Agnes Holstad

Does evolvability predict evolutionary divergence?

May 2020

Master's thesis

2020

Master's thesis

Agnes Holstad

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Does evolvability predict evolutionary divergence?

Agnes Holstad

MSc in Biology
Submission date: May 2020
Supervisor: Christophe Pélabon
Co-supervisor: Thomas F. Hansen, Geir H. Bolstad, Øystein H. Opedal

Norwegian University of Science and Technology
Department of Biology

Acknowledgement

First, I would like to thank Christophe Pélabon, Thomas F. Hansen and the Centre for Advanced Study (CAS) for hosting me and letting me take part in the Evolvability project. I have learned more this year than I ever could have imagined, through the countless talks, journal clubs and discussions. This somewhat out of the ordinary project has given me a peak into the scientific field.

I would like to thank Christophe for being the best supervisor I could have asked for. Your door is always open, no question is too stupid, and the discussions both enlightening and encouraging. I want to thank Thomas for being like my second main supervisor. I have always felt welcome and have learned so much from you. Also, thanks to my co-supervisors Geir H. Bolstad and Øystein H. Opedal for all their help and guidance. And of course, thanks to all my supervisors for invaluable feedback on this thesis. When it comes to supervisors, I really hit the jackpot.

Finally, I would like to thank my friends and family for all their support and for making my time as a student in Trondheim and this last year in Oslo a wonderful experience. Specially, I am very thankful to Mia for being there with fantastic support, motivational coaching and care packages.

Abstract

Evolvability is the potential for evolution and can be defined as the ability of a population to produce and maintain genetic variation. Evolvability depends on the additive genetic variance that represents the part of the trait variance that is directly transmitted from parent to offspring. The absence of additive genetic variance may limit trait evolution and therefore constrain divergence among populations and species. To test this hypothesis, I performed a comparative analysis where I assessed if evolvability constrains evolutionary divergence, by analysing the relationship between populations and species divergence and evolvability in a broad range of traits. I searched the primary scientific literature for studies reporting additive genetic variances and trait means in order to calculate evolvability as additive genetic variance standardised by the trait mean square. Across all species and traits, evolvability predicted evolutionary divergence. Evolvability explained 30% of the divergence at the among-population level and 12% at the among-species level. The relationship between evolvability and evolutionary divergence remained unchanged within trait and organism categories. These results seem to support the genetic-constraint hypothesis and hence contribute to connecting the micro- and macroevolutionary timescales.

Keywords: Evolvability, genetic constraint, evolutionary divergence, microevolution, macroevolution, heritability

Sammendrag

Evolverbarhet er potensialet for evolusjon og kan defineres som en populasjons evne til å produsere og opprettholde genetisk variasjon. Evolverbarhet avhenger av den additive genetiske variansen, som representerer den delen av trekkvariansen som overføres direkte fra foreldre til avkom. Fraværet av additiv genetisk varians kan begrense evolusjon av trekk og derfor begrense divergens blant populasjoner og arter. For å teste denne hypotesen, utførte jeg en komparativ analyse der jeg undersøkte om evolverbarhet begrenser evolusjonær divergens. Dette ble gjort ved å analysere forholdet mellom populasjons- og artsdivergens og evolverbarhet i et bredt spekter av trekk. Jeg søkte i den primære vitenskapelige litteraturen etter studier som rapporterte additive genetiske varianser og gjennomsnitt av trekk for å beregne evolverbarhet som additiv genetisk varians standardisert av det kvadrerte trekkgjennomsnittet. På tvers av alle arter og trekk forklarte evolverbarhet noe av den evolusjonære divergensen. Evolverbarhet forklarte 30% av divergensen mellom populasjoner og 12% mellom arter. Forholdet mellom evolverbarhet og evolusjonsdivergens forble uendret innen ulike trekk-kategorier og organismer. Disse resultatene ser ut til å støtte hypotesen om genetiske begrensninger og dermed bidra til å koble sammen mikro- og makroevolusjon.

Table of contents

1 Introduction

Phenotypic evolution can be measured as changes in the population mean and distribution of a phenotype over generations. Such changes are studied at both micro- and macroevolutionary timescales. Microevolution is often defined as short term dynamics of phenotypes within populations on a timescale of one to ca. hundred generations (Arnold et al., 2001, Merilä et al., 2001, Hansen et al., 2011, Hansen, 2012). These dynamics include changes in allele frequencies, trait means and trait distributions, typically studied with methods and theory derived from population genetics or quantitative genetics (Falconer and Mackay, 1996, Conner and Hartl, 2004). In macroevolution, patterns of phenotypic change will be studied at the among-population and among-species level using comparative studies and the fossil record (Gould and Eldredge, 1977, Gould, 2002, Uyeda et al., 2011, Hansen, 2012, Voje et al., 2018).

The transition from micro- to macroevolutionary timescale has been considered by some authors as a non-continuous path, where evolution at the different timescales depends on partly independent processes (Eldredge and Gould, 1972, Gould and Eldredge, 1977, Stanley, 1979, Gould, 2002). Punctuated equilibria, the morphological stasis observed within linages and sudden appearance of new lineages on a geological timescale in the fossil record, tend to support this view. With punctuated equilibria, morphological change is concentrated in rapid speciation events (Eldredge and Gould, 1972, Gould and Eldredge, 1977). This contrasts with the neo-Darwinian gradualist view of evolution where morphological change is an ongoing process with similar rates within and among species. At the time this sparked a fierce debate in evolutionary theory which remains to this day (Charlesworth et al., 1982, Gingerich, 1984, Arnold et al., 2001, Voje et al., 2020). Recently the best example of evidence for punctuated equilibria in the bryozoan genus *Metrarabdotos* (Gould, 2002) was refuted by Voje et al. (2020) who found no difference in the strength of selection between anagenesis (evolution within linages) and cladogenesis (evolution between linages), and showed that evolution happened mainly in the direction with above average genetic variation in both scenarios (Voje et al., 2020).

Theoretical and empirical studies have suggested that changes in adaptive landscapes through time may provide mechanisms that explain patterns of punctuated equilibrium with gradual changes linking micro- and macroevolution. For instance, the morphological stasis within linages may be due to adaptation to stationary fluctuating optima generating no net evolution in the fossil record (Hansen and Martins, 1996, Estes and Arnold, 2007, Uyeda et al., 2011, Hansen, 2012, Voje, 2016, Voje et al., 2018). More directional, semi-permanent or sudden changes to the adaptive landscape may induce divergent selection in lineages resulting

in cladogenesis (Simpson, 1944, 1953, Hansen and Martins, 1996, Arnold et al., 2001, Hansen, 2012). Another approach to the continuous transition between the two timescales is the extrapolation of parameters describing variational properties within populations to explain patterns of macroevolution (Andersson, 1991, 1997, Schluter, 1996, Bégin and Roff, 2003, 2004, Bolstad et al., 2014, Houle et al., 2017, McGlothlin et al., 2018, Voje et al., 2020). However, how far the parameters describing microevolutionary processes can be extrapolated to explain patterns of macroevolution remains an open question (Chenoweth et al., 2010, Futuyma, 2010). For instance, can patterns of genetic variation within populations (i.e. a microevolutionary timescale) explain patterns of evolutionary divergence among population and species (i.e. a macroevolutionary timescale)? If this is the case, genetic constraints are present.

The genetic-constraint hypothesis predicts that genetic variational properties within a population limit phenotypic changes in response to random genetic drift or selection (Arnold, 1992, Futuyma, 2010). Adaptive evolution depends on the standing genetic variation in a population, the strength of selection and time (i.e. number of generations). Lande (1976, 1979) modelled this response to selection over a few generations as:

$$\Delta \bar{z} = V_A \beta,$$
(1.1)

where $\Delta \bar{z}$ is the change in trait mean, V_A is the additive genetic variance and β is the directional selection gradient. The selection gradient represents the slope of the fitness function at the populations mean. It is estimated as the slope of the regression of the individual relative fitness (w) on the individual trait values of the population ($\beta = \frac{d\,w(z)}{d\,z} = \frac{\text{cov}[w,z]}{\text{var}[z]}$). The additive genetic variance is the part of the phenotypic trait variance (V_P) that is due to directly transmittable genetic effects from generation to generation (Falconer and Mackay, 1996, Conner and Hartl, 2004). This additive variance (V_A) thus represents the potential for evolution (Houle, 1992, Hansen, 2006), and is assumed to be constant over the examined time period (Lande, 1976, Lande, 1979, Turelli, 1988).

The response to selection in one trait may be constrained by genetic correlation with other traits. The multivariate version of the Lande equation estimates the response to selection (direct and indirect) in a set of correlated traits

$$\Delta \bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$$
(1.2)

(Lande, 1979, Lande and Arnold, 1983), where $\Delta \bar{\mathbf{z}}$ is the response vector, the sum of the change in trait means ($\Delta \bar{z}_i$), $\mathbf{G}$ a matrix of the additive genetic variances (V_{A_i}) and covariances (cov_{ij}) and $\boldsymbol{\beta}$ is the vector of selection gradients (β_i) for each trait ($i = 1, 2, \dots, n$). Lande (1977a, 1979,

1980) argued that the G-matrix stays roughly stable through time due to a selection, mutation, and recombination balance. Empirical work and simulations have shown that some features of the G-matrix stay remarkably stable (Schluter, 1984, 2000, Shaw et al., 1995, Jones et al., 2003, 2007, Arnold et al., 2008, Hohenlohe and Arnold, 2008).

The univariate (eq. 1.1) and multivariate (eq. 1.2) version of the Lande equation demonstrates that adaptive evolutionary changes may be constrained by lack of genetic variation, in amount or direction. Futuyma (2010) presented different forms of genetic constraint, including (i) little or undetectable genetic variation for a trait, (ii) genetic correlations between traits with opposing directional selection, (iii) divergence along the genetic line of least resistance (g-max), i.e. the direction of the G-matrix with the highest amount of variation (Schluter, 1996), and (iv) G-matrices with fewer detectable principal components than traits.

The presence of genetic constraint would imply that the amount of genetic variation can predict the amount of divergence observed between populations and species. If the lack of genetic variation represents a constraint to evolution, then we expect macroevolutionary patterns (population and species divergence) to be partly explained by the available genetic variation (Arnold, 1992, Schluter, 1996, 2000, Bégin and Roff, 2003, 2004). Several studies have tested this hypothesis. Andersson (1991) showed that within- population genetic variation explained among-population variation in the plant *Crepis tectorum*. Similarly, Bolstad et al. (2014) found that the G-matrix predicts divergence among populations of two closely related *Dalechampia* (Euphorbiaceae) species. Among species of *Anolis* lizards, McGlothlin et al. (2018) found that the G-matrix predicted divergence. Houle et al. (2017) found that the mutational variance matrix, summarising the combined effects and accumulation of mutations over generations (Houle and Fierst, 2013), predicted divergence in Drosophilid taxa.

Studying the scaling between evolvability and divergence among populations and species may help us understanding better the mechanisms causing the divergence. Models of the rate of divergence under different evolutionary scenarios may give an improved interpretation of which evolutionary processes generate the observed scaling relationships between evolvability and divergence.

One way to model divergence over time is a neutral evolution model, which only assumes mutation and drift (Lynch and Hill, 1986, Lynch, 1990). This can be done with a Brownian-motion process with a flat fitness function and a high phylogenetic heritability (proportion of trait variance explained by relationships among taxa (Housworth et al., 2004)). This model predicts that the variance between two species has an expected rate of divergence

proportional to the mutational variance on a log-log scale (Lynch, 1990, Martins, 1994, Houle et al., 2017). Thus, they scale with a slope of 1 (Fig 1.1)

Other evolutionary models assume directional divergent selection, using a Brownian motion process and linear fitness function divergent for the different species with intermediate phylogenetic heritability. In this scenario the rate of divergence scales with the mutational variance with a slope of 2 on a log-log scale (Houle et al., 2017). Bolstad et al. (2014) showed that for mean-scaled additive genetic variance (i.e. evolvability sensu Hansen et al. (2003b)) the scaling relationship with among-population variance is equivalent to the expectation for the mutational variance (Fig. 1.1).

A third approach to modelling lineage divergence is to assume an Ornstein-Uhlenbeck (OU) process. Unlike the Brownian-motion process the OU-process has a "pull parameter". This makes the OU-process capable of considering the mutational target and mutational average under drift and considering various scenarios where species are tracking fluctuating optima (Martins, 1994, Hansen and Martins, 1996, Hansen, 1997, Hansen et al., 2008). When populations track moving optima, the scaling relationship between evolvability and among-population variance depends on the movement of the optima and the rate of adaptation (Bolstad et al., 2014). There is no detectable scaling relationship when the rate of adaptation of the populations is either too slow (unable to track) or too fast (perfectly tracked) compared to the movement of the optima. However, with rates of adaptation in the same order of magnitude as the movement of the optima where the populations manage, a scaling relationship between evolvability and divergence lies between 0 and 1 on a log-log scale. If the optima move fast, but are still in the range of where the populations manage to track them, the scaling relationship will be close to 1 (Bolstad et al., 2014).
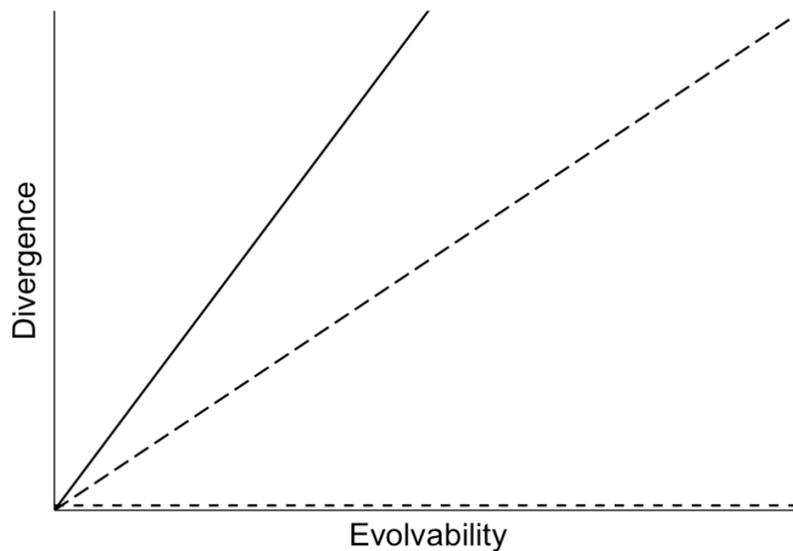
Figure 1.1: The predicted scaling relationship of divergence between species or populations and evolvability for a trait on a log-log scale. Solid line: Divergent directional selection, slope = 2, represents the upper limit. Long dashed line: Genetic drift, slope = 1. Short dashed line: Stabilising selection, slope = 0. Any mixture of modes of evolution with time yields a slope somewhere between 0 and 2.

As mentioned, different studies have tested aspects of the hypothesis that genetic constraint is important in the divergence of organisms. Most of these studies have been restricted to one species or a few closely related species and there is currently no general analysis of this pattern observed in specific groups of organisms. In this study I aim to test the generality of the genetic constraint hypothesis. To do so, I analyse a dataset containing a broad range of traits and organisms. The hypothesis is tested by asking to what extent genetic variation predicts divergence among populations and species. Specifically, I ask: (i) Does genetic variation predict evolutionary divergence? (ii) Does timescale affect the prediction? (iii) Are there differences among organisms and traits? (iv) How does evolutionary divergence scale with evolvability?

2 Methods

To assess the relationship between evolvability and evolutionary divergence, I collected estimates of additive genetic variance and trait means from the primary scientific literature. Testing this relationship in a broad range of traits and organisms requires standardisation of the additive genetic variance estimates. Similarly, a meaningful comparison between evolvability and divergence need to be based on comparable scale types.

There are different ways to standardise additive genetic variance, either proportional to the total variance or proportional to the mean (Houle, 1992, Hansen et al., 2003b, Hansen et al., 2011). The phenotypic variance depends on the phenotypic mean of the entity and the variance is more comparable on a proportional scale. For example, an increase of five grams in body size does not affect a mouse and an elephant equally, because the change in the mouse is much larger in proportion of its body size. Hansen et al. (2003b) standardised the Lande equation by the trait mean,

$$\frac{\Delta \bar{z}}{\bar{z}} = \frac{V_A}{\bar{z}^2} \beta \bar{z},$$

(2.1)

where the mean-scaled additive genetic variance is the evolvability ($e_\mu = \frac{V_A}{\bar{z}^2}$). The mean-scaled selection gradient ($\beta_\mu = \beta \bar{z}$) becomes an elasticity because it equals the proportional change in fitness per proportional change in trait mean (Hereford et al., 2004). Evolvability can be interpreted as the expected proportional change in the trait mean over the strength of selection.

$$e_\mu = \frac{\left(\frac{\Delta \bar{z}}{\bar{z}}\right)}{(\beta \bar{z})} = \frac{V_A}{\bar{z}^2}$$

(2.2)

If the trait is fitness itself, then $\beta \bar{z} = 1$, and $e_\mu \times 100$ can be interpreted as the percent potential evolutionary change for a trait that is unconstrained by correlations with other traits (Hansen et al., 2003b, Hansen et al., 2011).

Another common way to standardise the additive genetic variance is by dividing V_A by the total phenotypic variance (V_P),

$$h^2 = \frac{V_A}{V_P}.$$

(2.3)

This is called the heritability (h^2) and stems from the Breeders equation,

$$\Delta \bar{z} = h^2 S,$$

(2.4)

where S is the selection differential, that is, the difference between the population mean ($\bar{z}$) before and after selection ($\bar{z}^*$) (Lush, 1937). Heritability has been criticized as a measure of evolvability (Houle, 1992, Hansen et al., 2003b, 2011, Hereford et al., 2004, Hansen and Houle,

2008, Wilson, 2008). Phenotypic variance consists of the sum of the additive genetic variance (V_A), the trait variance due to epistatic interactions (V_I), the variance due to dominance interactions (V_D), the environmental variance (V_E) and some residual variance (V_R). Hence,

$$h^2 = \frac{V_A}{V_A + V_I + V_D + V_E + V_R},$$

$$(2.5)$$

and dividing V_A by V_P obscures the measure of V_A, because V_P is itself a function of V_A (Hereford et al., 2004). Consider the case where $V_{AA} + V_D + V_E + V_R = 0$ for a set of traits, then $h^2 = 1$ for all the traits, regardless of the amount of V_A. Furthermore, measures of V_{AA}, V_D and V_E are shown to be correlated with V_A (Hansen et al., 2011). Thus, the scale becomes stretched for high values of V_A, and V_A is therefore largely uncorrelated with h^2 (Lande and Arnold, 1983, Hansen et al., 2003b, Hereford et al., 2004, Hansen et al., 2011).

2.1   Data collection

I searched the primary scientific literature to build a meta-database for this comparative study. I only included studies where populations or species had diverged naturally. This excluded studies where divergence resulted from artificial selection, populations kept and bred in the lab (except for the generations in common garden for estimating genetic divergence) or populations constructed from inbred lines. The statistics required were a minimum of one genetic-variance estimate per trait with its corresponding mean (given in the form of one of the rows under "Genetic-varince estimates" in Table 2.1), and the phenotypic mean for two or more populations or species (Table 2.1).

Table 2.1: The different options of statistics required for the genetic-variance estimate and the divergence estimate per trait.

