1 Title

Coheritability and Coenvironmentability as Concepts for Partitioning the Phenotypic Correlation

Jorge Vasquez-Kool

2 **Affiliation:**

Coheritability , Coenvironmentability

Jorge Vasquez-Kool

Keywords: phenotypic correlation, coheritability, coenvironmentability, bivariate, quantitative traits.

Abstract

Central to the study of joint inheritance of quantitative traits is the determination of the degree of association between two phenotypic characters, and to quantify the relative contribution of shared genetic and environmental components influencing such relationship. One way to approach this problem builds on classical quantitative genetics theory, where the phenotypic correlation ($r_{P_{x,y}}$) between two traits is modelled as the sum of a genetic component called the coheritability ($h_{x,y}$), which reflects the degree of shared genetics influencing the phenotypic correlation, and an environmental component, namely the coenvironmentability ($e_{x,y}$) that accounts for all other factors that exert influence on the observed trait-trait association. Here a mathematical and statistical framework is presented on the partition of the phenotypic correlation into these components. I describe visualization tools to analyze $r_{P_{x,y}}$, $h_{x,y}$ and $e_{x,y}$ concurrently, in the form of a three-dimensional (3DHER-plane) and a two-dimensional (2DHER-field) plots. A large data set of genetic parameter estimates (heritabilities, genetic and phenotypic correlations) was compiled from an extensive literature review, from which coheritability and coenvironmentability were derived, with the object to observe patterns of distribution, and tendency. Illustrative examples from a diverse set of published studies show the value of applying this partition to generate hypotheses proposing the differential contribution of shared genetics and shared environment to an observed phenotypic relationship between traits.

Introduction

A fundamental aspect in the study of heredity is to investigate associations among traits at the phenotypic level, and to determine the degree shared genetics and common environmental influences shape such associations. Many traits are analyzed jointly in genetic studies in the hope of providing greater statistical power to detect associations to causal genetic factors (Melton et al. 2010, Cheng et al. 2013, Jia and Jannick 2012). Understanding how ceratin traits relate to disease risk is of primary concern in clinical medicine (Wellman et al. 2013, Oren et al. 2015, Barabási et al. 2010) and in animal and plant breeding (Sölkner et al. 2008). Most phenotypes are the result of a complex interaction of multiple genetic and environmental factors (Lander and Schork 1994), and the ability to assay these characters presents a unique opportunity to explore mechanisms underlying concerted inheritance.

The phenotypic correlation $r_{P_{x,y}}$ between two traits, say x and y, has been extensively used to quantify the relationship between observable characters. Its partition between a genetic component and an environmental component was originally worked out by Hazel (1943) as a function of heritabilities (h^2) of the traits and correlations (genetic $r_{A_{x,y}}$, environmental $r_{E_{x,y}}$) between them (see equation [5] below). The term coheritability, coined by Nei (1960) as a descriptor of the genetic component, is a measure of joint inheritance of two traits that representes the genetic contribution to their phenotypic correlation. This contrasts with coenvironmentability, $e_{x,y}$ that quantifies the contribution of all other sources of variation, excluding additive genetic factors, influencing the phenotypic correlation. However, the coheritability as a concept and measure has remained largely unrecognized. In the 1996 edition

of *Principles of Quantitative Genetics* the term coheritability was reintroduced more formally (Falconer and MacKay 1996, p. 317) in the context of correlated response to selection. Yet a proper treatment of its mathematical and statistical properties is still lacking. Despite its simple formulation, there is a persistent confusion in the scientific literature of what coheritability means or how should it be calculated. Some have denoted coheritability as the mere ratio of additive genetic to phenotypic covariances (de Reggi 1972, Janssens 1979) without sufficient theoretical support, while others employ the term coheritability (or co-heritability) to refer to different *ad hoc* measures of coinheritance (Sae-Lim 2015, Hoskens et al. 2018), or colloquially coinheritance applies to any observed joint transmission of traits (Wambua et al. 2006, Höblinger et al. 2009). Uncertainty is compounded when coheritability is conflated with genetic correlation (Yang et al. 2016, Traglia et al. 2017, Yin et al. 2017), or is referred by various names, such as 'coefficient of genetic prediction' (Baradat 1976), 'correlative heritability' (Chen et al. 2003), 'genetic contribution to the phenotypic correlation' (Posthuma et al. 2003), 'standardized genetic covariance' (Rao and Rice 2005), 'bivariate heritability' (DeStephano et at. 2009), 'Endophenotype Ranking Value' (Glahn et al. 2012), and 'proportion of phenotypic correlation due to genes' (Wu et al. 2010, 2013, Muñoz el al. 2018).

