Properties of phenotypic plasticity in discrete threshold traits

Running title: Plasticity in threshold traits

Jane M. Reid[1,2,*] and Paul Acker[1]

[1]Centre for Biodiversity Dynamics, Institutt for Biologi, NTNU, Trondheim, Norway.

[2]School of Biological Sciences, University of Aberdeen, UK.

*Corresponding author: jane.m.reid@ntnu.no

Author contributions

J.M.R. conceptualised and wrote the manuscript, with substantial conceptual and editing input from P.A.

Acknowledgements

We thank Yimen Araya-Ajoy, Russ Lande and Jon Wright for helpful comments and discussions. J.M.R. and P.A. were supported by the Norwegian Research Council through the Centre for Biodiversity Dynamics (grant 223257), NTNU, and University of Aberdeen.

Data accessibility

There are no primary data associated with this Perspective paper.

Abstract

Forms of phenotypic plasticity in key traits, and forms of selection on and genetic variation in such plasticity, fundamentally underpin phenotypic, population dynamic and evolutionary responses to environmental variation and directional change. Accordingly, numerous theoretical and empirical studies have examined properties and consequences of plasticity, primarily considering traits that are continuously distributed on observed phenotypic scales with linear reaction norms. However, many environmentally sensitive traits are expressed as discrete alternative phenotypes and are appropriately characterised as quantitative genetic threshold traits. Here, we highlight that forms of phenotypic plasticity, genetic variation and inheritance in plasticity, and outcomes of selection on plasticity, could differ substantially between threshold traits and continuously distributed traits (as are typically considered). We thereby highlight theoretical developments that are required to rationalise and predict phenotypic and micro-evolutionary dynamics involving plastic threshold traits, and outline how intrinsic properties of such traits could provide relatively straightforward explanations for apparently idiosyncratic observed patterns of phenotypic variation. We summarise how key quantitative genetic parameters underlying threshold traits can be estimated, and thereby set the scene for embedding dynamic discrete traits into theoretical and empirical understanding of the role of plasticity in driving phenotypic, population and evolutionary responses to environmental variation and change.

Keywords: Cryptic genetic variation; gene by environment interaction; phenotypic plasticity; quantitative genetics; reaction norm; threshold trait.

Introduction

Plasticity, defined as the ability of single genotypes to produce different phenotypes in different environments, is expected to substantially shape phenotypic, population dynamic and evolutionary responses to environmental variation and change (Pigliucci 2005; Ghalambor et al. 2007; Wennersten and Forsman 2012; Ehrenreich and Pfennig 2016; Chevin and Hoffmann 2017; Kelly 2019). Specifically, plasticity directly allows short-term phenotypic responses within and/or across generations, which could currently be adaptive, non-adaptive or maladaptive (e.g. Agrawal et al. 1999; Ghalambor et al. 2007; Beaman et al. 2016; Donelson et al. 2018; Arnold et al. 2019a). Theory also shows how plasticity could affect persistence of populations experiencing directional environmental change (Chevin et al. 2010), and facilitate longer-term evolutionary adaptation to novel environments through rapid expression and evolution of plastic responses followed by genetic assimilation (Lande 2009, 2015). Rationalising current forms of plasticity, and quantifying drivers of and constraints on future evolutionary changes, will consequently be central to understanding and predicting phenotypic, population dynamic and evolutionary outcomes across multiple spatial and temporal scales (Gavrilets and Scheiner 1993a; Scheiner 1993; Chevin et al. 2010; Lande 2014; Murren et al. 2014; Ehrenreich and Pfennig 2016; Chevin and Hoffmann 2017; Kelly 2019).

In general, such outcomes depend on numerous properties of focal systems, including i) forms and magnitudes of plasticity in ecologically-relevant traits that affect fitness; ii) resulting selection acting on such traits and on trait plasticity; and iii) components of additive genetic (co)variation involving plasticity, and resulting patterns of inheritance and potential for plasticity evolution and genetic assimilation. These properties arise and exert their effects in the contexts

of environmental variability and predictability across episodes of trait development and expression versus selection (Scheiner 1993; DeWitt et al. 1998; de Jong 2005; Lande 2009, 2014, 2015; Chevin et al. 2010; Ashander et al. 2016; Chevin and Hoffmann 2017; Donelson et al. 2018). Accordingly, numerous empirical studies aim to quantify forms of plasticity in focal traits in experimental or wild populations, and to estimate components of selection and underlying additive genetic (co)variances (Scheiner 1993; Pigliucci 2005; Charmantier and Gienapp 2014; Murren et al. 2014; Hendry 2016; Chevin and Hoffmann 2017; Arnold et al. 2019a; Kelly 2019).

However, many recent theoretical and empirical studies, and conceptual reviews, invoke a key assumption that relationships between expressed phenotypes and underlying environmental variables are linear on observed phenotypic scales (i.e. linear phenotypic reaction norms, e.g. Nussey et al. 2007; Lande 2009, 2014, 2015; Dingemanse et al. 2010; Chevin et al. 2010; Chevin and Hoffmann 2017; Arnold et al. 2019a), with extensions to consider polynomial functions (Gavrilets and Scheiner 1993a,b; Scheiner 1993; de Jong 2005; Morrissey and Leifting 2016). Such linearities are commonly invoked to facilitate mathematical, statistical and/or verbal tractability, not necessarily because there is strong evidence or belief that phenotype-environment relationships will typically be straightforwardly linear in nature (noted by Gavrilets and Scheiner 1993b; Nussey et al. 2007; Lande 2009, 2014; Chevin et al. 2010; Dingemanse et al. 2010; Chevin and Hoffmann 2017). Indeed, many ecologically important traits that could substantially affect population responses to varying and changing environments show markedly non-linear phenotypic changes, which also do not directly conform to simple polynomial functions (Valladares et al. 2006; Chevin et al. 2010; Murren et al. 2014; Beaman et al. 2016; Arnold et al. 2019b).

Not least, many important morphological, behavioural and life-history traits expressed across the natural world show two (or more) approximately discrete phenotypic states, whose expression partly depends on current and/or previous environments (West-Eberhard 1989; Scheiner 1993; Roff 1996). For example, such traits encompass dichotomous movements, mating behaviours and reproductive parameters alongside alternative morphological developments. Given some polygenic basis, such traits can be appropriately conceptualised as quantitative genetic 'threshold traits', where an underlying 'liability' translates into expression of alternative discrete phenotypes when above versus below a threshold (Falconer and Mackay 1996; Roff 1996). Such threshold traits are well known to have distinctive intrinsic genetic and micro-evolutionary properties compared to standard quantitative traits that are continuously distributed on observed phenotypic scales. For example, they can maintain substantial 'cryptic' genetic variation that is not typically phenotypically expressed, and phenotypic heritability, selection intensity and evolutionary responses all depend on phenotype frequencies (Gianola 1982; Falconer and Mackay 1996 Ch.18; Roff 1994a, 1996, 1998a; Lynch and Walsh 1998 Ch.25; Moorad and Linksvayer 2008). However, the ways in which threshold traits could also show distinctive intrinsic forms and properties of phenotypic plasticity, and hence foster distinctive patterns of short-term phenotypic variation and longer-term micro-evolutionary outcomes in varying and changing environments, have received surprisingly little explicit attention. Since rapid plastic expression of alternative discrete phenotypes could dramatically affect population outcomes, and even flip populations between alternative ecological states, such properties and their consequences should be fully incorporated into modern treatments of plasticity and its implications.

Accordingly, we highlight multiple intrinsic properties of threshold traits that could cause distinctive patterns of phenotypic plasticity, and drive distinctive phenotypic and micro-evolutionary outcomes, that differ substantially from those involving traits that are continuously distributed on observed phenotypic scales. We thereby outline key attributes that should be incorporated into (eco-)evolutionary theory for dichotomous traits and quantified empirically, and summarise how such advances could be achieved. To provide necessary context for our focus on threshold traits and facilitate comparisons, we first synthesise core principles of plasticity in continuously distributed traits, as are currently widely envisaged. Key terminologies are summarised in Table 1.

Concepts of plasticity in continuous traits

Plasticity in continuously distributed quantitative traits is commonly conceptualised and quantified in two different ways by differing groups of researchers. First, experimental evolutionary biologists commonly measure plasticity by splitting families or clones across different (known) environments and measuring resulting mean phenotypes in each environment (Figure 1A). Plasticity is therefore typically viewed as *among-individual* phenotypic variation within families or genotypes and hence within lineages, often focussing on traits that are effectively expressed once as the culmination of initial development (here termed *developmental plasticity*, Table 1, Figure 1A, elsewhere termed fixed, irreversible or 'one-shot' plasticity, Scheiner 1993, 2002; van Buskirk and Steiner 2009; Auld et al. 2010; Lande 2015; Hendry 2016). Such developmental plasticity is also the primary focus of most theory on plasticity evolution and its implications, which typically considers relationships between environments of trait

development versus selection and assumes non-zero additive genetic variation in developmental reaction norm slopes (e.g. Gavrilets and Scheiner 1993a; Lande 2009, 2015; Ashander et al. 2016). Such genetic variation, and potential covariation with developmental reaction norm intercepts, can be evaluated empirically through quantitative genetic experiments that split sufficient families across environments (Figure 1A, Scheiner 1993; Pigliucci 2005; Arnold et al. 2019b), and/or through experimental evolution (Scheiner 2002).

Second, in wild animal evolutionary ecology and behavioural ecology, where families or clones often cannot be so readily created or manipulated, plasticity is typically measured as changing phenotypes of single individuals when exposed to different environments (Figure 1B). It is consequently viewed as longitudinal *within-individual* phenotypic variation, evidenced by individual-level reaction norm slopes that differ from zero. It therefore necessarily concerns flexible traits that individuals express multiple times (termed *labile plasticity*, Table 1, Figure 1B, Nussey et al. 2007; Dingemanse et al. 2010; Dingemanse and Wolf 2013; Charmantier and Gienapp 2014; Lande 2014, 2015; Hendry 2016).

Since developmental plasticity and labile plasticity can induce differing short-term phenotypic dynamics and longer-term evolutionary dynamics (e.g. Lande 2009, 2014, 2015, 2019), yet could co-occur and interact, both ideally need to be combined into a single holistic conceptual framework that encompasses overall phenotypic variation resulting from initial development and subsequent environmental impacts (Beaman et al. 2016). This can in principle be achieved through quantitative genetic approaches that fully consider interacting environmental and genetic effects on overall individual reaction norms, encompassing intercepts

(i.e. elevations), slopes and intercept-slope and slope-slope covariances across developmental and labile plasticity (Figure 1C, Supporting Information S1).

Here, individual reaction norm intercepts can reflect the outcome of developmental plasticity, resulting in lasting effects on an individual's mean phenotype as shaped by its developmental reaction norm and early-life environmental experience (red, Figure 1C). The developmental reaction norm intercept and slope could in turn show additive genetic (co)variation, as commonly considered in theoretical and empirical studies focussing solely on developmental plasticity (Figure 1A). Meanwhile, variation in individual reaction norm slopes (blue, Figure 1C) could reflect direct additive genetic effects, representing genetic variation in labile plasticity, and/or permanent environmental effects including further legacies of developmental plasticity. For example, an individual's mean phenotype, as shaped by developmental plasticity, could affect its ability to express labile plasticity. Complex genetic and/or environmental intercept-slope covariances for overall reaction norms could then arise, for example if individuals with low phenotypic intercepts typically have shallower (or steeper) phenotypic reaction norm slopes than individuals with higher intercepts (Figure 1C, Supporting Information S1, Nussey et al. 2007; Dingemanse and Wolf 2013).

In simple cases with linear reaction norms for both developmental and labile plasticity (e.g. Figure 1A,B), the full quantitative genetic architecture of the overall individual reaction norm (e.g. Figure 1C) can then be conceptualised as a three-parameter G-matrix (i.e. additive genetic variance-covariance matrix). Here, the three elementary parameters comprise the intercept and slope of the reaction norm for developmental plasticity (which together shape the intercept of the overall individual reaction norm), and the slope of the reaction norm for labile plasticity

(Supporting Information S1). These three parameters could be positively or negatively genetically correlated (or uncorrelated), potentially implying composite constraints on the form of additive genetic variation available to allow evolutionary responses to selection on any characteristic of the overall individual reaction norm (e.g. Gomulkiewicz and Kirkpatrick 1992; Nussey et al. 2007; Walsh and Blows 2009). For example, if reaction norm slopes for developmental and labile plasticity are strongly positively or negatively correlated, then evolution of both could be exacerbated or inhibited given particular regimes of selection.

This three-trait conceptualisation of overall plasticity combines and advances current standard conceptualisations of linear reaction norms, which typically focus on either developmental plasticity or labile plasticity but not both. Such conceptualisations effectively generate two-parameter models that consider single intercepts and slopes (Supporting Information S1). Yet, the basic premise that environmentally-induced phenotypic variation in continuously distributed quantitative traits is adequately described through linear (or polynomial) reaction norms for developmental and/or labile plasticity expressed on observed phenotypic scales implies multiple intrinsic properties of genetic variation in, selection on and inheritance of plasticity (Table 2). These properties are fundamental to mathematical and verbal theory regarding the implications of plasticity for population dynamic and evolutionary outcomes, and have motivated and underpinned numerous empirical examinations.

Key properties of plasticity in continuous traits

First, ongoing evolution of the degree of phenotypic plasticity (whether developmental and/or labile) in continuously distributed traits is envisaged to involve evolution of reaction norm slopes.

Such evolution in turn requires non-zero additive genetic variation in slope (and/or in other parameters describing polynomial reaction norm shapes, Gavrilets and Scheiner 1993a,b; Scheiner 1993, 2002; de Jong and Gavrilets 2000; Nussey et al. 2007; Lande 2009, 2014; Ashander et al. 2016; Arnold et al. 2019b). This implies existence of gene-by-environment interactions on the observed phenotypic scale (i.e. non-zero GxE shaping developmental and/or labile plasticity), meaning that genetic effects on phenotypes vary among environments (Figure 1C, Table 2). Given linear reaction norms, non-zero GxE can readily cause substantially greater additive genetic variation in phenotypes in more extreme (i.e. unusual) environments. This is especially likely given some canalization of mean phenotype (and hence little additive genetic variation) in typical environments that a population has previously frequently experienced, due to long-term stabilizing selection (e.g. Nussey et al. 2007; Lande 2015). Such effects generate potential for rapid evolutionary responses to selection following environmental change (Lande 2009; Ashander et al. 2016). However, there will not necessarily be non-zero GxE in any particular system, and quantifying the form and magnitude of GxE has been a major empirical endeavour in experimental systems (Scheiner 1993, 2002; Pigliucci 2005) and, increasingly, in free-living wild populations (Nussey et al. 2007; Charmantier and Gienapp 2014; Ramakers et al. 2018; Kelly 2019; Supporting Information S2).