Genetic estimates	min. 1 population/species
Evolvability	$e_\mu = \frac{V_A}{\bar{z}^2}$
Additive genetic variance, trait mean	$V_A, \bar{z}$
Coefficient of variation	$CV_A = \frac{\sqrt{V_A}}{\bar{z}}$
Heritability, phenotypic variance, trait mean	$h^2 = \frac{V_A}{V_P}, V_A, \bar{z}$
Additive genetic variance on log-transformed trait values	$V_A[\log(z_i)]$
Divergence	**≥ 2 population/species**
Trait mean	$\bar{z}$

Estimates of additive genetic variances and coefficients of variation are occasionally erroneous in the scientific literature (Garcia-Gonzalez et al., 2012). The genetic parameters were therefore carefully checked for proper estimation. Additive genetic variances that were estimated on natural log-transformed trait values were included directly as an estimate of evolvability (Table 2.1, row 6). This is because a logarithmic transformation is asymptotically equivalent to a proportion of the mean (Lynch, 1990). When

$$\frac{\text{var}[z_i]}{\bar{z}^2} < 1, \qquad \text{var}[\log(z_i)] \simeq \frac{\text{var}[z_i]}{\bar{z}^2}. \qquad \textbf{(2.6)}$$

The requirement of two or more populations per trait was the largest constraint to finding studies and greatly reduced the amount of useable studies. In some cases, different published papers were combined for one trait measure. However, only when I was certain that the trait was measured in the exact same way in the different papers (ref. Houle et al. (2011)). These papers were difficult to find and combine by search in a literature data base. Data were for instance published over several papers, occasionally with many years apart (e.g. Arnold (1988), Arnold and Phillips (1999) and Hohenlohe and Arnold (2008)) or conducted by different groups (e.g. clutch size, Van Noordwijk et al. (1981), Flux and Flux (1982), Gibbs (1988)). An overview of used studies, and those combined for one or several trait measures, is given in Table B1 in Appendix B.

The basis for collection of trait measures were (i) studies gathered by Bolstad, Hansen and Pélabon for a review on genetic constraints in 2014, (ii) studies used in Hansen et al. (2011), Bolstad et al. (2014), Matthews et al. (2019), Noble et al. (2019) and Opedal (2019), and (iii) evolvability and divergence data on plants collected by Ø.H. Opedal as part of other ongoing projects.

I also conducted a search with ISI Web of Science (date: 04/12/2019) with the key words "additive genetic variance" and "divergence" over all collections and all years. This was done to retrieve papers reporting additive genetic variance, heritability, coefficient of additive genetic variance or evolvability measures for different populations or species. The search returned 89 papers, of which 75 were not already known. However, only two of these 75 papers had sufficient information to be included in the analysis.

2.2 Issues with scale

The standard deviation expressed as a proportion of the mean, i.e. the coefficient of variation ($CV = \frac{\sqrt{V}}{\bar{z}}$), is a dimensionless, comparable and easily interpretable measure of variation (Lewontin, 1966, Lande, 1977b, Houle, 1992, Pélabon et al., 2020). However, the

standardisation by the trait mean limits the use of CV (and mean-scaled evolvability ($e_\mu = CV^2$)) to traits on a ratio or log-interval scale, where the mean and zero are meaningful and order, ratios and differences (only ratio scale) are preserved with multiplication by a constant (Lewontin, 1966, Yablokov, 1974, Hansen et al., 2011, Houle et al., 2011, Pélabon et al., 2020). This excludes traits such as laying date, temperature in degrees Celsius, colour (either as nominal (name) or interval (RGB domain)) and regression slopes such as reaction norms (Pélabon et al., 2020).

Houle (1992) and Hansen et al. (2011) found that life history traits are more evolvable than morphological traits, and Opedal (2019) showed that vegetative traits have twice as high evolvabilities as floral traits. These differences may have biological meaning. For instance, vegetative traits may be under fluctuating or diversifying selection due to varying environment and resource allocation. Floral traits, on the other hand, are functionally linked to the reproductive process of the plants and may be under strong stabilizing selection to ensure e.g. pollinator fit with little size variation. However, it is important to be conscious of the proportionality between the mean and standard deviation when assigning biological meaning to these differences.

The CV represents a measure of variation proportional to the mean. This allows for comparison of variation independently of the effect of the mean when the standard deviation increases proportionally with the mean. If this is not the case, differences in CV may be due to specific properties of the traits that effect the relationship between the standard deviation and the mean. Such properties may be the dimensionality of the trait or levels of measurement error (Pélabon et al., 2020).

The dimensionality of the trait is important when considering non-proportionality between the mean and the standard deviation. Length, area (length^2) and volume (length^3) measures vary according to their dimensions (1, 2, and 3 respectively) and the correlation between these dimensions (Lande, 1977b). A volume has a CV up to three times and an area up to times the CV of a length measure.

$$CV_{volume} \leq 3\overline{CV}_{length} \ and \ CV_{area} \leq 2\overline{CV}_{length} \tag{2.7}$$

A perfect correlation between the dimensions of an area or volume, i.e. a constant shape with varying size, represents the upper boundaries of the equation (2.7) (Lande, 1977b). For evolvability this implies that a volume has up to 3^2, and an area up to 2^2 the evolvability of a length measure due to the different dimensions. The dimensionality of the trait was thus carefully considered in this study.

Another common case of non-proportionality between the mean and the standard deviation involves non-normal or log-normal distributions of the trait. For instance, in a trait with a count measure such as clutch size in birds, the mean and standard deviation are not expected to scale proportionally. Birds with smaller clutch sizes have higher CV (Pélabon et al., 2020) as a consequence of this non-proportionality. However, this may also have biological meaning, e.g. that it is harder to evolve to the double clutch size in a clutch size of 9 than a clutch size of 2 eggs. I therefore included traits with expected Poisson distribution. In traits with binomial distributions, like probability of survival ($0 \leq p \geq 1$), variance is highest for $p = 0.5$ and zero for $p = 0$ and $p = 1$. The CV's therefore approach infinity and zero for mean probability of survival (p) close to zero and one respectively. Due to this caveat, I excluded traits with binomial distribution from the analysis.

2.3 Analyses

2.3.1 The variables

To perform the analysis, I computed one evolvability and one divergence estimate per trait from the collected trait means and genetic-variance estimates (Table 2.1). To ensure that the estimates grouped per trait were estimated from the exact same measure of the trait, a universal unique identifier (UUID) was assigned per trait. In traits where two sexes were measured, one UUID was assigned to each sex, to avoid confusing sexual dimorphism with divergence. To make a separation of timescale, I computed divergence among populations and among species separately. If the same trait was measured among populations and species, different UUID's were assigned to the among-population and among-species measurements.

The divergence between populations or species was computed as the variance of the log-transformed trait means ($\bar{z}_i$) for each trait:

$$D_M = \mathrm{var}[\log(\bar{z}_i)], \qquad i = 1, 2, \dots, n,$$

(2.8)

where n is the total number of trait means, i.e. the number of populations or species measured, per trait. The D_M denotes mean-standardised divergence, since the log-transformation is almost equivalent to a mean-standardisation (eq. 2.6). Hence, the divergence measure is a dimensionless variance proportional to the mean square, comparable across traits and species, on the same scale as evolvability (eq. 2.2). One estimate of evolvability was used per unique trait in the analyses. If evolvability estimates were available for more than one population or species, I used the mean evolvability ($\bar{e}_\mu$) to avoid the complication of correlated estimation errors with several x-variable estimates per y-variable estimate.

$$\bar{e}_\mu = \text{E}\left[\frac{V_{A_i}}{\bar{z}_i^2}\right], \qquad i = 1, 2, \dots, n \tag{2.9}$$

Here E[ ] denotes the expected value, i.e. the mean. The mean was not weighted by the number of families in the estimation of the additive genetic variance, because an adequate number of families depends on the method and/or experimental design. An overview of the methods used for obtaining additive genetic variance in the different studies is given in Table B1 in Appendix B. Although the inverse squared standard error of the additive genetic variance, which is independent of the estimation method could be used as a weighing factor, only 32% of the additive genetic variance estimates were reported with their standard error (see Table 3.1). Thus, a non-weighted mean was used.

To test whether the relationship between evolvability and divergence is sensitive to the choice of standardization, I also assessed the relationship between heritability and evolutionary divergence. I needed a population and species divergence measure compatible with heritability. Consequently, a variance-standardised divergence measure (D_V) on the scale of the heritabilities was computed.

$$D_V = \frac{\text{var}[\bar{z}_i]}{\text{E}[V_{P_i}]}, \qquad i = 1, 2 \dots, n \tag{2.10}$$

Here $\text{E}[V_{P_i}]$ is the estimated mean of the phenotypic variances of the n different populations or species measured.

2.3.2 Random-regression models

To analyse the relationship between divergence and evolvability, I used random-regression models following Laird and Ware (1982) and Pinheiro et al. (2001). Divergence was fitted as the response variable with evolvability as fixed effect and species as a random effect.

$$\mathbf{D_M} = \mathbf{e_\mu}\boldsymbol{\alpha} + \mathbf{S}\boldsymbol{u} + \boldsymbol{\varepsilon}, \tag{2.11}$$

where divergence ($\mathbf{D_M}$) is the $N \times 1$ response vector of the n_j traits in species j of the total number of species J, and $N = \sum_j^J n_j$. Evolvability ($\mathbf{e_\mu}$) is the $N \times 2$ design matrix, with a column of ones and a column of evolvability observations ($e_{\mu_{ij}}$), for the vector $\boldsymbol{\alpha}$ which consists of the two fixed-effect coefficients, intercept and slope ($\boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}$). Species ($\mathbf{S}$) is the $N \times (2 \times J)$ design matrix of zeros and ones for the vector $\boldsymbol{u}$ of random effect coefficients, the intercepts and slopes ($\boldsymbol{u} = \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} = 2 \times J$), and the residuals ($\boldsymbol{\varepsilon}$) an $N \times 1$ vector. The random

effect ($\boldsymbol{u}$) and residuals ($\boldsymbol{\varepsilon}$) are assumed to be independent of each other, normally distributed, with means of zero

$$\text{cov}[\boldsymbol{u}, \boldsymbol{\varepsilon}] = 0, \qquad \boldsymbol{u} \sim \text{N}(0, \boldsymbol{\psi}), \qquad \boldsymbol{\varepsilon} \sim \text{N}(0, \mathbf{R}),$$

and positive-semidefinite variance matrices

$$\boldsymbol{\psi} = \text{var}\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} = \begin{bmatrix} \sigma_0^2 & \sigma_{01}^2 \\ \sigma_{01}^2 & \sigma_1^2 \end{bmatrix}, \qquad \mathbf{R} = \sigma^2 \mathbf{I}$$

where the intercepts (u_{0j}) and slopes (u_{1j}) of the different species are assumed to be correlated with a covariance σ_{01}^2. $\mathbf{I}$ is the identity matrix of the observations independent on the individual level, conditional on $\boldsymbol{\alpha}$ and $\boldsymbol{u}$. The models were fitted with restricted maximum likelihood using the "lme4" R-package (Bates et al., 2015). Divergence and evolvability were log-transformed prior to model fitting. Thus, traits with zero divergence or evolvability were excluded from the analysis. The best fit of the random effect (species), i.e. if the random effect could affect both slope and intercept or only intercept, was determined with model selection using Akaike information criterion (Akaike, 1974). Marginal (R_m^2) and conditional (R_c^2)) coefficients of determination were computed for the random-regression models following Nakagawa and Schielzeth (2013) with the R-package "MuMIn" (Barton, 2009). The marginal R_m^2 is the variance explained by the fixed effects over the total variance, while the conditional R_c^2 is the variance explained by both the random and fixed effects over the total variance. (Nakagawa and Schielzeth, 2013, Nakagawa et al., 2017).

The random-regression model (eq. 2.11) was fitted for among-population and among-species separately, as described under "2.2.3.1. The variables". The random effect was the lowest shared taxonomic group for the divergence measure in question, i.e. species for the among-population variance and genus for the among-species variance. Estimates of evolvability from one study share the experimental design (sample size, estimation procedure) and may in theory be non-independent. Contrary to classical meta-analysis, the study or published paper does not represent here a valid indication of non-independence between data points because several studies can be used for the same trait (see Table B1). Using species (or genus) as a source of non-independence takes the experimental design into account along with the non-independence due to potential phylogenetic inertia (Hansen and Bartoszek, 2012).

The objective of this analysis is to quantify how well evolvability predicts divergence. Thus, the models were kept simple with evolvability as the only fixed effect. To account for the non-proportionality between the standard deviation and the mean due to dimensionality (eq. 2.7), the random-regression model (eq. 2.11) was fitted for each trait category, dimension or type. The trait categories with a sufficient number of traits to fit the model were morphological

and life history traits and the dimension or types were linear, count and ratio (i.e. shape). Physiological, growth, mass/volume, area, time and complex traits were not analysed separately (plotted in Appendix A, Figs. A1 and A2). Plants and animals were also separated to evaluate if there are differences due to very distant shared ancestry or functionality or due to methodological differences in their study (see discussion).

I also checked if the distribution of evolvabilities (along the x-axis) differ between species, and if this affects the observed relationship between evolvability and divergence. If species differ in the distributions of evolvabilities, the overall scaling relationship (the slope) should be affected when the species are mean centred on the species mean evolvability. I mean centred within species data by taking the residuals of $\log(e_\mu)$ as a function of species, $r = \varepsilon_i[\log(e_\mu) \sim sp]$. The random-regression model (eq. 2.11) was then fitted with these residuals ($\mathbf{r}$) as the $N \times 2$ model matrix that is used to estimate the $\boldsymbol{\alpha}$ vector of the fixed-effect coefficients.

To assess whether the relationship between evolvability and divergence depends on the choice of standardization, I also modelled the relationship between heritability and population and species divergence. I fitted a similar random-regression model (eq. 2.11) with heritability as the fixed effect model matrix. The mean- ($\mathbf{D_M}$) and variance-standardised ($\mathbf{D_V}$) divergence were both fitted as response vectors. Heritability estimates above one and below zero were set to one and zero, respectively. Both mean- and variance-standardised divergence were log-transformed, while heritability was kept on the original scale. Heritability estimates were not specifically targeted in the literature search (ref. "2.1.2 Data collection" and Table 2.1). Heritability was not set as a requirement, and it was only entered when h^2 or V_P were available alongside evolvability. Hence, the analyses presented for heritability are done on a subset of estimates in the meta-database.

2.3.3 Attenuation bias

Measurement error in the primary data may influence the patterns seen in the meta-analysis (Hansen and Bartoszek, 2012, Morrissey, 2016). Specifically, measurement error in the x-variable (evolvability) may cause an attenuation bias in the regression slope between divergence and evolvability (Hansen and Bartoszek, 2012).

Only 32% of the retrieved additive genetic variances came with reported standard errors. Hence, the uncertainty was not directly included in the model. However, the available standard

error estimates were used to assess the average attenuation bias in the slope of different regressions between divergence and evolvability.

I calculated the error variance of the mean evolvability ($\log(\bar{e}_\mu)$) per trait on log-scale using the error variance of the additive genetic variance estimates as:

$$\text{SE}_{ij}^2[\log(\bar{e}_\mu)] = \frac{\text{E}[\frac{\text{SE}^2[V_A]_i}{\bar{z}_i^4}]}{\text{E}[e_{\mu_i}^2]}, \qquad i = 1, 2, \ldots, n_j, \qquad j = 1, 2, \ldots, k \qquad \textbf{(2.12)}$$

Here n_j is the number of populations or species per trait j of the total number of traits k. The steps to deriving equation (2.12) is given in Appendix C. I consider here the error variance of the trait mean to be negligible compared to error variance of the additive genetic variance estimate.

The regression slopes (α_1) of evolvability and divergence were corrected with a reliability ratio, K, following equation (2.13) from Hansen and Bartoszek (2012).

$$\hat{\alpha}_{1_{corrected}} = \frac{\hat{\alpha}_1}{K}, \qquad \textbf{(2.13)}$$

where the reliability ratio is the ratio between the true and observed variance of the predictor variable (Hansen and Bartoszek, 2012), estimated as one minus the error variance over the observed variance (relative error) as,

$$K = 1 - \frac{\text{SE}^2[\log(\bar{e}_\mu)]}{\text{var}[\log(\bar{e}_\mu)]}. \qquad \textbf{(2.14)}$$

A correction with the reliability ratio (K) is justified (i.e. improves the accuracy of the estimated slope) when the relative error of the slope is under 30% for a limited range of K. For a relative error under 10% and a reliability ratio between 0.11 and 0.98, the accuracy of the slope would be improved by a correction (eq. 2.13) (Hansen and Bartoszek, 2012).

2.3.4 Effect of environmental variation

In this study I aim to test genetic constraints to evolutionary divergence. Here, evolutionary divergence refers to phenotypic divergence among populations or species that is due to changes in the underlying genetic properties (e.g. allele frequencies) of the population. The divergence of populations and species may, however, not solely reflect genetic divergence but may also result from phenotypic plasticity. To test whether phenotypic plasticity affects the estimated divergence, I compared the mean among-population and among-species variance in traits that were measured in the field or in a controlled environment (common garden experiments) for both plants and animals. I also included the environment in the random-regression model (eq.

2.11.) of evolvability and divergence as a covariate to see if it affected the observed relationship. The models were only fitted on the traits where the populations per trait were measured in one kind of environment.

2.3.5 Sources of bias

2.3.5.1 The number of populations

The number of populations or species measured may influence the amount of divergence estimated. I assessed if the divergence was affected by fitting a regression with the divergence estimates over the number of populations or species measured per trait. If there is a correlation between divergence and the number of measured populations or species, it would only confound the pattern between evolvability and divergence if the evolvability measure also is correlated with the number of populations or species measured. I therefore fitted a regression of evolvability over the number of populations or species measured per trait.

2.3.5.2 Publication bias

The meta-data retrieved from the primary scientific literature may be subject to publication bias. Publication bias occurs when the true effect differs from the estimated effect reported by the published studies. This often results from the under-reporting / publication of studies where the focal parameter is statistically non-significant (Rosenthal, 1979, Palmer, 2000, Whitlock and Schluter, 2009, Nakagawa and Santos, 2012). Because larger sample size lead to higher precision in the estimate, a funnel plot of the effect size and sample size is a common tool for detecting publication bias. Plotting the assessed effect size against the sample size should therefore reveal a funnel shape, with less heterogeneity between the effect sizes with higher sample size. Deviation from this shape may indicate a publication bias (Palmer, 2000, Whitlock and Schluter, 2009).

To assess if there is a publication bias present in the meta-database, funnel plots were made for evolvability and heritability. The number of families in the experimental design (i.e. the estimation of the genetic parameters) in the studies was used as a proxy for precision of the study (Palmer, 2000). For evolvability and heritability there is not one expected effect size, but a range of theoretically expected effect sizes. I therefore fitted a regression of evolvability (and heritability) on the number of families. Publication bias would be indicated by a negative slope, because small evolvabilities or heritabilities that are not significant due to small number of families in the experimental design are not published.

16

3 Results

Out of the 302 studies initially considered, 57 met the requirements for inclusion (Table 2.1). Table 3.1 summarises the used studies, and the estimates retrieved from these. An overview of the used studies is given in Table B1 (Appendix B).

Table 3.1: The number of used studies and estimates retrieved from these studies.