This study substantially builds up on previous work (Hazel 1943, Lerner 1950, Searle 1961, Yamada 1968, Bedard et al. 1971, Plomin and DeVries 1979), and connects to modern investigations (Gui et al. 2017, Pick et al. 2016, Gianola et al 2015) in the attempt to clarify the nature of coheritability and coenvironmentability as constituent of the phenotypic correlation. I explore this topic through theoretical arguments and through the analyses of data compiled

from a diverse and large number of studies. The objectives of this paper were: (1) to present a theoretical background on the mathematical and statistical properties of phenotypic correlation, coheritability and coenvironmentability. (2) Analyze their distribution, dispersion and tendency. (3) Model the relationship between the phenotypic correlation and $(h_{x,y},\ e_{xy})$ and $\left(r_{A_{x,y}}, r_{E_{x,y}}\right)$. (4) Present illustrative examples on the application of the $r_{P_{x,y}}$ decomposition in order to gain insight in different biological problems. Further statistical and mathematical details as well as illustrative examples are provided in the Supplementary Information (SI) document accompanying this paper.

Theoretical Background

The components of the sample phenotypic correlation

In the framework of quantitative genetics, an individual's phenotypic value P is modelled as the sum of an additive genetic value A and an environmental value E, thus $P = A + E$. This simple decomposition also applies to the phenotypic covariance $C_{P_{x,y}}$, which as a measure of linear association between the phenotypic values of two characters x and y, results from the sum of an additive genetic covariance $C_{A_{x,y}}$, and a term including all residual genetic and non-genetic factors, namely the environmental covariance $C_{E_{x,y}}$, thus

$$C_{P_{x,y}} = C_{A_{x,y}} + C_{E_{x,y}} \quad [1]$$

To standardize, both sides of equation [1] are divided by the geometric mean of the phenotypic variances of each trait (i.e., which could be construed as the joint *bivariate phenotypic variability* of the traits x and y),

$$\frac{C_{P_{x,y}}}{\sqrt{V_{P_x}V_{P_y}}} = \frac{C_{A_{x,y}}}{\sqrt{V_{P_x}V_{P_y}}} + \frac{C_{E_{x,y}}}{\sqrt{V_{P_x}V_{P_y}}} \quad [2]$$

This expression can be summarized as follows:

$$r_{P_{x,y}} = h_{x,y} + e_{x,y} \quad [3]$$

The term $r_{P_{x,y}} = C_{P_{x,y}} / \sqrt{V_{P_x}V_{P_y}}$ is the phenotypic correlation between the phenotypic values of characters x and y; it measures the linear association between the two observable characters.

The coheritability, defined by Nei (1960), as

$$h_{x,y} = \frac{C_{A_{x,y}}}{\sqrt{V_{P_x}V_{P_y}}} \quad [4]$$

is the ratio of the genetic covariance on the bivariate phenotypic variability, is the component of the phenotypic correlation attributed to shared genetic effects, and thus reflects the extent that joint genetic influences have on the observed association of the characters.

The coenvironmentability (broad-sense), $e_{x,y} = C_{E_{x,y}} / \sqrt{V_{P_x}V_{P_y}}$, is the component of the phenotypic correlation defined as the ratio of the residual covariance (i.e. non-additive genetic,

environmental, GxE interactions) to the bivariate phenotypic variability. It represents the joint influence of all factors that are not accounted by additive genetic factors that exert influence on the observed relationship between the traits.

Since covariances can be expressed in terms of their respective variances V and correlations r, such that $C_{\blacksquare_{x,y}} = \sqrt{V_{\blacksquare_x} V_{\blacksquare_y}}\, r_{\blacksquare_{x,y}}$ (where ■ is either of the subscripts P, A or E). Equation [2] can be re-written as:

$$\frac{\sqrt{V_{P_x} V_{P_y}}\, r_{P_{x,y}}}{\sqrt{V_{P_x} V_{P_y}}} = \frac{\sqrt{V_{A_x} V_{A_y}}\, r_{A_{x,y}}}{\sqrt{V_{P_x} V_{P_y}}} + \frac{\sqrt{\left(V_{P_x} - V_{A_x}\right)\left(V_{P_x} - V_{A_x}\right)}\, r_{E_{x,y}}}{\sqrt{V_{P_x} V_{P_y}}}$$

