y < 0$ in eq. 3.E). We discuss these pathways more generally in section 6.3.

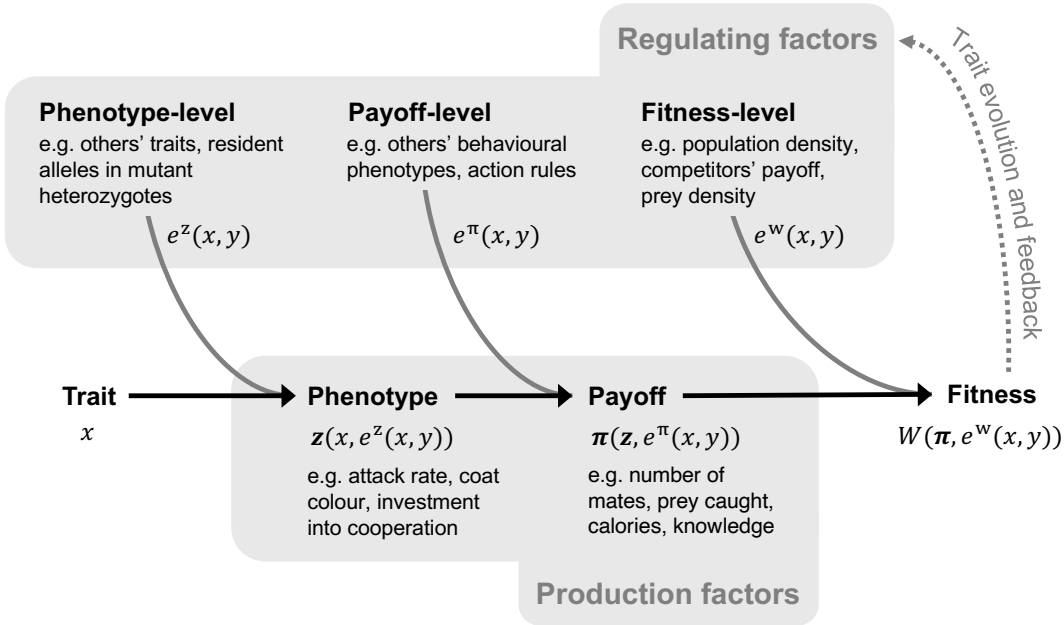

Figure 1: **Production and regulating fitness factors.** Diagram showing the construction of fitness, from genetic trait value x through production factors: phenotype and payoff. These various layers and their regulating factors offer different opportunities for population feedback as trait evolution leads to changes in regulating factors. Section 3.2 for details.

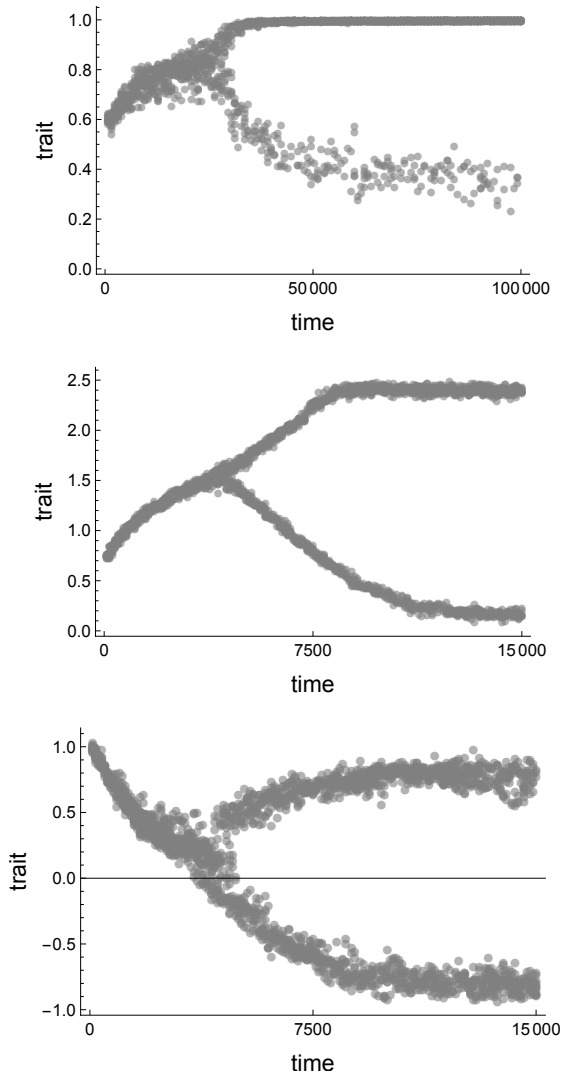

Figure 2: **Illustration of evolutionary branching and long term evolution under individual-based simulations for a sample of novel polymorphisms**. Top panel is for the fitness model eq. (17), which leads to the coexistence of an immortal morph investing all resources into survival and another one expressing a mixed strategy (the trait y is here investment into survival, with fraction $1-y$ invested into fecundity, see Appendix D.2 for the corresponding formal analysis). Middle panel is for payoff model eq. (29), which leads to the coexisting of one morph investing a lot into contests ("Hawk") and another one investing moderately ("Dove") (the trait y here is investment into contest, see Appendix E.2.1 for the corresponding formal analysis). Bottom panel is for the payoff model eqs. (32)–(34), which leads to the coexisting of optimists and pessimists (the trait y here is a self esteem bias, with $y > 0$ describing optimism and $y < 0$ pessimism, see Appendix E.2.2 for the corresponding formal analysis). In each case, reproduction follows a Wright-Fisher process (lottery model) with $N = 10000$ individuals. Mutations occur with probability 0.015 and effect sizes are sampled from a Normal distribution with mean 0 and standard deviation 0.01 (trait values are capped between zero and one in the top panel).

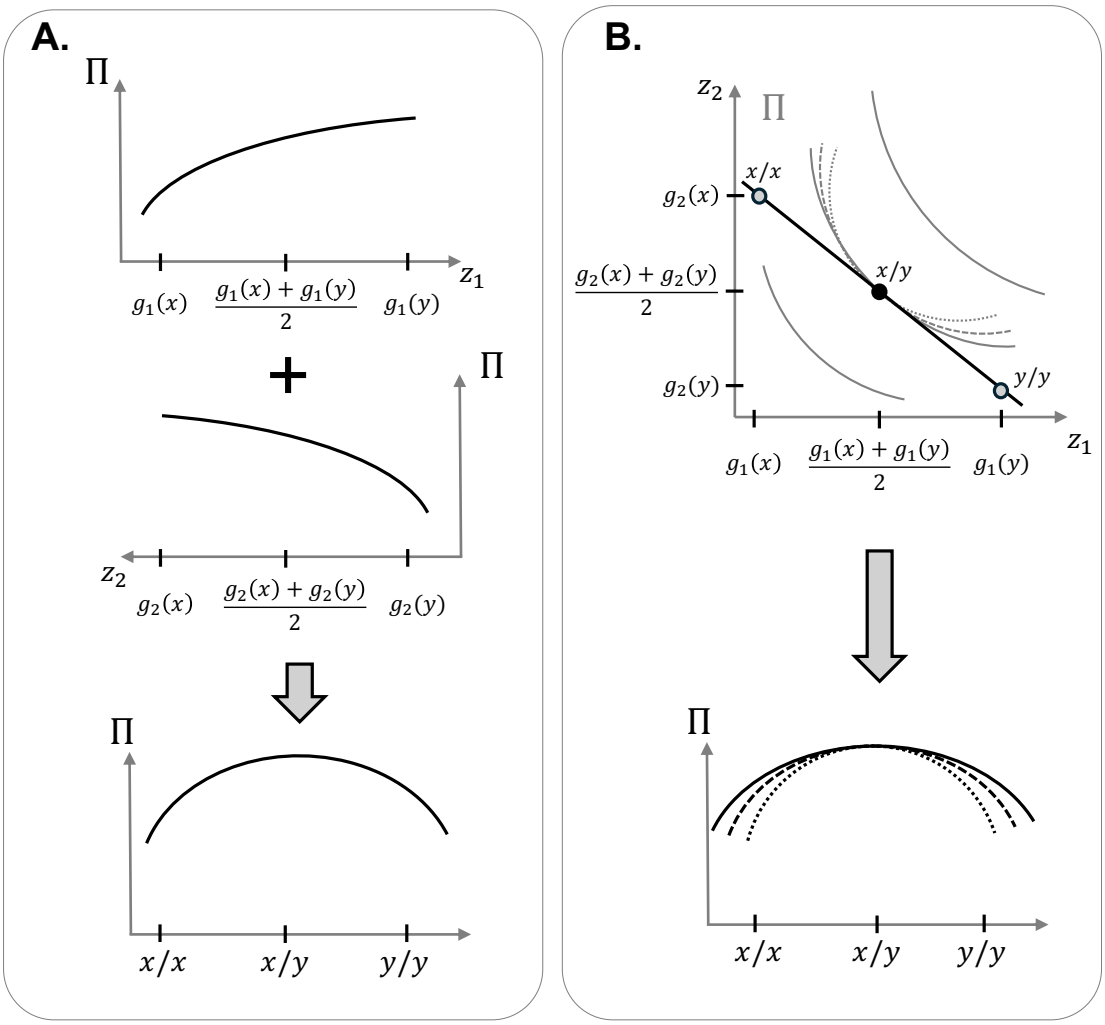

Figure 3: **A graphical illustration of overdominance origin under pleiotropic constraints. A**. Overdominance can arise when the payoff function $\Pi(z_1, z_2)$ decelerates with both phenotypes increasing (first two terms of eq. 43). The top and middle graphs show the payoff $\Pi(z_1, z_2)$ increasing in a decelerating manner with both z_1 (top) and z_2 (middle). When the payoff increases with both phenotypes, z_1 and z_2 must change in opposite directions near a singular trait value (as implied by eq. 42). The direction of the x-axes are reversed in the top and middle graphs to ensure genotype alignment: x/x homozygotes on the left, x/y heterozygotes in the centre, and y/y homozygotes on the right. Summing the effects across both phenotypes to compute the total payoff for each genotype leads to overdominance, with heterozygotes (x/y) achieving the highest payoff (captured by the two first terms in eq. 43). **B**. Overdominance is also favoured when phenotypes have complementary effects on payoff (last term of eq. 43). The top graph shows a case where eq. (43) is negative and payoff increases with both phenotypes. Grey curves are payoff isoclines, which are convex owing to eq. (43) being negative (solid, dashed, dotted lines indicate increasing strength of complementarity). According to eq. (42), z_1 and z_2 vary in opposite directions near the singular trait value y^*, necessarily moving along tangents to a payoff contour (black lines, following the genotypes shown on graph). When mutant and resident allelic values x and y diverge in opposite directions away from y^*, this yields heterozygotes with higher payoff than either homozygote. The effect is stronger under complementarity effects (compare black line and grey curves in top panel and see bottom panel). This increase in mutant fitness under opposing allelic changes is a signature of NTDS.