Collected	*n*
Studies	57
Additive genetic variance (with SE)	1043 (332)
Trait mean (with SE)	2686 (1603)
Heritability (with SE)	904 (224)
Unique traits	409

The majority of traits included were linear measurements of morphological traits. Plants, and especially eudicots, were best represented (Table 3.2). Table 3.2 presents how the unique traits were distributed among taxa, trait dimension or type and trait categories in the meta database. A more extensive view of taxa is given in Table B1.

Table 3.2: The number of unique traits (*n*) distributed among taxa, trait dimension/type and trait category.

Taxa	*n*	Trait dimension/type	*n*	Trait category	*n*
Monocots	9	Linear	238	Morphological	346
Eudicots	282	Count	90	Life history	35
Crustacea	9	Ratio[1]	40	Physiological	25
Insecta	33	Mass/volume	15	Growth	3
Amphibia	2	Area	11		
Reptilia	22	Time	6		
Mammalia	39	Complex[2]	9		
Aves	13				

1: Dimensionless e.g. shape traits with mm^2/mm^2. 2: Complex dimensions, e.g. area based photosynthetic rate ($\mu mol\ CO_2/m^2\ s$).

3.1 Genetic variation and evolutionary divergence

The slope of the regression models of evolvability and divergence on a log-log scale represents the scaling relationship of divergence with evolvability.

Across all species and traits, evolutionary divergence scaled positively with increasing evolvability (Fig. 3.1). The slope ±SE of the linear regression on a log-log scale is 0.45 ±0.06 with an R^2 of 12%. Thus, evolvability does predict part of the divergence among populations and species.
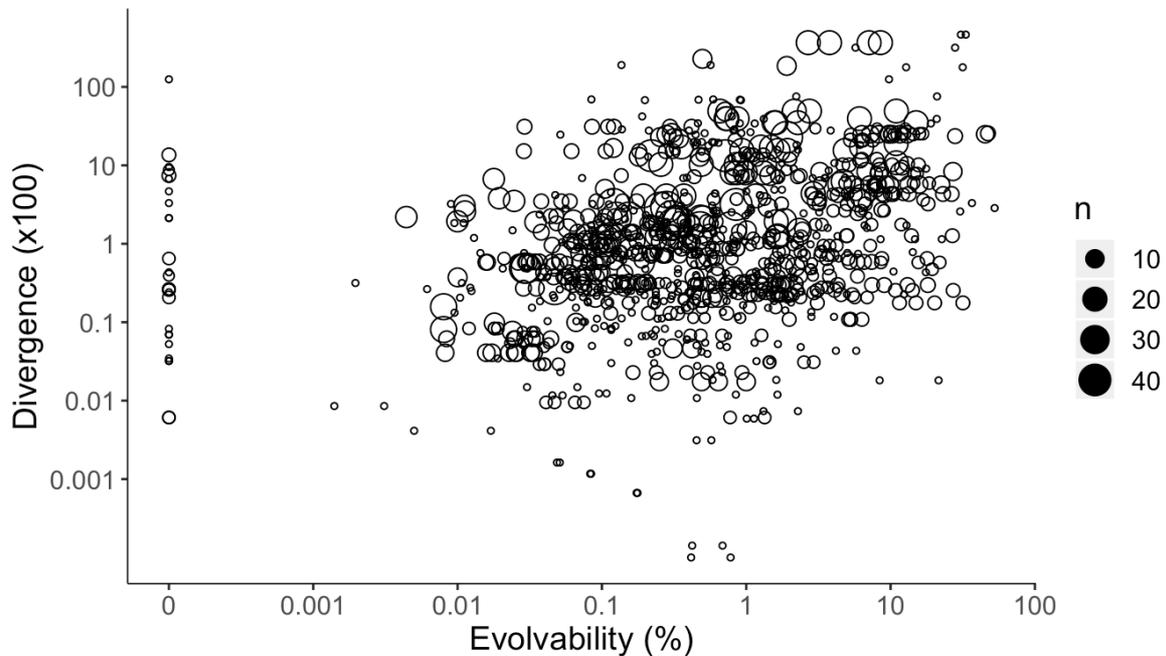
Figure 3.1: Among-population and among-species variance (divergence) as a function of evolvability on a log-log scale, with n = number of populations/species means per trait. All evolvability estimates are plotted, but the regression was fitted to the mean estimate per trait. Zeros are plotted as 0.0001%, but were not included in the linear regression. $\log(D_M) = 0.20$ (±0.11) + 0.45 (±0.06) $\log(e_\mu)$, $R^2 = 12\%$

In contrast, heritability explains little or none of the variation in evolutionary divergence among populations and species across all traits and organisms. For the phenotypic variance-standardised divergence measure (eq. 2.10), heritability explains 2% of both the among-population and -species variance (Fig. 3.2). For the mean-standardised divergence measure (eq. 2.8) heritability explains 0% (fig 3.3).
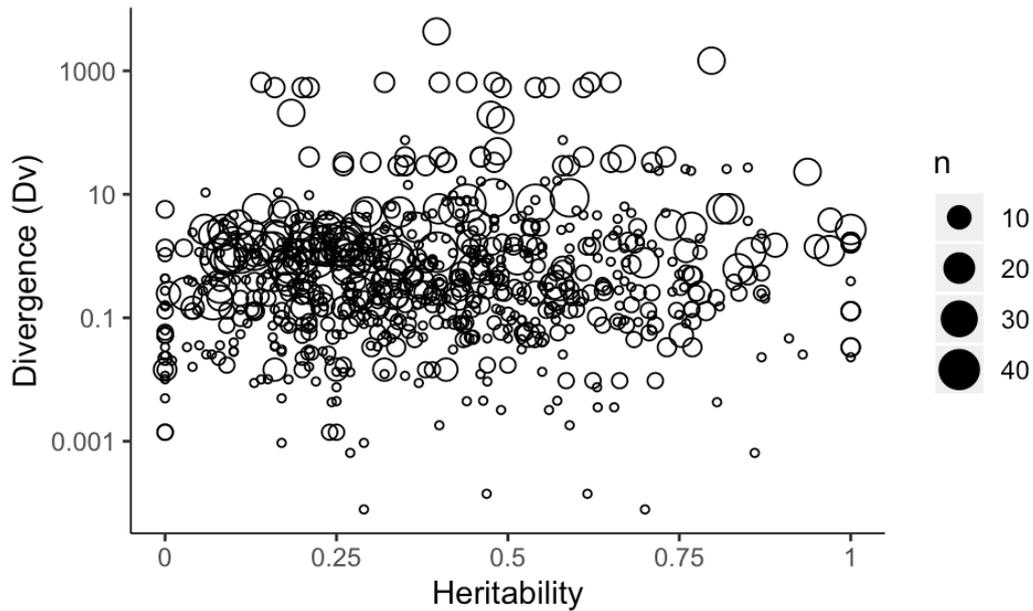
Figure 3.2: Divergence standardised by the mean phenotypic variance (log-scale) as a function of heritability (original scale), with *n* = number of populations/species means per trait. All heritability estimates are plotted, but the regression was fitted to the mean estimate per trait. $\log(D_V)$ = -1.49 (±0.30) + 1.82 (±0.67) h^2, R^2 = 2%.
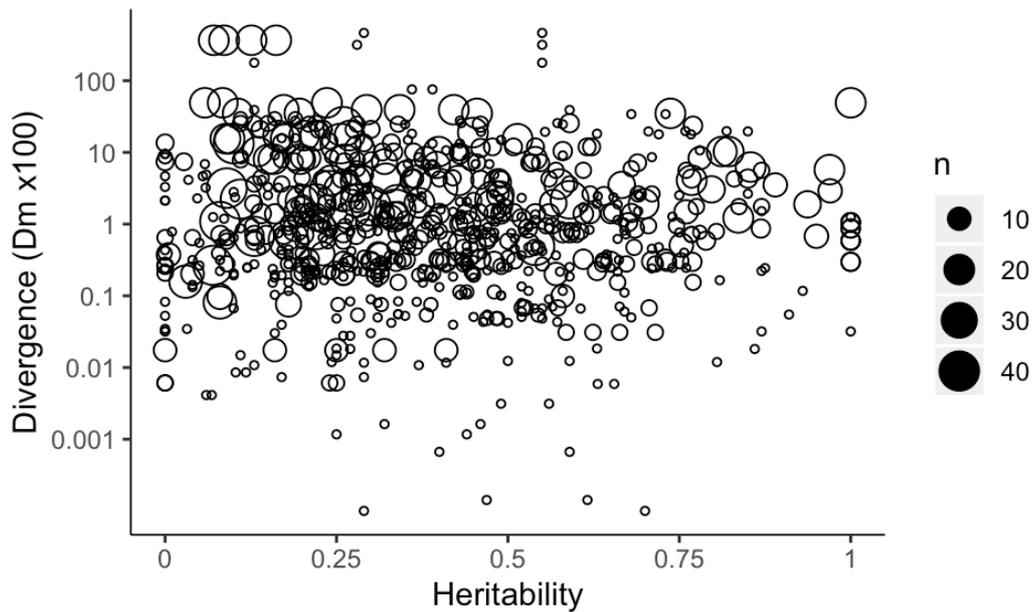
Figure 3.3: Mean-standardised divergence (log-scale) as a function of heritability (original scale), with *n* = number of populations/species means per trait. All heritability estimates are plotted, but the regression was fitted to the mean estimate per trait. $\log(D_M)$ = -0.41 (±0.27) + 0.57 (±0.62) h^2, R^2 = 0%.

3.1.1 Effect of timescale

When analysing the data separately at the among-population and among-species levels, more variation was explained at the among-population ($R_m^2 = 30\%$, Fig. 3.4 A) than at the among-species level ($R_m^2 = 12\%$, Fig. 3.4 B). The scaling relationship was also steeper among populations (slope $= 0.74 \pm 0.08$) than among species (slope $= 0.55 \pm 0.13$).

For heritability I found no qualitive difference between the among-population and among-species level (Table 3.4).
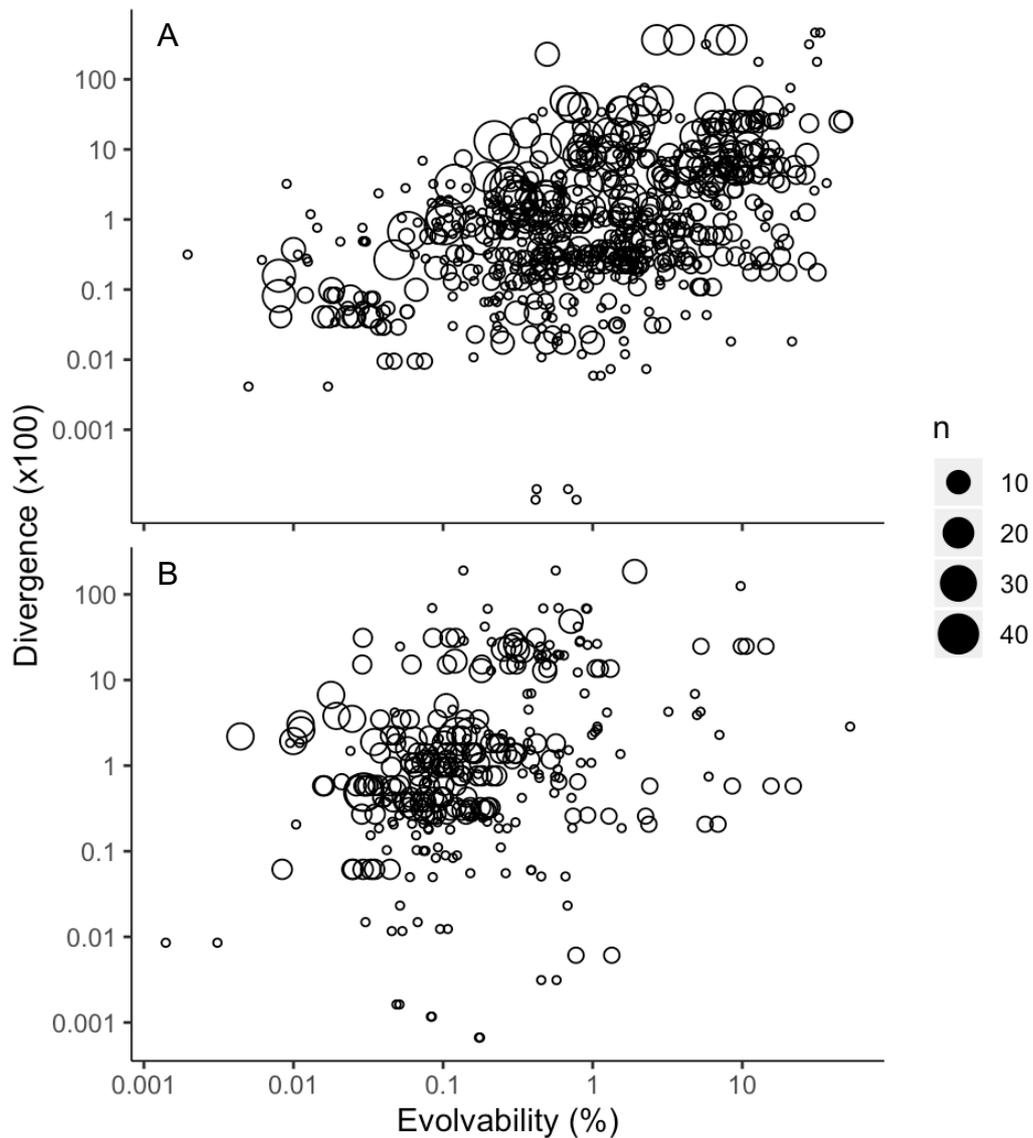
Figure 3.4: Mean-standardised divergence as a function of evolvability on a log-log scale, with *n* = number of populations means per trait. All evolvability estimates are plotted, but the model was fitted to the mean estimate per trait. A: Among-population variance, log(D_M) = -0.11 (±0.27) + 0.74 (±0.08) log(e_μ), R_m^2 = 30%. B: Among-species variance, log(D_M) = 1.38 (±0.51) + 0.55 (±0.13) log(e_μ), R_m^2 = 12%.

I found no difference in the mean divergence among populations and among species (difference in mean divergence on log-scale of 0.34 ± 0.25, *t*-test, p = 0.17). However, when separating plants and animals the mean divergence among species was higher than the mean divergence among populations for plants, but not for animals (Fig. 3.5).
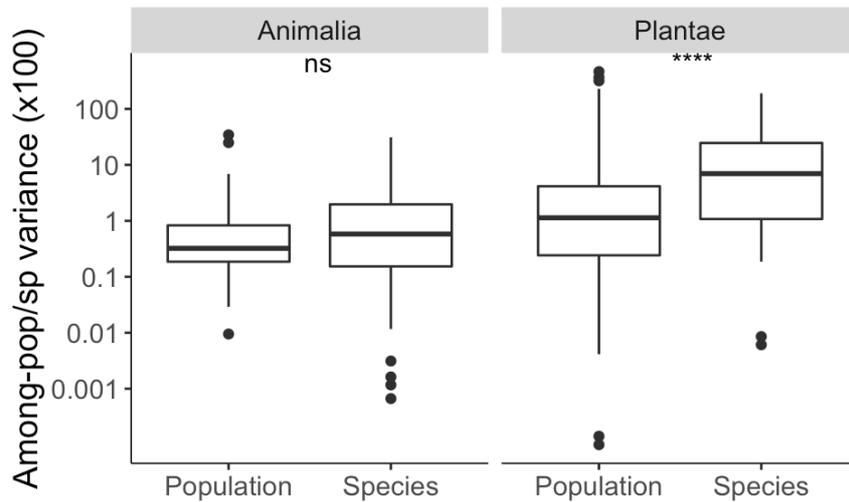
Figure 3.5: Difference in among-population and among-species variance on log-scale for animals and plants. Means compared with a *t*-test, ns: p > 0.05, ****: p < 0.0001

3.1.2 Trait category, type and dimensionality

When analysing the trait categories separately, I found similar scaling relationship between among-population variance and evolvability on a log-log scale (fig 3.6). The slope was 0.68 ±0.19 with an R_m^2 of 29% for life history traits and 0.85 ±0.08 ($R_m^2 = 36\%$) for morphological traits. Excluded trait categories are plotted in Figure A1.

These trait categories may consist of traits with several dimensionalities or from different types, where for instance morphological traits are area, mass/volume, ratio (i.e. shape) and count traits. When I analysed the different trait dimensionality or types separately, the relationship between evolvability and divergence persisted. I found similar scaling relationship and explanatory power for among-population variance and evolvability across the different dimensionalities and types. The slope (log-log scale) of among-population variance on evolvability for linear traits was 0.87 ± 0.19 with a R_m^2 of 26%. For count traits the slope was 0.48 ±0.12 with a R_m^2 of 21% and for ratio traits the slope was 0.80 ±0.14 with the $R_m^2 = 51\%$ (Fig. 3.7, Table 3.3). Trait dimensions or types excluded from the analyses are plotted in Figure A2.