Second, directional selection for or against phenotypic plasticity is envisaged to effectively exert directional selection on the magnitude of the reaction norm slope (either directly, or indirectly through covariance with intercepts and hence trait means, Figure 1D, Table 2, Scheiner 1993, 2002). Selection for increased plasticity (i.e. favouring steeper slopes) can in principle arise if increased phenotypic sensitivity to environmental conditions increases fitness,

irrespective of whether the sign of the optimal reaction norm slope is positive or negative (Arnold et al. 2019a). Meanwhile, selection against plasticity (i.e. favouring shallower slopes) can arise if there are fitness costs of expressing different or changing phenotypes or simply of maintaining physiological capacities to do so (often termed induced versus constitutive costs, DeWitt et al. 1998; van Buskirk and Steiner 2009; Auld et al. 2010; Chevin et al. 2010; Arnold et al. 2019a). Overall stabilising selection around an optimal degree of plasticity, and hence around an optimal reaction norm slope, could potentially result.

If there is initially non-zero GxE (and hence additive genetic variation in reaction norm slopes), then directional selection could cause evolutionary changes in slope and hence in the degree of phenotypic plasticity. Yet, directional or stabilising selection could gradually erode additive genetic variation and diminish GxE, reducing subsequent evolutionary responses (Figure 1D, van Buskirk and Steiner 2009; Ehrenreich and Pfennig 2016; Saltz et al. 2018; but see Lande 2009). Further studies have consequently postulated mechanisms by which genetic variation in plasticity and associated evolutionary potential, and polymorphisms comprising plastic and canalized (i.e. non-environmentally responsive) individuals, can be maintained (Wolf and Weissing 2010; Dingemanse and Wolf 2013; Saltz et al. 2018). For example, combinations of negative frequency-dependent selection on plastic versus canalized phenotypes, and positive feedbacks that reduce costs of repeated trait expression, have been invoked to explain ongoing coexistence of responsive and unresponsive lineages (i.e. polymorphism in labile plasticity, Wolf et al. 2008; Wolf and Weissing 2010). In principle, additive genetic variation in reaction norm slopes could be maintained through disruptive selection, where individuals with slopes of either

241 extreme have consistently higher fitness than individuals with intermediate slopes, but such
242 selection seems unlikely to be commonplace (Saltz et al. 2018; Supporting Information S2).

243 Third, reaction norm parameters, and resulting forms of phenotypic plasticity, are
244 expected to adhere to basic principles of additive genetic inheritance given sexual reproduction.
245 Consequently, matings between environmentally canalized parents (i.e. with reaction norm
246 slopes for developmental and/or labile plasticity that are close to zero) will typically produce
247 canalized offspring (i.e. also with expected reaction norm slopes close to zero), as could matings
248 between parents with slopes of similar magnitudes but opposite signs (Figure 1E, Table 2).
249 However, highly phenotypically canalized parents are unlikely to produce highly plastic offspring
250 (i.e. with very steep reaction norm slopes), although such offspring could occasionally arise
251 through sampling variance around the expectation. Indeed, population-wide canalization across
252 environments can be viewed as fixation of genetic variants that increase robustness to
253 environmental variation (Ehrenreich and Pfennig 2016). Additional sources of new genetic
254 variation in reaction norm slopes, whether from mutation or gene flow stemming from
255 immigration given local adaptation, are then required to generate new developmental and/or
256 labile phenotypic plasticity in currently phenotypically canalized populations (e.g. Saltz et al.
257 2018).

258 The above properties and principles of genetic variation in, selection on and inheritance
259 of phenotypic plasticity in continuously distributed traits (Table 2) explicitly or implicitly underpin
260 much theory regarding the expression, evolutionary dynamics and consequences of plasticity.
261 But, such properties could be qualitatively different for threshold traits, implying that the
262 dynamics and consequences of plasticity might also differ substantially.

Concepts of plasticity in threshold traits

The concept of a threshold trait (or 'threshold character') is long established in quantitative genetics, spanning animal breeding, medicine and diverse areas of evolutionary biology (Wright 1934; Gianola 1982; Falconer and Mackay 1996 Ch.18; Roff 1996, 1998a; Lynch and Walsh 1998 Ch.25; Tomkins and Hazel 2007; Moorad and Linksvayer 2008; Grossen et al. 2010; Pulido 2011; Dodson et al. 2013; Hadfield 2015). In brief, each individual is envisaged to have an underlying 'liability', defined as a latent (i.e. unobserved) continuously distributed quantitative trait which can have multiple genetic and environmental components, and which causes expression of alternative discrete phenotypes when above versus below a threshold value (Table 1, Figure 2A). Expression of the alternative phenotypes is envisaged to be deterministic conditional on the liability (i.e. outcomes are not necessarily directly probabilistic, Supporting Information S1). Substantial genetic and/or environmental variation in liability (on either side of the threshold) can consequently exist without causing any phenotypic variation.

To date, properties of threshold traits have commonly been considered for dichotomous traits that are effectively expressed once through differential development, resulting in permanently different discrete phenotypes (e.g. skeletal structures; morphological defences against predation; winged (i.e. dispersive) versus wingless forms; facultative diapause; sex; and fixed dichotomous reproductive and/or life-history strategies, among others, e.g. Falconer and Mackay 1996; Roff 1994a, 1996; Agrawal et al. 1999; Ostrowski et al. 2000; Bégin and Roff 2002; Hazel et al. 2004; Tomkins and Hazel 2007; Grossen et al. 2010; Dodson et al. 2013; Snell-Rood et al. 2018; Debes et al. 2020). Here, phenotypic outcomes commonly depend partly on initial

285 environmental conditions, representing developmental plasticity (Scheiner 1993; Roff 1996;
286 Bégin and Roff 2002; Snell-Rood et al. 2018).

287 However, as with many continuously distributed traits, many threshold traits are re-
288 expressed on timeframes of hours, days, seasons or years, potentially allowing labile plasticity
289 manifested as reversible individual expression of alternative phenotypes when exposed to
290 different environments (i.e. within-individual phenotypic variation, Table 1). Obvious examples
291 include dichotomous behaviours, including diel movements and dominant versus subordinate
292 mating tactics (e.g. Harrison et al. 2017; Crocker-Buta and Leary 2018); seasonal migration versus
293 year-round residence (Brodersen et al. 2014; Reid et al. 2018, 2020); and diverse aspects of
294 reproduction including breeding versus skipping, production of twins versus one offspring, or
295 divorce or extra-pair mating versus mate fidelity in iteroparous species (Falconer and Mackay
296 1996; Duthie et al. 2016; Germain et al. 2018). Key properties of plasticity in such threshold traits
297 could differ quite fundamentally from those for continuously distributed traits where linear (or
298 polynomial) reaction norms are manifested on observed phenotypic scales (Table 2). These
299 properties and differences result from forms of genetic and environmental variation in
300 underlying liability, and their non-linear translation into discrete observed phenotypes (Figure
301 2A).

302 Viewed most simply, each individual has an inherited additive genetic value for liability,
303 plus some set of environmental effects on liability which could reflect current and/or previous
304 environmental conditions (Roff 1996). These environmental effects can be envisaged to result
305 from liability-scale reaction norms that shape responses to environmental variation, where
306 reaction norm slopes could show additive genetic variation (i.e. generating GxE in liability).

Indeed, the whole structure of reaction norm intercepts, slopes and covariances that acts on the observed phenotypic scale for continuously distributed traits (Figure 1C) can be envisaged to similarly act on the latent liability scale for threshold traits (Figure 2C). The threshold itself can then be viewed as a higher-level reaction norm (or 'developmental map') that deterministically translates the combined genetic and environmental effects on the continuously distributed latent liability into discrete observed phenotypic outcomes (Figure 2A,C, Supporting Information S1, Roff 1996; Lynch and Walsh 1998 Ch.11). This conceptual structure distinguishes plasticity occurring on the latent liability scale (i.e. latent plasticity) from plasticity occurring on the observed phenotypic scale (i.e. phenotypic plasticity, Table 1). It has multiple interesting implications for the expression, form and maintenance of phenotypic plasticity in relation to underlying genetic variation; for the consequences of selection for or against phenotypic plasticity; and for patterns of phenotypic inheritance (Table 2).

Key properties of plasticity in threshold traits

Expression of phenotypic plasticity and underlying genetic variation

Individuals and lineages with initial liabilities (i.e. liability intercepts) near the threshold are more likely to be phenotypically plastic (i.e. express different discrete phenotypes at different times, Figure 2B). This is because any subsequent small environmental deviation in liability, whether due to predictable environmental responses shaped by liability-scale reaction norms and/or random micro-environmental deviations, is relatively likely to push individuals across the threshold (in either direction) and hence induce expression of the alternative phenotype (Figure 2B,C,E). In contrast, individuals and lineages with initial liabilities far from the threshold (in either

direction) are likely to be phenotypically canalized, since even sizeable subsequent environmental deviations in liability are unlikely to induce phenotypic shifts (Figure 2B,C,E). Substantial latent plasticity could then exist in the form of liability-scale reaction norm slopes, but have zero phenotypic effect and hence be decoupled from expression of phenotypic plasticity.

Individuals' expressions of either alternative phenotype, and of labile phenotypic plasticity (i.e. switching between discrete phenotypes), can consequently be intrinsically positively correlated within and across time periods on different scales (e.g. days, seasons, years). This is because individuals with liability intercepts far from or near to a fixed threshold on one occasion will also have such intercepts subsequently. Sequences of small environmental deviations, or of predictable environmental changes, will consequently have phenotypic effects (or lack of effects) that are consistent across occasions within individuals. High individual repeatability of expression of each discrete phenotype, alongside high individual repeatability of phenotypic plasticity, can therefore readily emerge (Figure 2E). Population-wide repeatability of expression of each discrete phenotype will also be positively associated with phenotype frequency, such that the most frequently expressed phenotype is the most highly repeatable. This is because, given the standard quantitative genetic assumption that the population distribution of liabilities is Gaussian, the most frequent phenotype will encompass more individuals whose liabilities are further from the threshold and hence less likely to be phenotypically plastic (Figure 2B).

Consequently, if there is non-zero additive genetic variation in liability intercept with a distribution of values that spans or approaches the threshold, the basic implication is that there

will be non-zero genetic variation in expression of phenotypic plasticity (Figure 2C,D). Specifically, genotypes that place an individual's liability intercept near to or far from the threshold are likely to effectively cause plastic and canalized phenotypes respectively, given some regime of environmental variation. Resulting genetic variation in phenotypic plasticity does not necessarily require genetic variation in any reaction norm slope, or hence require non-zero GxE on the liability scale (Figure 2D, Table 2, although such GxE could also exist, Figure 2C). Rather, genetic and environmental effects that are straightforwardly additive on the liability intercept can induce non-additive variation in environmental responses in observed discrete phenotypes (e.g. Gianola 1982). This is because any particular additive genetic or environmental increment on liability might or might not cause a change in phenotype, or hence cause expression of phenotypic plasticity, depending on the pre-existing liability value. Further, even with GxE and substantial resulting variation in linear reaction norm slopes on the latent liability scale, extreme environments will not necessarily cause increased genetic variation in phenotype, and might even decrease such variation if all individuals are pushed to the same side of the threshold (Supporting Information S3).

Effective genetic variation in phenotypic plasticity might consequently be commonplace, or even ubiquitous, in polymorphic labile threshold traits. This is because such traits are expected to maintain substantial additive genetic variation in liability. This in turn is because selection on genetic variants that do not immediately cause overall liability values to cross the threshold (or hence cause any change in phenotype) is expected to be weak, even given strong directional selection on phenotype, and especially if liability is highly polygenic (Roff 1994a, 1996, 1998a,b). Intrinsic negative frequency-dependent selection on underlying genetic variants can also arise,

373 further contributing to a relatively high expected mutation-selection balance, alongside effects
374 of drift in small populations (Roff 1998a). Resulting maintenance of substantial additive genetic
375 variation has, implicitly, been shown for liability-scale reaction norm intercepts (Roff 1994a,
376 1998a), but similar processes might also affect liability-scale reaction norm slopes. This is because
377 considerable genetic variation in slope could exist yet have little or no phenotypic effect, and
378 hence be effectively shielded from selection (Figure 2C, Supporting Information S3).

379 Given some regime of stochastic and/or predictable environmental variation, populations
380 could therefore readily comprise mixtures of highly phenotypically plastic and highly canalized
381 individuals (Figures 2C,D,E, Table 2). No further explanations, such as negative frequency-
382 dependent selection on expression of plasticity coupled with positive feedbacks that reduce
383 expression costs (Wolf et al. 2008; Wolf and Weissing 2010), or complex mechanisms driving
384 evolution of multiple distinct tactics (Engqvist and Taborsky 2016), are necessarily required
385 (Hazel et al. 2004, although such effects could additionally apply). Hence, while cross-
386 environment phenotypic canalization of continuously distributed traits effectively requires
387 fixation of genetic variants that reduce reaction norm slopes and make traits robust to
388 environmental variation (Ehrenreich and Pfennig 2016), such fixation is not required to generate
389 a highly phenotypically canalized threshold trait. Co-existence of canalized and phenotypic plastic
390 individuals, and transitions between apparent canalization and plasticity given environmental
391 changes, are consequently relatively straightforward to explain at both population and individual
392 levels, and could rapidly occur without need for complex mechanisms or substantial mutation or
393 gene flow.

394

Selection

The occurrence of directional selection for or against expression of labile phenotypic plasticity in a threshold trait (i.e. higher or lower fitness in individuals that sequentially express both alternative phenotypes) could imply that individuals with liability intercepts close to the threshold have higher or lower fitness respectively than individuals with liability intercepts further away. Directional selection for or against phenotypic plasticity could consequently induce stabilising or disruptive selection on liability intercept if the distribution of liabilities substantially spans the threshold (Figure 3A,B), or directional selection if the distribution only marginally spans the threshold. The form of selection on liability induced by selection on phenotypic plasticity could therefore depend on the liability distribution and hence change with the progress of evolution, even if the form of selection on expression of phenotypic plasticity does not change at all.