$$r_{P_{x,y}} = \sqrt{h_x^2\, h_y^2}\; r_{A_{x,y}} + \sqrt{(1 - h_x^2)(1 - h_y^2)}\; r_{E_{x,y}} \qquad [5]$$

where h_x^2 and h_y^2 denote the narrow-sense heritabilities of the traits x and y. $r_{E_{x,y}}$ is the environmental correlation, and $r_{A_{x,y}}$ is the genetic correlation, defined as

$$r_{A_{x,y}} = \frac{C_{A_{x,y}}}{\sqrt{V_{A_x} V_{A_y}}} \qquad [6]$$

Equation [5], originally introduced by Hazel (1943, p. 480) in the context of selection indexes, models the phenotypic correlation as the sum of a weighed genetic correlation, namely the coheritability, which can be expressed as

$$h_{x,y} = \sqrt{h_x^2\, h_y^2}\, r_{A_{x,y}} \qquad [7]$$

and a weighed environmental correlation, i.e., the coenvironmentability,

$$e_{x,y} = \sqrt{(1 - h_x^2)(1 - h_y^2)}\, r_{E_{x,y}} \qquad [8]$$

The expressions of coheritability given in Equations [4] and [7] are equivalent. Further equivalent can be done formulations using algebraic summations or by applying a path analytic method (or structural equation model) (Supplementary Information section 3.2). There are two more terms associated to the decomposition of the genetic covariance (Equation 1) and which involve terms of covariation of the breeding values a trait with the environmental values of the other (e.g. C_{A_x,E_y} , C_{A_y,E_x}). Some would opt to ignore these terms or consider them negligible, which is not different than including them in the environmental covariance term (Supplementary Information section 2.4-2.6). Their importance and influence deserve further study. From the relationships presented, it is clear that the phenotypic correlation's magnitude and sign is directly influenced by the coheritability and coenvironmentability.

The domains of coheritability, coenvironmentability and phenotypic correlation

The description of the phenotypic correlation as the sum of the coheritability and coenvironmentability imposes an intrinsic linear relationship between the three variables. Since the sum of $h_{x,y}$ and $e_{x,y}$ must yield $r_{P_{x,y}}$, and the Cauchy-Schwarz inequality proves that the phenotypic correlation is bound to the domain $[-1, +1]$. This implies that if either of the variables $h_{x,y}$ or $e_{x,y}$ becomes zero, then the other variable would become equal to the phenotypic correlation. Therefore, $h_{x,y}$ and $e_{x,y}$ each must also have a domain within $[-1, +1]$

with the added condition that the sum of their absolute values cannot exceed 1 ($|h_{x,y}| + |e_{x,y}| \leq 1$). For instance, if one of the heritabilities is unity, or the $r_{E_{x,y}} = 0$, then $e_{x,y} = 0$, then equation [4] becomes $r_{P_{x,y}} = h_{x,y}$. Similarly, if one the heritabilities or the genetic correlation is zero, then $h_{x,y} = 0$, and it results in $r_{P_{x,y}} = e_{x,y}$. All this implies that both coheritability and coenvironmentability can be subject to the same inferential statistical methods as those designed for the assessment of correlations (Supplementary Information section 5.1).

The environmental correlation $r_{E_{x,y}}$, whose factors may remain unspecified, could be calculated as a residual derived from [4]:

$$r_{E_{x,y}} = \frac{r_{P_{x,y}} - h_{x,y}}{\sqrt{(1-h_x^2)(1-h_y^2)}} \qquad [\ 8\ ]$$

and then use it to determine the (broad-sense) coenvironmentability using equation [6]. Note that equation [7] will introduce dependency and collinearity if applied to a regression model involving the phenotypic correlation. A superior method would estimate the environmental correlation independent of the phenotypic correlation and coheritability, and relate to specified environmental factors. In this case, the coheritability and the narrow-sense coenvironmentability would not necessarily add up to $r_{P_{x,y}}$ and should be an indication of that significant genotype x environment interaction terms are present.