I found no relationship of among-population variance and heritability when considering only linear traits (Table 3.4, Figs. A4 and A5).
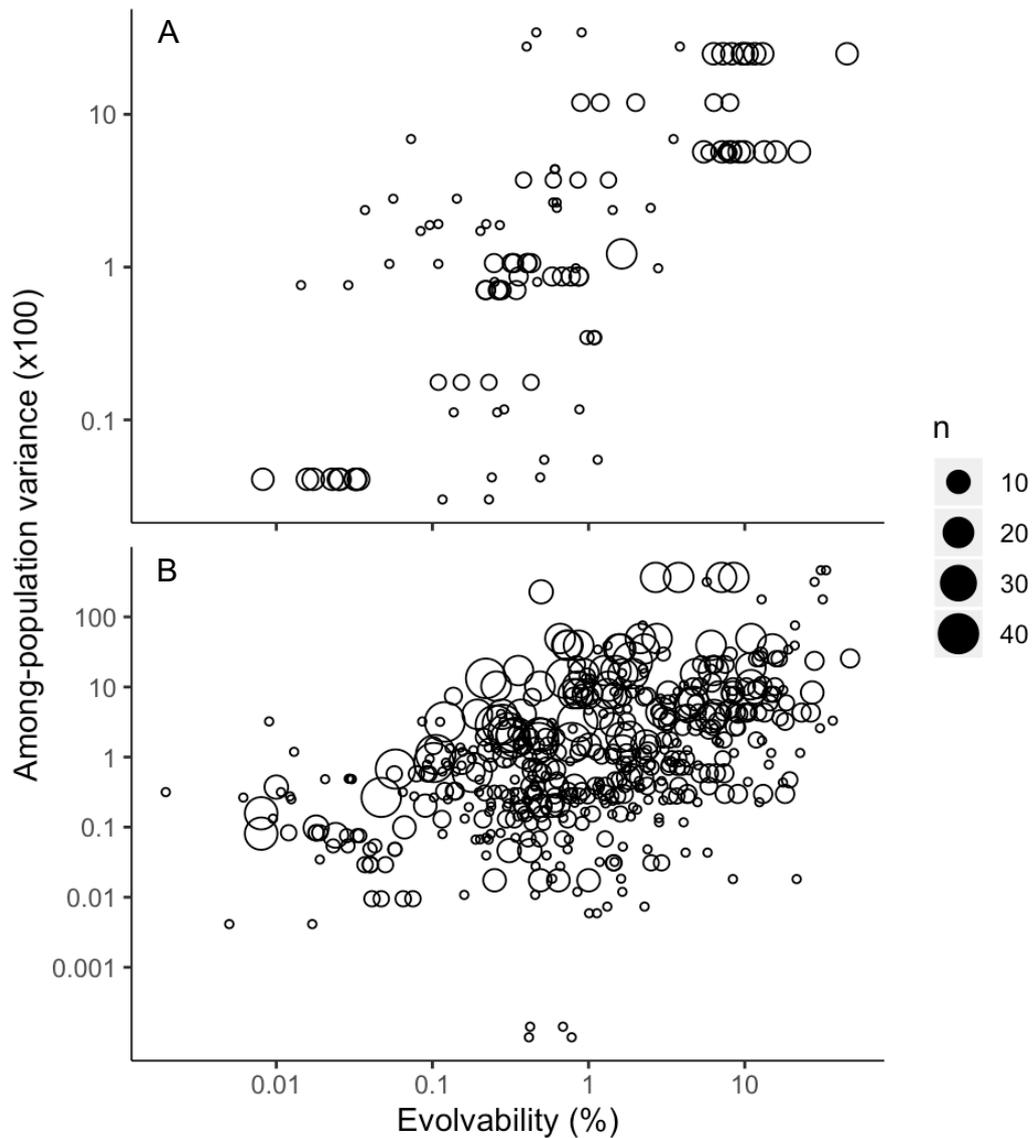
Figure 3.6: Among-population variance as a function of evolvability on a log-log scale, with _n_ = number of populations means per trait. All evolvability estimates are plotted, but the model was fitted to the mean estimate per trait. A: Life history traits, $\log(D_M)$ = 0.56 ($\pm$0.33) + 0.68 ($\pm$0.19) $\log(e_\mu)$, R_m^2 = 29%. B: Morphological traits, $\log(D_M)$ = -0.26 ($\pm$0.30) + 0.85 ($\pm$0.08) $\log(e_\mu)$, R_m^2 = 36%.
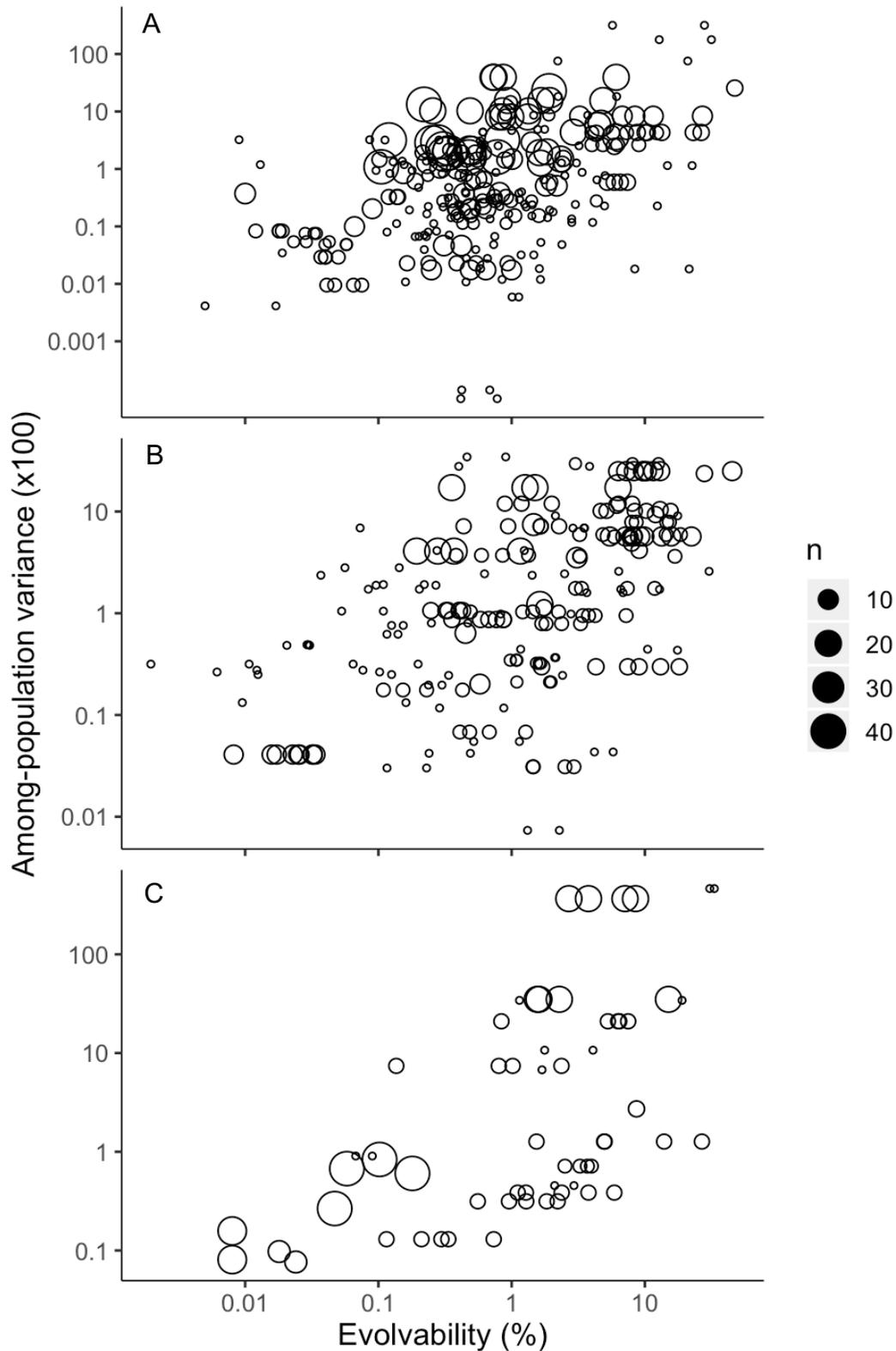
Figure 3.7: Among-population variance as a function of evolvability on a log-log scale, with *n* = number of populations means per trait. All evolvability estimates are plotted, but the model was fitted to the mean estimate per trait. A: Linear traits, $\log(D_M) = -0.45$ (± 0.36) $+ 0.87$ (± 0.19) $\log(e_\mu)$, $R_m^2 = 26\%$. B: Count traits, $\log(D_M) = 0.22$ (± 0.28) $+ 0.48$ (± 0.12) $\log(e_\mu)$, $R_m^2 = 21\%$. C: Ratio traits, $\log(D_M) = 0.68$ (± 0.64) $+ 0.80$ (± 0.14) $\log(e_\mu)$, $R_m^2 = 51\%$.

3.1.3 Comparison of taxa

When analysing data on plants and animals separately, I found scaling relationships between evolvability and divergence at the among-population level in both type of organisms but only for plants at the among species level (Table 3.3). The average evolvability on log-scale for animals was -6.53 ± 0.09, which equals 0.15%. For plants the average was 1.21% (-4.42 ± 0.06 on log-scale). The difference was strongly significant (t-test, $p < 2 \times 10^{-16}$). I found the same difference in the mean divergence, where plants were more divergent on average than where animals (plants: mean $= 0.30 \pm 0.13$, animals: mean $= -0.86 \pm 0.21$, t-test, $p < 5 \times 10^{-6}$). Despite this difference, the scaling relationship between evolvability and divergence was similar, consistent with similar levels of genetic constraints on evolutionary divergence. In this meta-database plant studies among populations were more common (80%) than among species, while for animal studies among populations were less common (37%, Table 3.3, Fig. A3). This may have an impact to the differences I observed.

The intercepts and slopes of the scaling relationships between divergence and evolvability varied among species, as quantified by the variance of the intercepts and slopes of the random-effect term (Table 3.3).

When species were centred on mean evolvability, the scaling relationship of among-population variance and evolvability became slightly steeper (Table 3.3). The slope of evolvability and among-population variance on log-log scale for all traits increased by 8% when mean-centring. For the linear traits the slope increased with 10% when mean-centring. This indicates some difference in distribution of evolvabilities between species.

3.1.4 Effect of environmental variation

I found an effect of the type of environment in which measurements for plants where the mean divergence (log-scale) was higher for traits measured in the field than for traits measured in a controlled environment (Fig 3.8). The difference was statistically significant between all environments, with the least difference between the field and controlled indoor environments. For animals there was no difference in divergence between traits measured in the field and traits measured in controlled indoor environments (Fig. 3.8)

24
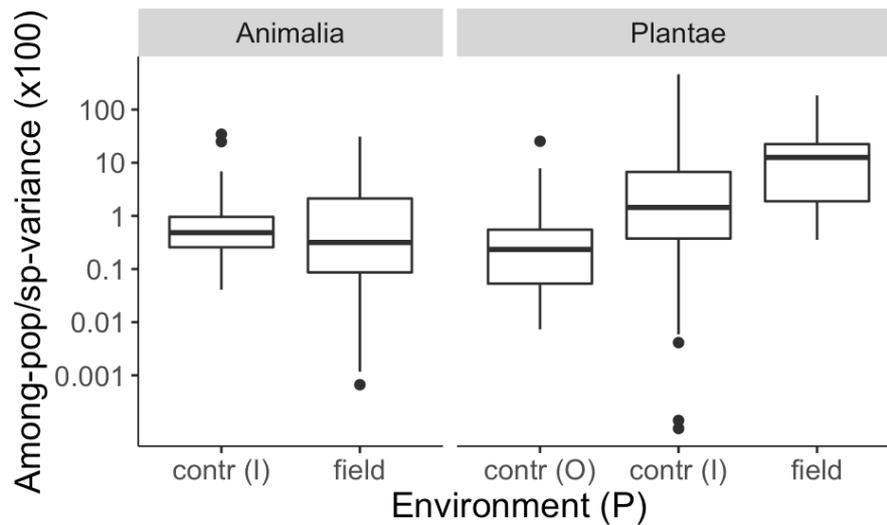
Figure 3.8: Among-population and -species variance on log-scale as a function of the environment where the phenotypic (P) trait measures were taken. I = indoor, O = outdoor. Mean ±SE (log-scale) for animals: controlled (I) = -0.58 ±0.32, field = -1.09 ±0.27, for plants: controlled (O) = -1.54 ±0.40, controlled (I) = 0.37 ±0.14 and field = 1.92 ±0.45. Difference tested with a *t*-test, p<0.004 between controlled (I) and field and controlled (O) and field for plants. For animals there was no significant difference (p = 0.2).

When I included the type of environment in which measures were recorded as a covariate in the random-regression models, there was a slight increase in the precision of the model for both plants and animals. For animals the R_m^2 was 49% for among-population variance as a function of evolvability on log-log scale with environment as a covariate (Fig. 3.9). However, the model including the environment and the one without were not conclusively different (ΔAICc = 0.68) in the model selection and the simplest model had the lowest AICc. Similarly, for plants the simplest model without environment had the lowest AICc, but it was not conclusively different from the one including the environment (ΔAICc = 1.46). For the plants the R_m^2 was 31% for among-population variance as a function of evolvability on log-log scale with environment as a covariate (Fig. 3.10). The model allowing for change in slopes between environments was not considered in either animals (ΔAICc = 2.34 to the simplest model) or plants (ΔAICc = 2.65).
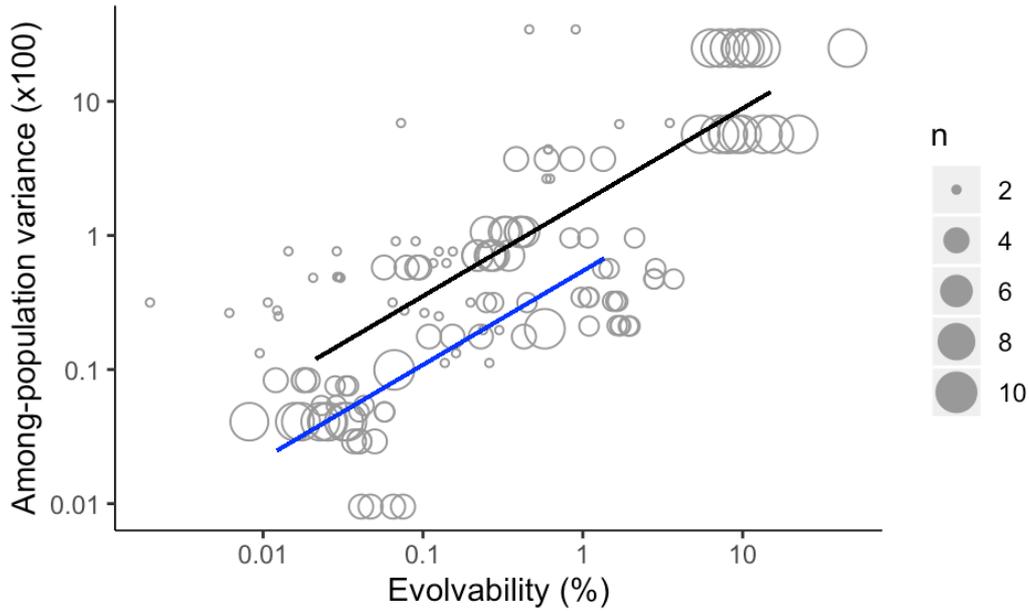
Figure 3.9: Animals: Among-population variance as a function of evolvability on a log-log scale with the environment for the phenotypic measures as a covariate (R^2_m = 49%), *n* = number of population means per trait. All evolvability estimates are plotted, but the model was fitted to the mean estimate per trait. Black line: controlled environment, $\log(D_M) = 0.57\ (\pm0.53) + 0.70\ (\pm0.11)\ \log(e_\mu)$, n_t traits = 34. Blue line: field, $\log(D_M) = -0.61\ (\pm0.82) + 0.70\ (\pm0.11)\ \log(e_\mu)$, n_t traits = 9.
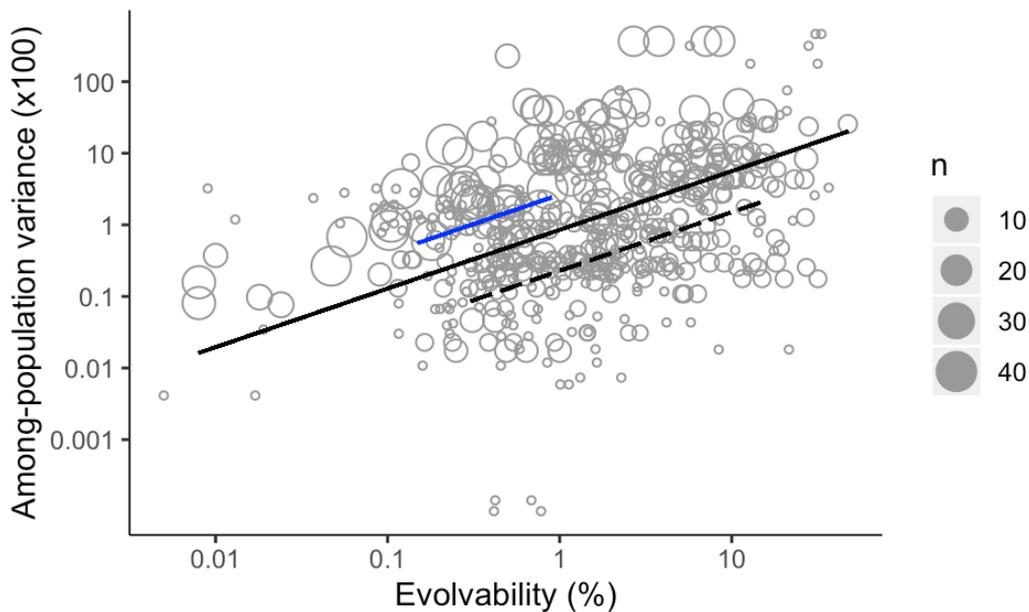
Figure 3.10: Plants: Among-population variance as a function of evolvability on a log-log scale with the environment for the phenotypic measures as a covariate (R^2_m = 31%), *n* = number of population means per trait. All evolvability estimates are plotted, but the model was fitted to the mean estimate per trait. Black solid line: controlled indoor environment, $\log(D_M) = -0.16\ (\pm0.33) + 0.82\ (\pm0.08)\ \log(e_\mu)$, n_t traits = 193. Black dashed line: controlled outdoor environment, $\log(D_M) = -1.48\ (\pm0.89) + 0.82\ (\pm0.08)\ \log(e_\mu)$, nt traits = 30. Blue line: field, $\log(D_M) = 0.96\ (\pm1.60) + 0.82\ (\pm0.08)\ \log(e_\mu)$, n_t traits = 4.

3.1.5 Attenuation bias

The error variance in the mean evolvability estimates per trait on log-scale are presented in Figure 3.11 as percent of the total variance in evolvability. The mean relative error variance was 9.1%. This represents the mean relative error of the regression slopes of evolvability and divergence. The reliability ratio was $K = 0.91$, and the slopes were therefore corrected (eq. 2.13) to improve the accuracy of the slope estimates. The estimated and corrected slopes are presented in Table 3.3.
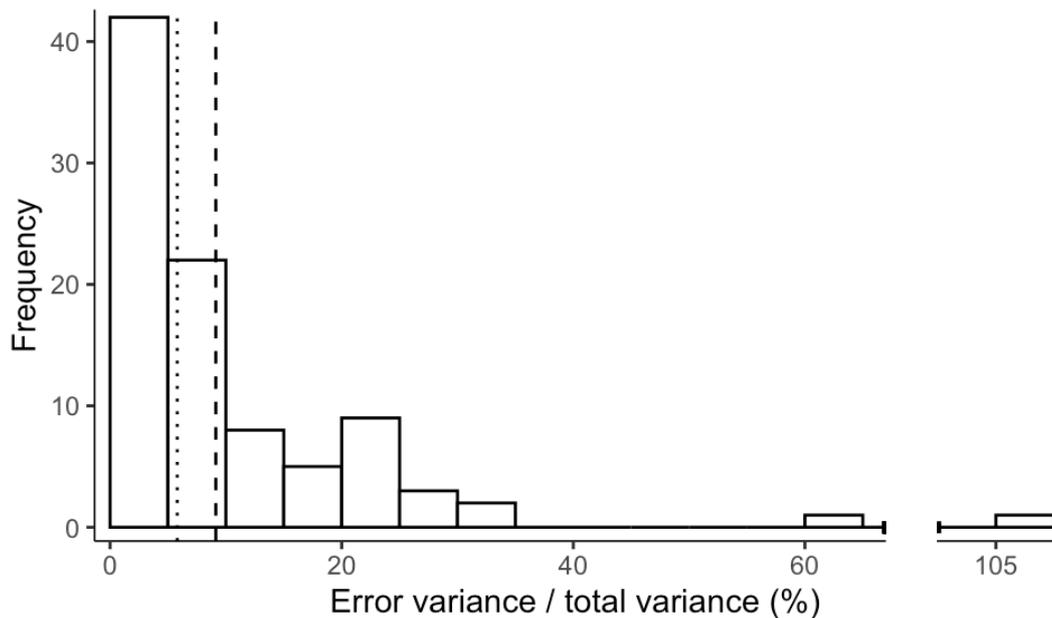
Figure 3.11: Frequency distribution of the error variance of the mean evolvability estimates per trait in percent of the total variance in evolvability on log-scale. Dotted line: median = 5.8%. Dashed line: mean = 9.1%. The error variance of the zero evolvability estimates are plotted, but not included in the estimation of the mean and median.