Appendix A Uninvadable coalition existence

In Box 1, we state that when $\mathcal{X}$ is compact and the payoff function $\rho(x,p)$ is continuous in p, then a mixed Nash equilibrium satisfying eq. (5) is guaranteed to exist and thus an uninvadable coalition. This follows from applying established results to the best-reply correspondence $b(q) = \arg\max_{p\in\Delta(\mathcal{X})} \overline{\rho}(p,q)$ as is usually done in game theory for normal form games (Fudenberg and Tirole, 1991; Mas-Colell et al., 1995). When $\mathcal{X}$ is non-empty and compact, so will be $\Delta(\mathcal{X})$ (Alipantris and Border, 2006, Th. 15.11). By Berge's maximum theorem (Alipantris and Border, 2006, Th. 17.31), the best-reply $b(q)$ then defines an upper hemi-continuous and non-empty compact-valued correspondence. Because $b(q)$ is convex-valued (since $\overline{\rho}(p,q)$ is linear in p) and $\Delta(\mathcal{X})$ is a convex set (since the mixture of two probability distribution is a probability distribution), the conditions of Glicksberg's fixed point theorem are satisfied, and therefore the set of equilibrium point $p^* \in b(p^*)$ is non-empty and compact (Alipantris and Border, 2006, Th. 17.55). In biological terms, there is at least one uninvadable coalition in any evolutionary model that satisfy our main biological assumptions (see section 2), unless the set of traits $\mathcal{X}$ is unbounded. This uninvadable coalition may consist of an adaptive polymorphism or be a singleton trait, say x^*, in which case the uninvadable trait value satisfies $x^* \in \arg\max_{x\in\mathcal{X}} \rho(x,x^*)$ for invasion fitness $\rho(x,y)$ of a mutant x in a monomorphic resident population y. This latter case corresponds to a symmetric Nash equilibrium in pure strategies and such an equilibrium exists if $\rho(x,y)$ is continuous, quasi-concave in its first argument, and $\mathcal{X}$ is compact (e.g. Fudenberg and Tirole, 1991, p. 34).

Appendix B Dimorphic coexistence near a singularity

Under gradual evolution, evolutionary convergence towards an invadable monomorphism y^* results in the population subsequently spiting into two distinct morphs leading to their coexistence in a protected polymorphism (Eshel et al., 1997; Geritz et al., 1998). This coexistence can be understood by considering two morphs, say $x_1 = y+\delta \in \mathcal{X}$ and $x_2 = y-\delta\epsilon \in \mathcal{X}$, that are weakly diverged around y with $|\delta| \ll 1$ where $\epsilon > 0$ is assumed to be of order $O(1)$. These morphs will coexist if they can mutually invade one another, that is when $\rho(x_1,x_2) > 1$ and $\rho(x_2,x_1) > 1$ (e.g., Hofbauer and Sigmund, 1998; Turelli et al., 2001). Second order Taylor expanding $\rho(x_1,x_2)$ and $\rho(x_2,x_1)$ in δ around $\delta = 0$, and using the condition $\rho(y,y) = 1$ differentiated twice with respect to y, we get

$$\begin{aligned}\rho(x_1,x_2) &= 1-\delta(1+\epsilon)\left[\epsilon\delta I(y) - \left(S(y) + \frac{\delta(1-\epsilon)}{2}H(y)\right)\right] + O(\delta^3)\\ \rho(x_2,x_1) &= 1-\delta(1+\epsilon)\left[\delta I(y) + \left(S(y) + \frac{\delta(1-\epsilon)}{2}H(y)\right)\right] + O(\delta^3),\end{aligned} \tag{A-1}$$

where $S(y)$, $H(y)$ and $I(y)$ are given in eq. (6). The term $(S(y) + \delta(1-\epsilon)H(y)/2)$ has opposite effects on the invasion fitness of each morph: increasing fitness of x_2 and diminishing that x_1. Any mechanism that

makes this term closer to zero can therefore be seen as an equalizing mechanism in the jargon of coexistence theory (Chesson, 2000b). Equalizing mechanisms reduce inherent fitness differences among morphs thus favouring coexistence. From that perspective, directional selection leading to a singular trait value y^* such that $S(y^*) = 0$ can be seen as an equalizing mechanism.

The term $I(y)$, in contrast, influences the invasion fitness of each morph in the same way. Any mechanism that makes $I(y)$ more negative thus acts as a stabilizing mechanism, favouring mutual invasion. When $I(y) > 0$, mutual invasion is not possible, i.e. $I(y) < 0$ is a necessary condition for mutual invasion. Additionally, when $\epsilon = 0$ (such that the two morphs are symmetrically distributed around y) and $y = y^*$ is singular, then $I(y) < 0$ also becomes a sufficient condition for mutual invasion.

Appendix C Oligomorphic coexistence

We here extend the analysis of the (in)stability of monomorphism (section 2.3) to the case of oligomorphism. This is to (i) complement the analysis of the main text, (ii) show how the decomposition of invasion fitness eq. (10) can be applied to the oligomorphic case, (iii) allow us to characterize the uninvadable polymorphism for examples in the main text, and (iv) display some of the complexities involved in characterizing an uninvadable and convergence stable oligomorphism.

Appendix C.1 No class structure nor environmental fluctuations

We start by describing the conditions for the uninvadability and convergence stability of an oligomorphism for the case of absence of (i) class structure and (ii) between generation fluctuating environments (and discuss relaxing these assumptions in Appendix C.2). For this case, we can replace μ by p in eq. (2) and since we consider only a resident population with a discrete number n of segregating morphs we set $p = \mathbf{p} = (p_1, p_2, ..., p_n)$ to denote the probability distribution of these morphs in the population ($\sum_{i=1}^{n} p_i = 1$), with support $\mathcal{P}(\mathbf{p}) = \mathbf{y} = (y_1, y_2, ..., y_n)$. We then denote by $\tilde{\rho}(x, \mathbf{y}, \mathbf{p})$ the invasion fitness of type x in a type $\mathbf{y}$ population with morph frequency distribution $\mathbf{p}$.

Appendix C.1.1 Uninvadable oligomorphism

At an equilibrium of the dynamics of morph frequencies, each morph $y_i \in \mathcal{P}(\mathbf{p})$ must have invasion fitness equal to one so that the equilibrium frequency vector $\mathbf{p}$ must satisfy:

$$\tilde{\rho}(y_i, \mathbf{y}, \mathbf{p}) = 1 \qquad \text{for all } y_i \in \mathcal{P}(\mathbf{p}). \tag{A-2}$$

Assume such an equilibrium exists and that it is hyperbolic (a hyperbolic equilibrium means that the Jacobian matrix of the maps $\tilde{\rho}(y_i, \mathbf{y}, \mathbf{p})p_i$ of the dynamical system of morph frequencies $i = 1, .., n$ has eigenvalues with non-zero real parts, Hirsch et al., 2004). Hyperbolicity guarantees that (i) the equilibrium $\mathbf{p}(\mathbf{y})$ of eq. (A-2) can be written as a continuous function of $\mathbf{y}$ and (ii) that any of the finite number of fixed point of eq. (A-2) retain their stability properties upon changes of morph $\mathbf{y}$ values around a singularity $\mathbf{y}^*$ (from the implicit function theorem Sydsaeter et al., 2008, p.84 and Principle 3 of Karlin and McGregor, 1972). Then, the invasion fitness of a mutant $x \in \mathcal{X}$ can be written as

$$\rho(x, \mathbf{y}) = \tilde{\rho}(x, \mathbf{y}, \mathbf{p}(\mathbf{y})). \tag{A-3}$$

From this, the three coefficients of selection of eq. (6) generalize to

$$S_i(\mathbf{y}) = \left.\frac{\partial \rho(x, \mathbf{y})}{\partial x}\right|_{x=y_i}, \quad H_i(\mathbf{y}) = \left.\frac{\partial^2 \rho(x, \mathbf{y})}{\partial x^2}\right|_{x=y_i} \text{ and } \quad I_{ij}(\mathbf{y}) = \left.\frac{\partial^2 \rho(x, \mathbf{y})}{\partial x \partial y_j}\right|_{x=x_i}. \tag{A-4}$$

A sufficient condition for an oligomorphic resident population with support $\mathcal{P}(\mathbf{p}^*) = \mathbf{y}^*$ and equilibrium frequency $\mathbf{p}^*$ to be locally uninvadable with trait values in the interior of the state space is thus that for all $i \in \{1, 2, ..., n\}$:

$$S_i(\mathbf{y}^*) = 0 \tag{A-5}$$

$$H_i(\mathbf{y}^*) < 0. \tag{A-6}$$

A necessary condition for local uninvadability is that eq. (A-6) holds with weak inequality. An oligomorphism satisfying only condition (A-5) is referred to as a singular coalition, and those satisfying both conditions (A-5) and (A-6) as locally uninvadable coalitions. The concept of a locally uninvadable coalition is standard, e.g. eq. (17) in Vincent et al. (1996), eq. (15) in Geritz et al. (1998), eq. (36) in Sasaki and Dieckmann (2011), or eqs. (11) in Kisdi and Geritz (2016).

Appendix C.1.2 Convergence stable oligomorphism and saturation

When conditions (A-5) and (A-6) hold, a coalition with morphs $\mathbf{y}^*$ and morphs frequencies $\mathbf{p}^*$, if established, will no longer change. But will such a pair ($\mathbf{y}^*$, $\mathbf{p}^*$) establish through gradual evolution; namely, is ($\mathbf{y}^*$, $\mathbf{p}^*$) an attractor of the evolutionary dynamics? Convergence stability here means that for any coalition $\mathbf{y} \neq \mathbf{y}^*$ in a small neighborhood of $\mathbf{y}^*$, the evolutionary dynamics should gradually converge to ($\mathbf{y}^*$, $\mathbf{p}^*$) by small step mutations. In a population of morphs $\mathbf{y} \neq \mathbf{y}^*$ within some neighborhood of the uninvadable coalition, the associated equilibrium morph frequencies $\mathbf{p} \neq \mathbf{p}^*$ will differ from that of the uninvadable coalition. If $\mathbf{y}^*$ is favoured in such a population, the frequency of morphs should eventually be attracted to $\mathbf{p}^*$, which requires that this is a local asymptotically stable fixed point of morph frequency dynamics. We therefore require

that an uninvadable coalition $\mathbf{y}^*$ is locally convergence stable in the sense that it is a phenotypic attractor and that $\mathbf{p}^*$ is a local asymptotically stable point of morph frequency dynamics (implying that if any of the regulating variables are extensive variables, then they must also reside at asymptotically stable fixed points of their local dynamics).