Visualization of $\{\boldsymbol{h_{xy}}, \boldsymbol{e_{xy}}, \boldsymbol{r_{P_{x,y}}}\}$

The three-dimensional coheritability-coenvironmentability-phenotypic correlation plane, 3DHER) plane is a Cartesian three-dimensional space employed to represent these three variables (Figure 1), where a single datum is defined by its coordinates $(h_{x,y}, e_{x,y}, r_{P_{x,y}})$. Spherical coordinates can also be used to describe the behavior of $h_{x,y}$, $e_{x,y}$, and $r_{P_{x,y}}$ (Supplementary Information section 6.3) and serve as a complementary way to provide further insight. A first impression from observing this graph is that the data lies on a virtual slanted plane with zero volume. Mathematical theory tells that this is due to the existence of an intrinsic linear relationship among the three variables. Further corroboration that the possible values fall on a plane with zero volume is given by the fact that the determinant of a matrix involving any three distinct data points is zero (Supplementary Information section 6.4).

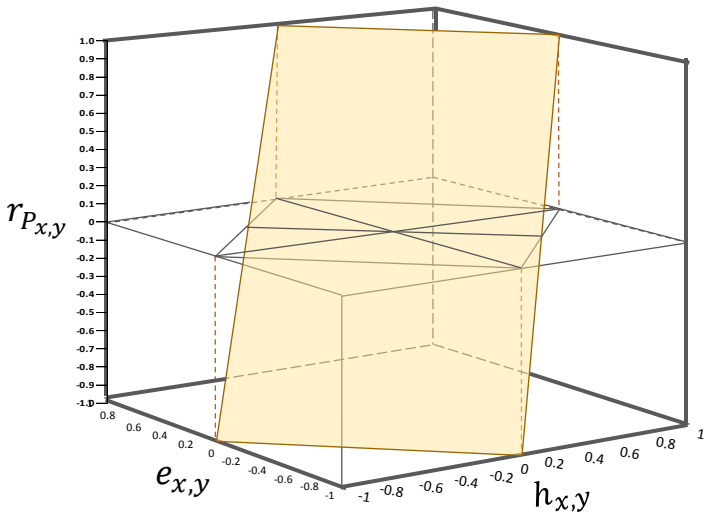


Figure 1. The three-dimensional $h_{xy} \cdot e_{xy} \cdot r_{P_{x,y}}$-plane (3DHER-plane) presented in (A) the Cartesian coordinate system. The plane represents the area where the data of coheritability, coenvironmentability and phenotypic correlation satisfy the equation $r_{P_{x,y}} = h_{x,y} + e_{x,y}$, where $r_{P_{x,y}}, h_{x,y}, e_{x,y} \in [-1, +1]$.

The two-dimensional coheritability-coenvironmentability-phenotypic correlation (2DHER) field is the orthogonal projection of the 3Dplane onto the $h_{x,y} \cdot e_{x,y}$-surface when $r_{P_{x,y}} = 0$ (see Figures 1), this graph retains information of the three variables (Figure 2, Supplementary Information section 6.4). This field represents an area bound by the relationship $|h_{x,y}| + |e_{x,y}| = 1$ (i.e. its borders are demarcated by the lines $h_{x,y} + e_{x,y} = 1$, $h_{x,y} + e_{x,y} = -1$, $h_{x,y} - e_{x,y} = 1$, and $h_{x,y} - e_{x,y} = -1$). The domain of the phenotypic correlation can be superimposed on it, knowing that the elements of a data point $(h_{x,y}, e_{x,y})$ must add to $r_{P_{x,y}}$. The field is a continuum domain of $h_{x,y}$, $e_{x,y}$ and $r_{P_{x,y}}$, that accounts for all sign combinations among these three variables (excluding incongruous combinations consisting of $h_{x,y}$ and $e_{x,y}$ having the same sign, yet summing up to an $r_{P_{x,y}}$ with a different sign). For the sake of clarity, the field can be divided by tracing the lines $h_{x,y} = 0$, $e_{x,y} = 0$, $r_{P_{x,y}} = 0$. The six triangular partitions thus produced are labeled with the letter S followed by a subscript having a signed numeral, which is the sign of the coheritability (which in turn is conferred by the genetic correlation), and the numeral serves as a dummy indicator (Figure 1). The domain for variables $h_{x,y}$, $e_{x,y}$ and $r_{P_{x,y}}$ is specific for each partition (Figure 2B). Partitions that share a reciprocal angle become reciprocal partitions whose labels possess the same numeral by have different