Eras of directional selection on phenotypic plasticity that cause directional, stabilising or disruptive selection on the liability intercept could then respectively reduce additive genetic variation, or increase additive genetic variation and even generate bimodal distributions (Figure 3A,B). Unlike with continuously distributed traits, evolution of phenotypic plasticity (whether developmental and/or labile) could then proceed solely through evolution of distributions of reaction norm intercepts, irrespective of slopes (Table 2).

But, complex forms of liability-scale GxE that facilitate or prevent expression of phenotypic plasticity could also conceivably evolve. For example, individuals with liability intercepts close to or far from the threshold could in principle have liabilities that are differentially sensitive to environmental variation, implying evolution of some form of intercept-

slope covariance for the liability-scale reaction norm (Figure 3C,D). Directional selection for or against expression of phenotypic plasticity could therefore potentially generate non-zero liability-scale GxE rather than eradicate it, which is opposite to basic expectations for continuously distributed traits (Figure 1D, Table 2). The presence of non-zero liability-scale GxE could then serve to prevent expression of phenotypic plasticity rather than necessarily define variation in its presence.

Since the intercept of the overall individual liability-scale reaction norm could depend on the form of developmental plasticity (Figure 2C), such outcomes will depend on the full quantitative genetic architecture of the liability-scale G-matrix, including genetic covariances between reaction norm slopes for developmental versus labile plasticity. Evolution of such complex liability-scale quantitative genetic architectures (Figure 3) would likely require consistently strong selection for or against phenotypic plasticity. But, such selection might be expected to be stronger for labile threshold traits than for continuously distributed traits, since there could be substantial induced and/or constitutive costs, or benefits, associated with switching between alternative phenotypic states.

Phenotypic inheritance

Threshold traits could also generate distinctive patterns of apparent inheritance of phenotypic plasticity or canalization, manifested as parent-offspring resemblance (Table 2). Such patterns could arise because the threshold trait structure translates additive genetic effects on liability into non-additive genotypic effects on observed phenotypes (Gianola 1982; Lynch and Walsh 1998 Ch.25).

Specifically, reproduction between parents with different highly canalized discrete phenotypes, resulting from liability intercepts far from the threshold on opposite sides, could readily generate highly phenotypically plastic offspring. This is because resulting offsprings' liability intercepts could be close to the threshold, meaning that their phenotypes change in response to any small environmental deviation. Maintenance (or elimination) of phenotypic plasticity will therefore partly depend on the degree of assortative (or disassortative) mating with respect to canalized parental phenotypes. If both alternative phenotypes are expressed in a population, then phenotypic plasticity could be hard to eradicate, even given very strong selection against it, unless there is strong assortative mating. Equally, given standard inheritance of liability with Mendelian sampling variance (reflecting segregation and recombination), it could be hard to eradicate expression of both (relatively) canalized phenotypes even given strong selection for phenotypic plasticity. Genetic variation for plasticity (and for canalization) in threshold traits is consequently unlikely to be purged as readily as could be the case for continuously distributed traits (e.g. van Buskirk and Steiner 2009; Saltz et al. 2018). Rather, selection for or against phenotypic plasticity could potentially drive indirect selection on the mating system, where optimal mate choice given any regime of selection on plasticity could depend on the degree of phenotypic plasticity of any focal individual.

Finally, if females and males have different mean liability intercepts and hence show different phenotype frequencies (i.e. phenotypic sexual dimorphism), then deceptive patterns of phenotypic variation could arise, which could be incorrectly interpreted to imply sex-specific genetic architectures or other causes of phenotypic variation (e.g. Lynch and Walsh 1998 Ch.25). For example, individuals of whichever sex has mean liability intercept closer to the threshold

could be more phenotypically plastic, even without any sex-specific effects on liability-scale reaction norm slopes. Evidence of apparent sex-specific genetic dominance reversals at large-effect loci could also arise, even in the absence of such effects. This is because genetic effects that are purely additive on the liability scale, and identical in both sexes, can become non-additive on the observed phenotypic scale, with forms that differ between the sexes if mean liabilities differ (Supporting Information S4).

Discussion

The potential for threshold traits to show distinctive properties of phenotypic plasticity compared to traits that are continuously distributed on observed phenotypic scales, including differing patterns of expression, genetic variation, selection and apparent inheritance, has numerous implications for rationalising and predicting phenotypic, population and evolutionary dynamics in the contexts of environmental variation and change. Core premises of theory concerning continuously distributed traits include that genetic variation in phenotypic plasticity equates to genetic variation in reaction norm slopes, constituting GxE; that directional selection for or against plasticity will cause evolutionary changes that respectively increase or decrease reaction norm slopes; that fixation of phenotypic canalization or of some degree of plasticity implies erosion of GxE; and that observed degrees of plasticity are directly inherited following standard (additive) quantitative genetic expectations (Table 2). Such properties underpin mathematical and verbal explorations of how combinations of plasticity and micro-evolution can allow mean values of continuously distributed traits to re-match changing environmentally-determined optima, and thereby 'rescue' declining populations by restoring the previous

ecological status quo (Nussey et al. 2007; Lande 2009; Ashander et al. 2016; Chevin and Hoffmann 2017; Arnold et al. 2019a; Kelly 2019). But, such properties and outcomes do not necessarily apply to threshold traits, where inherited reaction norms are postulated to act on latent liability scales.

For such traits, genetic variation in phenotypic plasticity does not necessarily require genetic variation in liability-scale reaction norm slopes; evolution of the degree of phenotypic plasticity could potentially occur through changing intercepts rather than necessarily slopes; strong directional selection on phenotypic plasticity could conceivably generate rather than erode complex forms of liability-scale GxE; and substantial phenotypic plasticity could readily re-emerge in offspring of highly phenotypically canalized parents, generating persistent population-wide variation in plasticity (Table 2). New theory will consequently be required to rationalise and predict joint phenotypic, population and evolutionary dynamics involving plastic threshold traits, and thereby discern how such traits could act to rescue populations by flexible switching between alternative phenotypic and ecological states rather than by restoring the status quo.

Some progress can be achieved through mathematical theory, aiming to formalise basic principles that are currently outlined verbally (Table 2) and identify conditions under which they apply (e.g. Hazel et al. 2004). For example, population frequencies of phenotypically plastic versus canalized individuals, and resulting opportunities for selection, will presumably depend on relative magnitudes of additive genetic means and (co)variances in liability-scale reaction norm intercepts and slopes alongside micro-environmental variances. Evolved intercepts and slopes, and resulting environmental variances, will in turn affect the impact of selection and the consequent maintenance of (cryptic) genetic variation. However, the fact that existing

505 evolutionary theory involving plasticity explicitly focuses on linear phenotypic reaction norms to
506 facilitate mathematical tractability (e.g. Lande 2009, 2019; Ashander et al. 2016) implies that
507 dynamics involving threshold traits will be less readily tractable. Advances may consequently
508 require numerical or individual-based simulations, which can more readily encompass key
509 features such as intrinsically dynamic impacts of selection and genetic (co)variances in relation
510 to phenotype frequencies, and dependence of plasticity on forms of assortative mating.
511 Simulations have previously been used to examine basic properties of threshold traits, such as
512 the maintenance of substantial additive genetic variation in liability through selection-mutation-
513 drift balance despite strong directional selection (Roff 1998a), evolution of threshold traits from
514 continuously varying traits (Chevin and Lande 2013), and interacting effects of indirect selection,
515 environmental variation and assortative mating (de Zoeten and Pulido 2020). However, such
516 approaches have not yet been applied to explicitly consider evolutionary dynamics of threshold
517 trait plasticity.
518
519 **Applicability of the threshold trait model**
520 Justification for developing theory for dynamic threshold traits requires that key traits that show
521 discrete alternative phenotypes do reasonably conform to the threshold trait model. While latent
522 liabilities cannot, by definition, be directly observed, the basic adequacy of the threshold trait
523 model has been well demonstrated through quantitative genetics studies based in animal
524 breeding and medicine, alongside experiments in evolutionary biology, which show that
525 observed patterns of phenotypic variation and micro-evolution concur with theoretical
526 predictions (Gianola 1982; Falconer and Mackay 1996; Roff 1994b, 1996; Lynch and Walsh 1998).

However, such work has predominantly focused on initial development, not on labile phenotypic plasticity (but see Negussie et al. 2012). Formalising theory and predictions for labile plasticity could therefore provide opportunities to evaluate the applicability of the threshold model to dynamic discrete traits in wild populations, and highlight pertinent parameters that need to be estimated. Such theory could then provide relatively straightforward explanations for complex observed patterns of phenotypic variation that are otherwise puzzling or require some specific or idiosyncratic explanation.

As one broad example, dichotomous expression of seasonal migration versus year-round residence in partially migratory systems is a critical trait in the context of understanding population responses to environmental variation and change. Here, any change in phenotype frequencies, whether due to micro-evolution, plasticity and/or micro-evolution of plasticity, will directly alter seasonal population dynamics and distributions (Reid et al. 2018). Migration versus residence has long been conceptualised as a threshold trait, underpinned by combinations of genetic and environmental effects (Berthold 1988; Pulido 2011; Dodson et al. 2013). Classic work on blackcaps (*Sylvia atricapilla*) used breeding designs and artificial selection to show that liability for migration (as expressed in captivity) is highly heritable and positively genetically correlated with measures of migratory activity, and hence that phenotypic expression of migration or residence could be rapidly eliminated and regained (Berthold 1988; Pulido 1996, 2011; de Zoeten and Pulido 2020). Recent analyses of partially-migratory European shags (*Phalacrocorax aristotelis*) revealed patterns of among-individual and within-individual phenotypic variation (i.e. labile plasticity) that are also highly consistent with verbal predictions from the threshold trait model. Specifically, focal populations comprise mixtures of individuals that are non-breeding

549 season migrants or residents (representing two alternative phenotypes), alongside individuals
550 that switch between resident and migrant states part-way through the non-breeding season
551 which can be viewed as expressing within-season phenotypic plasticity (Reid et al. 2020; Acker et
552 al. in review). High between-year repeatability is evident, such that individuals that migrate,
553 remain resident or switch within one non-breeding season are highly likely to do the same again
554 in subsequent non-breeding seasons, with the expected positive associations between
555 phenotype frequency and repeatability (Reid et al. 2020; Acker et al. in review). Yet, between-
556 year plasticity also occurs and is correlated with within-season plasticity, such that individuals
557 that switch between residence and migration within any one non-breeding season are more likely
558 to switch to residence or full-season migration in subsequent non-breeding seasons. This concurs
559 with the expected correlation in plasticity across temporal scales (Acker et al. in review). These
560 analyses also revealed episodes of selection for or against within-season plasticity, that varied
561 substantially among years in relation to environmental conditions. Such fluctuating selection
562 implies variable costs of plasticity, and resulting episodes of directional, stabilising and disruptive
563 selection on underlying liability (Reid et al. 2020; Acker et al. in review).

564 Other partially-migratory systems also show mixtures of phenotypically canalized and
565 plastic residents and migrants, with different levels of individual variation and repeatability
566 across timescales, for example regarding diel feeding movements in burbot (*Lota lota*, Harrison
567 et al. 2017), and seasonal migration versus residence in roach (*Rutilus rutilus*, Brodersen et al.
568 2014). Such patterns are commonly discussed in the broad context of the threshold trait model
569 (e.g. Dodson et al. 2013), but not explicitly tested against qualitative or quantitative predictions.
570 For example, resident elk (*Cervus elaphus*) were more likely to switch to migration across years

than vice versa in a population where residence was the more frequent tactic, and some individuals switched repeatedly (Eggeman et al. 2016). These patterns concur with expectations given the threshold trait model, yet were not explicitly considered in that context.

Further, apparent non-additive genotypic effects on migration (and associated age at reproductive maturity) have been documented in Atlantic salmon (*Salmo salar*), and interpreted as sex-specific genetic dominance reversals at a large-effect locus which serve to resolve sexual conflict (Barson et al. 2015; Czorlich et al. 2018). However, since migration timing shows substantial sexual dimorphism, similar patterns of phenotypic variation could readily arise given a predominantly quantitative genetic threshold trait with sex-specific mean liability intercepts and sex-independent co-dominance at the large-effect locus, and hence without requiring any sex-specific dominance reversal (Supporting Information S4). Finally, while assortative mating has been suggested to facilitate evolution of distinct migratory types within sympatric breeding populations (Bearhop et al. 2005; de Zoeten and Pulido 2020), the implication that such mating could reduce phenotypic plasticity in offspring has not been discussed. There is consequently considerable scope for tighter quantitative interpretation, or re-interpretation, of diverse empirical results in the context of the threshold trait model, regarding seasonal migration, and regarding numerous other labile dichotomous traits.

Estimation of key parameters

Given that observed patterns of phenotypic variation broadly conform to expectations for threshold traits, the next challenge is to explicitly estimate key quantitative genetic parameters comprising additive genetic (co)variances in intercepts and slopes of liability-scale reaction

norms. Such estimates have sometimes been obtained by considering other relevant and directly observable continuously distributed phenotypic traits as proxies of the (unobserved) liability. For example, juvenile hormone titres have been interpreted as liability to develop winged versus wingless forms in hemimetabolous insects (Roff 1994b), and body size has been interpreted as liability to produce particular morphological forms or enact particular reproductive strategies in fish (Dodson et al. 2013). However, such observable proxy traits are unlikely to simply equate to liability, and are best viewed as additional traits that could be genetically correlated with liability (Roff and Fairbairn 2007; Debes et al. 2020). In general, such correlated traits could induce or constrain multivariate evolution, including evolution of overall plasticity (e.g. Lande 2019). Resulting indirect selection on threshold trait liabilities could have particularly important effects, by effectively exposing otherwise cryptic genetic variation to selection (e.g. de Zoeten and Pulido 2020).

Indeed, it is not necessary to measure any observable proxy trait in order to infer key quantitative genetic parameters for the latent liability underlying a threshold trait (Falconer and Mackay 1996). Rather, parameters representing liability-scale reaction norm intercepts and slopes can in principle be directly estimated given observations of dichotomous phenotypes expressed by relatives across environments. This can be achieved by utilizing statistical machinery that is now well established in the form of generalized linear mixed models (de Villemereuil et al. 2016). Here, a specified function links observations of discrete phenotypes to underlying latent values, which could in turn be modelled as any function (i.e. reaction norm) of environmental variables or other covariates of interest. Such formulations of the threshold trait model are well established in statistical and quantitative genetics theory, and are explicitly

615 enacted using a probit link (i.e. representing an inverse cumulative normal distribution, Gianola
616 1982; Hadfield 2015; de Villemereuil et al. 2016; Germain et al. 2018; Debes et al. 2020). Yet, the
617 implications and interpretations of such analyses for concepts of latent versus phenotypic
618 plasticity are still surprisingly infrequently explicitly discussed.