In a population at $\mathbf{y}$ close to an uninvadable $\mathbf{y}^*$ with local asymptotically stable $\mathbf{p}^*$, the selection gradient $S_i(\mathbf{y})$ is sufficient to describe the direction of long-term evolution. Specifically, $S_i(\mathbf{y})$ determines whether a mutant $y_i + \delta$ arising from morph y_i with $|\delta| \ll 1$ in population state $\mathbf{y}$ will go extinct or substitute it ancestral type, leading to a new polymorphic population state $(y_1, y_{i-1}, y_i + \delta, y_{i+1}, ..., y_n)$ (Cai and Geritz, 2020; Priklopil and Lehmann, 2020, 2021). Since mutations occur independently in each morph, the rate of trait change of morph i is given by $v_i(\mathbf{y}(t))S_i(\mathbf{y})$, where $v_i(\mathbf{y}(t)) > 0$ is some measure of genetic variance in the subpopulation of morph i. More often than not, the variances $v_i(\mathbf{y}(t))$ differ among different morphs (due to e.g. different subpopulation sizes of morphs or selection, Sasaki and Dieckmann, 2011). Therefore, a necessary condition for a polymorphic singularity satisfying eq. (A-5) to be convergence stable, i.e. to be attracted towards $\mathbf{y}^*$ regardless of the scaling factor, is that the Jacobian matrix

$$\mathbf{J}(\mathbf{y}^*) = \begin{pmatrix} \dfrac{\partial S_1(\mathbf{y})}{\partial y_1} & \cdots & \dfrac{\partial S_1(\mathbf{y})}{\partial y_n} \\ \vdots & & \vdots \\ \dfrac{\partial S_n(\mathbf{y})}{\partial y_1} & \cdots & \dfrac{\partial S_n(\mathbf{y})}{\partial y_n} \end{pmatrix}_{\mathbf{y}=\mathbf{y}^*} \tag{A-7}$$

has all its eigenvalues with negative real parts. Sufficient conditions for convergence stability include (i) that $\mathbf{J}(\mathbf{y}^*)$ is a negative definite matrix (i.e., all eigvenvalues of the symmetric part of $\mathbf{J}^*_{\mathbf{y}}(\mathbf{s})$ are negative in which case the oligomorphic singularity is said to be strongly convergence stable (Apaloo and Butler, 2009; Kisdi and Geritz, 2016) and (ii) that $\mathbf{J}(\mathbf{y}^*)$ is a sign stable matrix (e.g. Quirk and Ruppert, 1965, Theorem 5, May, 1973, p. 71), in which case convergence stability can be predicated upon qualitative properties of the signs of the entries of $\mathbf{J}^*_{\mathbf{y}}(\mathbf{s})$.

In the same way as eq. (8) decomposes convergence stability into quadratic selection and trait-dependent selection, the Jacobian eq. (A-7) can be decomposed as

$$\mathbf{J}(\mathbf{y}^*) = \mathbf{H}(\mathbf{y}^*) + \boldsymbol{I}(\mathbf{y}^*), \tag{A-8}$$

where $\mathbf{H}(\mathbf{y}^*)$ is a diagonal matrix with (i,i)-element $H_i(\mathbf{y}^*)$ capturing disruptive selection on morphs i (eq. A-4), while matrix $\boldsymbol{I}(\mathbf{y}^*)$ has (i,j)-element given by $I_{ij}(\mathbf{y}^*)$, which is the effect of a change in the trait of morph j on selection in morph i. As such, matrix $\boldsymbol{I}(\mathbf{y}^*)$ captures trait-dependent selection in an oligomorphic population. For an oligomorphism $\mathbf{y}^*$ to be locally convergence stable yet invadable (so that $H_i(\mathbf{y}^*) > 0$ for some i), it is necessary that $\boldsymbol{I}(\mathbf{y}^*)$ contains negative entries and thus that some NTDS occurs. To see this, suppose that all non-diagonal elements of $\boldsymbol{I}(\mathbf{y}^*)$ are non-negative. In this case, i.e., $\mathbf{J}(\mathbf{y}^*)$ is a Metzler matrix. We then know that for $\mathbf{J}(\mathbf{y}^*)$ to have all eigenvalues with negative real parts, it is necessary that $\mathbf{J}(\mathbf{y}^*)$ has

only negative diagonal elements (Briat, 2017, Lemma 3.1). This entails that it is necessary that $I_{ii} < 0$ for any trait i where $H_i(\mathbf{y}^*) > 0$. This latter condition also holds if each morph evolves independently of each other (Geritz et al., 1998; Sasaki and Dieckmann, 2011).

Appendix C.1.3 Vanishing NTDS and model in terms of regulating and production factors

The expression of invasion fitness in terms of production and regulating factor for a monomorphic population (eq. 10) immediately extends to the oligomorphic case by setting

$$\tilde{\rho}(x, \mathbf{y}, \mathbf{p}) = W\big(\boldsymbol{\pi}(\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x, \mathbf{y}, \mathbf{p})), \mathbf{e}^{\pi}(x, \mathbf{y}, \mathbf{p})), \mathbf{e}^{\mathrm{w}}(x, \mathbf{y}, \mathbf{p})\big), \tag{A-9}$$

where the regulating factors are now expressed in terms of the traits in the population and their frequency distribution, i.e., the tuple $(\mathbf{y}, \mathbf{p})$. For the case of ecologically mediated interaction (section 4) invasion fitness reduces to $\tilde{\rho}(x, \mathbf{y}, \mathbf{p}) = W(\mathbf{z}(x), \mathbf{e}(\mathbf{y}, \mathbf{p}))$ and the population trait feedbacks only act through the regulating variables. Interestingly, Kisdi and Geritz (2016, Proposition 6) have shown that for this case and when the number of oligomorphs in the uninvadable coalition matches the number regulating variables, then trait-dependent selection vanishes; that is

$$\boldsymbol{I}(\mathbf{y}^*) = \mathbf{0}, \tag{A-10}$$

where $\mathbf{0}$ is a matrix of zeroes. Then, the criteria for convergence stability collapses to that of local uninvadability (recall eq. A-8), and in the terminology of Kisdi and Geritz (2016), the resulting adaptive polymorphism is said to be saturated.

Appendix C.2 Class structure and environmental fluctuations

In the presence of class structure (*sensu* Box 3), but still no environmental fluctuations, the broad line of argument described in Appendix C.1.1–Appendix C.1.2 still applies, but then eq. (A-2) needs to be amended by adding an auxiliary systems of equations to describe the equilibrium allele frequency in each class (see Lion et al., 2022; Bukkuri and Brown, 2023 for work on the stability of oligomorphism in class structured populations).

In the presence of between generation environmental fluctuations, there will no longer be an evolutionary equilibrium in morph values $\mathbf{y}^*$ associated with some deterministic morph frequency equilibrium $\mathbf{p}^*$. This frequency has to be replaced by some stationary ergodic probability distribution μ_{p}^* over morph frequencies. While the selection coefficients for monomorphic and polymorphic population, eq. (A-3), can be extended to such case and used to characterize directional selection, uninvadability, trait-dependent selection, and concergence stability, we are aware of no example where this has been achieved analytically (see Ellner and Sasaki, 1996 for work on a concrete characterization of the stability of an oligomorphism under between

generation environmental fluctuations and Ellner, 1996 for extending this to class structured population).

Appendix D Ecologically mediated interactions

Appendix D.1 Interaction coefficient

We derive here eqs. (14)–(15) of the main text. We first substitute eq. (11) into eq. (6), which after factoring out common terms gives

$$I(y^*) = \left[\frac{\mathrm{d}z_1(y)}{\mathrm{d}y}\frac{\mathrm{d}e_1(y)}{\mathrm{d}y}\left(\frac{\partial^2 W}{\partial z_1 \partial e_1} + \left(\frac{\mathrm{d}z_2(y)}{\mathrm{d}y}/\frac{\mathrm{d}z_1(y)}{\mathrm{d}y}\right)\frac{\partial^2 W}{\partial z_2 \partial e_1}\right) + \frac{\mathrm{d}z_1(y)}{\mathrm{d}y}\frac{\mathrm{d}e_2(y)}{\mathrm{d}y}\left(\frac{\partial^2 W}{\partial z_1 \partial e_2} + \left(\frac{\mathrm{d}z_2(y)}{\mathrm{d}y}/\frac{\mathrm{d}z_1(y)}{\mathrm{d}y}\right)\frac{\partial^2 W}{\partial z_2 \partial e_2}\right)\right]_{x=y=y^*}. \quad \text{(A-11)}$$

Then, using eqs. (12)–(13) to substitute $\left(\frac{\mathrm{d}z_2(y)}{\mathrm{d}y}/\frac{\mathrm{d}z_1(y)}{\mathrm{d}y}\right)\Big|_{x=y=y^*} = -\left(\partial W/\partial z_1\right)/(\partial W/\partial z_2)|_{x=y=y^*}$ and $\left(\frac{\mathrm{d}e_2(y)}{\mathrm{d}y}/\frac{\mathrm{d}e_1(y)}{\mathrm{d}y}\right)\Big|_{x=y=y^*} = -\left(\partial W/\partial e_1\right)/(\partial W/\partial e_2)|_{x=y=y^*}$ produces the desired result by noting that

$$K(y^*) = \left[\left(\frac{\partial^2 W(\mathbf{z},\mathbf{e})}{\partial z_1 \partial e_1} - M_{\mathrm{z}}(y^*)\frac{\partial^2 W(\mathbf{z},\mathbf{e})}{\partial z_2 \partial e_1}\right) - M_{\mathrm{e}}(y^*)\left(\frac{\partial^2 W(\mathbf{z},\mathbf{e})}{\partial z_1 \partial e_2} - M_{\mathrm{z}}(y^*)\frac{\partial^2 W(\mathbf{z},\mathbf{e})}{\partial z_2 \partial e_2}\right)\right]_{x=y=y^*} \quad \text{(A-12)}$$

$$= \left[\left(\frac{\partial}{\partial z_1} - M_{\mathrm{z}}(y^*)\frac{\partial}{\partial z_2}\right)\left(\frac{\partial}{\partial e_1} - M_{\mathrm{e}}(y^*)\frac{\partial}{\partial e_2}\right)W(\mathbf{z},\mathbf{e})\right]_{x=y=y^*}. \quad \text{(A-13)}$$

Appendix D.2 Within generation heterogeneity in production factors

Substituting eq. (17) into eq. (12), shows that any singularity y^* must solve

$$\frac{\mathrm{d}z_1(x)}{\mathrm{d}x}\bigg|_{x=y=y^*} + \frac{(2-m)(1-z_1(y^*))}{z_2(y^*)}\frac{\mathrm{d}z_2(x)}{\mathrm{d}x}\bigg|_{x=y=y^*} = 0, \quad \text{(A-14)}$$

which expresses the trade-off between investing in survival (first term) and reproduction (second term). Using $z_1(x) = x^{1/2}$ and $z_2(x) = (1-x)^{\beta}$ in eq. (A-14) reveals a unique singular level of reproductive effort

$$y^* = \frac{1}{(1-4\beta+2\beta m)^2}, \quad \text{(A-15)}$$

and using eq. (6) produces

$$J(y^*) = -\frac{m\left(1+y^*+2\sqrt{y^*}-2y^*\beta(2-m)\right)}{4\left(\sqrt{y^*}+1\right)^2(y^*)^{3/2}} \quad \text{(A-16)}$$

$$H(y^*) = \frac{m\Big(y^*(1+4\beta(1-m)) - (y^*)^{3/2}\left(4\beta(1+\beta m(2m-3))-1\right) - \sqrt{y^*} - 1\Big)}{4\left(1-\sqrt{y^*}\right)\left(\sqrt{y^*}+1\right)^2(y^*)^{3/2}}.$$

It can be checked that there is a range of parameters values in m and β such that $J(y^*) < 0$ and $H(y^*) > 0$, whereby dimorphism is favored by gradual evolution at y^*. This is for instance the case when $m = 0.5$ and $\beta = 0.7$.