188 signs (i.e. there are three pairs of reciprocal partitions, namely S_{+1} and S_{-1}; S_{+2} and S_{-2}; S_{+3}
189 and S_{-3}). Notice that positive phenotypic correlations are found in partitions S_{-3}, S_{+1}, S_{+2},
190 and the negative phenotypic correlations in S_{+3}, S_{-1}, S_{-2}. Positive coheritabilities (and
191 implicitly genetic correlations) are found in partitions S_{+1}, S_{+2}, S_{+3}; and negative
192 coheritabilities in partitions S_{-1}, S_{-2}, S_{-3}. Lastly, partition, S_0, contain all data with at least
193 one of the three variables ($h_{x,y}$, $e_{x,y}$, $r_{P_{x,y}}$) equal to zero and therefore lie on one of the diving
194 lines. The S_{+1} and S_{-1} are reciprocal partitions where all the three variables ($h_{x,y}$, $e_{x,y}$, $r_{P_{x,y}}$)
195 have the same sign, either positive (S_{+1}) or negative (S_{-1}), and each occupy an area equal to
196 0.5. Both are demarcated by the lines $h_{x,y} = 0$ and $r_{P_{x,y}} = 0$, and line $h_{x,y} + e_{x,y} = 1$ in the
197 case of S_{+1}, and line $h_{x,y} + e_{x,y} = -1$ for S_{-1}. The area is equally divided by the line
198 $h_{x,y} = e_{x,y}$ separating data that satisfy either $h_{x,y} < e_{x,y}$ or $h_{x,y} > e_{x,y}$.

199 The reciprocal partitions S_{+2} and S_{-2} are characterized by $r_{P_{x,y}}$ and $h_{x,y}$ having the same sign,
200 positive in the case of S_{+2}, and negative in the case of S_{-2}. In these partitions, the magnitude
201 of coheritability is larger than coenvironmentability, and therefore shows the preponderant
202 influence of a common genetic component upon the phenotypic correlation. These partitions
203 occupy an area of 0.25 each and are demarcated by the lines $r_{P_{x,y}} = 0$ and $e_{x,y} = 0$, and either
204 line $h_{x,y} - e_{x,y} = 1$ (for S_{+2}) or $h_{x,y} - e_{x,y} = -1$ (for S_{-2}). Note that the
205 coenvironmentability is bound by the interval $[0, +0.5]$ in S_{-2}, and $[-0.5, 0]$ for S_{+2}. The
206 reciprocal partitions S_{+3} and S_{-3} are characterized by $r_{P_{x,y}}$ and $h_{x,y}$ having different signs, due
207 to the fact that the phenotypic and genetic correlation differ in direction. Each of these

208 partitions cover an area of 0.25, and are delimited by $r_{P_{x,y}} = 0$ and $h_{x,y} = 0$, and line

209 $h_{x,y} - e_{x,y} = 1$ (for S_{+3}) or line $h_{x,y} - e_{x,y} = -1$ (for S_{-3}). In these partitions, the magnitudes

210 of $h_{x,y}$ is smaller than the one for coenvironmentability, indicating that there is an overriding

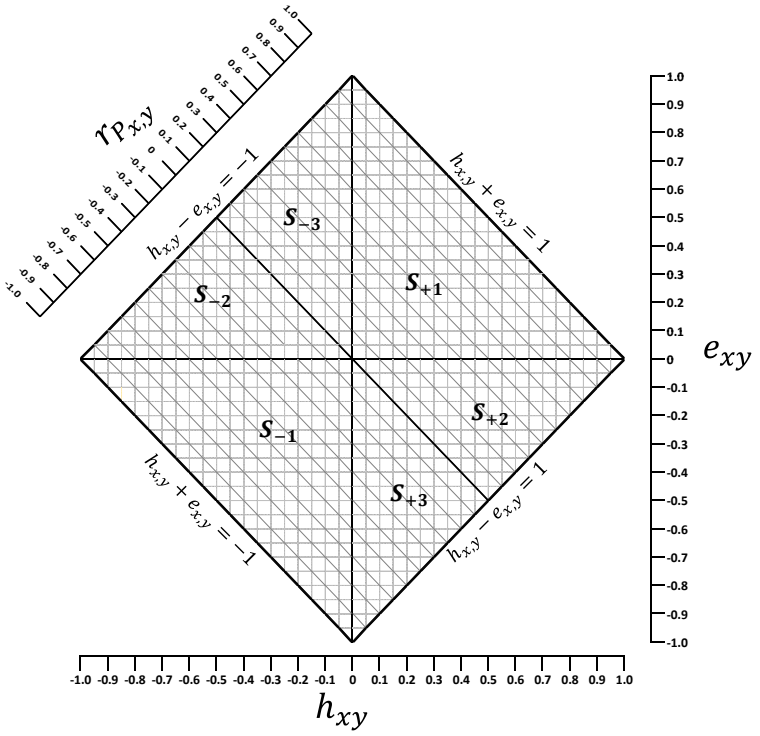
211 environmental effect on the phenotypic correlation. The line diving equally partitions S_{+1} and