619 Rather, evolutionary ecologists interested in environment-dependent expression of
620 dichotomous traits have reformulated the standard threshold trait model into the '(latent)
621 environmental threshold model' (Hazel et al. 2004; Tomkins and Hazel 2007; Pulido 2011; Buoro
622 et al. 2012; Buzatto et al. 2015). Here, the focal evolving trait is defined as the point on some
623 known environmental axis (or 'cue') at which each individual or lineage's phenotype changes
624 state, envisaged as a genetically controlled threshold value. Hence, in contrast to the standard
625 threshold trait model, the environmental threshold model assumes that the environmental axis
626 exerts equal effects across individuals and lineages, while the threshold value can show additive
627 genetic variation and evolve (Tomkins and Hazel 2007; Buoro et al. 2012; Buzatto et al. 2015).
628 The standard threshold trait model and the environmental threshold model are equivalent (i.e.
629 re-parameterisations of the same conceptual model) in the special case where there is zero
630 liability-scale GxE given the standard threshold model, and hence a cross-environment genetic
631 correlation of 1 (i.e. zero variation in latent plasticity, Roff 1994b; Hazel et al. 2004; Tomkins and
632 Hazel 2007). The point at which an individual or lineage's liability crosses the fixed threshold (i.e.
633 its observed threshold value), is then directly related to its liability intercept (e.g. Figure 2D).
634 Consequently, the environmental threshold model formulation has motivated empirical
635 investigations of whether the condition of zero GxE in liability (or in some proxy trait) is
636 approximately valid (Roff 1994b).

637 However, if one ambition is to rationalise and predict evolutionary dynamics of plasticity,
638 then it seems highly restrictive to formulate analyses and conceptual developments through a
639 model that assumes zero liability-scale GxE. In fact, there is no need to make such an assumption,
640 which arises because the environmental threshold model considers environmental and genetic
641 effects on separate orthogonal axes of cues versus thresholds. This structure can be eliminated
642 by considering joint genetic and environmental effects on the same latent liability, by formulating
643 liability-scale reaction norms that could include GxE (Figure 2C). Indeed, the possibility of both
644 environmental and genetic effects on liability was already fully embedded in the original
645 quantitative genetic threshold trait model, for example yielding theory on heritability on latent
646 liability (and observed phenotypic) scales (Dempster and Lerner 1950; Lynch and Walsh 1998
647 Ch.25). Here, the heritability represents the ratio of genetic to total variance, and the total
648 variance typically includes a substantial component of environmental variance (Roff 1996).
649 Formulation of a distinct (latent) environmental threshold model is consequently not really
650 necessary, and has invoked additional restrictive assumptions (Roff 1994b; Hadfield 2015).
651 Meanwhile, some papers that are written in terms of the environmental threshold model in fact
652 approximately implement the standard threshold model, including interacting latent-scale
653 genetic and environmental effects (e.g. Ostrowski et al. 2000).
654
655 **Interpretations and implications**
656 Given that intrinsic properties of phenotypic plasticity and plasticity evolution could differ
657 markedly between threshold traits and continuously distributed traits (Table 2), it is important
658 that traits of biological interest are not mis-conceptualised (e.g. true threshold traits treated as

659 continuously distributed traits), at least without considering the implications of such conversions.
660 It might seem implausible that true discrete and continuous traits could be confused or inter-
661 changed. However, given labile phenotypic plasticity, the distinction between the two may not
662 always be clear and may be a matter of biological interpretation.

663 For example, the seasonal timing of key life-history events, such as breeding, is now a
664 major focus for theoretical and empirical work on plasticity and evolutionary dynamics in relation
665 to environmental change (Nussey et al. 2007; Lof et al. 2012; Inouye et al. 2019; Radchuk et al.
666 2019). Here, the focal trait is commonly defined, measured and analysed as the observed date of
667 event (e.g. egg laying in birds, parturition in mammals, flowering date in plants), which is
668 continuously distributed (e.g. Arnold et al. 2019a,b; Inouye et al. 2019; de Villemereuil et al.
669 2020). However, biologically, these observed events represent the dates on which individuals
670 switch between approximately discrete phenotypic states (e.g. 'non-breeding' and 'breeding').
671 They could consequently be viewed as manifestations of a threshold trait with labile plasticity,
672 where most or all individuals cross the threshold for reproduction at some point during a season.
673 Quantitative genetic analyses that treat observed date as the focal trait are then effectively using
674 the environmental threshold model, implicitly invoking the assumption of zero liability-scale GxE.
675 This could hinder appropriate evolutionary inference, especially when the objective in studying
676 breeding date is to test for, or evaluate the implications of, non-zero GxE (e.g. Nussey et al. 2007;
677 Charmantier and Gienapp 2014). Further, 'date' is unlikely to be the primary driving
678 environmental variable; indeed mean breeding dates commonly vary substantially among years,
679 implying that liability-scale reaction norms shaping the transition to breeding respond to other
680 environmental variable(s) (e.g. Lof et al. 2012; Inouye et al. 2019; de Villemereuil et al. 2020).

Micro-evolutionary analyses that treat observed event dates as continuously distributed traits, if in fact the true underlying biology is a threshold trait with labile plasticity, might consequently give incomplete or erroneous predictions of evolutionary dynamics, at least under some circumstances. Indeed, breeding dates often do not show expected micro-evolutionary responses to observed directional selection, which has been suggested to reflect additional environmental or genetic constraints (de Villemereuil et al. 2020). Yet, while studying phenology in terms of time to event and associated reaction norms rather than simply observed events has been advocated (Gienapp et al. 2005; Inouye et al. 2019), the merits and practicalities of treating events as manifestations of a plastic threshold trait have scarcely been explicitly considered.

Overall, therefore, future empirical ambitions should be to fully exploit recent conceptual and empirical developments in evolutionary quantitative genetics to explicitly estimate key parameters defining liability-scale reaction norms and resulting expression of threshold traits, and estimate genetic covariances with other traits and components of fitness. Such analyses will require substantial data on discrete phenotypes of relatives and non-relatives distributed across representative environments, and will therefore require appropriate multivariate axes of environmental variation driving developmental and/or labile plasticity to be measured or imposed. Such advances will be challenging, but will be necessary in order to predict how complex dynamics of discrete traits can drive phenotypic, population and evolutionary responses to environmental variation and directional change.

Literature cited

Acker, P., F. Daunt, S. Wanless, S.J. Burthe, M.A. Newell, M.P. Harris, C. Gunn, R. Swann, A. Payo-Payo, and J.M. Reid. Hierarchical variation in phenotypic flexibility across timescales and associated survival selection shape the dynamics of partial seasonal migration. In review.

Agrawal, A.A., C. Laforsch, and R. Tollrian. 1999. Transgenerational induction of defences in animals and plants. Nature 401:60-63.

Arnold, P.A., L.E.B. Kruuk, and P.A. Nicotra. 2019b. How to analyse plant phenotypic plasticity in response to a changing climate. New Phytol. 222:1235-1241.

Arnold, P.A., A.B. Nicotra, and L.E.B. Kruuk. 2019a. Sparse evidence for selection on phenotypic plasticity in response to temperature. Phil. Trans. R. Soc. B 374:20180185.

Ashander, J., L.M. Chevin, and M. Baskett. 2016. Predicting rescue via evolving plasticity in stochastic environments. Proc. R. Soc. B 283:20161690.

Auld, J.R., A.A. Agrawal, and R.A. Relyea. 2010. Re-evaluating the costs and limits of adaptive phenotypic plasticity. Proc. R. Soc. B 277:503-511.

Barson, N.J., T. Aykanat, K. Hindar, M. Baranski, G.H. Bolstad, P. Fiske, C. Jacq, +15 co-authors. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature 528:405-408.

Beaman, J.E., C.R. White, and F. Seebacher. 2016. Evolution of plasticity: Mechanistic link between development and reversible acclimation. Trends Ecol. Evol. 31:237-249.

Bearhop, S., W. Fiedler, R.W. Furness, S.C. Votier, S. Waldron, J. Newton, G.J. Bowen, P. Berthold, and K. Farnsworth. 2005. Assortative mating as a mechanism for rapid evolution of a migratory divide. Science 310:502–504.

725 Bégin, M., and D.A. Roff. 2002. The common quantitative genetic basis of wing morphology and
726 diapause occurrence in the cricket *Gryllus veletis*. Heredity 89:473-479.
727 Berthold, P. 1988. Evolutionary aspects of migratory behaviour in European warblers. J. Evol. Biol.
728 1:195-209.
729 Brodersen, J., B.B. Chapman, P.A. Nilsson, C. Skov, L.-A. Hansson, and C. Brönmark. 2014. Fixed
730 and flexible: coexistence of obligate and facultative migratory strategies in a freshwater fish.
731 PLoS One 9:e90294.
732 Buoro, M., O. Gimenez, and E. Prévost. 2012. Assessing adaptive phenotypic plasticity by means
733 of conditional strategies from empirical data: the latent environmental threshold model.
734 Evolution 66:996-1009.
735 Buzatto, B.A., M. Buoro, W.N. Hazel, and J.L. Tomkins. 2015. Investigating the genetic
736 architecture of conditional strategies using the environmental threshold model. Proc. R. Soc.
737 B 282:20152075.
738 Charmantier, A., and P. Gienapp. 2014. Climate change and timing of avian breeding and
739 migration: evolutionary versus plastic changes. Evol. Appl. 7:15-28.
740 Chevin, L.-M., and A.A. Hoffmann. 2017. Evolution of phenotypic plasticity in extreme
741 environments. Phil. Trans. R. Soc. B 372:20160138.
742 Chevin, L.-M., and R. Lande. 2013. Evolution of discrete phenotypes from continuous norms of
743 reaction. Am. Nat. 182:13-27.
744 Chevin, L.-M., R. Lande, and G.M. Mace. 2010. Adaptation, plasticity and extinction in a changing
745 environment: towards a predictive theory. PLoS Biol. 8:e1000357.

746 Crocker-Buta, S.P., and C.J. Leary. 2018. Bidirectionality of hormone-behavior relationships and
747 satellite-caller dynamics in green treefrogs. Behav. Ecol. 29:948-957.
748 Czorlich, Y., T. Aykanat, J. Erkinaro, P. Orell, and C.R. Primmer. 2018. Rapid sex-specific evolution
749 of age at maturity is shaped by genetic architecture in Atlantic salmon. Nature Ecol. Evol.
750 2:1800-1807.
751 de Jong, G. 2005. Evolution of phenotypic plasticity: patterns of plasticity and the emergence of
752 ecotypes. New Phytol. 166:101-118.
753 de Jong, G., and S. Gavrilets. 2000. Maintenance of genetic variation in phenotypic plasticity: the
754 role of environmental variation. Genet. Res. 76:295-304.
755 de Villemereuil, P., A. Charmantier, D. Arlt, P. Bize, P. Brekke, L. Brouwer, A. Cockburn, +27 co-
756 authors. 2020. Fluctuating optimum and temporally variable selection on breeding date in
757 birds and mammals. Proc. Ntl. Acad. Sci. 117:31969-31978.
758 de Villemereuil, P., H. Schielzeth, S. Nakagawa, and M. Morrissey. 2016. General methods for
759 evolutionary quantitative genetic inference from generalized mixed models. Genetics
760 204:1281-1294.
761 de Zoeten, T., and F. Pulido. 2020. How migratory populations become resident. Proc. R. Soc. B.
762 287:20193011.
763 Debes, P.V., N. Piavchenko, J. Erkinaro, and C.R. Primmer. 2020. Genetic growth potential, rather
764 than phenotypic size, predicts migration phenotype in Atlantic salmon. Proc. R. Soc. B
765 287:20200867.
766 Dempster, E.R., and I.M. Lerner. 1950. Heritability of threshold characters. Genetics 35:212-236.

767 DeWitt, T.J., A. Sih, and D.S. Wilson. 1998. Costs and limits of phenotypic plasticity. Trends Ecol.
768 Evol. 13:77-81.
769 Dingemanse, N.J., A.J.N. Kazem, D. Réale, and J. Wright. 2010. Behavioural reaction norms:
770 animal personality meets individual plasticity. Trends Ecol. Evol. 25:81-89.
771 Dingemanse, N.J., and M. Wolf. 2013. Between-individual differences in behavioural plasticity
772 within populations: causes and consequences. Anim. Behav. 85:1031-1039.
773 Dodson, J.J., N. Aubin-Horth, V. Thériault, and D.J. Páez. 2013. The evolutionary ecology of
774 alternative migratory tactics in salmonid fishes. Biol. Rev. 88:602-625.
775 Donelson, J.M., S. Salinas, P.L. Munday, and L.N.S. Shama. 2018. Transgenerational plasticity and
776 climate change experiments: where do we go from here? Glob. Change Biol. 24:13–34.
777 Duthie, A.B., G. Bocedi, and J.M. Reid. 2016. When does female multiple mating evolve to adjust
778 inbreeding? Effects of inbreeding depression, direct costs, mating constraints, and polyandry
779 as a threshold trait. Evolution 70:1927-1943.
780 Eggeman, S.L., M. Hebblewhite, H. Bohm, J. Whittington, and E.H. Merrill. 2016. Behavioural
781 flexibility in migratory behaviour in a long-lived large herbivore. J. Anim. Ecol. 85:785-797.
782 Ehrenreich, I.M., and D.W. Pfennig. 2016. Genetic assimilation: a review of its potential proximate
783 causes and evolutionary consequences. Annals Bot. 117:769-779.
784 Engvist, L., and M. Taborsky. 2016. The evolution of genetic and conditional alternative
785 reproductive tactics. Proc. R. Soc. B 283:20152945.
786 Falconer, D.S., and T. Mackay. 1996. Introduction to quantitative genetics. Prentice Hall.
787 Gavrilets, S., and S.M. Scheiner. 1993a. The genetics of phenotypic plasticity. V. Evolution of
788 reaction norm shape. J. Evol. Biol. 6:31–48.