Given branching occurs, let us now consider evolution in a polymorphic population with trait values $\mathcal{P}(\mathbf{p}) = \mathbf{y} = (y_1, y_2, ..., y_n)$ (as described in Appendix C). Generalizing eq. (17) to the polymorphic case, the invasion fitness of type x in polymorphic population is then given by

$$\tilde{\rho}(x, \mathbf{y}, \mathbf{p}) = z_1(x) + (1 - z_1(x))\frac{(1-m)z_2(x)}{(1-m)z_2(x) + m\sum_{i=1}^{n} e_2(y_i)p_i} + m\sum_{i=1}^{n} p_i \frac{(1 - e_1(y_i))z_2(x)}{\sum_{i=1}^{n} e_2(y_i)p_i}. \tag{A-17}$$

Because there are only two regulating factors in this model, let us focus on a dimorphism and applying eq. (A-2), the equilibrium frequency $p_1(y_1, y_2)$ of morph 1 can be computed analytically, but the expression is complicated and not illuminating so we refrain from giving it here. Applying eq. (A-4), and evaluating the selection gradients $S_1(y_1, y_2)$ and $S_2(y_1, y_2)$ on morph y_1 and y_2 using eq. (A-17), we were not able to find analytical expressions for the singular coalition $\mathbf{y}^* = (y_1^*, y_2^*)$. But numerical analysis of the selection gradients shows that one morph is favored to take the value of one, say $y_1^* = 1$, whereby all resources are channelled into survival. At this point, the selection gradient on the second morphs vanishes, i.e. $\lim_{y_2 \to 1} S_1(y_1, y_2) = 0$ (or $\lim_{y_1 \to 1} S_2(y_1, y_2) = 0$), which results in a line of neutral equilibria for the second morph. However, in the evolutionary process with mutations, the exact value of one is never reached, whereby selection becomes stabilizing on the second morph resulting in the coexistence of one (close to) immortal morph and another investing a fraction of resources into reproduction (see Fig. 2).

Appendix D.3 Between generation heterogeneity in production factors

Substituting fitness model eq. (19) with $z_i(x) = e^{-(x-o_i)^2/(2\sigma^2)}$ into eqs. (6)–(7) shows that there is a unique singular strategy

$$y^* = qo_1 + (1-q)o_2, \tag{A-18}$$

and

$$\begin{aligned} J(y^*) &= -\frac{2m(2-m)(1-s)}{\sigma^2} \\ H(y^*) &= \frac{2m(2-m)(1-s)}{\sigma^4}\left(2(1-\rho)\rho(o_1-o_2)^2\left(\frac{(2-m)^2 s - (1-m)^2}{2-m}\right) - \sigma^2\right). \end{aligned} \tag{A-19}$$

This shows that y^* is convergence stable for $s \in [0, 1)$ and that evolutionary branching can occur under a certain range of parameter values. The larger s is, the more likely polymorphism is to emerge, while an decrease in m inhibits branching. We do not attempt to characterize the resulting oligomorphism here, because the presence of between-generation environmental fluctuations make this problem prohibitively challenging (see Ellner and Sasaki, 1996 for the case when $m = 1$).

Appendix E Socially mediated interactions

Appendix E.1 Payoff based gradients

We here derive condition (23). For this, first substitute eq. (21) into the selection gradient of eq. (6) and use the chain rule to see that it can be written in terms of the payoff gradient as

$$S(y) = k(y) \left. \frac{\partial \pi(x,y)}{\partial x} \right|_{x=y}, \tag{A-20}$$

where $k(y) = \partial W(\pi, e(y))/\partial \pi|_{\pi=\pi(y,y)} > 0$ by assumption. Then, any singular strategy y^* satisfying $S(y^*)$ satisfies

$$\left. \frac{\partial \pi(x,y)}{\partial x} \right|_{x=y=y^*} = 0. \tag{A-21}$$

Using this equality after having substituted eq. (21) into the second-order selection coefficients of eqs. (6)–(7), we find that

$$\begin{aligned} H(y^*) &= k(y^*) \left. \frac{\partial^2 \pi(x,y)}{\partial x^2} \right|_{x=y=y^*} \\ I(y^*) &= k(y^*) \left. \frac{\partial^2 \pi(x,y)}{\partial x \partial y} \right|_{x=y=y^*} \\ J(y^*) &= k(y^*) \left[\frac{\partial^2 \pi(x,y)}{\partial x^2} + \frac{\partial^2 \pi(x,y)}{\partial x \partial y} \right]_{x=y=y^*}. \end{aligned} \tag{A-22}$$

Hence, the Hessian, interaction and Jacobian coefficients can all directly be computed from the payoff function. The necessary condition $I(y^*) = k(y^*) I_\pi(y) < 0$ for protected polymorphism can then be reduced to ineq. (23), since $k(y^*) > 0$.

Appendix E.2 Examples: contests and preference evolution

Appendix E.2.1 Contest evolution

Substituting payoff model eq. (29) into eqs. (A-21)-(A-22) and using $a(x) = e^{\omega x}$ along $C(a(x)) = x^2$, we get a unique singular point

$$y^* = \frac{(N-1)v\omega}{2N^2} \tag{A-23}$$

and

$$\begin{aligned} J(y^*) &\propto -2 < 0 \\ H(y^*) &\propto \frac{(N-1)(N-2)v\omega^2}{N^3} - 2. \end{aligned} \tag{A-24}$$

Thus, when $N \leq 2$, the singular trait value is always convergence stable and uninvadable. When $N > 2$, one can find parameter values such that $H(y^*) > 0$. Thus, there is a range of parameter values where a dimorphism can be favored by gradual evolution at y^*.

Given that branching occurs, let us consider evolution in a polymorphic population $\mathcal{P}(\mathbf{p}) = \mathbf{y} = (y_1, y_2, ..., y_n)$ (as described in Appendix C). Assuming that invasions fitness in a monomophic population is $\rho(x, y) = \Pi(a(x), b(y))/\Pi(a(y), b(y))$ with payoff given by eq. (29), i.e., a Wright-Fisher (or lottery) reproductive process (Ewens, 2004), then the invasion fitness of type x in a polymorphic population is

$$\tilde{\rho}(x, \mathbf{y}, \mathbf{p}) = \frac{\left(\frac{a(x)}{a(x)+(N-1)\sum_{i=1}^{n} a(y_i)p_i} v\right) - C(a(x))}{\sum_{j=1}^{n} p_j \left[\left(\frac{a(y_j)}{a(y_j)+(N-1)\sum_{i=1}^{n} a(y_i)p_i} v\right) - C(a(y_j))\right]}. \tag{A-25}$$

Assuming a dimorphism, $n = 2$, the singular coalition(s) can then be evaluated numerically given parameter values for N, ω, and v and checked for convergence stability and locally uninvadability. For instance, when $N = 5$, $\omega = 2$, and $v = 10$, we find that there is unique (strongly) convergence stable and uninvadable symmetric coalition at $(y_1^* = 0.16, y_2^* = 2.4)$ [singular coalitions always occur in reciprocal pairs, Geritz et al., 1999]. These results are confirmed by individual-based simulations (see Fig. 2) and it can be checked that this polymorphism is saturated, i.e., eq. (A-10) is satisfied.

Appendix E.2.2 Preference evolution: dispositions

Here, we analyze the model of dispositions of section 5.3.2 but extend the payoff of this model to

$$\pi(x, y) = \Pi(a(x, y), a(y, x)) - \eta x^4, \tag{A-26}$$

where $a(x, y)$ is given by eq. (34) and the term ηx^4 represents a cost of overconfidence, which increases with x and is controlled by the parameter $\eta \geq 0$. This payoff implements the main text model eqs. (31)–(32) but adds a cost to dispositions. This prevents the open ended evolution of exaggerated trait expression. The right hand side of eq. (A-26) is therefore more general than the right hand side of eq. (24). As a result the first-order condition for model eq. (A-26) is not given directly by eq. (25). Nevertheless the synergy coefficient corresponding to eq. (A-26) is equivalent to that obtained from model eqs. (31)–(32).

Substituting payoff model eq. (A-26) along equilibrium actions (34) into eq. (22) yields a unique feasible singular strategy y^*, which can be expressed as

$$y^* = \frac{\alpha\beta^2}{4\gamma^2 + 2\beta\gamma - \beta^2} + \eta\, h(\alpha, \beta, \gamma, \eta) \tag{A-27}$$

for some complicated function $h(\alpha, \beta, \gamma, \eta)$ of the parameters. When $\eta = 0$ and $\gamma = 1$, y^* reduces to the level of disposition found by Heifetz et al. (2007a, p. 267). The Jacobian and the Hessian coefficients at this

singular point are obtained from eq. (A-22) to read as

$$J(y^*) \propto -\frac{\beta^2 - 2\gamma^2 - 2\gamma(\beta+\gamma)}{(\beta-2\gamma)(\beta+2\gamma)^2} - 12\eta\,(y^*)^2 \tag{A-28}$$
$$H(y^*) \propto \frac{4\gamma\left(\beta^2 - 2\gamma^2\right)}{\left(\beta^2 - 4\gamma^2\right)^2} - 12\eta\,(y^*)^2.$$

Straightforward analysis of these equations reveals that there is a range of parameter values such that $J(y^*) < 0$ and $H(y^*) > 0$ and so evolutionary branching will occur. For instance, this is the case when $\alpha = 0.2$, $\beta = 1.6$, $\gamma = 1$, and $\eta = 0.45$.

Given that branching occurs, let us consider evolution in a polymorphic population $\mathcal{P}(\mathbf{p}) = \mathbf{y} = (y_1, y_2, ..., y_n)$. Assuming a Wright-Fisher process, the invasion fitness of type x in a polymorphic population is

$$\tilde{\rho}(x, \mathbf{y}, \mathbf{p}) = \frac{\sum_{i=1}^{n}\left[(\alpha - \beta a(y_i, x) - \gamma a(x, y_i))a(x, y_i) - \eta x^{\nu}\right] p_i}{\sum_{j=1}^{n} p_j \left[\sum_{i=1}^{n}\left[(\alpha - \beta a(y_i, y_j) - \gamma a(y_j, y_i))a(y_j, y_i) - \eta y_j^{\nu}\right] p_i\right]}. \tag{A-29}$$

Focusing on a dimorphism, $n = 2$, and solving $\tilde{\rho}(y_1, \mathbf{y}, \mathbf{p}) = \tilde{\rho}(y_2, \mathbf{y}, \mathbf{p})$ for p_1 by setting $p_2 = 1 - p_1$, we find a unique complicated analytical expression for $p_1(y_1, y_2)$. Nevertheless, this expression allows us to compute the selection gradients $S_1(\mathbf{y})$ and $S_2(\mathbf{y})$ analytically. By solving $S_1(\mathbf{y}^*) = 0$ and $S_2(\mathbf{y}^*) = 0$ for $\mathbf{y}^* = (y_1^*, y_2^*)$, we find a single singular coalition such that

$$y_1^* = -\frac{\sqrt{\gamma}\sqrt{\beta^2 - 2\gamma^2}}{\sqrt{\eta}\,(\beta^2 - 4\gamma^2)} \qquad \text{and} \qquad y_2^* = \frac{\sqrt{\gamma}\sqrt{\beta^2 - 2\gamma^2}}{\sqrt{\eta}\,(\beta^2 - 4\gamma^2)}, \tag{A-30}$$

and that leads us to

$$p_1^* = \frac{\frac{\alpha\beta\sqrt{\eta}\left(\beta^2 - 4\gamma^2\right)}{\sqrt{\gamma}\sqrt{\beta^2 - 2\gamma^2}} - \frac{2\alpha\sqrt{\gamma}\sqrt{\eta}\left(\beta^2 - 4\gamma^2\right)}{\sqrt{\beta^2 - 2\gamma^2}} + \beta}{2\beta}. \tag{A-31}$$

Computing the Hessian for each morph (eq. 7), we find that

$$H_1(\mathbf{y}^*) = H_2(\mathbf{y}^*) = -\frac{8\gamma\eta\left(\beta^2 - 4\gamma^2\right)^2\left(\beta^2 - 2\gamma^2\right)}{\gamma\left(2\gamma\left(-3\beta^2\gamma^2 + \beta^4 + 2\gamma^4\right) - \alpha^2\eta(\beta - 2\gamma)^4(\beta + 2\gamma)^2\right)}. \tag{A-32}$$

Computing the Jacobian matrix $\mathbf{J}(\mathbf{y}^*)$ (eq. A-8) shows that its two eigenvalues take the same value as the Hessian coefficients eq. (A-32). Hence, whenever the coalition is locally uninvadable, it is also strongly convergence stable and the polymorphism is saturated, i.e. eq. (A-10) is satisfied. This is the case for instance with the aforementioned parameter values ($\alpha = 0.2$, $\beta = 1.6$, $\gamma = 1$, and $\eta = 0.45$), which leads to a symmetric uninvadable coalitions with trait values ($y_1^* = 0.77, y_2^* = -0.77$), which is confirmed by individual-based simulations (see Fig. 2).