212 S_{-1} separates the instances where $h_{x,y} < e_{x,y}$ (at the side of Partitions S_{-2} and S_{-3}), and

213 where $h_{x,y} > e_{x,y}$ (at the side of S_{+2}, S_{+3}).

214

A



B

$r_{P_{xy}}$	h_{xy}	e_{xy}				$r_{P_{xy}}$	h_{xy}	e_{xy}
$-1 < r_{P_{xy}} < 0$	$-1 < h_{xy} < 0$	$-1 < e_{xy} < 0$	S_{-1}		S_{+1}	$0 < r_{P_{xy}} < 1$	$0 < h_{xy} < 1$	$0 < e_{xy} < 1$
$-1 < r_{P_{xy}} < 0$	$-1 < h_{xy} < 0$	$0 < e_{xy} < 0.5$	S_{-2}		S_{+2}	$0 < r_{P_{xy}} < 1$	$0 < h_{xy} < 1$	$-0.5 < e_{xy} < 0$
$0 < r_{P_{xy}} < 1$	$-0.5 < h_{xy} < 0$	$0 < e_{xy} < 1$	S_{-3}		S_{+3}	$-1 < r_{P_{xy}} < 0$	$0 < h_{xy} < 0.5$	$-1 < e_{xy} < 0$

Figure 2. The two-dimensional $h_{x,y}.e_{x,y}.r_{P_{x,y}}$-field (2DHER-field). This is the result of the projection of the 3D plane onto the $h_{x,y} \cdot e_{x,y}$ surface. (A) Partitions are demarcated by tracing the lines $r_{P_{x,y}} = 0$, $h_{x,y} = 0$, and $e_{x,y} = 0$. Partitions are labeled with a S symbol whose subscript positive (+) or negative (–) correspond to the sign of the genetic correlation followed by a numeral. Partitions denoted with the same numeral but with different sign are reciprocal. (B) The domain interval of $h_{x,y}$, $e_{x,y}$, and $r_{P_{x,y}}$ by partition. The equation $|h_{x,y}| + |e_{x,y}| = 1$ defines the boundary lines of the field, which are explicitly written at its sides. The total area of the field is 2. Partitions S_{+1} and S_{-1} occupy each an area of 0.5, while partitions S_{+2}, S_{-2}, S_{+3},

and S_{-3} each cover an area equal to 0.25. The line $h_{x,y} = e_{x,y}$ distinguishes all instances where $h_{x,y} < e_{x,y}$ and $h_{x,y} > e_{x,y}$.

Relationship between genetic correlation and heritabilities

The bivariate correlations and univariate heritabilities are independent, random variables, variation in one does not lead to a concomitant effect on the other. The weights expressed as functions of the heritabilities, however, are inversely related: as the geometric mean of the heritabilities $\sqrt{h_x^2\, h_y^2}$ increases, there is a concomitant decrease of $\sqrt{(1-h_x^2)(1-h_x^2)}$. However, larger heritabilities do not necessarily imply larger influence of the genetic component. For instance, if the heritabilities of the traits are $h_x^2 = 0.7$, $h_y^2 = 0.8$, and the correlations $r_{A_{x,y}} = 0.15$, $r_{E_{x,y}} = 0.75$, then the resulting coheritability $h_{x,y} = 0.111$, despite the large heritabilities of the traits, is smaller than the coenvironmentability $e_{x,y} = 0.1836$. On the other hand, a larger genetic correlation does not necessarily translate into a larger genetic influence either. For example, let us say that $h_x^2 = 0.2$, $h_y^2 = 0.3$, $r_{A_{x,y}} = 0.615, r_{E_{x,y}} = 0.344$. The coheritability becomes $h_{x,y} = 0.15$, which is smaller than the coenvironmentability $e_{x,y} = 0.25$, despite that the latter has a low environmental correlation. This fact may bring into reconsideration methods that attempt to map the degree of genetic influence on the phenotypic correlation based on the mere comparison of $r_{P_{x,y}}$ and $r_{A_{x,y}}$. Rank between these statistics is not preserved when using the totality of the data.