789 Gavrilets, S., and S.M. Scheiner. 1993b. The genetics of phenotypic plasticity. VI. Theoretical
790 predictions for directional selection. J. Evol. Biol. 6:49-68.
791 Germain, R.R., M.E. Wolak, and J.M. Reid. 2018. Individual repeatability and heritability of divorce
792 in a wild population. Biol. Let. 14:20180061.
793 Ghalambor, C.K., J.K. McKay, S.P. Carroll, and D.N. Reznick. 2007. Adaptive versus non-adaptive
794 phenotypic plasticity and the potential for contemporary adaptation in new environments.
795 Func. Ecol. 21:394-407.
796 Gianola, D. 1982. Theory and analysis of threshold characters. J. Anim. Sci. 54:1079-1096.
797 Gienapp, P., L. Hemerik, and M.E. Visser. 2005. A new statistical tool to predict phenology under
798 climate change scenarios. Glob. Change Biol. 11:600-606.
799 Gomulkiewicz, R., and M. Kirkpatrick. 1992. Quantitative genetics and the evolution of reaction
800 norms. Evolution 46:390–411.
801 Grossen, C., S. Neuenschwander, and N. Perrin. 2010. Temperature-dependent turnovers in sex-
802 determination mechanisms: a quantitative model. Evolution 65:64-78.
803 Hadfield, J.D. 2015. Increasing the efficiency of MCMC for hierarchical phylogenetic models of
804 categorical traits using reduced mixed models. Meth. Ecol. Evol. 6:706-714.
805 Harrison, P.M., L.F.G. Gutowsky, E.G. Martins, D.A. Patterson, S.J. Cooke, and M. Power. 2017.
806 Partial diel migration: A facultative migration underpinned by long-term inter-individual
807 variation. J. Anim. Ecol. 86:1246-1256.
808 Hazel, W., R. Smock, and C.M. Lively. 2004. The ecological genetics of conditional strategies. Am.
809 Nat. 163:888-900.

Hendry, A.P. 2016. Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. Heredity 107:25–41.

Inouye, B.D., J. Ehrlén, and N. Underwood. 2019. Phenology as a process rather than an event: from individual reaction norms to community metrics. Ecol. Monog. 89:e01352.

Kelly, M. 2019. Adaptation to climate change through genetic accommodation and assimilation of plastic phenotypes. Phil. Trans. R. Soc. B 374:20181076.

Lande, R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. J. Evol. Biol. 22:1435-1446.

Lande, R. 2014. Evolution of phenotypic plasticity and environmental tolerance of a labile quantitative character in a fluctuating environment. J. Evol. Biol. 27:866–875.

Lande, R. 2015. Evolution of phenotypic plasticity in colonizing species. Mol. Ecol. 24:2038-2045.

Lande, R. 2019. Developmental integration and evolution of labile plasticity in a complex quantitative character in a multiperiodic environment. Proc. Ntl. Acad. Sci. 116:11361-11369.

Lof, M.E., T.E. Reed, J.M. McNamara, and M.E. Visser. 2012. Timing in a fluctuating environment: environmental variability and asymmetric fitness curves can lead to adaptively mismatched avian reproduction. Proc. R. Soc. B 279:3161–3169.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer.

Moorad, J.A., and T.A. Linksvayer. 2008. Levels of selection on threshold characters. Genetics 179:899-905.

Morrissey, M.B., and M. Liefting. 2016. Variation in reaction norms: statistical considerations and biological interpretation. Evolution 70:1944-1959.

831 Murren, C.J., H.J. Maclean, S.E. Diamond, U.K. Steiner, M.A. Heskel, C.A. Handelsman, C.K.
832 Ghalambor, J.R. Auld, H.S. Callahan, D.W. Pfennig, R.A. Relyea, C.D. Schlichting, and J.
833 Kingsolver. 2014. Evolutionary change in continuous reaction norms. Am. Nat. 183:453-467.
834 Negussie, E., I. Strandén, S. Tsuruta, and E.A. Mäntysaari. 2012. Longitudinal threshold model
835 analysis of clinical mastitis using linear splines. Livestock Sci. 149:173-179.
836 Nussey, D.H., A.J. Wilson, and J.E. Brommer. 2007. The evolutionary ecology of individual
837 phenotypic plasticity in wild populations. J. Evol. Biol. 20:831-844.
838 Ostrowski, M.-F., P. Jarne, and P. David. 2000. Quantitative genetics of sexual plasticity: the
839 environmental threshold model and genotype-by-environment interaction for phallus
840 development in the snail *Bulinus truncatus*. Evolution 54:1614-1625.
841 Pigliucci, M. 2005. Evolution of phenotypic plasticity: where are we going now? Trends Ecol. Evol.
842 20:481–486.
843 Pulido, F. 2011. Evolutionary genetics of partial migration – the threshold model revis(it)ed. Oikos
844 120:1776-1783.
845 Pulido, F., P. Berthold, and A.J. van Noordwijk. 1996. Frequency of migrants and migratory activity
846 are genetically correlated in a bird population: evolutionary implications. Proc. Ntl. Acad. Sci.
847 93:14642-14647.
848 Radchuk, V., T. Reed, C. Teplitsky, M. van de Pol, A. Charmantier, C. Hassall, +59 co-authors. 2019.
849 Adaptive responses of animals to climate change are most likely insufficient. Nature Comms.
850 10:3109.

851 Ramakers, J.J.C., A. Culina, M.E. Visser, and P. Gienapp. 2018. Environmental coupling of
852 heritability and selection is rare and of minor evolutionary significance in wild populations.
853 Nature Ecol. Evol. 2:1093-1103.
854 Reid, J.M., M. Souter, S.R. Fenn, P. Acker, A. Payo-Payo, S.J. Burthe, S. Wanless, and F. Daunt.
855 2020. Among-individual and within-individual variation in seasonal migration covaries with
856 subsequent reproductive success in a partially-migratory bird. Proc. R. Soc. B 287:20200928.
857 Reid, J.M., J.M.J. Travis, F. Daunt, S.J. Burthe, S. Wanless, and C. Dytham. 2018. Population and
858 evolutionary dynamics in spatially-structured seasonally-varying environments. Biol. Rev.
859 93:1578-1603.
860 Roff, D.A. 1994a. Evolution of dimorphic traits: effect of directional selection on heritability.
861 Heredity 72:36-41.
862 Roff, D.A. 1994b. The evolution of dimorphic traits: predicting the genetic correlation between
863 environments. Genetics 136:395-401.
864 Roff, D.A. 1996. The evolution of threshold traits in animals. Q. Rev. Biol. 71:3-35.
865 Roff, D.A. 1998a. Evolution of threshold traits: the balance between directional selection, drift
866 and mutation. Heredity 80:25-32.
867 Roff, D.A. 1998b. The maintenance of phenotypic and genetic variation in threshold traits by
868 frequency-dependent selection. J. Evol. Biol. 11:513-529.
869 Roff, D.A., and D.J. Fairbairn. 2007. The evolution and genetics of migration in insects. BioScience
870 57:155-164.
871 Saltz, J.B., A.M. Bell, J. Flint, R. Gomulkiewicz, K.A. Hughes, and J. Keagy. 2018. Why does the
872 magnitude of genotype-by-environment interaction vary? Ecol. & Evol. 8:6342-6353.

873 Scheiner, S.M. 1993. Genetics and evolution of phenotypic plasticity. Annu. Rev. Ecol. Evol. Syst.
874 24:35-68.
875 Scheiner, S.M. 2002. Selection experiments and the study of phenotypic plasticity. J. Evol. Biol.
876 15:889-898.
877 Snell-Rood, E.C., M.E. Kobiela, K.L. Sikkink, and A.M. Shephard. 2018. Mechanisms of plastic
878 rescue in novel environments. Annu. Rev. Ecol. Evol. Syst. 49:331-354.
879 Tomkins, J.L., and W. Hazel. 2007. The status of the conditional evolutionarily stable strategy.
880 Trends Ecol. Evol. 22:522-528.
881 Valladares, F., D. Sanchez-Gomez, and M.A. Zavala. 2006. Quantitative estimation of phenotypic
882 plasticity: bridging the gap between the evolutionary concept and its ecological applications.
883 J. Ecol. 94:1103-1116.
884 van Buskirk, J., and U.K. Steiner. 2009. The fitness costs of developmental canalization and
885 plasticity. J. Evol. Biol. 22:852–860.
886 Walsh, B., and M.W. Blows. 2009. Abundant genetic variation + strong selection = multivariate
887 genetic constraints: a geometric view of adaptation. Annu. Rev. Ecol. Evol. Syst. 40:41-59.
888 Wennersten, L., and A. Forsman. 2012. Population-level consequences of polymorphism,
889 plasticity and randomized phenotype switching: a review of predictions. Biol. Rev. 87:756–
890 767.
891 West-Eberhard, M.J. 1989. Phenotypic plasticity and the origins of diversity. Annu. Rev. Ecol. Evol.
892 Syst. 20:249-278.
893 Wolf, M., G.S. van Doorn, and F.J. Weissing. 2008. Evolutionary emergence of responsive and
894 unresponsive personalities. Proc. Ntl. Acad. Sci. 105:15825-15830.

Wolf, M., and F.J. Weissing. 2010. An explanatory framework for adaptive personality differences. Phil. Trans. R. Soc. B 365:3959-3968.

Wright, S. 1934. The results of crosses between inbred strains of guinea pigs, differing in number of digits. Genetics 19:537-551.

Table 1. Summary of key terminology.

Term	Definition
Phenotypic plasticity	Ability of single genotypes to produce different phenotypes in different environments, resulting in directly observable phenotypic variation across environments.
Developmental plasticity	Plasticity in traits that are permanently and non-reversibly expressed once in an individual's life, through environment-dependent initial development. Directly evident when individuals from the same lineage are exposed to different initial environmental conditions, resulting in permanent *among-individual* variation.
Labile plasticity	Plasticity in traits that are repeatedly and hence potentially reversibly re-expressed multiple times through an individual's life. Directly evident when individuals experience changing environmental conditions, resulting in longitudinal *within-individual* variation.
Threshold trait	A trait that shows two (or more) discrete phenotypic states, where phenotypic expression directly depends on the value of an underlying 'liability' relative to a threshold.
Liability	A latent quantitative genetic trait that can comprise genetic and environmental components, and that translates deterministically into expression of alternative discrete phenotypes when above versus below a threshold.

Latent plasticity	Environmentally-induced variation in liability. Such latent plasticity is not necessarily manifested as observable phenotypic plasticity, depending on whether it causes the liability to cross the threshold. Latent plasticity can be developmental and/or labile.

935 **Table 2.** Summary of key differences between typically assumed properties of traits that are
936 continuously distributed on observed phenotypic scales (left column) versus properties of
937 discrete threshold traits (right column), regarding (A) genetic variation in phenotypic plasticity,
938 (B) phenotypic variation in phenotypic plasticity; (C) selection on phenotypic plasticity; and (D)
939 apparent inheritance of phenotypic plasticity.
940

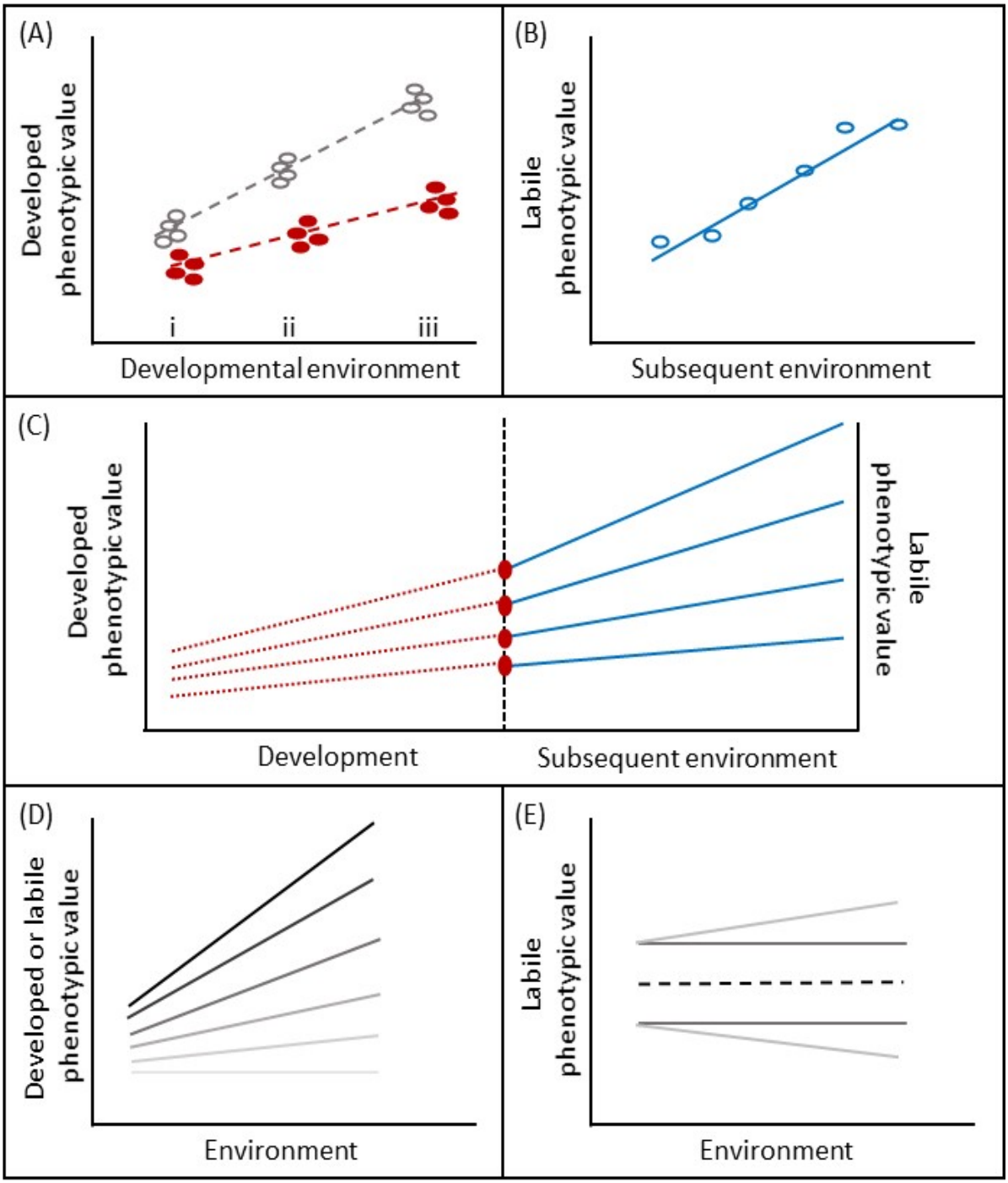
	Continuously distributed traits	Threshold traits
(A)	• Genetic variation in phenotypic plasticity requires genetic variation in reaction norm slopes, and hence requires gene-by-environment interactions (GxE) on the observed phenotypic scale.	• Genetic variation in phenotypic plasticity could result from genetic variation in liability-scale reaction norm intercepts. It does not necessarily require genetic variation in reaction norm slopes, or hence require gene-by-environment interactions (GxE), on the liability scale.
	• Such GxE, and hence genetic variation in phenotypic plasticity, will not necessarily be present.	• If there is genetic variation in liability intercept, and phenotypic variation, there is likely to be genetic variation in phenotypic plasticity.
	• GxE could be eroded by directional or stabilising selection on phenotypic plasticity, and hence on reaction norm slope. Long-term maintenance of additive genetic variation (and hence phenotypic variation) in phenotypic plasticity may therefore require further explanations.	• Substantial additive genetic variation in intercepts and/or slopes of liability-scale reaction norms could potentially be maintained even given consistent strong directional selection. Genetic variation in phenotypic plasticity may therefore be revealed given environmental variation or change, and require little further explanation.