3.1.6 Correlation of heritability and evolvability

Heritabilities and evolvabilities were essentially uncorrelated (Fig. 3.12), as expected from the contrasting relationships of each to divergence. Evolvability explained 4% of the variation in heritability.
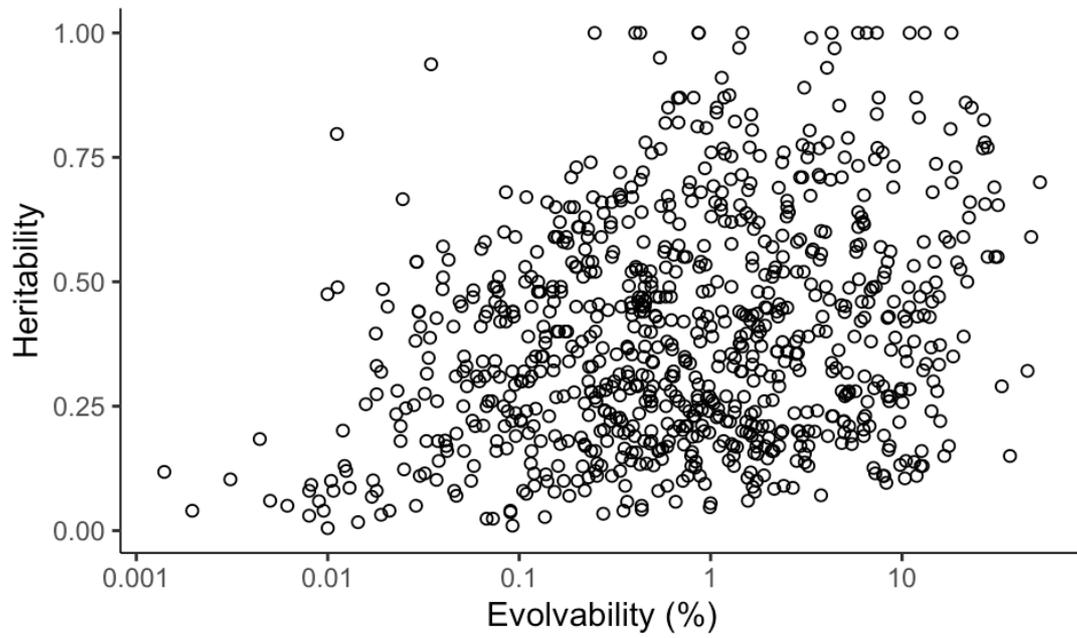
Figure 3.12: Estimates of heritability plotted over evolvability (log-scale). The correlation $r = 0.19 \pm 0.03$ and $R^2 = 4\%$ with evolvability on original scale. For evolvability on log-scale $r = 0.27 \pm 0.03$ and $R^2 = 7\%$.

Table 3.3: Summary of the parameters estimated from the random-regression models for the mean-standardised (D_M) among-species and among-population variance predicted by evolvability on a log-log scale. MCE stands for mean-centred on species mean evolvability.

	$\hat{\alpha}_0 \pm SE$	$\hat{\alpha}_1 \pm SE$	$\hat{\alpha}_1$ corrected	var[$\hat{u}_{0j}$]	var[$\hat{u}_{1j}$]	R^2_m	R^2_c	N	J
Response vector = among-species variance (D_M), fixed effect = evolvability, random effect = genus									
All traits	1.38 ±0.51	0.55 ±0.13	0.61	2.32	—	12%	50%	130	13
Animals	0.69 ±0.73	0.34 ±0.23	0.37	1.72	—	3%	37%	73	8
Plants	1.73 ±0.86	0.65 ±0.16	0.72	3.44	—	13%	61%	57	5
Response vector = among-population variance (D_M), fixed effect = evolvability, random effect = species									
All traits	-0.11 ±0.27	0.74 ±0.08	0.81	1.89	0.05	30%	66%	270	32
Animals	0.24 ±0.46	0.70 ±0.11	0.77	0.87	—	43%	67%	43	6
Plants	-0.25 ±0.31	0.80 ±0.08	0.88	2.11	—	28%	64%	227	26
Linear traits	-0.45 ±0.36	0.87 ±0.19	0.96	2.46	0.34	26%	73%	131	26
Count traits	0.22 ±0.28	0.48 ±0.12	0.53	1.20	—	21%	53%	78	27
Ratio traits	0.68 ±0.64	0.80 ±0.14	0.88	3.34	—	51%	94%	24	9
Life history traits	0.56 ±0.33	0.68 ±0.19	0.75	0.22	—	29%	34%	32	16
Morphological traits	-0.26 ±0.30	0.85 ±0.08	0.94	2.24	—	36%	71%	219	28
MCE all traits	-0.15 ±0.26	0.80 ±0.09	0.88	1.80	0.06	22%	61%	270	32
MCE linear traits	-0.63 ±0.35	0.96 ±0.25	1.06	2.74	0.70	14%	70%	131	26

Abbreviations: $\hat{\alpha}_0$ (SE) = intercept of the fixed effect ± standard error, $\hat{\alpha}_1$ (SE) = slope of the fixed effect ± standard error, $\hat{\alpha}_1$ corrected = attenuation-corrected slope of the fixed effect, var[$\hat{u}_{0j}$] = variance in intercepts of the random effect, var[$\hat{u}_{1j}$] = variance in slopes of the random effect, R^2_m = the marginal coefficient of determination, R^2_c = the conditional coefficient of determination, N = the number of traits, J = the number of species or genera.

Table 3.4: Summary of the parameters estimated from the random-regression models for the mean-standardised (D_M, log scale) or variance-standardised (D_V, log scale) among-species and among-population variance predicted by heritability (original scale).

	$\hat{\alpha}_0$ (±SE)	$\hat{\alpha}_1$ (±SE)	var[$\hat{u}_{0j}$]	R_m^2	R_c^2	N	J
Response vector = among-species variance (D_M), fixed effect = heritability, random effect = genus							
All traits	-1.40 (±0.76)	3.17 (±1.41)	1.13	6%	27%	84	9
Response vector = among-population variance (D_M), fixed effect = heritability, random effect = species							
All traits	-0.36 (±0.38)	0.92 (±0.64)	2.08	1%	43%	267	32
Linear traits	-0.84 (±0.54)	1.28 (±0.91)	3.08	1%	54%	130	26
Response vector = among-species variance (D_V), fixed effect = heritability, random effect = genus							
All traits	0.50 (±1.18)	1.62 (±1.65)	6.53	1%	60%	81	8
Response vector = among-population variance (D_V), fixed effect = heritability, random effect = species							
All traits	-2.16 (±0.35)	2.49 (±0.54)	1.76	7%	53%	235	30
Linear traits	-1.74 (±0.46)	1.56 (±0.77)	2.03	3%	54%	122	24

Abbreviations: $\hat{\alpha}_0$ (SE) = intercept of the fixed effect ± standard error, $\hat{\alpha}_1$ (SE) = slope of the fixed effect ± standard error, $\hat{\alpha}_1$ corrected = attenuation corrected slope of the fixed effect, var[$\hat{u}_{0j}$] = variance in intercepts of the random effect, R_m^2 = the marginal coefficient of determination, R_c^2 = the conditional coefficient of determination, N = the number of traits, J = the number of species or genera.

3.2 Sources of bias

3.2.1 The number of populations

Estimates of population and species divergence tended to increase with the number of populations or species measured per trait (Fig. 3.13) The weighted mean by the number of populations was higher than the mean and there was an increase in the linear regression (slope = 0.63 ±0.13, R^2 = 6%). This means that the divergence measure is not independent of the number of populations or species that are measured per trait. Evolvability, on the other hand, showed no increase with the number of populations or species measured per trait ($\log(e_\mu)$ = -0.36 (±0.17) - 0.12 (±0.10) $\log(n)$, R^2 = 0%), and is therefore independent of the number of populations or species measured.
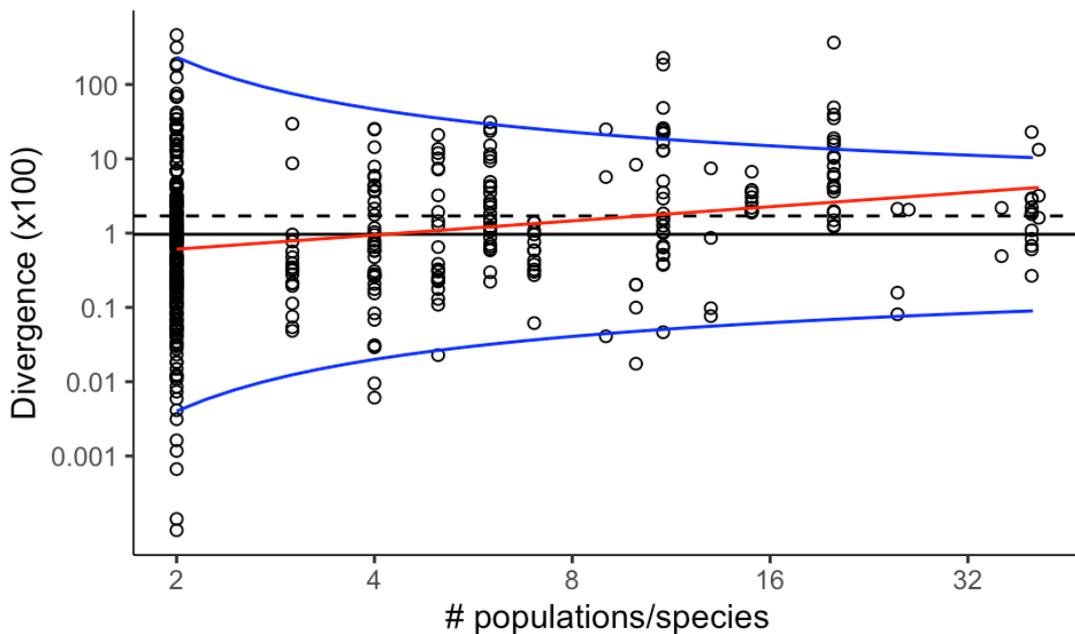
Figure 3.13: Divergence (D_M) as function of the number of populations or species (n) measured per divergence estimate (i.e. per trait) on a log-log scale. Black horizontal line: mean = 0.96 ±0.11, dashed horizontal line: weighted mean = 1.70, blue lines: 95% confidence intervals of the mean. Red line: $\log(D_M)$ = -0.94 (±0.22) + 0.63 (±0.13) $\log(n)$, R^2 = 6%

3.2.2 Publication bias

Evolvabilities did not depend on the number of families included in the breeding design (Fig. 3.14). The number of families explained none of the variance in evolvability estimates (R^2 = 0%), and there was no difference in the mean and the weighted mean. Similarly, there was no relationship between heritability and the number of families in the experimental design

($R^2 = 0\%$, Fig. 3.15). Notice that 2% of the heritability estimates were outside the expected theoretically range of $0 \leq h^2 \geq 1$.
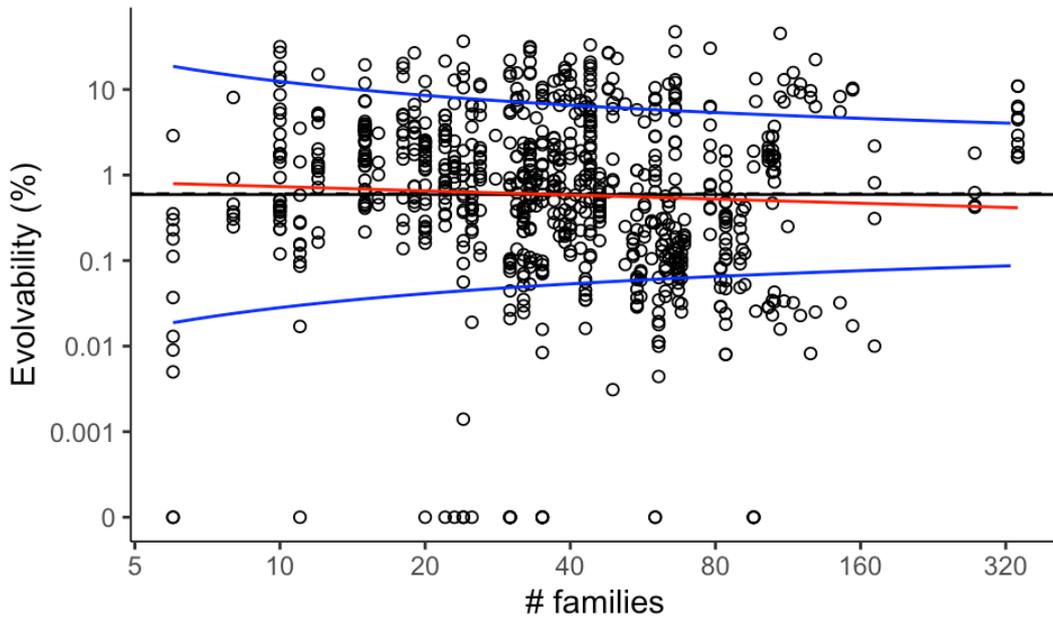
Figure 3.14: Evolvability over the number of families (n) in the experimental design (i.e. estimation of V_A) on a log-log scale. Black horizontal line: Mean = -0.52 ±0.08 (= 0.6%), dashed horizontal line: weighted mean = 0.50, blue lines: 95% confidence intervals of the mean. Red line: $\log(e_\mu) = 0.06\ (\pm0.41) - 0.16\ (\pm0.11)\ \log(n)$, $R^2 = 0\%$. Zeros included as 0.0001% in the regression.
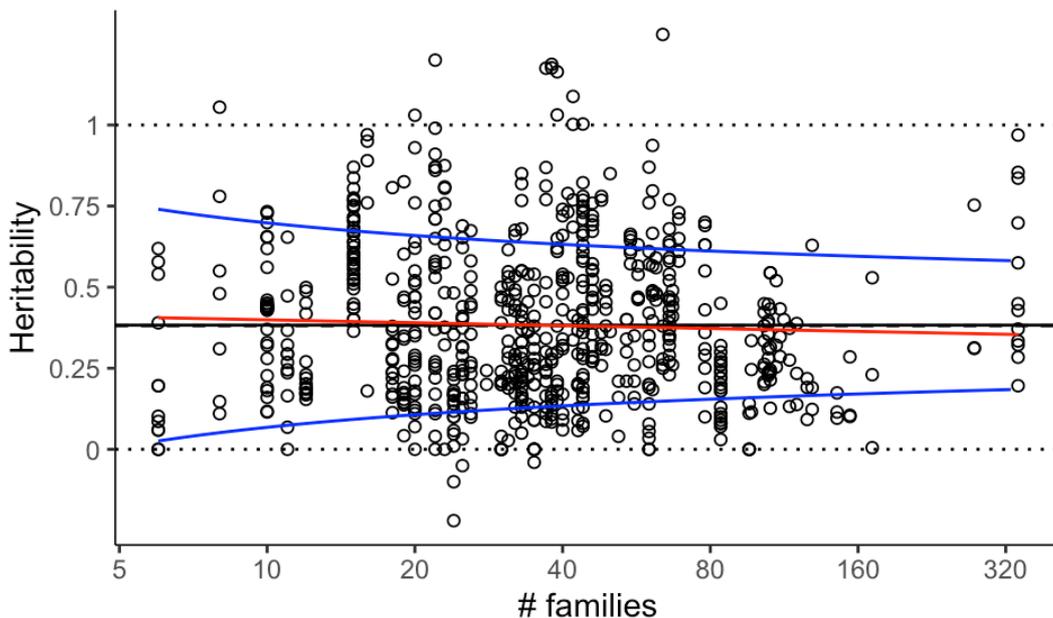
Figure 3.15: Heritability (original scale) over the number of families (n, log-scale) in the experimental design (i.e. estimation of V_A). Black horizontal line: mean = 0.38, dashed horizontal line: weighted mean = 0.38 ±0.01, dotted horizontal lines: theoretical range of h^2, blue lines: 95% confidence intervals of the mean. Red line: $h^2 = 0.43\ (\pm0.04) - 0.01\ (\pm0.01)\ \log(n)$, $R^2 = 0\%$. One value of $h^2 > 2$ excluded.

4   Discussion

In this study, evolvability measured as mean-standardised genetic variance predicted evolutionary divergence among populations and among species. These results confirm the various case studies (Andersson, 1991, Bolstad et al., 2014, Houle et al., 2017, McGlothlin et al., 2018) and theory (Arnold, 1992, Schluter, 1996, Blows and Hoffmann, 2005), and suggests that genetic constraint to evolution may be a general case. This supports the idea that parameters describing microevolutionary processes may be important in describing patterns of divergence among species and populations on the macroevolutionary timescale. Similar patterns were found among all organisms and trait groups assessed (Table 3.3).

An alternative interpretation to the pattern of increasing divergence with increasing evolvability, is that selection shapes genetic variation and among populations and species divergence similarly. Consider an adaptive ridge in a phenotype-fitness landscape. Trait's whose variation aligns with the ridge will be under weak selection, and will possibly exhibit high genetic variance and may be highly diverged along the ridge. Trait variation orthogonal to the adaptive ridge will experience strong stabilising selection that may decrease the genetic variation and the divergence among populations and species. Selection would then generate the correlation between evolvability and divergence. This would be hard to distinguish from genetic constraints, since the history of the traits examined here is not known. Although there are suggested cases where selection may be shaping genetic variation to fit the fitness landscape (Pavlicev et al., 2011, Jones et al., 2014), there is also evidence where genetic variation does not align with the direction of selection (Blows et al., 2004). Theory also suggests that there is no strong relationship between selection and direction of genetic variation (Armbruster et al., 2014). Thus, I believe the relationship between evolvability and genetic divergence observed in this study may be explained to some extent by genetic constraints. If selection is shaping genetic variation (and the G-matrix) and the divergence between populations and species, this would also provide a link between micro- and-macroevolution.

The current data confirms the weak relationship between heritability and evolvability (Hansen et al., 2011, Houle, 1992, Opedal, 2019, Opedal et al., 2017), and also demonstrate that heritabilities fail as a predictor of divergence (Table 3.4). These results are consistent with the findings in Houle (1992), Hansen et al. (2011), Opedal et al. (2017) and Opedal (2019), who report similarly low correlations. The evolvabilities and heritabilities from these studies originate from a range of different traits and organisms in varying environments. Hoffmann et al. (2016), on the other hand, found higher correlations ($r = 0.37$ ($R^2 = 14\%$) on original scale

and $r = 0.54$ ($R^2 = 29\%$) on log-scale) between heritability and evolvability. In their analysis heritability and evolvability trait estimates originated form livestock, where the environments are considered to be relatively constant across the different commercial farms. This suggests that with constant environments heritability is closer to a measure of genetic variance available for evolution, however the error is still large. Heritability would therefore likely fail to predict divergence even in more constant environments. In this study, the environments are assumed to be highly variable due to the estimates originating from natural populations for a wide range of traits and organisms. Thus, the environment stretches the scale of heritability, so it does not show the true amount of genetic variance and fails to predict evolutionary divergence.

4.1   Genetic correlations and conditional evolvability

The genetic variance in a univariate trait may be bound up in correlation with other traits (eq. 1.2). The amount of genetic variance available for evolution in one trait is less than the observed amount if other correlated traits are kept constant. Genetic correlation may limit the amount of genetic variance available for a trait to evolve (Hansen et al., 2003a) and this may affect the observed relationship between evolvability and divergence by including more noise in the predictor variable. Despite this caveat, I found a relationship between evolvability and divergence with evolvability predicting up to 51% of the divergence observed among populations. This might suggest that the amount of genetic variation may pose a constraint to evolution even on a macroevolutionary timescale.