Appendix E.2.3 Preference evolution: other-regarding motivations

Here, we consider an alternative model of preference, in which the evolving trait x measures an individual's other-regard. We assume that the utility functions for a pair of interacting individual, one mutant with action a and trait x and a resident with action b and trait y, are given by

$$\begin{aligned} u_x(a,b) &= \Pi(a,b) + x\Pi(b,a) \\ u_y(b,a) &= \Pi(b,a) + y\Pi(a,b). \end{aligned} \tag{A-33}$$

Each utility is the sum of own material payoff to the actor plus the payoff to the interaction partner and thus represents an other-regarding preference. The weight of other-regard is the evolving traits. To close the model, we postulate the linear-quadratic payoff eq. (32) of the main text. The necessary first-conditions for a Nash equilibrium under this game are that $\partial u_x(a,b)/\partial a = 0$ and $\partial u_y(b,a)/\partial b = 0$, which can be expressed as

$$\begin{aligned} a &= \frac{\alpha}{2\gamma} - b\frac{\beta(1+x)}{2\gamma} \\ b &= \frac{\alpha}{2\gamma} - a\frac{\beta(1+y)}{2\gamma}. \end{aligned} \tag{A-34}$$

Solving for the equilibrium actions $a = a(x,y)$ and $b = a(y,x)$ yields

$$a(x,y) = \frac{\alpha(\beta - 2\gamma + \beta x)}{\beta^2(x+1)(y+1) - 4\gamma^2}. \tag{A-35}$$

This is indeed a Nash equilibrium action since the solution is unique and satisfies the second-order condition $\partial^2 u_x(a,b)/\partial a^2 = -2\gamma \leq 0$ and $\partial^2 u_y(b,a)/\partial b^2 = -2\gamma \leq 0$.

For this model, the corresponding singular strategy, response coefficient and action interaction coefficients are

$$y^* = \lambda(y^*) = -\frac{\beta}{\beta + 2\gamma} \quad \text{and} \quad A(y^*) = \frac{8\gamma^2(\beta+\gamma)}{\alpha(\beta+2\gamma)^2}. \tag{A-36}$$

Since $\beta > 0$ and $\gamma > 0$, the singular strategy is always convergence stable and $\partial^2\Pi(a,b)/(\partial a \partial b) < 0$, $\partial\Pi(a,b)/(\partial b) < 0$, $[\partial^2\Pi(a,b)/\partial a^2 + \partial^2\Pi(a,b)/\partial b^2] = -2\gamma < 0$, and $\partial\Pi/\partial a + \partial\Pi/\partial b = -\alpha\beta/2\gamma < 0$. This entails that the two component of eq. (27) result in negative trait-dependent selection:

$$K_\pi(y^*) = -\frac{2\beta^2(\beta+\gamma)}{(\beta+2\gamma)^2} \qquad K_\lambda(y^*) = -\frac{4\beta\gamma(\beta+\gamma)}{(\beta+2\gamma)^2} \tag{A-37}$$

Polymorphism in other-regard can thus evolve and it can be checked that there is a range of parameter values where evolutionary branching occurs, leading to the open-ended evolution of a positive $x > 0$ and negative $x < 0$ other-regarding morph unless a cost is added to the expression of the trait.

Appendix F Genetically mediated interactions

Appendix F.1 No effect of diploidy on the instability of monomorphism

We here show eq. (37). For this, first substitute eq. (36) into the selection gradient of eq. (6) and use the chain rule to show that it can be written as

$$S(y) = \tilde{k}(y) \left. \frac{\partial W(z(x,y), z(y,y))}{\partial z(x,y)} \right|_{x=y}, \tag{A-38}$$

where $\tilde{k}(y) = \partial z(x,y)/\partial x|_{x=y} \neq 0$ by assumption. Thus, any singular strategy y^* satisfying $S(y^*) = 0$ has to satisfy

$$\left. \frac{\partial W(z(x,y), z(y,y))}{\partial z(x,y)} \right|_{x=y=y^*} = 0. \tag{A-39}$$

Using this equality along $\partial z(x,y)/\partial x|_{x=y} = \partial z(x,y)/\partial y|_{x=y}$ by substituting eq. (36) into the second-order coefficients of eqs. (6)–(7), we find that

$$J(y^*) = \tilde{k}(y^*)^2 \left[\frac{\partial^2 W(z(x,y), z(y,y))}{\partial z(x,y)^2} + \frac{\partial^2 W(z(x,y), z(y,y))}{\partial z(x,y) \partial z(y,y)} \right]_{x=y=y^*} \tag{A-40a}$$

$$H(y^*) = \tilde{k}(y^*)^2 \left. \frac{\partial^2 W(z(x,y), z(y,y))}{\partial z(x,y)^2} \right|_{x=y=y^*}. \tag{A-40b}$$

Hence, for $J(y^*) < 0$ and $H(y^*) > 0$ to hold, requires that

$$\left. \frac{\partial^2 W(z(x,y), z(y,y))}{\partial z(x,y) \partial z(y,y)} \right|_{x=y=y^*} < 0, \tag{A-41}$$

i.e. there is NTDS at the phenotypic level.

Because in eqs. (A-39)–(A-41) all derivatives are phenotypic and these phenotypes are evaluated at $z(y^*, y^*) = z^*$, the allelic value y^* is in fact not needed to perform the local evolutionary analysis, which can be carried out by focusing only on the phenotype value z^* solving

$$\left. \frac{\partial W(z, \bar{z})}{\partial z} \right|_{z=\bar{z}=z^*} = 0, \tag{A-42}$$

and this leads to condition (37) of the main text.

The above shows that under our assumptions, diploidy does not affect the conditions for the gradual emergence of polymorphism for evolutionary branching. However, diploidy does affect the conditions for mutual invasibility of alleles x and y. From eq. (A-1) the mutual invasibility of two weakly and symmetrically diverged

allelic values around y^* requires that

$$\left.\frac{\partial^2 W(z(x,y),z(y,y))}{\partial x \partial y}\right|_{x=y=y^*} < 0, \tag{A-43}$$

which holds if and only if

$$2\left.\frac{\partial^2 W(z(x,y),z(y,y))}{\partial z(x,y)\partial z(y,y)}\right|_{x=y=y^*} + \left.\frac{\partial^2 W(z(x,y),z(y,y))}{\partial z(x,y)^2}\right|_{x=y=y^*} < 0. \tag{A-44}$$

This, in particular, captures the notion that with diploidy, two alleles on either side of a local fitness peak (i.e., either side of a singular value y^* such that $\partial^2 W(z(x,y),z(y,y))/(\partial z(x,y)\partial z(y,y)) = 0$ and $\partial^2 W(z(x,y),z(y,y))/\partial z(x,y)^2 < 0$ at $x = y = y^*$) can be maintained owing to overdominance: i.e. where the intermediate heterozygote expresses a fitter phenotype. However, such polymorphism is necessarily short-lived on an evolutionary timescale as any allele that codes for the optimal phenotype in homozygotic form will invade and fix.

Appendix F.2 Average heterozygote advantage

We here detail an example leading to heterozygote advantage through temporal averaging over different contexts by considering a scenario envisaged by Gillespie (1978), where the environment changes each generation, being in state 1 with probability q and in state 2 with probability $1-q$ such that invasion fitness is given by

$$\rho(x,y) = \left(\frac{z_1(x,y)}{z_1(y,y)}\right)^q \left(\frac{z_2(x,y)}{z_2(y,y)}\right)^{1-q}, \tag{A-45}$$

which takes the form of eq. (39) and (40) with $\Pi(z_1,z_2) = z_1^q z_2^{1-q}$ and $e(y) = z_1(y,y)^q z_2(y,y)^{1-q}$ under the assumption of a Wright-Fisher process: $\rho(x,y) = \pi(x,y)/\pi(y,y)$ where $\pi(y,y) = e(y)$. The interpretation is that z_1 and z_2 are the physiological activity of a gene in environments 1 and 2, respectively; and we assume that activity increases with allelic value in environment 1, say $g_1(y) = y^{u_1}$, and decreases with allelic value in environment 2, say $g_2(y) = (1-y)^{u_2}$ (with $0 < y < 1$).

Using $\Pi(z_1,z_2) = z_1^q z_2^{1-q}$ in eq. (40) when substituting eq. (41) along $g_i(x) = e^{-\omega(x-o_i)^2}$ and eq. (39) into eqs. (6)–(7) shows that there is a unique singular strategy

$$y^* = \frac{qu_1}{qu_1 + (1-q)u_2} \tag{A-46}$$

and that

$$K_\pi(y^*) = -\frac{(qu_1 + (1-q)u_2)^2}{4(1-q)q} < 0, \tag{A-47}$$

which displays NTDS. Further,

$$J(y^*) - \frac{(q(u_1-u_2)+u_2)^3}{2q(1-q)u_1u_2} \tag{A-48}$$
$$H(y^*) \propto \frac{q\left((u_1-u_2)+u_2\right)^2\left(2qu_1+u_2\left(2-2q-u_1\right)\right)}{2q(1-q)u_1u_2}.$$

Thereby, provided $u_i > 2$ and $u_j/(2u_i) > (1-q)/(u_i-2q)$ (for $i \neq j$), evolutionary branching will occur and the coexistence of two morphs: one that increases activity in environment 1 and another in environment 2, leading to heterozygote advantage on average over the two environments. See also Siljestam and Rueffler (2024) for a conceptually similar model that leads to poly-allelic polymorphism when there are more than two biological functions (or environments) over which to average.