	• With non-zero GxE and linear reaction norms and stabilizing selection on phenotypes in typical environments, additive genetic variation in phenotype can readily be greater in extreme environments, facilitating rapid evolutionary responses to selection.	• Genetic variation in phenotype will not necessarily increase, and could easily decrease towards zero, in extreme environments.
(B)	• A persistent mixture of highly phenotypically plastic and strongly environmentally canalized genotypes or individuals requires some further explanation (e.g. disruptive or negative frequency-dependent selection on reaction norm slope). • The degree of phenotypic plasticity or canalization defined by the reaction norm slope can, in principle, be independent of the reaction norm intercept, and hence of population mean phenotype.	• A persistent mixture of highly phenotypically plastic and strongly environmentally canalized genotypes or individuals is likely. It requires little further explanation beyond some mechanism maintaining variation in liability that spans the threshold, and resulting expression of each alternative phenotype. • The mean degree of canalization (and hence repeatability) of a particular discrete phenotype is linked to the mean liability-scale reaction norm intercept, and is expected to increase with phenotype frequency.
(C)	• Consistent directional selection for or against phenotypic plasticity will cause consistent directional selection on reaction norm slope.	• Consistent directional selection for or against phenotypic plasticity could potentially cause stabilising or disruptive selection on liability intercept. The form of selection on liability could change as evolution progresses.

	• Such selection could gradually erode GxE.	• Such selection could potentially drive evolution of complex forms of liability-scale GxE and intercept-slope covariances.
(D)	• Reproduction between highly canalized parents, or parents with opposite reaction norm slopes, will typically result in highly phenotypically canalized offspring. Highly canalized parents cannot readily produce highly phenotypically plastic offspring. • Mutation or gene flow would be required to regenerate plasticity in canalized populations.	• Reproduction between highly phenotypically canalized parents with different fixed phenotypes can readily result in highly phenotypically plastic offspring. Selection for or against plasticity may consequently impose indirect selection on mate choice. • Phenotypic plasticity (or canalization) could be hard to eradicate even given strong selection against it.

Figure 1. Illustrations of phenotypic plasticity in continuously distributed quantitative traits conceptualised as: (A) the change in mean phenotype when different individuals from split families or clones are exposed to different discrete environments (labelled i-iii), typically focussing on developmental plasticity. Red and grey points depict individuals from two different lineages. Dashed lines depict the lineage reaction norm for developmental plasticity, assumed to be effectively linear. Here, the grey lineage has a higher reaction norm intercept (and hence mean phenotypic value) and also a steeper reaction norm slope, implying greater developmental plasticity; (B) the change in phenotype when a single individual is exposed to different environments within a continuous range of variation, representing labile plasticity. Blue points depict repeat observations of the individual, and the solid line depicts the individual reaction norm for labile plasticity inferred from linear regression of phenotype on environment, with micro-environmental deviations; (C) overall reaction norm intercepts, slopes and their covariance measured among and within four individuals. Red points depict individuals' developed phenotypes in a reference environment (as in A, shown here as the intercepts of individuals' overall reaction norms, but could be standardised to a mean environment). Red dotted lines depict the development process, shaped by the intercept and slope of the underlying reaction norm for developmental plasticity. Blue lines depict individuals' reaction norms for labile plasticity (as in B). Non-parallel lines could indicate gene-by-environment interactions (GxE, Supporting Information S2). Slopes of reaction norms for developmental and labile plasticity could also be correlated. (D) Scenario of directional selection on reaction norm slope, where lineages or individuals with steep positive slopes have higher (or lower) fitness than those with shallow slopes, with relative fitness denoted by differing shades of grey. Lower fitness lineages

would gradually be selected out, reducing GxE. In the depicted example, directional selection on slope could also cause or be caused by indirect selection through correlated intercepts. (E) Expected patterns of additive genetic inheritance, where pairs of parents with different reaction norm intercepts and slopes close to zero (dark grey lines) or slopes with opposite signs (light grey lines) produce offspring with intermediate intercepts and slopes close to zero (dashed line).



981

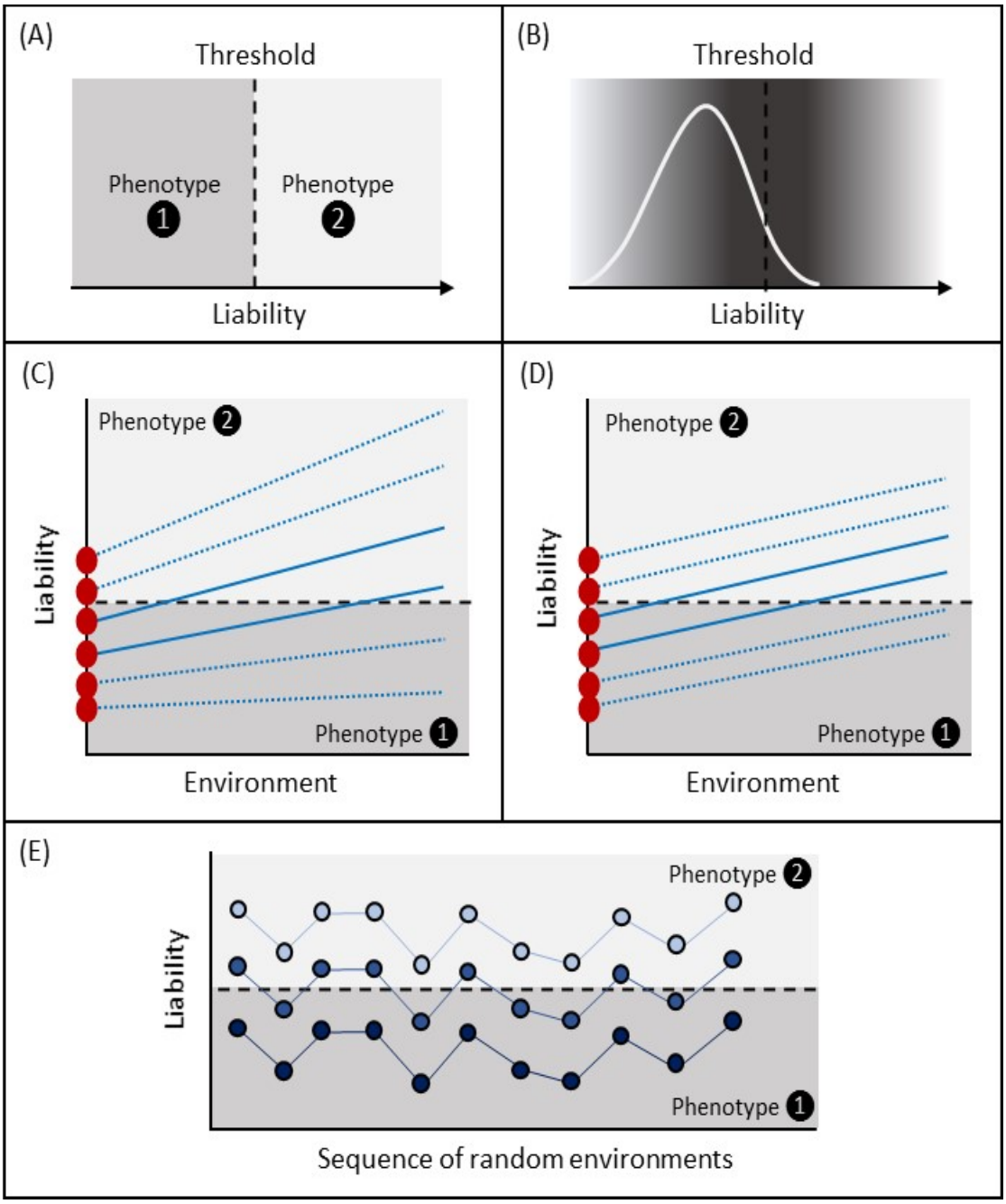
982

983

984

Figure 2. (A) Basic concept of a threshold trait, where liability values below and above a threshold (dashed line) translate deterministically into discrete phenotypes ❶ and ❷ (zones of lighter and darker grey shading) respectively. This concept can be extended to traits with multiple discrete phenotypes, given further thresholds. (B) Individuals with initial liabilities near the threshold (dark grey shading) are likely to be phenotypically plastic, while individuals with initial liabilities far from the threshold (light grey shading) are likely to be phenotypically canalized. Given an approximately Gaussian distribution of liabilities centred away from the threshold (white curve) the more frequently expressed phenotype will include more individuals with liabilities further from the threshold, which are therefore more likely to be phenotypically canalized (i.e. repeatable). (C and D) Individual liabilities in relation to environmental variation (i.e. liability-scale reaction norms), comprising intercepts (red points, which can reflect outcomes of developmental plasticity in liability) and slopes (blue lines, representing labile plasticity in liability). These attributes are analogous to those for continuous phenotypic traits (Figure 1C). Black dashed lines demarcate thresholds distinguishing expression of discrete phenotypes ❶ and ❷. Individuals with reaction norms that cross the threshold within the observed range of environmental variation show phenotypic plasticity (solid blue lines), while individuals with reaction norms that do not cross the threshold are phenotypically canalized (dotted blue lines). There might or might not be among-individual variation in reaction norm slopes for labile plasticity in liability, representing liability-scale GxE (C versus D). (E) Illustration of high repeatability of both alternative phenotypes and of phenotypic plasticity. Dark, mid and light blue points represent liability values, and resulting observed phenotypes (❶ and ❷), across a timeseries (sequence of occasions) for three different individuals with different liability

intercepts. All three individuals experience similar regimes of environmental variation across occasions and hence show similar variation in liability, representing responses through similar liability-scale reaction norm slopes. Nevertheless, the light and dark blue individuals are repeatable (i.e. canalized) for phenotypes ❶ and ❷ respectively, while the mid blue individual is repeatably phenotypically plastic (i.e. regularly switches between alternative phenotypes between consecutive occasions).



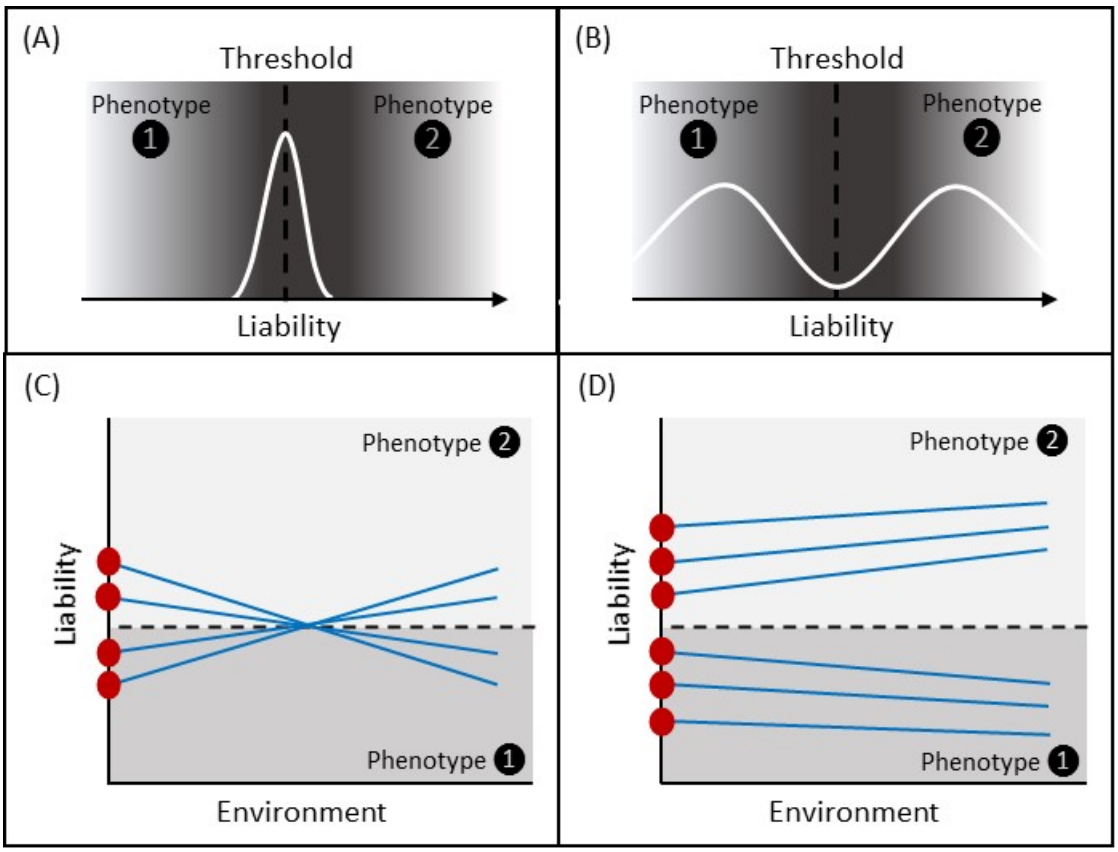
1014

1015

1016

1017

Figure 3. Potential consequences of selection (A and C) for and (B and D) against expression of labile phenotypic plasticity in a threshold trait for evolution of liability-scale reaction norm (A and B) intercepts and (C and D) slopes. On A and B, dark and light grey shading respectively indicate zones where individuals are likely to be phenotypically plastic or canalized respectively. White curves depict potential distributions of liability-scale reaction norm intercepts following strong selection (A) for and (B) against phenotypic plasticity, resulting in stabilising and disruptive selection on intercept respectively. On C and D, phenotypic plasticity occurs when an individual's liability-scale reaction norm crosses the threshold (black dashed line), causing changed expression of phenotype ❶ versus ❷ (darker versus lighter grey zones). Reaction norm intercepts and slopes (depicted by red points and blue lines) could become (C) negatively or (D) positively correlated due to selection for or against phenotypic plasticity. Intercepts might in turn be shaped by the intercept and slope of the reaction norm for developmental plasticity, and reaction norms for labile and/or developmental plasticity could also be non-linear.