The ability of a trait to evolve independently of other traits is the measured by its conditional evolvability (Hansen et al., 2003a). Conditional evolvability incorporates genetic correlations with other traits under opposing or stabilizing selection, that may constrain the trait in question to respond to selection. It may be measured as the response to a unit strength of selection in a trait keeping all other traits constant (Hansen, 2003, Hansen et al., 2003a, Hansen and Houle, 2008). Conditional evolvability is therefore a measure of potential for evolution unconstrained by genetic correlations. Further studies are necessary to better understand how genetic correlations affects the observed genetic constraints on evolutionary divergence. I expect that even stronger genetic constraint to divergence among populations or species may be found.

4.2   Timescale

The relationship between divergence and evolvability was weaker with a higher intercept and a shallower slope at the among-species than at the among-population level. I found no

difference in the mean divergence among species and among populations, which may seem counterintuitive. This pattern is, however, consistent with the "blunderbuss pattern" found by (Uyeda et al., 2011). In this study the authors showed that the divergence among populations and species does not increase with time until one million years has passed since a common ancestor (Uyeda et al., 2011). One explanation to why I don't observe any difference in the average divergence among populations and among species, may therefore be that the species compared for trait divergence in this study are closely related. It is plausible that less than a million years have passed since a common ancestor. The difference observed between the among-population and among-species levels in slopes and intercepts of the scaling relationship between evolvability and divergence can therefore not be explained just by the difference in the amount of divergence observed between the two levels. It is possible, however, that enough time has passed at the among-species level that the species have had time to reach their optima. Some divergence would therefore be present even when there is little genetic variance due to a longer timescale. The ability to detect a scaling relationship between evolvability and divergence hence decreases with increasing timescale.

There was an effect of timescale on divergence and evolvability, when only plants were considered (Fig. 3.5). I found higher divergence on average at the among-species level. However, there was a higher proportion of traits measured in the field among species ($n = 20$, 35%) than among populations ($n = 4$, 0.02%) for plants and the traits measured in the field also showed higher divergence than traits measured in a controlled environment (Fig. 3.8). The effect of timescale on the average divergence in plants may therefore be confounded with the effect of environment.

Even though additive genetic variance (and the G-matrix) have features that stay quite stable over time (Schluter, 1984, Shaw et al., 1995, Jones et al., 2003, 2007, Arnold et al., 2008, Hohenlohe and Arnold, 2008), the G-matrix itself may be subject to evolution. There is no guarantee for the consistency of the G-matrix, as new patterns in mutation and selection may change its size, shape and orientation (Turelli et al., 1988, Barton and Turelli, 1989, Turelli and Barton, 1994, Shaw et al., 1995, Jones et al., 2007). These changes in the G-matrix will result in changes in amount of additive genetic variance for a trait in the univariate case. The accuracy with which evolvability predicts divergence may therefore decrease as the timescale increases.

In this study I considered only two discrete measures of timescale by comparing divergence among populations within species with divergence among species. This assumes that longer time has passed since divergence of populations within a species than the divergence between populations from different species. Although this holds in general, there may be cases

where the timescale is similar at both levels. Considering the phylogeny and the generations passed since divergence would improve the analyses. This would not be easy, however, because phylogeny is generally not estimated similarly at the species and population levels, and there are very few studies that estimate phylogeny among populations.

4.3   Selection and random genetic drift

The different evolutionary forces operating on the traits affects the predicted scaling relationship between divergence and evolvability (fig 1.1). The scaling relationships found in this study ranged from 0.53 to 1.06 (after correction) among populations. This is compatible with evolutionary processes dominated by genetic drift and populations tracking fluctuating optima with the rate in the same order of magnitude as movement in the optima (prediction: slope of 1 and 0-1 respectively (Bolstad et al., 2014, Houle et al., 2017)). However, the predicted scaling relationships of divergence and evolvability are based on one specific process of evolution for a pair of diverging species or populations. These predictions represent upper limits, when the mode of evolution remains constant. A combination of evolutionary forces operating at different times would yield scaling relationships somewhere between 2 and 0 depending on how long the different forces operate. Hence, in practice it is difficult to distinguish between e.g. drift and directional selection followed by stabilising selection. In this study the evolutionary forces at work presumably differ both in time for the same trait and between different traits. Thus, concluding on a dominating mode of evolution may be difficult.

Even though I cannot conclude on a dominating mode of evolution, the scaling relationships between evolvability and divergence found for different trait types and organisms were surprisingly stable. The variance in slopes (scaling relationships, after correction) between the different subgroups analysed was 0.02 at the among-population level and 0.04 for both among-population and among-species level. The mean scaling relationship for evolvability and among-population variance was 0.85 ± 0.05. Perhaps this is a general predicted scaling relationship when the mode of evolution varies among traits and organisms. Would the same mean scaling relationship be found if this study was repeated?

4.4   Effect of environmental variation

The R_m^2 increased from 43% to 49% for the relationship between evolvability and among-population variance in animals when the environment in which traits were recorded was included as a cofactor. There is a slight effect of the environment, with higher divergence between the populations measured in the field. This may indicate that some of the divergence

36

measured in the field also include some phenotypic plasticity. However, the slopes did not differ between the environments. This means that even though the field measures may include some phenotypic plasticity, the relationship between evolvability and divergence holds. Which may be expected due to the positive relationship observed between genetic and residual variation (Hansen et al., 2011, Opedal et al., 2017). I found no difference in the divergence between the field and a controlled environment for animals, when both among-population and among-species variance was considered (Fig. 3.8). Therefore, phenotypic plasticity does not seem to affect the observed divergence to a great extent in animals.

For plants there was considerable difference in the mean divergence between the field and controlled environments, where the field measures showed more divergence. Consistent with the tendency for plants to be plastic in response to the environment (Pélabon et al., 2011). I also found similar patterns for plants as in animals, where the precision increased when both evolvability and environment predicted among-population variance, and only the intercept differed among environments. It has been suggested that plasticity is also aligned with directions of high plasticity (Noble et al., 2019), which is consistent with these results.

4.5 Comparison of traits and taxa

There were some differences among plants and animals. For instance, at shorter timescale (among populations) plants and animals show similar strong scaling relationships between evolvability and divergence. On a longer timescale (among species), the scaling relationship between divergence and evolvability differed. For animals, evolvability did not predict divergence among species. There were, however, proportionally more studies among species in animals than in plants, and vice versa for the studies among populations in the meta-database (Table 3.4, Fig. A4). If the amount of data for the two levels would have been proportionally equal for both plants and animals, the patterns could perhaps have been more similar at the among-species level.

There was consistency of the relationship between evolvability and divergence across different trait dimensionalities and categories. This suggest that it is not spurious correlation between trait dimensionality, type or category and evolvability and divergence that is causing the observed relationship between evolvability and divergence. Though the scaling relationship of evolvability and divergence was mostly similar across the different trait dimensions or trait types, there were slight differences. Specific analyses of the mean and standard deviation relationships may be necessary to further understand whether these differences result from the

properties of the measurement or from real biological effects. The question is also if the number of traits available for analyses might have influenced the difference observed.

4.6   Sources of bias

4.6.1   Number of populations

The divergence measure increases with the number of populations or species measured (R^2 = 6%). The estimation method of the divergence measure (eq. 2.8) that is a variance estimate of the log-transformed trait means, does not yield a bias towards a larger number of populations. It is possible that the bias comes from the expectation of catching more of the variance in a trait when more populations from different environments are sampled. This would only confound the observed pattern between evolvability and divergence if evolvability also increased with the number of populations sampled. I found no such increase (R^2 = 0%). Thus, the slight bias in the divergence measure is not an issue to the analyses and does not affect the conclusion of evolvability predicting divergence among populations and species.

4.6.2   Publication bias

There was no detectable publication bias in the evolvability and heritability measures (Fig. 3.14 and 3.15), contrary to Palmer (2000) who found a publication bias in heritability estimates. In Palmer (2000) 166 heritabilities form Weigensberg and Roff (1996) were examined, while I had 711 estimates for heritability that reported the number of families in the experimental design. I may therefore have captured a larger part of the published heritabilities. I also included heritabilities estimated from published additive genetic- and phenotypic variances, that were not initially published as heritability. This was because I wanted to examine if there was a publication bias in additive genetic variance measured as both heritability and evolvability. Thus, my results are not directly comparable to Palmer (2000).

Theoretically we expect a range of genetic variances in the wild for different traits and organisms. Often more than one trait is assessed per study. Hence, studies estimating additive genetic variance are more likely to be published than a study testing a general effect size if some estimates are small and not statistically significant due to small sample size. The non-significance may be interpreted as biologically meaningful (i.e. no genetic variance for the trait) and/or the study may be published with several genetic variance estimates of which only a few are non-significant.

It may be noticed that the number of families in the experimental design is not necessarily directly linked to the power of the estimation of additive genetic variance to detect significance. This varies across the different experimental designs (Palmer, 2000). The number of families is, however, an indication of the power to detect significance and is the same variable Palmer (2000) used.

4.7 Conclusion and future directions

In this study I found that evolvability predicted evolutionary divergence among populations with surprisingly high precision considering the many caveats to this analysis. The pattern was general, with similar scaling relationships between evolvability and divergence among traits and organisms. Evolvability was proven a measure of evolutionary potential, comparable over different traits, organisms and environments. This may suggest that there is genetic constraint to evolutionary divergence and that the microevolutionary process can to some extant explain macroevolution, however, the accuracy decreases with time.

Future studies could consider the genetic correlations that may be obscuring the univariate pattern of divergence, using conditional evolvability and the G-matrix. However, estimates of G may be sparser in the primary scientific literature. It also raises the question of what traits are genetically correlated and how many traits is enough to understand the genetic constraints also posed by genetic correlations. In the future it would be interesting to understand and make predictions for how several evolutionary forces acting over different points in time and on different traits would affect the scaling relationship between divergence and evolvability. Lastly, further investigation may be done to how the evolution of evolvability (and the G-matrix) over time affects genetic constraints, and if the G-matrix tends to align with the adaptive landscape. Are some aspects of selection constant enough to produce the consistency observed in the G-matrix?

5 References

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Contr.,* 19**,** 716-723.

Andersson, S. 1991. Quantitative genetic variation in a population of *Crepis tectorum* subsp. *pumila* (Asteraceae). *Biol. J. Linn. Soc. Lond.,* 44**,** 381-393.

Andersson, S. 1997. Genetic constraints on phenotypic evolution in *Nigella* (Ranunculaceae). *Biol. J. Linn. Soc. Lond.,* 62**,** 519-532.

Armbruster, W. S., Pélabon, C., Bolstad, G. H. & Hansen, T. F. 2014. Integrated phenotypes: understanding trait covariation in plants and animals. *Phil. Trans.R. Soc. B,* 369**,** 20130245

Arnold, S. J. 1988 Quantitative genetics and selection in natural populations: microevolution of vertebral numbers in the garter snake *Thamnophis elegans*. *In:* Weir, B. S., Eisen, E. J., Goodman, M. M. & Namkoong G. (eds.) Proceedings of the second international conference on quantitative genetics. Sunderland: Sinauer Associates Inc., 619-638.

Arnold, S. J. 1992. Constraints on phenotypic evolution. *Am. Nat.,* 140**,** S85-S107

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

Arnold, S. J., Pfrender, M. E. & Jones, A. G. 2001. The adaptive landscape as a conceptual bridge between micro-and macroevolution. *Genetica,* 112/113**,** 9-32

Arnold, S. J. & Phillips, P. C. 1999. Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the garter snake, *Thamnophis elegans*. *Evolution,* 53**,** 1516-1527.

Baer, C. F. & Lynch, M. 2003. Correlated evolution of life-history with size at maturity in Daphnia pulicaria: patterns within and between populations. *Genet. Res. (Camb.),* 81**,** 123-132.

Baker, R. H. & Wilkinson, G. S. 2003. Phylogenetic analysis of correlation structure in stalk-eyed flies (*Diasemopsis*, Diopsidae). *Evolution,* 57**,** 87-103.

Barrett, S. C. & Shore, J. S. 1987. Variation and evolution of breeding systems in the *Turnera ulmifolia* L. complex (Turneraceae). *Evolution,* 41**,** 340-354.

Barton, K. 2009. MuMIn: multi-model inference. R package version 1.43.6. *http://r-forge. r-project. org/projects/mumin/.*

Barton, N. H. & Turelli, M. 1989. Evolutionary quantitative genetics: how little do we know? *Annu. Rev. Genet.,* 23**,** 337-370.

Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.,* 67**,** 1-48.

Bégin, M. & Roff, D. A. 2003. The constancy of the G matrix through species divergence and the effects of quantitative genetic constraints on phenotypic evolution: a case study in crickets. *Evolution,* 57**,** 1107-1120.

Bégin, M. & Roff, D. A. 2004. From micro-to macroevolution through quantitative genetic variation: positive evidence from field crickets. *Evolution,* 58**,** 2287-2304.

Billington, H., Mortimer, A. & Mcneilly, T. 1988. Divergence and genetic structure in adjacent grass populations. I. Quantitative genetics. *Evolution,* 42**,** 1267-1277.

Bissell, E. & Diggle, P. 2010. Modular genetic architecture of floral morphology in *Nicotiana*: quantitative genetic and comparative phenotypic approaches to floral integration. *J. Evol. Biol.,* 23**,** 1744-1758.

Blows, M. W., Chenoweth, S. F. & Hine, E. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. *Am. Nat.,* 163**,** 329-340.

Blows, M. W. & Hoffmann, A. A. 2005. A reassessment of genetic limits to evolutionary change. *Ecology,* 86**,** 1371-1384.

Bolstad, G. H., Hansen, T. F., Pélabon, C., Falahati-Anbaran, M., Pérez-Barrales, R. & Armbruster, W. S. 2014. Genetic constraints predict evolutionary divergence in *Dalechampia* blossoms. *Phil. Trans. R. Soc. B,* 369**,** 20130255.

Bonnin, I., Prosperi, J.-M. & Olivieri, I. 1997. Comparison of quantitative genetic parameters between two natural populations of a selfing plant species, *Medicago truncatula* Gaertn. *Theor. Appl. Genet.,* 94**,** 641-651.

Carr, D. E. & Fenster, C. B. 1994. Levels of genetic variation and covariation for *Mimulus* (Scrophulariaceae) floral traits. *Heredity,* 72**,** 606-618.

Carter, M. E. B. & Murdy, W. H. 1986. Divergence for sexual and asexual reproductive characters in *Talinum mengesii* (Portulacaceae). *Bull. Torrey Bot. Club,*113**,** 259-267.

Caruso, C. M. 2004. The quantitative genetics of floral trait variation in *Lobelia*: potential constraints on adaptive evolution. *Evolution,* 58**,** 732-740.

Caruso, C. M., Maherali, H., Mikulyuk, A., Carlson, K. & Jackson, R. B. 2005. Genetic variance and covariance for physiological traits in *Lobelia*: are there constraints on adaptive evolution? *Evolution,* 59**,** 826-837.

Caruso, C. M., Peterson, S. B. & Ridley, C. E. 2003. Natural selection on floral traits of *Lobelia* (Lobeliaceae): spatial and temporal variation. *Am. J. Bot.,* 90**,** 1333-1340.

Charlesworth, B., Lande, R. & Slatkin, M. 1982. A neo-Darwinian commentary on macroevolution. *Evolution,* 36**,** 474-498

Charlesworth, D. & Mayer, S. 1995. Genetic variability of plant characters in the partial inbreeder *Collinsia heterophylla* (Scrophulariaceae). *Am. J. Bot.,* 82**,** 112-120.

Charmantier, A., Kruuk, L., Blondel, J. & Lambrechts, M. 2004. Testing for microevolution in body size in three blue tit populations. *J. Evol. Biol.,* 17**,** 732-743.

Chenoweth, S. F., Rundle, H. D. & Blows, M. W. 2010. The contribution of selection and genetic constraints to phenotypic divergence. *Am. Nat.,* 175**,** 186-196.

Cheverud, J. M. 1996. Quantitative genetic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. *J. Evol. Biol.,* 9**,** 5-42.

Colautti, R. I. & Barrett, S. C. 2011. Population divergence along lines of genetic variance and covariance in the invasive plant *Lythrum salicaria* in eastern North America. *Evolution,* 65**,** 2514-2529.

Conner, J. K. & Hartl, D. L. 2004. *A primer of ecological genetics*, Sunderland: Sinauer Associates Inc.

Coyne, J. A. & Beecham, E. 1987. Heritability of two morphological characters within and among natural populations of *Drosophila melanogaster. Genetics,* 117**,** 727-737.

Culley, T. M., Dunbar-Wallis, A. K., Sakai, A. K., Weller, S. G., Mishio, M., Campbell, D. R. & Herzenach, M. 2006. Genetic variation of ecophysiological traits in two gynodioecious species of *Schiedea* (Caryophyllaceae). *New Phytol.,* 169**,** 589-601.

Delahaie, B., Charmantier, A., Chantepie, S., Garant, D., Porlier, M. & Teplitsky, C. 2017. Conserved G-matrices of morphological and life-history traits among continental and island blue tit populations. *Heredity,* 119**,** 76-87.

Delesalle, V. A. & Mazer, S. J. 1995. The structure of phenotypic variation in gender and floral traits within and among populations of *Spergularia marina* (Caryophyllaceae). *Am. J. Bot.,* 82**,** 798-810.

Eldredge, N. & Gould, S. J. 1972. Punctuated equilibria: an alternative to phyletic gradualism. *In:* Schopf, T. & Thomas, J. (eds.) *Models of Paleobiology.* San Francisco: Freeman Cooper.

Elle, E. 1998. The quantitative genetics of sex allocation in the andromonoecious perennial, *Solanum carolinense* (L.). *Heredity,* 80**,** 481-488.

Estes, S. & Arnold, S. J. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. *Am. Nat.,* 169**,** 227-244.

Evans, A. & Marshall, M. 1996. Developmental instability in *Brassica campestris* (Cruciferae): fluctuating asymmetry of foliar and floral traits. *J. Evol. Biol.,* 9**,** 717-736.

Falconer, D. & Mackay, T. 1996. *Introduction to quantitative genetics,* Essex, UK: Longman.

Fenster, C. & Carr, D. 1997. Genetics of sex allocation in *Mimulus* (Scrophulariaceae). *J. Evol. Biol.,* 10**,** 641-661.

Flux, J. & Flux, M. 1982. Artificial selection and gene flow in wild starlings, *Sturnus vulgaris*. *Naturwissenschaften,* 69**,** 96-97.

Futuyma, D. J. 2010. Evolutionary constraint and ecological consequences. *Evolution,* 64**,** 1865-1884.

Garcia-Gonzalez, F., Simmons, L. W., Tomkins, J. L., Kotiaho, J. S. & Evans, J. P. 2012. Comparing evolvabilities: common errors surrounding the calculation and use of coefficients of additive genetic variation. *Evolution,* 66**,** 2341-2349.

Gibbs, H. L. 1988. Heritability and selection on clutch size in Darwin's medium ground finches (*Geospiza fortis*). *Evolution,* 42**,** 750-762.

Gingerich, P. D. 1984. Punctuated equilibria-where is the evidence? *Syst. Zool.,* 33**,** 335-338.