Appendix F.3 Trait-dependent kin selection: social-dispersal syndrome

As an example of relatedness modulated NTDS, let us consider a situation where the trait constitutes a form of helping that increases the fecundity of one's neighbour, and that its costs are over dispersal such that an individual who invests more resources to help one's neighbour has fewer resources to commit to its dispersing offspring. Trait expression therefore has positive indirect fitness benefits (via fecundity) and positive effects on relatedness (via dispersal), especially when rare. For simplicity, we assume a Moran type life cycle whereby exactly one individual dies per life-cycle iteration in each patch, such that individual fitness takes the form

$$w(x,x_\mathrm{n},y) = \frac{1}{2} + \frac{1}{2}\left[\frac{(1-z_1(x))z_2(x_\mathrm{n})}{\left[(1-z_1(x))z_2(x_\mathrm{n}) + (1-z_1(x_\mathrm{n}))z_2(x)\right]/2 + z_1(y)z_2(y)} + \frac{z_1(x)z_2(x_\mathrm{n})}{z_2(y)}\right], \tag{A-49}$$

where $z_1(x)$ is the probability of dispersal of an individual expressing trait value x (which we assume is a decreasing function) and $z_2(x_\mathrm{n})$ is the fecundity of an individual whose neighbour expresses trait x_n (an increasing function of x_n). Therefore, the first term within square bracket in eq. (A-49) is the philopatric component of fitness, i.e. the expected number of offspring that establish in the natal patch, and the second term is the dispersal component, i.e. the expected number of offspring that establish in other patches. Using Box 2 of Alger and Lehmann, 2023, we obtain that relatedness for this model can be expressed as

$$r(x,y) = \frac{(1-z_1(x))\Big((1-z_1(x))z_2(x) + z_1(y)z_2(y)\Big)}{z_2(x)\Big((1-z_1(x))^2 + z_1(y) - z_1(y)^2\Big) + 2z_1(y)z_2(y)\Big(1-z_1(x)+z_1(y)\Big)}. \tag{A-50}$$

Strictly speaking, this is a relatedness proxy, whereby the right-hand side of eq. (44) becomes the basic reproductive number of the mutant, which can be used to predict evolutionary outcomes, since the basic reproductive number is sign equivalent to invasion fitness, e.g., Metz et al. (2008) and see Mullon et al. (2016) for these connections under limited dispersal.

Analysis of this model does not give rise to a simple expression for the interaction coefficient $I(y^*)$ from

which useful information can be extracted, but a straightforward numerical approach with $z_1(x) = 1 - x$ and $z_2(x_\mathrm{n}) = 3x_\mathrm{n}^2 + 0.3x_\mathrm{n} + 1.05$ shows that at $y^* = 0.38$, we have $I(y^*) < 0$ while $I_\mathrm{q}(y^*) < 0 < I_\mathrm{w}(y^*)$, i.e. the population experiences NTDS due to relatedness effects only. Additionally, the population undergoes evolutionary branching leading to the coexistence of two differentiated morphs that show a social-dispersal syndrome: (i) one that helps and thus tends to remain philopatric (expressing $x \approx 0.65$); and (ii) another that does not help and disperses (expressing $x = 0$). Trait-dependent selection can be understood by considering the mutant and resident population both being more selfish and thus more prone to dispersal simultaneously. Such a mutant benefits by dispersing away and thereby avoiding interactions with other mutants. However, if the resident population simultaneously becomes more selfish and dispersal-prone, these benefits vanish because resident individuals help less anyway and also disperse. In that case, both mutant and resident lineages experience the same reduction in relatedness, so the mutant does not differ in its ability to interact preferentially with relatives. A handful models have reported the evolution of similar social-dispersal syndromes owing to kin selection (Mullon et al., 2018; Prigent and Mullon, 2023) but beyond these, there remain few analyses of how effects on relatedness can generate NTDS.

Appendix G Class structured populations

Appendix G.1 Within generation heterogeneity

Here, we show eq. (3.E) in Box 3. To that end, we first note that eq. (3.A) of that box can be written as

$$\rho(x,y) = 1 + \frac{L(x,y)}{V(x,y)}, \tag{A-51}$$

with

$$L(x,y) = \sum_{c'\in\mathcal{C}}\sum_{c\in\mathcal{C}} v_{c'}(y)\left[w_{c'c}(x,y) - w_{c'c}(y,y)\right] q_c(x,y) \tag{A-52}$$

$$V(x,y) = \sum_{c'\in\mathcal{C}}\sum_{c\in\mathcal{C}} v_{c'}(y) w_{c'c}(y,y) q_c(x,y), \tag{A-53}$$

(see Ohtsuki et al., 2020 for a proof), where $v_c(y)$ is the (neutral) reproductive value of an individual in class c in a resident population y defined to satisfy the recursion $v_c(y) = \sum_{c'\in\mathcal{C}} v_{c'}(y) w_{c'c}(y,y)$ (e.g., Taylor, 1990; Rousset, 2004). For completeness, note that one also has $q_{c'}(x,y)\rho(x,y) = \sum_{c\in\mathcal{C}} w_{c'c}(x,y) q_c(x,y)$, whereby $\rho(x,y)$ is the leading eigenvalue of the matrix with elements $w_{c'c}(x,y)$ that has leading right eigenvector with elements $q_c(x,y)$ [and summing the equation over all $c' \in \mathcal{C}$ produces eq. 3.A, see Lehmann et al., 2016; Ohtsuki et al., 2020 for more details on invasion fitness derivation and the interpretations of the weights $q_c(x,y)$].

The choice of scaling fitness effects with reproductive values ensures that under neutrality ($x = y$), one immediately has $V(y, y) = 1$ (as the term in squared brackets in eq. A-52 is zero and thus $L(y, y) = 0$, Ohtsuki et al., 2020). Using these equalities after differentiating $\rho(x, y)$ as given by eq. (A-51) with respect to both x and y, we obtain,

$$\frac{\partial^2 \rho(x,y)}{\partial x \partial y} = \frac{\partial^2 L(x,y)}{\partial x \partial y} - \frac{\partial L(x,y)}{\partial x}\frac{\partial V(x,y)}{\partial y} - \frac{\partial V(x,y)}{\partial x}\frac{\partial L(x,y)}{\partial y}. \tag{A-54}$$

Now note that $S(y) = \partial L(x, y)/\partial x$ and that by total differentiating $\rho(y, y)$ with respect to y one obtains that $\partial L(x, y)/\partial x = -\partial L(x, y)/\partial y$. Since at a singular point $S(y^*) = 0$, we have

$$I(y^*) = \left.\frac{\partial^2 \rho(x,y)}{\partial x \partial y}\right|_{x=y=y^*} = \left.\frac{\partial^2 L(x,y)}{\partial x \partial y}\right|_{x=y=y^*}. \tag{A-55}$$

Using eq. (A-52) for $L(x, y)$ we thus obtain

$$I(y^*) = \sum_{c' \in \mathcal{C}} \sum_{c \in \mathcal{C}} \left[\left(v_{c'}(y) \frac{\partial^2 w_{c'c}(x,y)}{\partial x \partial y} + \frac{\mathrm{d}v_{c'}(y)}{\mathrm{d}y} \frac{\partial w_{c'c}(x,y)}{\partial x} \right) q_c(y,y) \right.$$
$$\left. + v_{c'}(y) \frac{\partial w_{c'c}(x,y)}{\partial x} \left(\frac{\partial q_c(x,y)}{\partial y} - \frac{\partial q_c(x,y)}{\partial x} \right) \right]_{x=y=y^*}, \tag{A-56}$$

where the term $[-\partial w_{c'c}(x, y)/\partial x \times \partial q_c(x, y)/\partial x]$ comes from differentiation of $[-w_{c'c}(y, y)q_c(x, y)]$ in eq. (A-52) (i.e. due to scaling of fitness). Finally, by factoring out $q_c(y, y)$ and $v_{c'}(y)$ and using the property of the logarithm function, we obtain eq. (3.E) of the main text.

Appendix G.2 Between generation heterogeneity

We here show eq. (3.D) and eq. (3.F) of Box 3. For this, it is useful to express invasion fitness as $\rho(x, y) = e^{\phi(x,y)}$ where $\phi(x, y)$ is the invasion exponent of the mutant x in a resident y population (e.g. Tuljapurkar, 1990; Ferrière and Gatto, 1995). Owing to the fact that $a(y, y) = 0$, we obtain by direct differentiation

$$S(y) = \left.\frac{\partial \phi(x,y)}{\partial x}\right|_{x=y} \tag{A-57}$$

$$I(y) = S(y)^2 + \left.\frac{\partial^2 \phi(x,y)}{\partial x \partial y}\right|_{x=y} \tag{A-58}$$

and at a singularity $I(y^*) = \partial^2 \phi(x, y)/\partial x \partial y\big|_{x=y=y^*}$ since $S(y^*) = 0$. Thus, we can compute all derivative of invasion fitness directly from all corresponding derivatives of the growth rate, which facilitates calculations in what follows.

Substituting eq. (3.B) into $\phi(x, y) = \log(\rho(x, y))$, we get upon rearrangement

$$\phi(x,y) = \sum_{c \in \mathcal{C}} \log(w_c(x,y)) q_c(y), \tag{A-59}$$

which is the familiar expression for the mutant growth rate under between generation heterogeneity. Differentiating, we get

$$\left.\frac{\partial \phi(x,y)}{\partial x}\right|_{x=y} = \sum_{c\in\mathcal{C}} \frac{1}{w_c(y,y)} \left.\frac{\partial w_c(x,y)}{\partial x}\right|_{x=y} q_c(y) \tag{A-60}$$

and

$$\left.\frac{\partial^2 \phi(x,y)}{\partial x \partial y}\right|_{x=y} = \sum_{c\in\mathcal{C}} \left[\frac{1}{w_c(y,y)} \frac{\partial^2 w_c(x,y)}{\partial x \partial y} q_c(y) - \frac{1}{w_c(y,y)^2} \frac{\partial w_c(x,y)}{\partial x} \frac{\partial w_c(x,y)}{\partial y} q_c(y) \right.$$
$$\left. + \frac{1}{w_c(y,y)} \frac{\partial w_c(x,y)}{\partial x} \frac{\partial q_c(y)}{\partial y} \right]_{x=y}, \tag{A-61}$$

which, after re-arrangements by factoring out $q_c(y)$ and using the property of logarithm gives eq. (3.F).

Appendix G.3 Class-specific regulation and production factors

Following the same construction of fitness factors as in section 3.2, we can also express $w_c(x,y)$ and $q_c(x,y)$ in terms of fitness factors as follows by writing

$$w_{c',c}(x,y) = \tilde{w}(\boldsymbol{\pi}_c(\mathbf{z}_c(x, \mathbf{e}^{\mathrm{z}}_c(x,y)), \mathbf{e}^{\pi}_c(x,y)), \mathbf{e}^{\mathrm{w}}_c(x,y)), \tag{A-62}$$

where we have indexed by c the production and regulating factors by the class of the bearer of this fitness, since in class structured population they will be class specific as both the phenotypes expressed by an individual and its experienced environment may depend on the class in which the individual resides. The right hand side of eq. (A-62) can also be applied to fitness $w_c(x,y)$ for temporal heterogeneity.