(A)
Threshold
Phenotype 1
Phenotype 2
Liability
(B)
Threshold
Phenotype 1
Phenotype 2
Liability
(C)
Phenotype 2
Liability
Phenotype 1
Environment
(D)
Phenotype 2
Liability
Phenotype 1
Environment

1032

Supporting Information

Properties of phenotypic plasticity in discrete threshold traits

Jane M. Reid and Paul Acker

Appendix S1. Basic quantitative genetic formulation of a trait showing both developmental and labile plasticity

Basic formulation of developmental plasticity

In evolutionary quantitative genetics, researchers focusing on developmental plasticity (Table 1, sometimes termed 'fixed plasticity' or 'one-shot plasticity') commonly formulate the observed value of a focal trait as:

$$z_{Di} = g_{0Di} + g_{1Di}.\varepsilon_D + e_{Di} \qquad \text{Equation 1}$$

Here, z_{Di} is the developed (i.e. observed) trait value for individual i, and g_{0Di} is individual i's breeding value for the trait (i.e. the elevation, giving the expected developed phenotype in some reference environment). The slope g_{1Di} describes the change in breeding value when individual i is exposed to some early-life developmental environment ε_D, representing a reaction norm for developmental plasticity (assumed to be linear for simplicity). In addition, e_{Di} is a random deviate affecting development and the resulting trait value (e.g. Gavrilets and Scheiner 1993a; de Jong and Gavrilets 2000; Lande 2009, 2015; Chevin and Lande 2010). This deviate is itself a compound effect, which could reflect non-additive genetic effects and/or micro-environmental impacts, and could equally be envisaged as the single phenotypic outcome of any deviation in reaction norm

slope from the breeding value g_{1Di} (e.g. Tonsor et al. 2013). Across a population, g_{0D}, g_{1D} and e_D are typically assumed to be normally distributed with variances that can be estimated. The genetic parameters g_{0D} and g_{1D} could covary, but e_D is assumed to be independent of both.

Basic formulation of labile plasticity

Researchers focusing on labile plasticity (Table 1) commonly formulate observed values of a focal trait as:

$$z_{Lit} = b_{0Li} + b_{1Li}.\varepsilon_{Lt} + e_{Lit} \quad \text{Equation 2}$$

Here, z_{Lit} is the observed labile trait value for individual i at time t, and b_{0Li} is individual i's intercept or elevation for the focal trait (again typically standardised to some reference environment). Meanwhile, b_{1Li} is individual i's reaction norm slope for labile plasticity and describes the change in trait value when the individual is exposed to some post-development environment ε_{Lt} at time t, again representing a linear reaction norm for simplicity (e.g. Dingemanse et al. 2010; Arnold et al. 2019). In addition, e_{Lit} is a random deviate affecting post-development trait expression at time t, reflecting micro-environmental impacts.

Given repeated observations on individuals and their relatives, both the intercept and the slope of individuals' reaction norms for labile plasticity can be partitioned into additive genetic and broadly-defined environmental effects, giving:

$$z_{Lit} = (g_{0Li} + e_{0Li}) + (g_{1Li} + e_{1Li}).\varepsilon_{Lt} + e_{Lit} \quad \text{Equation 3}$$

Here, g_{0Li} and g_{1Li} are individual i's breeding values for the reaction norm intercept and slope, while e_{0Li} and e_{1Li} are individual i's environmental deviations for the intercept and slope, which

could include non-additive genetic effects and/or permanent environmental effects resulting from the individual's developmental environment. Across a population, reaction norm intercepts and slopes could therefore covary due to genetic covariance (i.e. between g_{0L} and g_{1L}) and/or due to environmental covariance (i.e. between e_{0L} and e_{1L}). Further, there could potentially be gene-by-environment interactions, such that additive genetic effects on the reaction norm slope for labile plasticity (g_{1L}) partly depend on the environmental deviation for the intercept (e_{0L}).

Joint formulation of both developmental and labile plasticity

The standard conceptualisation of labile plasticity (above) assumes that genetic and environmental effects on the reaction norm intercept are additive, with no gene-by-environment interactions. Yet in practice, the full process of initial development, as represented by a developmental reaction norm, could determine the intercept of the reaction norm for labile plasticity. Observed individual reaction norm intercepts could therefore reflect complex underlying additive genetic and environmental effects manifested through developmental plasticity (see also Figure 1C).

In the simplest form, such effects of developmental plasticity on the intercept of the reaction norm for subsequent labile plasticity can be considered by substituting equation 1 into the intercept of equation 3 to give:

$$z_{Lit} = z_{Di} + (g_{1Li} + e_{1Li}).\varepsilon_{Lt} + e_{Lit}$$

$$z_{Lit} = (g_{0Di} + g_{1Di}.\varepsilon_{D} + e_{Di}) + (g_{1Li} + e_{1Li}).\varepsilon_{Lt} + e_{Lit} \qquad \text{Equation 4}$$

Equation 4 represents an expression for the value of individual i's labile phenotype at time t that explicitly encompasses genetic effects on developmental plasticity. Now, the quantitative genetic architecture of trait z is underpinned by a three-parameter G-matrix, comprising additive genetic effects on the intercept and slope for developmental plasticity and the slope for labile plasticity and their covariances, such that:

$$\begin{bmatrix} \sigma^2_{g0D} & \sigma_{g0D,g1D} & \sigma_{g0D,g1L} \\ \sigma_{g1D,g0D} & \sigma^2_{g1D} & \sigma_{g1D,g1L} \\ \sigma_{g1L,g0D} & \sigma_{g1L,g1D} & \sigma^2_{g1L} \end{bmatrix}$$

Here, σ^2_x and $\sigma_{x,y}$ represent the variance in x and the covariance between x and y respectively. In addition, there could be gene-by-environment interactions such that additive genetic effects on the reaction norm slope for labile plasticity (g_{1L}) partly depend on the full environmental deviation for the intercept stemming from developmental plasticity (given by $g_{1D}.\varepsilon_D+e_D$). There can be no gene-by-environment interactions such that additive genetic effects on the intercept and slope for developmental plasticity (g_{0D} and g_{1D}) depend on environmental deviations for labile plasticity, since such deviations occur after development has been completed. However, the environmental deviation for the reaction norm slope for labile plasticity (e_{1L}) could also potentially depend on the environmental deviation stemming from developmental plasticity.

Extension from current typical treatments

This three-parameter conceptualisation of overall plasticity extends previous quantitative genetic conceptualisations of linear reaction norms, which are effectively two-parameter models. Specifically, numerous theoretical and empirical studies effectively consider intercepts and slopes of reaction norms for developmental plasticity, but do not consider slopes for labile plasticity at all (Gavrilets and Scheiner 1993a,b; Scheiner 2002; Lande 2009). Meanwhile, studies that consider labile plasticity typically consider the overall reaction norm intercept and slope as two parameters, and do not explicitly consider parameters describing the form of developmental plasticity that could underlie the intercept (e.g. Nussey et al. 2007; Dingemanse et al. 2010; Arnold et al. 2019a; but see Lande 2019). They thereby implicitly assume zero additive genetic variation in developmental reaction norm slope, such that developmental effects on overall reaction norm intercepts can be captured as purely additive genetic and environmental effects across individuals. This assumption may sometimes be biologically appropriate, and certainly generates tractable theoretical and statistical models (e.g. Lande 2014; Morrissey and Liefting 2016). However, it contradicts considerable empirical evidence that there is often non-zero additive genetic variance in reaction norm slopes for developmental plasticity (Scheiner 1993, 2002; Pigliucci 2005). It also excludes the possibility of explicitly considering non-zero genetic covariance between reaction norm slopes for developmental and labile plasticity. Fully understanding overall phenotypic and evolutionary dynamics involving plasticity is therefore ultimately likely to require joint consideration of both developmental and labile plasticity and any shared quantitative genetic basis. Equations 3 and 4 also indicate that estimates of g_{0D} and g_{1L} could potentially be biased in empirical quantitative genetic analyses where g_{1D} is not explicitly

modelled. Such effects, and the form of any bias, will depend on the form of the full G-matrix, and on patterns of developmental environmental experiences across relatives.

Application to threshold traits

The reaction norm formulations for developmental plasticity (Equation 1), labile plasticity (Equations 2 and 3) or both (Equation 4) could be envisaged to directly act on the observed phenotypic scale for any continuously distributed focal trait z, or on the latent liability scale of a threshold trait. In the latter case, alternative phenotypes P_1 and P_2 arise when an individual's value of z is less than or greater than some defined threshold value T (which is typically taken as zero). Formally:

If $z_i < T$ then P_1

If $z_i \geq T$ then P_2

Expression of the alternative phenotypes P_1 and P_2 is therefore deterministic given the liability value z_i. Consequently, unlike other conceptual models that directly consider dichotomous traits as outcomes of Bernoulli trials given some underlying probability, the threshold trait model does not necessarily directly envisage probabilistic outcomes or associated intrinsic phenotypic variances. Specifically, if the underlying environmental variable ε varies highly predictably, for example reflecting strong seasonality, then phenotypic expression of P_1 versus P_2 will also vary highly predictably conditional on the liability-scale reaction norm parameters. But, if the underlying environmental variable ε varies stochastically and can be taken as a random variable with some mean and variance, then each genotype can be envisaged as having a probability of

expressing phenotypes P_1 or P_2 at any time, which can be calculated given parameters describing reaction norms and environmental variation.

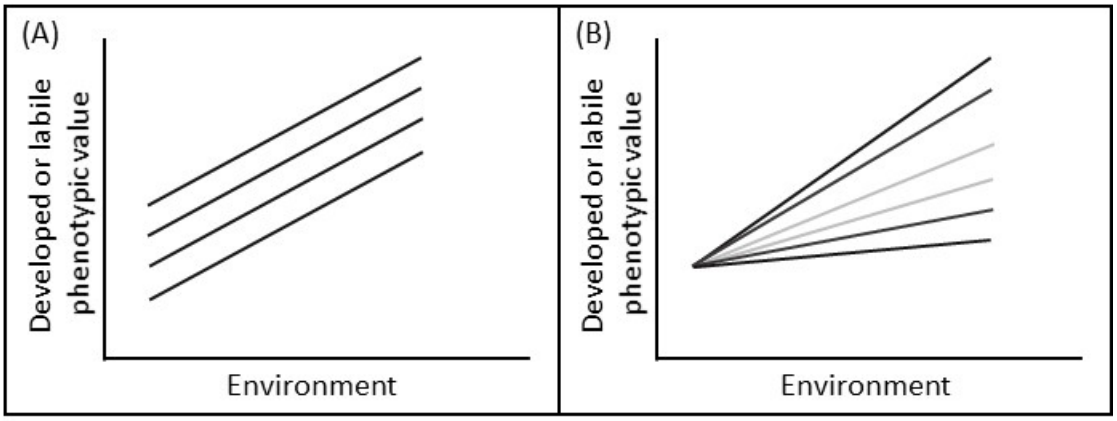
Complexities of conceptualisation and estimation

Equations 1-4 set out a basic framework to conceptualise overall individual reaction norms in terms of both developmental and labile plasticity. However, in practice, numerous complexities could arise in terms of conceptualising overall reaction norms, and estimating key parameters. Reaction norms for developmental and/or labile plasticity might not be simply linear, requiring extra parameters to describe environmental response functions (e.g. Gavrilets and Scheiner 1993; Arnold et al. 2019) which could in turn show genetic covariances. The key environmental variables driving developmental and labile plasticity (ε_D and ε_L) might or might not be the same and/or could be multivariate, yielding multi-dimensional plasticity (e.g. Westneat et al. 2019) that could affect both the intercept and slope of the overall reaction norm. The intercept (and hence elevation) could be mean-centred, as is standard statistical practice. However, it might make more biological sense to define the intercept as the initial trait expression following development. This could particularly apply if there is autocorrelation in successively expressed values of a labile trait, such that the current value depends partly on previous value(s) (as could arise due to multiple forms of cognitive learning and/or physiological memory). Such alternative definitions of intercepts could substantially affect estimates of intercept-slope covariances. Further, the slope of the reaction norm for labile plasticity might not be fixed across an individual's life but, for example, could be partly age-dependent and/or experience-dependent (e.g. Schaeffer 2011). Finally, regarding transformation of reaction norms from latent to observed

phenotypic scales for threshold traits, the form of the threshold function could potentially also vary and evolve over time (e.g. Chevin and Lande 2013). Future challenges will be to develop appropriate protocols for such analyses, which might need to include methods for reducing dimensionality.

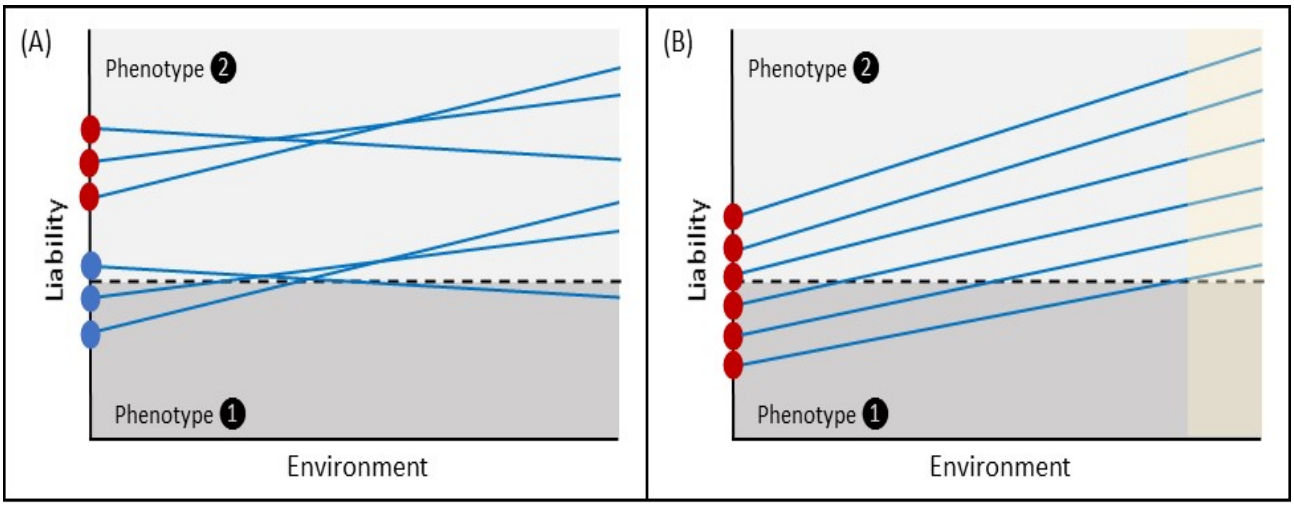
Appendix S2. Additional illustrations of forms of phenotypic plasticity in continuously distributed traits

Figure S1. (A) If linear reaction norms for developmental or labile plasticity in a continuously distributed trait are parallel with no genetic variation in slope, then there is no GxE and hence no immediate potential for further evolution of plasticity. [Note that in wild populations non-zero GxE could potentially exist even with zero phenotypic variation in slope, since genetic variation could in principle be completely masked by counteracting environmental variation (Nussey et al. 2007).] (B) Substantial genetic variation in reaction slopes, and mixtures of environmentally canalized and highly plastic individuals, could in principle be maintained by disruptive selection on slopes such that entities with slopes close to or far from zero have higher fitness (dark grey lines) than entities with intermediate slopes (light grey lines). Similarly, disruptive selection could in principle act to favour entities with very positive or negative reaction norm slopes against individuals with slopes close to zero. Either way, such disruptive selection could potentially generate bimodal distributions of phenotypes in extreme environments, but is likely to require some additional mechanism to favour alternative phenotypes.