Gould, S. J. 2002. *The structure of evolutionary theory*, Cambridge: Harvard University Press.

Gould, S. J. & Eldredge, N. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology,* 3**,** 115-151.

Grant, P. & Price, T. 1981. Population variation in continuously varying traits as an ecological genetics problem. *Am. Zool.,* 21**,** 795-811.

Hangartner, S., Lasne, C., Sgro, C. M., Connallon, T. & Monro, K. 2020. Genetic covariances promote climatic adaptation in Australian *Drosophila*. *Evolution.* 74**,** 326-337

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution,* 51**,** 1341-1351.

Hansen, T. F. 2003. Is modularity necessary for evolvability?: Remarks on the relationship between pleiotropy and evolvability. *Biosystems,* 69**,** 83-94.

Hansen, T. F. 2006. The evolution of genetic architecture. *Annu. Rev. Ecol. Evol. Syst.,* 37**,** 123-157.

Hansen, T. F. 2012. Adaptive landscapes and macroevolutionary dynamics. *In:* Svensson, E. & Calsbeek, R. (eds.) *The adaptive landscape in evolutionary biology.* Oxford: Oxford University Press., 205-226.

Hansen, T. F., Armbruster, W. S., Carlson, M. L. & Pélabon, C. 2003a. Evolvability and genetic constraint in *Dalechampia* blossoms: genetic correlations and conditional evolvability. *J. Exp. Zool. B Mol. Dev. Evol.,* 296**,** 23-39.

Hansen, T. F. & Bartoszek, K. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. *Syst. Biol.,* 61**,** 413-425.

Hansen, T. F. & Houle, D. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *J. Evol. Biol.,* 21**,** 1201-1219.

Hansen, T. F. & Martins, E. P. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution,* 50**,** 1404-1417.

Hansen, T. F., Pélabon, C., Armbruster, W. S. & Carlson, M. L. 2003b. Evolvability and genetic constraint in *Dalechampia* blossoms: components of variance and measures of evolvability. *J. Evol. Biol.,* 16**,** 754-766.

Hansen, T. F., Pélabon, C. & Houle, D. 2011. Heritability is not evolvability. *Evol. Biol.,* 38**,** 258-277.

Hansen, T. F., Pienaar, J. & Orzack, S. H. 2008. A comparative method for studying adaptation to a randomly evolving environment. *Evolution,* 62**,** 1965-1977.

Hereford, J., Hansen, T. F. & Houle, D. 2004. Comparing strengths of directional selection: how strong is strong? *Evolution,* 58**,** 2133-2143.

Herlihy, C. R. & Eckert, C. G. 2007. Evolutionary analysis of a key floral trait in *Aquilegia canadensis* (Ranunculaceae): genetic variation in herkogamy and its effect on the mating system. *Evolution,* 61**,** 1661-1674.

Hoffmann, A. A., Merilä, J. & Kristensen, T. N. 2016. Heritability and evolvability of fitness and nonfitness traits: lessons from livestock. *Evolution,* 70**,** 1770-1779.

Hohenlohe, P. A. & Arnold, S. J. 2008. MIPoD: a hypothesis-testing framework for microevolutionary inference from patterns of divergence. *Am. Nat.,* 171**,** 366-385.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics,* 130**,** 195-204.

Houle, D., Bolstad, G. H., Van Der Linde, K. & Hansen, T. F. 2017. Mutation predicts 40 million years of fly wing evolution. *Nature,* 548**,** 447-450.

Houle, D. & Fierst, J. 2013. Properties of spontaneous mutational variance and covariance for wing size and shape in Drosophila melanogaster. *Evolution,* 67**,** 1116-1130.

Houle, D., Pélabon, C., Wagner, G. P. & Hansen, T. F. 2011. Measurement and meaning in biology. *Q. Rev. Biol.,* 86**,** 3-34.

Housworth, E. A., Martins, E. P. & Lynch, M. 2004. The phylogenetic mixed model. *Am. Nat.,* 163**,** 84-96.

Jones, A. G., Arnold, S. J. & Bürger, R. 2003. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution,* 57**,** 1747-1760.

Jones, A. G., Arnold, S. J. & Bürger, R. 2007. The mutation matrix and the evolution of evolvability. *Evolution,* 61**,** 727-745.

Jones, A. G., Bürger, R. & Arnold, S. J. 2014. Epistasis and natural selection shape the mutational architecture of complex traits. *Nat. Commun.,* 5**,** 1-10.

Laird, N. M. & Ware, J. H. 1982. Random-effects models for longitudinal data. *Biometrics,* 38**,** 963-974.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution,* 30**,** 314-334.

Lande, R. 1977a. The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics,* 86**,** 485-498.

Lande, R. 1977b. On comparing coefficients of variation. *Syst. Zool.,* 26**,** 214-217.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution,* 33**,** 402-416.

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics,* 94**,** 203-215.

Lande, R. & Arnold, S. J. 1983. The measurement of selection on correlated characters. *Evolution,* 37**,** 1210-1226.

Lessells, C., Cooke, F. & Rockwell, R. 1989. Is there a trade-off between egg weight and clutch size in wild Lesser Snow Geese (*Anser c. caerulescens*)? *J. Evol. Biol.,* 2**,** 457-472.

Lewontin, R. C. 1966. On the measurement of relative variability. *Syst. Zool,* 15**,** 141-142.

Lush, J. 1937. *Animal Breeding Plans,* Ames: Iowa State College Press.

Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.,* 136**,** 727-741.

Lynch, M. & Hill, W. G. 1986. Phenotypic evolution by neutral mutation. *Evolution,* 40**,** 915-935.

Martin, S. L. & Husband, B. C. 2012. Whole genome duplication affects evolvability of flowering time in an autotetraploid plant. *PLoS One,* 7**,** e44784

Martins, E. P. 1994. Estimating the rate of phenotypic evolution from comparative data. *Am. Nat.,* 144**,** 193-209.

Matthews, G., Hangartner, S., Chapple, D. G. & Connallon, T. 2019. Quantifying maladaptation during the evolution of sexual dimorphism. *Proc. R. Soc. B,* 286**,** 20191372.

Mccleery, R., Pettifor, R., Armbruster, P., Meyer, K., Sheldon, B. & Perrins, C. 2004. Components of variance underlying fitness in a natural population of the great tit *Parus major*. *Am. Nat.,* 164**,** E62-E72.

Mcglothlin, J. W., Kobiela, M. E., Wright, H. V., Mahler, D. L., Kolbe, J. J., Losos, J. B. & Brodie Iii, E. D. 2018. Adaptive radiation along a deeply conserved genetic line of least resistance in *Anolis* lizards. *Evol. Lett.,* 2**,** 310-322.

Mcgoey, B. V. & Stinchcombe, J. 2018. Introduced populations of ragweed show as much evolutionary potential as native populations. *bioRxiv***,** 305540.

Merilä, J., Sheldon, B. & Kruuk, L. 2001. Explaining stasis: microevolutionary studies in natural populations. *Genetica,* 112/113**,** 199-222.

Morrissey, M. B. 2016. Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *J. Evol. Biol.,* 29**,** 1882-1904.

Nakagawa, S., Johnson, P. C. & Schielzeth, H. 2017. The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface,* 14**,** 20170213.

Nakagawa, S. & Santos, E. S. 2012. Methodological issues and advances in biological meta-analysis. *Evol. Ecol.,* 26**,** 1253-1274.

Nakagawa, S. & Schielzeth, H. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.,* 4**,** 133-142.

Noble, D. W., Radersma, R. & Uller, T. 2019. Plastic responses to novel environments are biased towards phenotype dimensions with high additive genetic variation. *PNAS***,** 201821066.

Opedal, Ø. H. 2019. The evolvability of animal-pollinated flowers: towards predicting adaptation to novel pollinator communities. *New Phytol.,* 221**,** 1128-1135.

Opedal, Ø. H., Bolstad, G. H., Hansen, T. F., Armbruster, W. S. & Pélabon, C. 2017. The evolvability of herkogamy: Quantifying the evolutionary potential of a composite trait. *Evolution,* 71**,** 1572-1586.

Palmer, A. R. 2000. Quasi-replication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. *Ann. Rev. Ecol. Syst.,* 31**,** 441-480.

Palo, J., O'hara, R. B., Laugen, A. T., Laurila, A., Primmer, C. R. & Merilä, J. 2003. Latitudinal divergence of common frog (*Rana temporaria*) life history traits by natural selection: evidence from a comparison of molecular and quantitative genetic data. *Mol. Ecol.,* 12**,** 1963-1978.

Pavlicev, M., Cheverud, J. M. & Wagner, G. P. 2011. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. *Proc. R. Soc. B,* 278**,** 1903-1912.

Pélabon, C., Armbruster, W. S. & Hansen, T. F. 2011. Experimental evidence for the Berg hypothesis: vegetative traits are more sensitive than pollination traits to environmental variation. *Funct. Ecol.,* 25**,** 247-257.

Pélabon, C., Hilde, C. H., Einum, S. & Gamelon, M. 2020. On the use of the coefficient of variation to quantify and compare trait variation. *Evol. Lett.,* In press.

Pinheiro, J. C., Liu, C. & Wu, Y. N. 2001. Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. *J. Comput. Graph. Stat.,* 10**,** 249-276.

Podolsky, R. H., Shaw, R. G. & Shaw, F. H. 1997. Population structure of morphological traits in *Clarkia dudleyana*. II. Constancy of within-population genetic variance. *Evolution,* 51**,** 1785-1796.

Puentes, A., Granath, G. & Ågren, J. 2016. Similarity in G matrix structure among natural populations of *Arabidopsis lyrata*. *Evolution,* 70**,** 2370-2386.

Ramírez-Valiente, J. A., Etterson, J. R., Deacon, N. J. & Cavender-Bares, J. 2018. Evolutionary potential varies across populations and traits in the neotropical oak *Quercus oleoides*. *Tree Physiol.,* 39**,** 427-439.

Rosenthal, R. 1979. The file drawer problem and tolerance for null results. *Psychol. Bull.,* 86**,** 638-641.

Schluter, D. 1984. Morphological and phylogenetic relations among the Darwin's finches. *Evolution***,** 921-930.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution,* 50**,** 1766-1774.

Schluter, D. 2000. *The ecology of adaptive radiation,* Oxford: Oxford University Press.

Schoen, D. J., Bell, G. & Lechowicz, M. J. 1994. The ecology and genetics of fitness in forest plants. IV. Quantitative genetics of fitness components in *Impatiens pallida* (Balsaminaceae). *Am. J. Bot.,* 81**,** 232-239.

Service, P. M. 2000. The genetic structure of female life history in *D. melanogaster*: comparisons among populations. *Genet. Res. (Camb.),* 75**,** 153-166.

Shaw, F. H., Shaw, R. G., Wilkinson, G. S. & Turelli, M. 1995. Changes in genetic variances and covariances: G whiz! *Evolution,* 49**,** 1260-1267.

Sheldon, B., Kruuk, L. & Merila, J. 2003. Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution,* 57**,** 406-420.

Shore, J. S. & Barrett, S. C. 1990. Quantitative genetics of floral characters in homostylous *Turnera ulmifolia* var. *angustifolia* Willd.(Turneraceae). *Heredity,* 64**,** 105-112.

Simpson, G. G. 1944. *Tempo and mode in evolution*, New York: Columbia University Press.

Simpson, G. G. 1953. *The Major Features of Evolution,* New York: Columbia University Press.

Stanley, S. 1979. *Macroevolution, pattern and process,* San Francisco: Freeman Cooper.

Turelli, M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution,* 42**,** 1342-1347.

Turelli, M. & Barton, N. H. 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? *Genetics,* 138**,** 913-941.

Turelli, M., Gillespie, J. H. & Lande, R. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. *Evolution,* 42**,** 1085-1089.

Uyeda, J. C., Hansen, T. F., Arnold, S. J. & Pienaar, J. 2011. The million-year wait for macroevolutionary bursts. *PNAS,* 108**,** 15908-15913.

Van Noordwijk, A., Van Balen, J. & Scharloo, W. 1981. Genetic and environmental variation in clutch size of the great tit (*Parus major*). *Neth. J. Zool.,* 31**,** 342-372.

Venable, D. L. & Alberto, B. M. 1989. Quantitative genetics of size, shape, life-history, and fruit characters of the seed-heteromorphic composite *Heterosperma pinnatum*. I. Variation within and among populations. *Evolution,* 43**,** 113-124.

Voje, K. L. 2016. Tempo does not correlate with mode in the fossil record. *Evolution,* 70**,** 2678-2689.

Voje, K. L., Di Martino, E. & Porto, A. 2020. Revisiting a Landmark Study System: No Evidence for a Punctuated Mode of Evolution in *Metrarabdotos*. *Am. Nat.,* 195**,** 899-917.

Voje, K. L., Starrfelt, J. & Liow, L. H. 2018. Model adequacy and microevolutionary explanations for stasis in the fossil record. *Am. Nat.,* 191**,** 509-523.

Walter, G. M., Aguirre, J. D., Blows, M. W. & Ortiz-Barrientos, D. 2018. Evolution of Genetic Variance during Adaptive Radiation. *Am. Nat.,* 191**,** E108-E128.

Weigensberg, I. & Roff, D. A. 1996. Natural heritabilities: can they be reliably estimated in the laboratory? *Evolution,* 50**,** 2149-2157.

Whitlock, M. C. & Schluter, D. 2009. *The analysis of biological data,* Greenwood Village: Roberts and Company Pub.

Widén, B. & Andersson, S. 1993. Quantitative genetics of life-history and morphology in a rare plant, *Senecio integrifolius*. *Heredity,* 70**,** 503-514.

Widén, B., Andersson, S., Rao, G. Y. & Widén, M. 2002. Population divergence of genetic (co) variance matrices in a subdivided plant species, *Brassica cretica*. *J. Evol. Biol.,* 15**,** 961-970.

Wilkinson, G. S. & Taper, M. 1999. Evolution of genetic variation for condition-dependent traits in stalk-eyed flies. *Proc. R. Soc. B,* 266**,** 1685-1690.

Wilson, A. 2008. Why h2 does not always equal VA/VP? *J. Evol. Biol.,* 21**,** 647-650.

Worley, A. & Barrett, S. 2001. Evolution of floral display in *Eichhornia paniculata* (Pontederiaceae): genetic correlations between flower size and number. *J. Evol. Biol.,* 14**,** 469-481.

Yablokov, A. 1974. *Variability of Mammals,* New Delhi: Amerind Publishing Co.