If, on the other hand, one of the variable is set to a specified value, it would effectively limit the range of possible values on the other variables. Questions such as, given the value of a phenotypic correlation, what would be the possible values of the coheritability and coenvironmentability?, or, condition on specified values of coheritability and coenvironmentability, what are the possible values of the heritabilities of the traits and the genetic and the environmental correlations? These topics are duly elaborated in the supplementary information (SI section 7.1 and 7.2) in an empirical manner that yield possible, not probable, values. This is primarily meant to lay the groundwork for more rigorous development.

The relationship of the magnitudes of the coheritability and of the genetic correlation.

If the heritabilities of the traits are each unity, then the coenvironmentability becomes zero, the coheritability equates the genetic correlation, and only under this extreme condition, $r_{P_{x,y}} = r_{A_{x,y}}$. In the majority of the cases, however, the absolute value of the coheritability would always be less than the absolute value of the genetic correlation. The geometric mean of the heritabilities, $\sqrt{h_x^2\, h_y^2}$, itself a decimal number between 0 and 1, multiplied to the genetic correlation results in the product (i.e. coheritability) with a smaller magnitude, such that the coheritability would move away from the $r_{A_{x,y}}$ towards the origin. For instance, if $\sqrt{h_x^2\, h_y^2} = 0.2$, $r_{A_{x,y}} = -0.4$, then $h_{x,y} = -0.08$. The latter is closer to zero than $r_{A_{x,y}}$. If $\sqrt{h_x^2\, h_y^2} = 0.35$,

$r_{A_{x,y}} = 0.5$, then $h_{x,y} = 0.175$, which also moves towards the zero direction Therefore, in terms of magnitude, the coheritability would be equal of less than the magnitude of the genetic correlation,

$$h_{x,y} \in \begin{cases} [r_{A_{x,y}}, 0) & if\ \ r_{A_{x,y}} < 0 \\ 0 & if\ \ r_{A_{x,y}} = 0 \\ (0, r_{A_{x,y}}] & if\ \ r_{A_{x,y}} > 0 \end{cases}$$

Inference

Details on the characteristics of the base population and the sample, as well as aspects concerning hypothesis testing and confidence intervals determination for coheritability, coenvironmentability and phenotypic correlation are duly elaborated in the Supplementary Information sections 1 and 5. In brief, consider a large, genetically-structured population whose individuals possess heritable traits (h_x^2, h_y^2), and with genetic $\rho_{A_{x,y}}$, environmental $\rho_{E_{x,y}}$, and phenotypic correlation $\rho_{P_{x,y}}$ between the traits, the latter being the sum of the population-level coheritability $H_{x,y}$ and coenvironmentability $E_{x,y}$ parameters $\rho_{P_{x,y}} = H_{x,y} + E_{x,y}$. It is, therefore, feasible to make propositions about these population parameters using data obtained from sampling. If a sample of size n is obtained from such population, one could hypothesize the values of the parameters $\rho_{P_{x,y}}$, $H_{x,y}$, and $E_{x,y}$, provided that they satisfy the relationship $\rho_{P_{x,y}} = \sqrt{h_x^2\, h_y^2}\ \rho_{A_{x,y}} + \sqrt{(1-h_x^2)(1-h_y^2)}\ \rho_{E_{x,y}}$. The distribution of the estimator of $\rho_{P_{x,y}}$ is the sample phenotypic correlation $r_{P_{x,y}}$ is construe to be the distribution of the correlation coefficients determined from m samples, all of size n, drawn from the original population (see Supplementary Information section 4 for derivation of sampling distribution of

283 the correlation coefficient, Fisher 1915, Hotelling 1959). Naturally, sample size is an important
284 consideration for estimation and hypothesis testing. Under a condition of moderate ($70 < n \leq$
285 100) to large sample sizes ($n > 100$), asymptotic properties of the statistics can be assumed
286 (Supplementary Information section 6). Other aspects of statistical inference on
287 $r_{P_{x,y}}, r_{A_{x,y}}, r_{E_{x,y}}, h_{x,y}$, and $e_{x,y}$ such as hypothesis testing, power, and confidence level of
288 parameters, as well as aspects of experimental design regarding determination of an adequate
289 sample size are presented in Supplementary Information section 5.