Appendix S3. Additional illustrations of forms of latent plasticity and phenotypic plasticity in threshold traits

Figure S2. (A) Further illustration that variation in phenotypic plasticity in a threshold trait can arise solely due to variation in liability-scale reaction norm intercepts rather than necessarily slopes. Two groups of individuals have different mean liability intercepts (red and blue), but identical slopes (blue lines). Nevertheless, all blue individuals are phenotypically plastic and express both alternative phenotypes across the observed range of environmental variation, while all red individuals are canalized for phenotype ❷ across the same range of variation. The presence of latent plasticity, and variation in latent plasticity, consequently does not necessarily translate into phenotypic plasticity. (B) Illustration that for a threshold trait with GxE manifested as variation in linear reaction norm slopes on the liability scale, phenotypic variation (and hence genetic variation in phenotype) can decrease rather than increase in extreme environments. Here, all individuals express phenotype ❷ once the environment becomes sufficiently extreme (yellow shaded zone).


(A)
Phenotype 2
Liability
Phenotype 1
Environment
(B)
Phenotype 2
Liability
Phenotype 1
Environment

Appendix S4. Illustration of apparent sex-specific genetic dominance reversal

Wright (1934) highlighted that spurious evidence of apparent genetic dominance can arise from observed phenotypic variation in a developmental threshold trait in offspring of crosses between different inbred parental lines. Specifically, patterns of phenotypic resemblance that are consistent with genetic dominance can arise without any strict dominance because mean offspring liability could fall above or below the threshold, generating different phenotypes, depending on the mean liabilities of the two parental lines. Offspring phenotypes will then resemble one parental line more than the other, even with purely additive genetic effects. Here, we illustrate circumstances where spurious evidence of sex-specific genetic dominance reversals at a large effect locus could potentially arise.

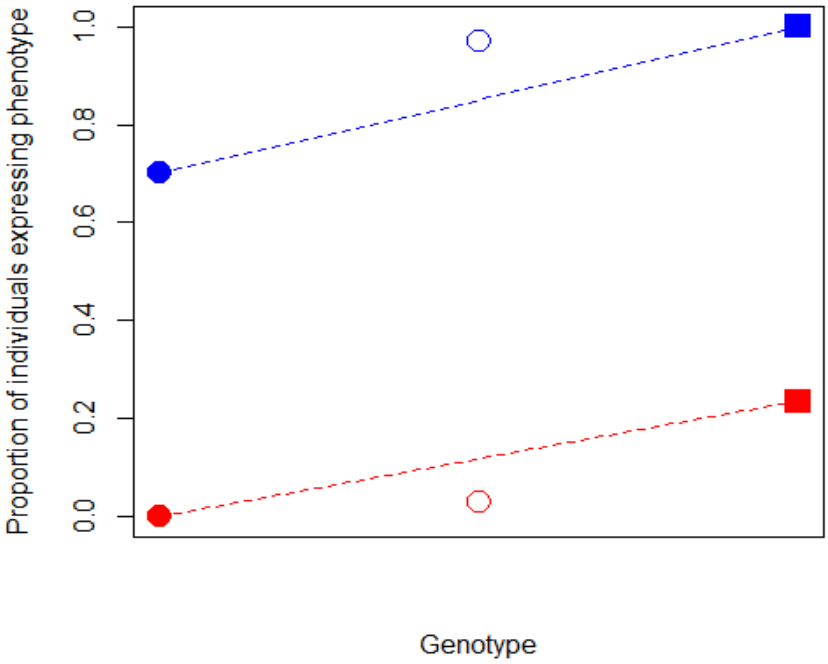
Consider a dichotomous threshold trait where the underlying liability is polygenic, and hence can be represented by a breeding value and environmental deviation, but is also affected by an additional locus of large effect with two alleles. Such a system has been broadly described for migration timing (and hence maturation) in Atlantic salmon. Here, observed patterns of phenotypic variation are consistent with sex-specific genetic dominance reversal at the large-effect locus, such that phenotypes of heterozygotes resemble phenotypes of opposite homozygotes in females and males. Such dominance reversals have been suggested to resolve sexual conflict over migration and maturation timing (Barson et al. 2015; Czorlich et al. 2018).

Yet, simple simulations that assume additive effects of alleles at the large-effect locus (i.e. representing sex-independent co-dominance) in combination with sex-specific mean breeding

value for liability can readily yield patterns of phenotypic variation that are consistent with sex-specific genetic dominance reversals, even though such reversals have not been simulated (Figure S3, Box S1). These patterns arise because the same additive effects of alleles at the large effect locus can have different effects on the phenotypic expression of the threshold trait in the two sexes, causing differences in mean phenotype in heterozygotes versus each homozygote due to the difference in mean breeding value relative to the threshold. Similar patterns could also arise given some forms of sex-independent partial dominance. Sexual conflict could then in fact be effectively resolved by sexual dimorphism, without any strict sex-specific dominance.

Further, apparent evidence of sex-specific dominance reversal can potentially appear or disappear depending on environmental effects on mean sex-specific phenotype. Apparent genetic architecture could therefore potentially differ between populations (e.g. Czorlich et al. 2018), even with no real difference in genetic architecture.

Figure S3. Proportions of individuals of each sex (blue and red) expressing a discrete phenotype given homozygosity for each allele (filled symbols) or heterozygosity (open symbols) at a large effect locus alongside sexual dimorphism in mean breeding value. Given the example parameterisations specified in Box S1, resulting patterns of phenotypic variation could be interpreted to imply sex-specific dominance reversal (i.e. where heterozygotes resemble opposite homozygotes in the two sexes), even though no such dominance reversal was simulated. Dashed lines link the sex-specific homozygotes, to help visualise the deviation of the heterozygote.



Box S1. Simple R Code for simulating a quantitative genetic threshold trait affected by a large-effect locus

```
#Code set with arbitrary parameter values, which can be adjusted

#Set number of individuals of each sex and genotype
n <- 200
#Set additive genetic and environmental variances in liability
#Assume they are the same in both sexes (which could be relaxed)
Va <- 0.16
Ve <- 0.01
```

```
#Set liability values for the two homozygotes at a large effect locus

#Allele 1 decreases liability to express the trait relative to allele 2

g.hom.1 <- -1

g.hom.2 <- 0


#Set liability value for a heterozygote at the large effect locus

#Value set half-way between the two homozygoytes effectively sets additive effects of each allele (i.e.

codominance)

#This could be adjusted to examine scenarios of partial dominance

g.het <- -0.5

#Assume that effects of heterozygosity are the same in both sexes

#(i.e. no explicit sex-specific dominance)


#Set mean breeding values for females and males

#Generates sexual dimorphism in mean liability

#Values arbitrarily set relative to a threshold at zero

fem.mean <- -0.3

male.mean <- 1.2


#Note that mean environmental deviations are currently taken as zero

#This could be adjusted, for example to consider effects of different environmental conditions


#Create females of each genotype - values for liability and phenotype as a threshold trait

#Female homozygote allele 1
```

```
fem.hom.1.liab <- rnorm(n, fem.mean, sqrt(Va)) + rnorm(n, 0, sqrt(Ve)) + g.hom.1

fem.hom.1.phen <- ifelse(fem.hom.1.liab<0, 0, 1)


#Female heterozygote

fem.het.liab <- rnorm(n, fem.mean, sqrt(Va)) + rnorm(n, 0, sqrt(Ve)) + g.het

fem.het.phen <- ifelse(fem.het.liab<0, 0, 1)


#Female homozygote allele 2

fem.hom.2.liab <- rnorm(n, fem.mean, sqrt(Va)) + rnorm(n, 0, sqrt(Ve)) + g.hom.2

fem.hom.2.phen <- ifelse(fem.hom.2.liab<0, 0, 1)


#Create males of each genotype - values for liability and phenotype as a threshold trait

#Male homozygote allele 1

male.hom.1.liab <- rnorm(n, male.mean, sqrt(Va)) + rnorm(n, 0, sqrt(Ve)) + g.hom.1

male.hom.1.phen <- ifelse(male.hom.1.liab<0, 0, 1)


#Male heterozygote

male.het.liab <- rnorm(n, male.mean, sqrt(Va)) + rnorm(n, 0, sqrt(Ve)) + g.het

male.het.phen <- ifelse(male.het.liab<0, 0, 1)


#Male homozygote allele 1

male.hom.2.liab <- rnorm(n, male.mean, sqrt(Va)) + rnorm(n, 0, sqrt(Ve)) + g.hom.2

male.hom.2.phen <- ifelse(male.hom.2.liab<0, 0, 1)
```

```
#Plot probability of trait expression (i.e. phenotypic mean) against genotype for each sex
#Given the set parameterisations, the pattern of phenotypic variation looks like there is sex-specific dominance, but in fact the simulations specify sex-independent codominance
plot(c(mean(fem.hom.1.phen), mean(fem.het.phen), mean(fem.hom.2.phen)) ~ c(1, 2, 3),
        col = "red", pch=c(19,1,15), cex = 2, ylim=c(0,1), xaxt="n",
        ylab = "Proportion of individuals expressing phenotype",
        xlab = "Genotype")
points(c(mean(male.hom.1.phen), mean(male.het.phen), mean(male.hom.2.phen)) ~ c(1, 2, 3),
        col = "blue", pch=c(19,1,15), cex = 2)
segments(x0=1, y0=mean(fem.hom.1.phen), x1=3, y1=mean(fem.hom.2.phen), col = "red", lty = 2)
segments(x0=1, y0=mean(male.hom.1.phen), x1=3, y1=mean(male.hom.2.phen), col = "blue", lty = 2)
```

Literature cited

Arnold, P.A., L.E.B. Kruuk, and P.A. Nicotra. 2019b. How to analyse plant phenotypic plasticity in response to a changing climate. New Phytol. 222:1235-1241.

Arnold, P.A., A.B. Nicotra, and L.E.B. Kruuk. 2019a. Sparse evidence for selection on phenotypic plasticity in response to temperature. Phil. Trans. R. Soc. B 374:20180185.

Barson, N.J., T. Aykanat, K. Hindar, M. Baranski, G.H. Bolstad, P. Fiske, C. Jacq, +15 co-authors. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature 528:405-408.

Dingemanse, N.J., A.J.N. Kazem, D. Réale, and J. Wright. 2010. Behavioural reaction norms: animal personality meets individual plasticity. Trends Ecol. Evol. 25:81-89.

Chevin, L.-M., and R. Lande. 2010. When do adaptive plasticity ad genetic evolution prevent extinction of a density-regulated population? Evolution 64:1143-1150.

Chevin, L.-M., and R. Lande. 2013. Evolution of discrete phenotypes from continuous norms of reaction. Am. Nat. 182:13-27.

Czorlich, Y., T. Aykanat, J. Erkinaro, P. Orell, and C.R. Primmer. 2018. Rapid sex-specific evolution of age at maturity is shaped by genetic architecture in Atlantic salmon. Nature Ecol. Evol. 2:1800-1807.

Gavrilets, S., and S.M. Scheiner. 1993a. The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. J. Evol. Biol. 6:31–48.

Gavrilets, S., and S.M. Scheiner. 1993b. The genetics of phenotypic plasticity. VI. Theoretical predictions for directional selection. J. Evol. Biol. 6:49-68.

Lande, R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. J. Evol. Biol. 22:1435-1446.

Lande, R. 2014. Evolution of phenotypic plasticity and environmental tolerance of a labile quantitative character in a fluctuating environment. J. Evol. Biol. 27:866–875.

Lande, R. 2015. Evolution of phenotypic plasticity in colonizing species. Mol. Ecol. 24:2038-2045.

Lande, R. 2019. Developmental integration and evolution of labile plasticity in a complex quantitative character in a multiperiodic environment. Proc. Ntl. Acad. Sci. 116:11361-11369.

Morrissey, M.B., and M. Liefting. 2016. Variation in reaction norms: statistical considerations and biological interpretation. Evolution 70:1944-1959.

Nussey, D.H., A.J. Wilson, and J.E. Brommer. 2007. The evolutionary ecology of individual phenotypic plasticity in wild populations. J. Evol. Biol. 20:831-844.

Pigliucci, M. 2005. Evolution of phenotypic plasticity: where are we going now? Trends Ecol. Evol. 20:481–486.

Schaeffer, L.R. 2011. Cumulative permanent environmental effects for repeated records animal models. J. Anim. Breed. Genet. 128:95-99.

Scheiner, S.M. 1993. Genetics and evolution of phenotypic plasticity. Annu. Rev. Ecol. Evol. Syst. 24:35-68.

Scheiner, S.M. 2002. Selection experiments and the study of phenotypic plasticity. J. Evol. Biol. 15:889-898.

Tonsor, S.J., T.W. Elnaccash, and S.M. Scheiner. 2013. Developmental instability is genetically correlated with phenotypic plasticity, constraining heritability, and fitness. Evolution 67:2923-2935.

Westneat, D.F., L.J. Potts, K.L. Sasser, and J.D. Shaffer. 2019. Causes and consequences of phenotypic plasticity in complex environments. Trends Ecol. Evol. 34:555-568.

Wright, S. 1934. The results of crosses between inbred strains of guinea pigs, differing in number of digits. Genetics 19:537-551.