Appendix A: Additional figures
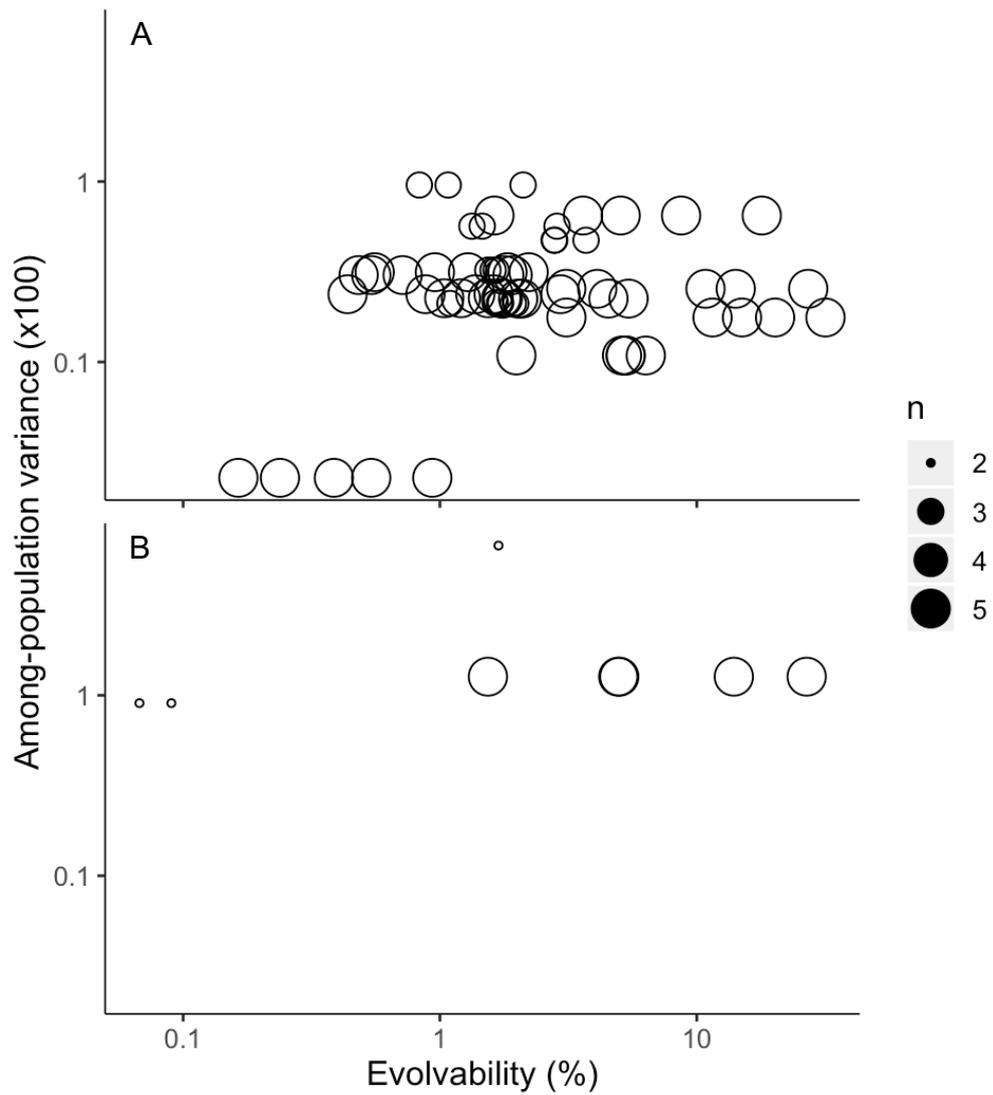
Figure A1: Trait categories excluded from the analysis due to low number of traits. Among-population variance plotted over evolvability on a log-log scale. *n* = **number of population means per trait. A: Physiological traits. B: Growth traits**
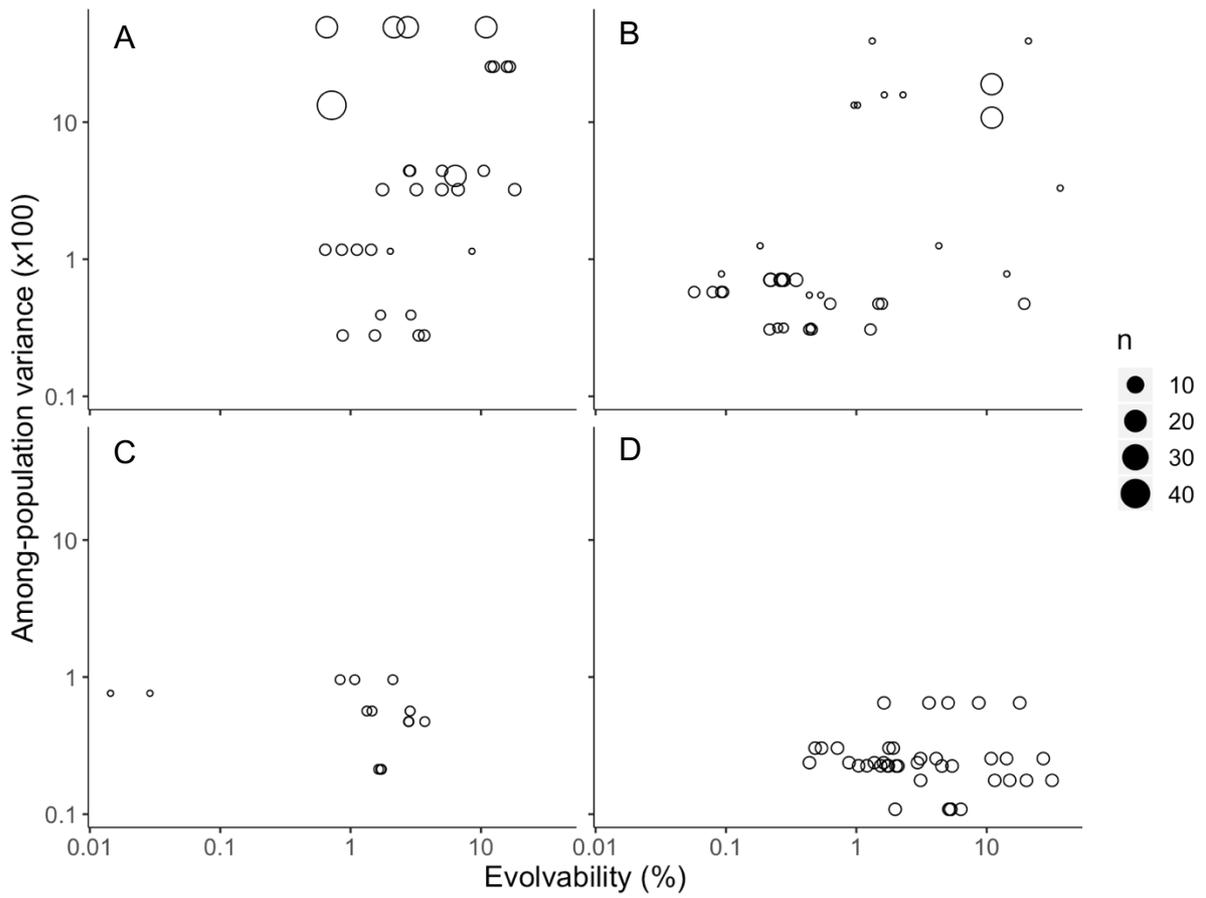
Figure A2: Trait dimensions and type excluded from the analysis due to low number of traits. Among-population variance plotted over evolvability on a log-log scale. *n* = number of population means per trait. A: Area traits. B: Mass/volume traits. C: Time traits. D: Complex traits.
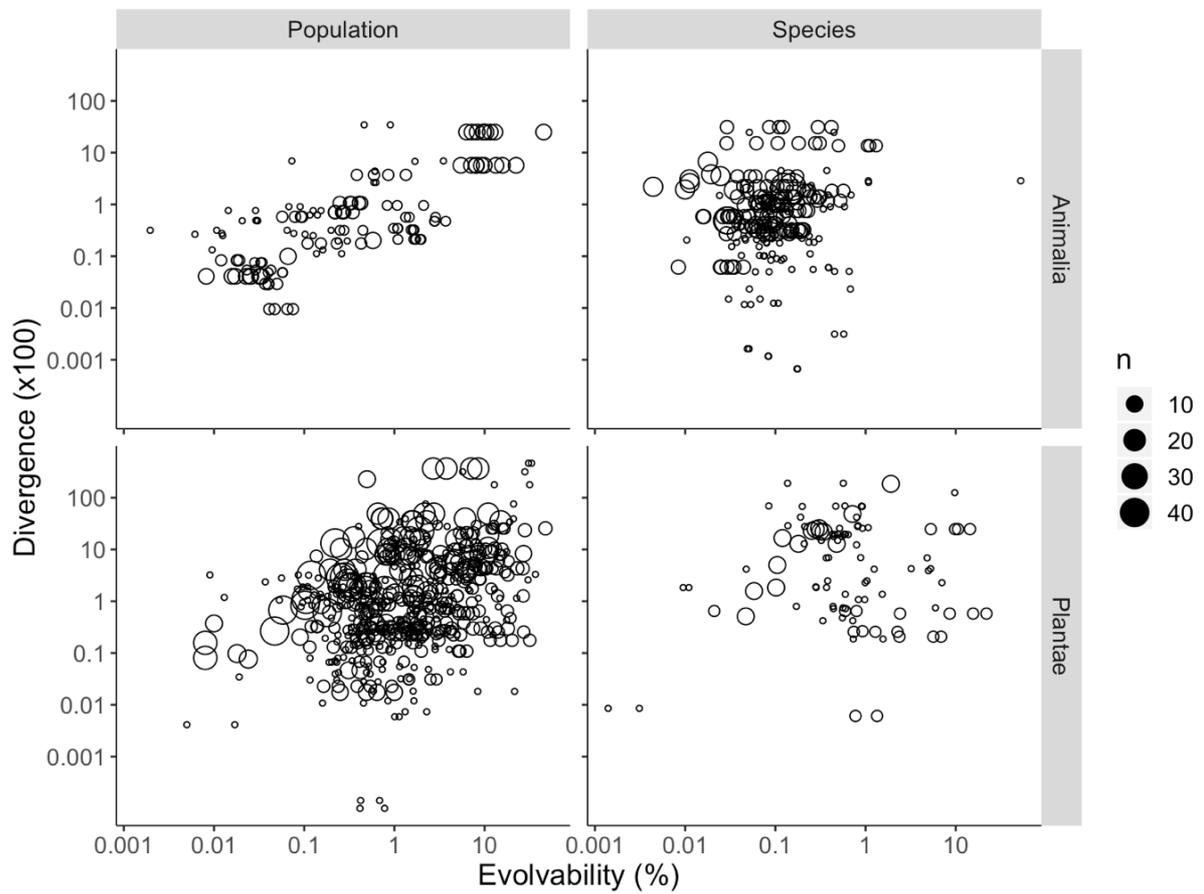
Figure A3: Left panels: Among-population variance over evolvability on log-log scale. Right panels: Among-species variance over evolvability on log-logscale. Upper panels: Animals. Lower panels: Plants. *n* = the number of populations/species means per trait. All evolvability estimates are plotted, but the model was fitted to the mean estimate per trait. Statistics in Table 3.3.
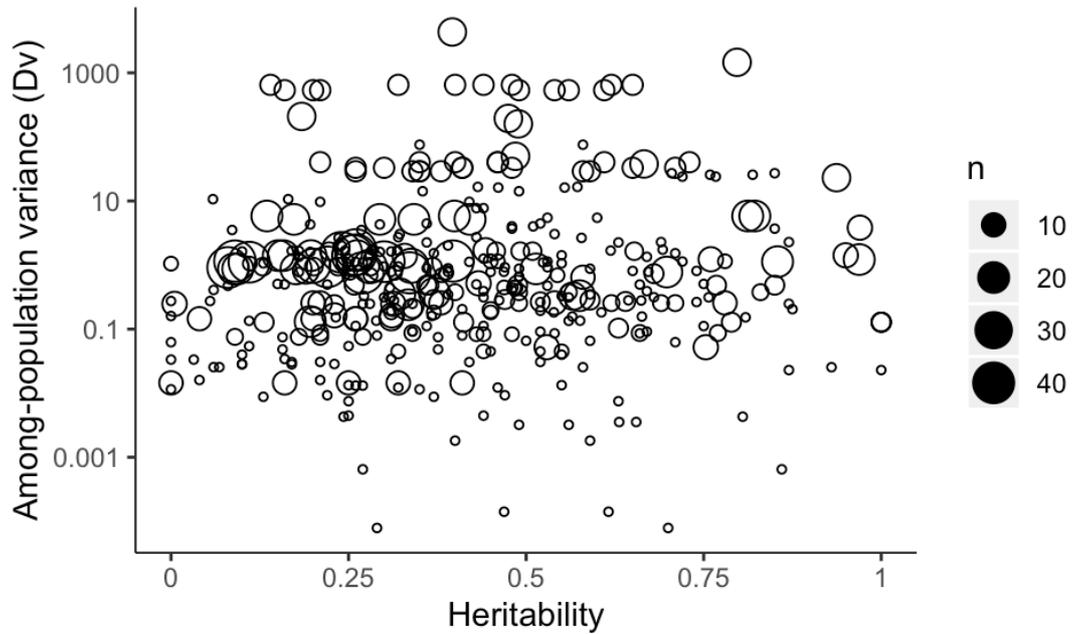
Figure A4: Linear traits. Variance-standardised divergence (log-scale) as a function of heritability (original-scale) for linear traits, with *n* = number of populations means per trait. All heritability estimates are plotted, but the model was fitted to the mean estimate per trait. $\log(D_V) = -1.74\ (\pm 0.46) + 1.56\ (\pm 0.77)\ h^2$, $R^2_m = 3\%$.
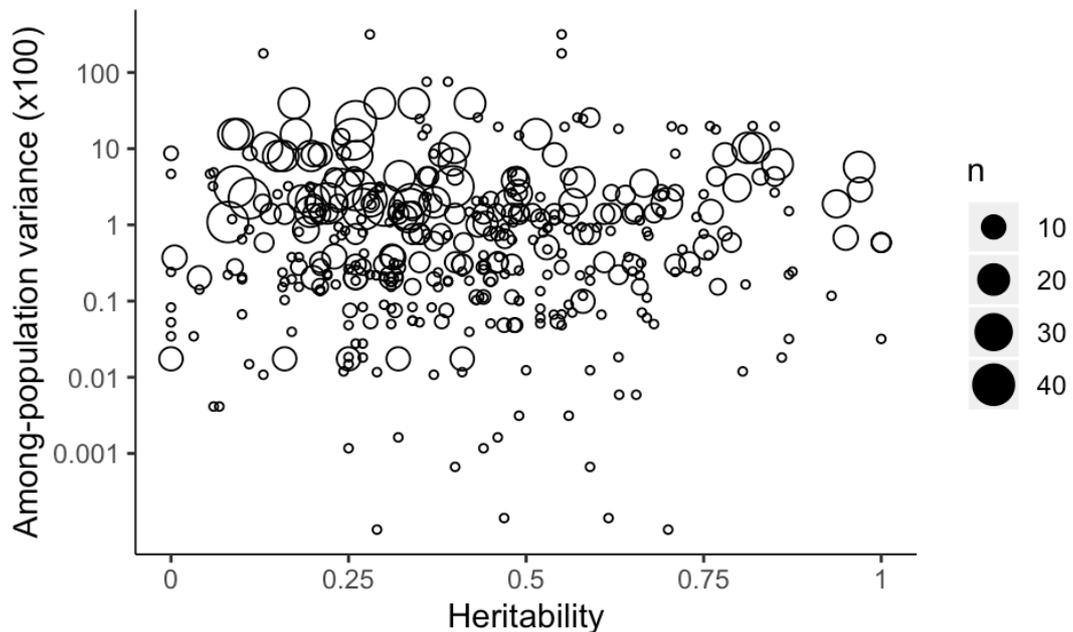
Figure A5: Linear traits. Mean-standardised divergence (log-scale) as a function of heritability (original-scale) for linear traits, with *n* = number of populations means per trait. All heritability estimates are plotted, but the model was fitted to the mean estimate per trait. $\log(D_V) = -0.84\ (\pm 0.54) + 1.28\ (\pm 0.91)\ h^2$, $R^2_m = 1\%$.

Appendix B: Table of included studies

Table B1: Overview of the used studies and a summary of the gathered data. Taxa is the lowest shared taxa of the populations/species in the study. Method is the breeding design and estimation of additive genetic variance (abbreviations at the bottom of the table). The summary includes the number of species, population and/or sex (n_{pop}) measured, and the number of traits (n_t), evolvabilities (n_e) and trait means (n_m) retrieved from the study. The grey background indicates studies that were grouped for one or more traits.

Taxa	Study	Method	n_{pop}	n_t	n_e	n_m
Monocots						
Eichhornia paniculata	Worley and Barrett (2001)	FS	2	3	12	12
Holcus lanatus	Billington et al. (1988)	NCII	2	6	12	12
Eudicots (basal group)						
Aquilegia canadensis	Herlihy and Eckert (2007)	FS	10	3	14	24
Nigella degenii	Andersson (1997)	POR	2	6	12	12
Eudicots (Rosids)						
Arabidopsis lyrata	Puentes et al. (2016)	FSHS	4	4	16	16
Brassica campestris	Evans and Marshall (1996)	FS	2	6	12	12
Brassica cretica	Widén et al. (2002)	FSHS	5	7	35	35
Dalechampia	Opedal et al., various field data	—	11	14	0	154
Dalechampia scandens	Bolstad et al. (2014)	D	1	4	4	4
Dalechampia scandens	Bolstad et al., Mexico	—	9	22	0	198
Dalechampia scandens	Hansen et al. (2003b)	D	4	24	24	96
Dalechampia scandens	Opedal et al., Costa Rica	—	17	18	0	426
Medicago truncatula	Bonnin et al. (1997)	FS	2	24	48	48
Quercus oleoides	Ramírez-Valiente et al. (2018)	FSHS	5	15	75	75
Lythrum salicaria	Colautti and Barrett (2011)	FS	20	12	12	240
Chamerion angustifolium	Martin and Husband (2012)	AS	2	1	2	2
Clarkia dudleyana	Podolsky et al. (1997)	NCI	11	4	8	44
Turnera ulmifolia	Barrett and Shore (1987)	—	10	4	0	40
Turnera ulmifolia	Shore and Barrett (1990)	FS	1	4	4	3
Eudicots (Superastrids)						
Schiedea	Culley et al. (2006)	PD	4	7	28	28
Spergularia marina	Delesalle and Mazer (1995)	FS	4	9	36	36
Talinum mengesii	Carter and Murdy (1986)	POR	11	5	5	55
Eudicots (Astrids)						
Ambrosia artemisiifolia	McGoey and Stinchcombe (2018)	FSHS	6	4	24	24
Heterosperma pinnatum	Venable and Alberto (1989)	FS	6	17	17	102

Table 1: Cont.

Taxa	Study	Method	n_{pop}	n_t	n_e	n_m
Senecio pinnatifolius	Walter et al. (2018)	FSHS	20	9	36	180
Senecio integrifolius	Widén and Andersson (1993)	FSHS	2	26	52	52
Impatiens pallida	Schoen et al. (1994)	FSHS	2	5	12	12
Lobelia	Caruso et al. (2005)	FSHS	2	7	14	14
Lobelia	Caruso et al. (2003)	—	2	6	0	12
Lobelia siphilitica	Caruso (2004)	FSHS	2	6	12	12
Mimulus	Carr and Fenster (1994)	POR	4	8	32	32
Mimulus	Fenster and Carr (1997)	POR	8	4	16	16
Collinsia heterophylla	Charlesworth and Mayer (1995)	FS	4	5	20	19
Nicotiana	Bissell and Diggle (2010)	FSHS	2	11	22	22
Solanum carolinense	Elle (1998)	D	3	5	15	15
Crustacea						
Daphnia pulicaria	Baer and Lynch (2003)	ACV	2	9	18	18
Insecta						
Diopsidae	Baker and Wilkinson (2003)	FSHS	15	9	9	135
Diopsidae	Wilkinson and Taper (1999)	FSHS	12	3	36	36
Drosophila melanogaster	Coyne and Beecham (1987)	POR	10	2	2	20
Drosophila melanogaster	Hangartner et al. (2019)	FSHS	6	8	24	24
Drosophila melanogaster	Service (2000)	FSHS	3	3	27	27
Gryllidae	Bégin and Roff (2004)	FS	4	5	20	20
Gryllus	Bégin and Roff (2003)	FS	3	5	15	15
Amphibia						
Rana temporaria	Palo et al. (2003)	FSHS	6	2	12	12
Reptilia						
Anolis	McGlothlin et al. (2018)	FSHS	7	8	56	56
Colubridae	Hohenlohe and Arnold (2008)	—	34	2	0	68
Thamnophis elegans	Arnold and Phillips (1999)	POR	4	6	24	24
Aves						
Cyanistes caeruleus	Charmantier et al. (2004)	PED	3	2	6	6
Cyanistes caeruleus	Delahaie et al. (2017)	PED	4	6	24	24
Geospiza	Grant and Price (1981)	POR	2	3	6	6

Table 1: Cont.

Taxa	Study	Method	n_{pop}	n_t	n_e	n_m
Anser caerulescens	Lessells et al. (1989)	POR	1	1	1	1
Ficedula albicollis	Sheldon et al. (2003)	POR	1	1	1	1
Geospiza fortis	Gibbs (1988)	POR	1	1	1	1
Parus major	McCleery et al. (2004)	POR	1	1	1	1
Parus major	Van Noordwijk et al. (1981)	POR	2	1	2	2
Sturnus vulgaris	Flux and Flux (1982)	POR	1	1	1	1
Mammalia						
Saguinus	Cheverud (1996)	PED	2	39	78	78

Abbreviations:
ACV = Among clone variance
AS = Artificial selection
D = Diallel, FS = Full-sibs/family variance
FSHS = Full-sib/half-sib breeding design
NCI = North Carolina I
NCII = North Carolina II
PD = Partial diallel
PED = Pedigree analysis
POR = Parent offspring regression

Appendix C: Derivation of the error variance equation

Error variance of evolvability

To convert the estimate of error variance (SE^2) of additive genetic variance (V_A) to an estimate of error variance for evolvability on log scale ($\log(e_\mu)$) the following operation was done. First, the error variance of evolvability is

$$SE^2[e_\mu] = SE^2\left[\frac{V_A}{\bar{z}^2}\right] = \frac{SE^2[V_A]}{\bar{z}^4},\tag{C1}$$

where the error variance of the trait mean ($\bar{z}$) is considered negligible. Then, the error variance of each estimate of $\log(e_\mu)$ is converted to the scale of V_A using a Taylor approximation (eq. C2) to the first derivative around each estimate of evolvability ($\hat{e}_\mu$).

$$SE^2[\log(\hat{e}_\mu)] \approx SE^2\left[\log(\hat{e}_\mu) + \frac{1}{\hat{e}_\mu}(e_\mu - \hat{e}_\mu)\right]\tag{C2}$$

$$SE^2\left[\frac{e_\mu}{\hat{e}_\mu}\right] = \frac{SE^2[e_\mu]}{\hat{e}_\mu^2} = \frac{\frac{SE^2[V_A]}{\bar{z}^4}}{\hat{e}_\mu^2}\tag{C3}$$

In equation (C2) the estimates $\log(\hat{e}_\mu)$ and $-\frac{\hat{e}_\mu}{\hat{e}_\mu}$ are constants and do not affect the variance measure. Hence, they are removed in C3. The random regression models (eq. 2.11) were fitted with the mean evolvability ($\log(\bar{e}_\mu)$) per trait. I assume each estimate of V_A is independent and therefore use the mean error variance per trait. The error variance of evolvability on log-scale becomes:

$$SE_{ij}^2[\log(\bar{e}_\mu)] = \frac{E\left[\frac{SE^2[V_A]_i}{\bar{z}_i^4}\right]}{E[e_{\mu_i}^2]}, \qquad i = 1, 2, \dots, n_j, \qquad j = 1, 2, \dots, k,\tag{2.12}$$

where n_j is the number of populations or species per trait j of the total number of traits k.