The distribution $q_c(x,y)$ represents the likelihood of being in a certain state. Since the occurrence of any state can depend on all factors affecting reproduction and survival across all states, we generally have

$$q_c(x,y) = \tilde{q}(\boldsymbol{\pi}(\mathbf{z}(x, \mathbf{e}^{\mathrm{z}}(x,y)), \mathbf{e}^{\pi}(x,y)), \mathbf{e}^{\mathrm{w}}(x,y)), \tag{A-63}$$

for some function $\tilde{q}$ mapping the collection of all fitness factors to probabilities of state occurrence, and where a vector $\mathbf{v}$ without subscript (where $\mathbf{v}$ could be $\boldsymbol{\pi}$, $\mathbf{z}$, $\mathbf{e}^{\mathrm{z}}$, $\mathbf{e}^{\pi}$, or $\mathbf{e}^{\mathrm{w}}$) concatenates the vectors with subscript, i.e. $\mathbf{v} = \{\mathbf{v}_c\}_{c\in\mathcal{C}}$. By substituting eqs. (A-62)–(A-63) into eq. (3.A), we can express invasion fitness as in eq. (10), i.e., as a mapping from all fitness factors to geometric growth ratio (and the same logic applies to $w_c(x,y)$ and $q_c(y)$ under temporal heterogeneity).

References

Ajar, E. 2003. Analysis of disruptive selection in subdivided populations. *BMC Evolutionary Biology* 3:22.

Akçay, E. and J. Van Cleve. 2012. Behavioral responses in structured populations pave the way to group optimality. *American Naturalist* 179:257–269.

Akçay, E. and J. V. Cleve. 2014. Pathways to social evolution: reciprocity, relatedness, and synergy. *Evolution* 68:2245–58.

Alger, I. 2023. Evolutionarily stable preferences. *Proceedings of the Royal Society of London Series B-Biological Sciences* 378.

Alger, I. and L. Lehmann. 2023. Evolution of semi-Kantian preferences in two-player assortative interactions with complete and incomplete Iiformation and plasticity. *Dynamic Games and Applications* 13:1288–1319.

Alipantris, C. D. and K. C. Border. 2006. Infinite Dimensional Analysis: a Hitchiker's Guide. Springer, Berlin, 3th edn.

Altenberg, L. and M. W. Feldman. 1987. Selection, generalized transmission and the evolution of modifier genes. I. The reduction principle. *Genetics* 117:559–572.

Altenberg, L., U. Liberman, and M. W. Feldman. 2017. Unified reduction principle for the evolution of mutation, migration, and recombination. *Proceedings of the National Academy of Sciences of the United States of America* 114:E2392–E2400.

Anderson, B. G., , S. Sojoudi, and M. Arcak. 2013. Evolutionary games on infinite strategy sets: convergence to Nash equilibria via dissipativity. arXiv:2312.08286v2.

André, J. B. and T. Day. 2007. Perfect reciprocity is the only evolutionarily stable strategy in the continuous iterated prisoner's dilemma. *Journal of Theoretical Biology* 247:11–22.

Apaloo, J. and S. Butler. 2009. Evolutionary stabilities in multidimensional-traits and several-species models. *Evolutionary Ecology Research* 11:637–650.

Armstrong, R. A. and R. McGehee. 1980. Competitive exclusion. *The American Naturalist* 115:151–170.

Arnold, L., V. M. Gundlach, and L. Demetrius. 1994. Evolutionary formalism for products of positive random matrices. *The Annals of Applied Probability* 4:859–901.

Arnold, S. J. 2023. Evolutionary Quantitative Genetics. Oxford University Press, Oxford, UK.

Avila, P. and C. Mullon. 2023. Evolutionary game theory and the adaptive dynamics approach: adaptation where individuals interact. *Philosophical Transactions of the Royal Society B* 378:615255.

Begon, M. and C. R. Townsend. 2021. Ecology: from Individuals to Ecosystems. Wiley, West Sussex, 5th edn.

Brännström, A., J. Johansson, and N. Von Festenberg. 2013. The hitchhiker's guide to adaptive dynamics. *Games* 4:304–328.

Briat, C. 2017. Sign properties of Metzler matrices with applications. *Linear Algebra and its Applications* 515:53–86.

Bukkuri, A. and J. S. Brown. 2023. Integrating eco- evolutionary dynamics into matrix population models for

structured populations: Discrete and continuous frameworks. *Methods in Ecology and Evolution* 14:1475–1488.

Bulow, J., J. Geanakoplos, , and P. Klemperer. 1985. Multimarket oligopoly: strategic substitutes and strategic complements. *Journal of Political Economy* 93:488–511.

Bürger, R. 2000. The Mathematical Theory of Selection, Recombination, and Mutation. John Wiley and Sons, New York.

Cai, Y. and S. A. H. Geritz. 2020. Resident-invader dynamics of similar strategies in fluctuating environments. *Mathematical Biology* 81:907–959.

Cant, M. 2012. Suppression of social conflict and evolutionary transitions to cooperation. *American Naturalist* 179:293–301.

Case, T. J. 2000. An Illustrated Guide to Theoretical Ecology. Oxford University Press, Oxford.

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

Charnov, E. L. 1993. Life History Invariants. Oxford University Press, Oxford, UK.

Charnov, E. L. 1997. Trade-off-invariant rules for evolutionarily stable life histories. *Nature* 387:393–4.

Chesson, P. L. 2000a. General theory of competitive coexistence in spatially-varying environments. *Theoretical Population Biology* 58:211–237.

Chesson, P. L. 2000b. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31:343–66.

Chesson, P. L. and R. R. Warner. 1981. Environmental variability promotes coexistence in lottery competitive systems. *American Naturalist* 117:923–943.

Christiansen, F. B. 1991. On conditions for evolutionary stability for a continuously varying character. *American Naturalist* 138:37–50.

Conitzer, V. and T. Sandholm. 2008. New complexity results about Nash equilibria. *Games and Economic Behavior* 63:621–641.

Crow, J. F. and M. Kimura. 1970. An Introduction to Population Genetics Theory. Harper and Row, New York.

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

de Mazancourt, C. and U. Dieckmann. 2004. Trade-off geometries and frequency-dependent selection. *The American Naturalist* 164:765–778.

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

Diekmann, O., M. Gyllenberg, and J. A. J. Metz. 2003. Steady-state analysis of structured population models. *Theoretical Population Biology* 63:309–338.

Doebeli, M., C. Hauert, and T. Killingback. 2004. The evolutionary origin of cooperators and defectors. *science* 306:859–862.

Ellner, S. 1996. Environmental fluctuations and the maintenance of genetic diversity in age or stage-structured populations. *Journal of Mathematical Biology* 58:103–127.

Ellner, S. and A. Sasaki. 1996. Patterns of genetic polymorphism maintained by fluctuating selection with overlapping generations. *Theoretical Population Biology* 50:31–65.

Eshel, I. 1983. Evolutionary and continuous stability. *Journal ot 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. *The American Naturalist* 124:631–640.

Eshel, I., U. Motro, and E. Sansone. 1997. Continuous stability and evolutionary convergence. *Journal of Theoretical Biology* 074:222–232.

Ewens, W. J. 2004. Mathematical Population Genetics. Springer-Verlag, New York.

Fairbairn, D. J. 2013. Odd couples: extraordinary differences between the sexes in the animal kingdom. Princeton University Press.

Ferrière, R. and M. Gatto. 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theoretical Population Biology* 48:126–171.

Fisher, R. A. 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Frank, S. A. 1998. Foundations of Social Evolution. Princeton University Press, Princeton, NJ.

Fudenberg, D. and J. Tirole. 1991. Game Theory. MIT Press, Massachusetts.

Geritz, S. A. H., E. Kisdi, G. Meszéna, and J. A. J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology* 12:35–57.

Geritz, S. A. H., E. van der Meijden, and J. A. J. Metz. 1999. Evolutionary dynamics of seed size and seedling competitive ability. *Theoretical Population Biology* 55:324–43.

Gigord, L. D., M. Macnair, and A. Smithson. 2001. Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid Dactylorhiza sambucina (L.) Soò. *Proceedings of the National Academy of Sciences* 98:6253–6255.

Gillespie, J. H. 1978. A general model to account for enzyme variation in natural populations. V. The SAS-CFF model. *Theoretical population biology* 14:1–45.

Gillespie, J. H. 1991. The Causes of Molecular Evolution. Oxford University Press, Oxford.

Gillespie, J. H. 2004. Population Genetics: a Concise Guide. Johns Hopkins University Press, Baltimore, Maryland.

Gross, M. R. 1996. Alternative reproductive strategies and tactics: diversity within sexes. *Trends in Ecology & Evolution* 11:92–98.

Grovers, J. P. 1997. Resource Competition. Chapman and Hall, New York.

Gyllenberg, M. and G. Meszéna. 2005. On the impossibility of coexistence of infinitely many strategies. *Mathematical Biology* 50:133–160.

Haigh, J. and J. Maynard Smith. 1972. Can there be more predators than prey? *Theoretical Population Biology* 3:290–299.

Hamilton, W. D. 1964. The genetical evolution of social behaviour, 1. *Journal of Theoretical Biology* 7:1–16.

Harris, T. E. 1963. The Theory of Branching Processes. Springer, Berlin.

Hartl, D. and A. G. Clark. 2007. Principles of Population Genetics. Sinauer, Massachusetts, 4th edn.

Heifetz, A., C. Shannon, and Y. Spiegel. 2007a. The dynamic evolution of preferences. *Economic Theory* 32:251–286.

Heifetz, A., C. Shannon, and Y. Spiegel. 2007b. What to maximize if you must. *Journal of Economic Theory* 133:31 – 57.

Heino, M., J. A. J. Metz, and V. Kaitala. 1998. The enigma of frequency-dependent selection. *Trends in Ecology and Evolution* pp. 367–370.

Hingu, D., K. S. Mallikarjuna Rao, and A. J. Shaiju. 2008. Evolutionary stability of polymorphic population states in continuous games. *Dynamic Games and Applications* 8:141–156.

Hirsch, M. W., S. Smale, and R. L. Devaney. 2004. Differential Equations, Dynamical Systems, and an Introduction to Chaos. Academic Press, San Diego, CA.

Hirshleifer, J. 1987. Conflict and rent-seeking success functions: ratio vs. difference models of relative success. *Public Choice* 63:101–112.

Hofbauer, J. and K. Sigmund. 1998. Evolutionary Games and Population Dynamics. Cambridge University Press.

Houston, A. I. and J. M. McNamara. 1999. Models of Adaptive Behaviour: An Approach Based on State. Cambridge University Press, Cambridge, UK.

Ito, H. and U. Dieckmann. 2007. A new mechanism for recurrent adaptive radiations. *The American Naturalist* 170:E96–E111.

Ito, H. and A. Sasaki. 2016. Evolutionary branching under multi-dimensional evolutionary constraints. *Journal of Theoretical Biology* 407:409–428.

Jehle, A. G. and P. J. Reny. 2011. Advanced Microeconomics Theory. Prentice Hall, Essex.

Karlin, S. and J. McGregor. 1972. Polymorphisms for genetic and ecological systems with weak coupling. *Theoretical Population Biology* 3:210–38.

Kisdi, E. 2006. Trade-off geometries and the adaptive dynamics of two co-evolving species. *Evolutionary Ecology Research* 8:959–973.

Kisdi, E. 2015. Construction of multiple trade-offs to obtain arbitrary singularities of adaptive dynamics. *Journal of Mathematical Biology* 70:1093–1117.

Kisdi, E. and S. A. H. Geritz. 1999. Adaptive dynamics in allele space: evolution of genetic polymorphism by small mutations in a heterogeneous environment. *Evolution* 53:993–1008.

Kisdi, E. and S. A. H. Geritz. 2016. Adaptive dynamics of saturated polymorphisms. *Journal of Mathematical Biology* 72:1039–1079.

Kondepuli, D. and I. Prigogine. 2015. Modern Thermodynamics: From Heat Engines to Dissipative Structures. Wiley, New York, second edn.

Křivan, V., R. Cressman, and C. Schneider. 2008. The ideal free distribution: A review and synthesis of the game-theoretic perspective. *Theoretical Population Biology* 73:403–425.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34:292–305.

Lehmann, L. 2022. Hamilton's rule, the evolution of behavior rules and the wizardry of control theory. *Journal of Theoretical Biology* 55.

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.

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

Leslie, P. H. and J. C. Gower. 1958. The properties of a stochastic model for two competing species. *Biometrika* 45:316–330.

Leslie, P. H. and J. C. Gower. 1960. The properties of a stochastic model for the predator-prey type of interaction between two species. *Biometrika* 47:219–234.

Lessard, S. 1990. Evolutionary stability: one concept, several meanings. *Theoretical Population Biology* 37:159–170.

Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. *The American Naturalist* 87:331–333.

Levin, S. A. 1970. Community equilibria and stability, and an extension of the competitive exclusion principle. *The American Naturalist* pp. 413–423.

Lewontin, R. C. 1974. The Genetic Basis of Evolutionary Change. Columbia University Press.

Liberman, U. 1988. External stability and ESS: criteria for initial increase of a new mutant allele. *Journal of Mathematical Biology* 26:477–485.

Lion, S. 2018. Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. *American Naturalist* 191:21–44.

Lion, S., M. Boots, and A. Sasaki. 2022. Multimorph eco-evolutionary dynamics in structured populations in structured populations. *American Naturalist* 200:345–371.

Lynch, M. and B. Walsh. 2018. Evolution and Selection of Quantitative Traits. Sinauer Associates, New York.

Mas-Colell, A., M. D. Whinston, and J. R. Green. 1995. Microeconomic Theory. Oxford Unversity Press, Oxford.

May, R. M. 1973. Stability and Complexity in Model Ecosystems. Princeton University Press, New Jersey.

Maynard Smith, J. 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge.

Maynard Smith, J. 1998. Evolutionary Genetics. Oxford University Press, Oxford, UK.

McNamara, J. M., C. E. Gasson, and A. I. Houston. 1999. Incorporating rules for responding into evolutionary games. *Nature* 401:368–71.

McNamara, J. M. and O. Leimar. 2020. Game Theory in Biology. Oxford University Press.

Meszéna, G., M. Gyllenberg, L. Pásztor, and J. A. J. Metz. 2006. Competitive exclusion and limiting similarity: a unified theory. *Theoretical Population Biology* 69:68–87.

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*, Mathematics and biosciences in interaction, pp. 193–231. Birkhäuser, Basel.

Metz, J. A. J., S. A. H. Geritz, G. Meszéna, F. J. A. Jacobs, and J. S. van Heerwaarden. 1996. Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction. In van Strien, S. J. and S. M. Verduyn Lunel (eds.), *Stochastic and Spatial Structures of Dynamical Systems*, pp. 183–231. North-Holland, Amsterdam.

Metz, J. A. J., S. D. Mylius, and O. Diekmann. 2008. When does evolution optimize? *Evolutionary Ecology Research* 10:629–654.

Metz, J. A. J., R. M. Nisbet, and S. A. H. Geritz. 1992. How should we define fitness for general ecological scenarios? *Trends in Ecology and Evolution* 7:198–202.

Michod, R. E. 1982. The theory of kin selection. *Annual Review of Ecology and Systematics* 13:23–55.

Mode, C. J. 1971. Multitype branching processes: theory and applications. No. 34. American Elsevier Publishing Company, New York.

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

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

Mullon, C. and L. Lehmann. 2018. Eco-evolutionary dynamics in metacommunities: ecological inheritance, helping within species, and harming between species. *American Naturalist* 192:664–685.

Mullon, C. and L. Lehmann. 2019. An evolutionary quantitative genetics model for phenotypic (co)variances under limited dispersal, with an application to socially synergistic traits. *Evolution* 73:1695–1728.

Nagylaki, T. 1992. Introduction to Population Genetics. Springer-Verlag, Heidelberg.

Narahari, Y. 2014. Game Theory and Mechanism Design. World Scientific, New Jersey.

Ohtsuki, H., C. Rueffler, J. Y. Wakano, K. Parvinen, and L. Lehmann. 2020. The components of directional and disruptive selection in heterogeneous group-structured populations. *Journal of Theoretical Biology* 507:1–16.

Oliveira, R. F., M. Taborsky, and H. J. Brockmann. 2008. Alternative reproductive tactics: an integrative approach. In *Alternative reproductive tactics: An integrative approach*, pp. 1–22. Cambridge University Press.

Otto, S. P. and T. Day. 2007. A biologist's Guide to Mathematical Modeling in Ecology and Evolution. Princeton University Press, Princeton, NJ.

Papadimitriou, C. 2014. Algorithms, complexity, and the sciences. *Proceedings of the National Academy of Sciences of the United States of America* 111:15881–15887.

Papadimitriou, C. H. 2007. The complexity of finding Nash equilibrium. In Nisan, N., T. Roughgarden, E. Tardos, and V. V. (eds.), *Algorithmic Game Theory*, pp. 29–52. Cambridge University Press.

Pásztor, L., Z. Botta-Dukát, G. Magyar, T. Czárán, and G. Meszéna. 2016. Theory-based ecology: a Darwinian approach. Oxford University Press, Oxford.

Pen, I. 2000. Reproductive effort in viscous populations. *Evolution* 54:293–297.

Prigent, I. and C. Mullon. 2023. The molding of intraspecific trait variation by selection under ecological inheritance. *Evolution* 77:1–18.

Priklopil, T. and L. Lehmann. 2020. Invasion implies substitution in ecological communities with class-structured populations. *Theoretical Population Biology* 134:36–52.

Priklopil, T. and L. Lehmann. 2021. Metacommunities, fitness and gradual evolution. *Theoretical Population Biology* 142:12–35.

Quirk, J. and R. Ruppert. 1965. Qualitative economics and the stability of equilibrium. *The Review of Economic Studies* 32:311–326.

Rand, D. A., H. B. Wilson, and J. M. McGlade. 1994. Dynamics and evolution: evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Philosophical Transactions of the Royal Society B-Biological Sciences* 343:261–83.

Ravigné, V., U. Dieckmann, and I. Olivieri. 2009. Live where you thrive: joint evolution of habitat choice and local adaptation facilitates specialization and promotes diversity. *The American Naturalist* 174:E141–E169.

Reeve, H. K. and P. W. Sherman. 1993. Adaptation and the goal of evolutionary research. *The Quarterly Review of Biology* 68:1–32.

Ronce, O. and D. Promislow. 2010. Kin competition, natal dispersal and the moulding of senescence by natural selection. *Proceedings of the Royal Society B-Biological Sciences* 277:3659–67.

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

Rousset, F. and O. Ronce. 2004. Inclusive fitness for traits affecting metapopulation demography. *Theoretical Population Biology* 65:127–141.

Sandholm, W. H. 2011. Population Games and Evolutionary Dynamics. MIT Press, Cambridge, MA.

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

Sasaki, A. and S. Ellner. 1995. The evolutionarily stable phenotype distribution in a random environment. *Evolution* 49:337–350.

Schmid, M., C. Rueffler, L. Lehmann, and C. Mullon. 2024. Resource variation within and between patches: where exploitation competition, local adaptation, and kin selection meet. *The American Naturalist* 203:E19–E34.

Schreiber, S. J. 2020. When do factors promoting genetic diversity also promote population persistence? A demographic perspective on Gillespie's SAS-CFF model. *Theoretical Population Biology* 133:141–149.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. Heterozygote advantage as a natural consequence of adaptation in diploids. *Proceedings of the National Academy of Sciences* 108:20666–20671.

Siljestam, M. and C. Rueffler. 2024. Heterozygote advantage can explain the extraordinary diversity of immune genes. *Elife* pp. 1–35.

Sinervo, B. and C. M. Lively. 1996. The rock–paper–scissors game and the evolution of alternative male

strategies. *Nature* 380:240–243.

Skaperdas, S. 1996. Contest success functions. *Economic Theory* 7:283–290.

Stephens, D. W. and J. R. Krebs. 1987. Foraging Theory. Princeton University Press, Princeton, NJ.

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.

Sydsaeter, K., P. Hammond, A. Seierstad, and A. Strøm. 2008. Further Mathematics for Economic Analysis. Prentice Hall, Essex, 2th edn.

Tanny, D. 1981. On multitype branching processes in a random environment. *Advanced Apllied Probability* 13:464–497.

Taylor, P. 1990. Allele-frequency change in a class-structured population. *American Naturalist* 135:95–106.

Taylor, P. D. 1989. Evolutionary stability in one-parameter models under weak selection. *Theoretical Population Biology* 36:125–143.

Taylor, P. D. and T. Day. 2004. Stability in negotiation games and the emergence of cooperation. *Proceedings of the Royal Society of London Series B-Biological Sciences* 271:669–674.

Taylor, P. D. and S. A. Frank. 1996. How to make a kin selection model. *Journal of Theoretical Biology* 180:27–37.

Tishkoff, S. A., R. Varkonyi, N. Cahinhinan, and et al. 2001. Haplotype diversity and linkage disequilibrium at human G6PD: Recent origin of alleles that confer malarial resistance. *Science* 293:455–462.

Tuljapurkar, S. 1990. Population Dynamics in Variable Environments. Springer-Verlag, New York.

Turelli, M., D. W. Schemske, and P. Bierzychudek. 2001. Stable two-allele polymorphisms maintained by fluctuating fitnesses and seed banks: protecting the blues in Linanthus parryae. *Evolultion* 55:1283–98.

Van Cleve, J. 2023. Evolutionarily stable strategy analysis and its links to demography and genetics through invasion fitness. *Philosophical transactions of the Royal Society B* 378:20210496.

Van Dooren, T. J. 1999. The evolutionary ecology of dominance-recessivity. *Journal of theoretical Biology* 198:519–532.

van Veelen, M. and P. Spreij. 2009. Evolution in games with a continuous action space. *Economic Theory* 39:355–376.

Vincent, T. L. and J. S. Brown. 2005. Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics. Cambridge University Press, Cambridge.

Vincent, T. L., M. V. Van, and B. S. Goh. 1996. Ecological stability, evolutionary stability and the ESS maximum principle. *Evolutionary Ecology* 10,:567–591.

Wakano, J. Y. and L. Lehmann. 2014. Evolutionary branching in deme-structured populations. *Journal of Theoretical Biology* 351:83–95.

West-Eberhard, M. J. 2003. Developmental Plasticity and Evolution. Oxford University Press.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.

Yamamichi, M., A. D. Letten, and S. J. Schreiber. 2023. Eco-evolutionary maintenance of diversity in fluctuating environments. *Ecology Letters* 26:S152–S167.