290

291 ## Simulation of bivariate $(\boldsymbol{h_{x,y}}, \boldsymbol{e_{x,y}})$ data

292 Simulation is fundamental to further an understanding on probabilities, estimation of the
293 sampling distribution of statistics, calculation of coverage probability of confidence intervals,
294 and to evaluate the robustness of statistical tests (Wicklin 2013). To facilitate these paths of
295 inquiry, a method to generate bivariate $(h_{x,y}, e_{x,y})$ observations is presented (Supplementary
296 Information section 10) based on a transformation of two random uniform random variables.

297

298

Materials and Methods

Data compilation and validation

An extensive search of data published in journal articles was carried out using bibliographic sources such as PubMed, Web of Science, GoogleScholar. The search was carried out using keywords: 'coheritability', 'genetic parameters', 'phenotypic and genetic correlations', 'correlated response', 'age-age correlations', and 'early selection', 'fitness trade-offs'. To be included in the compilation (see Flow chart in Supplementary Material, File D1), the reported data had to involve continuous traits and minimally include trait heritabilities and at least two correlations (phenotypic, genetic, or environmental). Information was also gathered about the organism studied, standard errors of the parameter estimates (if any), and bibliographic citation. Coheritabilities and coenvironmentabilities were then calculated using equations [4] to [7]. The data was categorized into partitions according to the criteria expounded in Theoretical Background and summarized in Figure 2.

To ensure validation, three criteria were used for outlier detection, failure to satisfy at least one of them led to its exclusion from the data set: The first criterion checked whether the values of heritabilities and correlations were within their domains ($h^2 \in [0,1]$, $r \in [-1,1]$). A second criterion ascertained whether a given datum was within the boundaries of the 2DHER-field by holding the relationship $|h_{x,y}| + |e_{x,y}| \leq 1$. Finally, the data had to satisfy the relationship $|r_{P_{x,y}}| + D \leq 1$, (D, disparity index, see below) to indicate that the values of $r_{P_{x,y}}$ and $r_{A_{x,y}}$ were coherent to the relationship expressed in equation [4]. Analyses of descriptive and

inferential statistical analyses were carried out using SAS/STAT Software (SAS Institute, Cary, NC).

Analysis of count data

It was of particular interest to see if the occupancy of the data in partitions (S_o , S_{+1} , S_{+2} , S_{+3} , S_{-1} , S_{-2} , S_{-3}) came from independent trials. If each of n independent trials result in placing the data to one of the k partitions and the probability that a (collected) datum belongs to a given partition, is the same in every trial, then the count of the data in the partitions would follow a multinomial probability distribution. A goodness-of-fit test for multinomial distribution was therefore conducted to evaluate a null hypothesis proposing that the observed count of data points in each partition was the result of expected proportions assigned to each of the partitions.

Empirical Distribution

Cognizant that the data came from disparate sources (organisms, populations, traits, experimental goals), it was not the intention to conduct a formal metaanalysis, but to simply visualize the general occupancy, tendency and dispersion of the phenotypic correlation, coheritability and coenvironmentability as scatter plot on the 3DHER-plane and 2DHER-field. Histograms were drawn to visualize the frequency and variability of the data.

Modeling the phenotypic correlation

Though it was not the aim of this work to create a predictive statistical model, it was nevertheless of interest to observe how $r_{P_{x,y}}$, as a scalar, dependent variable, related to either $(h_{x,y}, e_{x,y})$ or $(r_{A_{x,y}}, r_{E_{x,y}})$ acting as regressors (or explanatory variables) in a multiple regression model. This exercise allowed also to check the relationship between regression parameters, and the consistency of regression equation using data as a whole and by partition. The models were:

$$\text{Model 1:}\quad r_{P_{x,y_i}} = \boldsymbol{\beta_{o1}} + \boldsymbol{\beta_{1h}} \cdot h_{x,y_i} + \boldsymbol{\beta_{2e}} \cdot e_{x,y_i} + \epsilon_{1i}$$

$$\text{Model 2:}\quad r_{P_{x,y_i}} = \boldsymbol{\beta_{o2}} + \boldsymbol{\beta_{1r}} \cdot r_{A_{x,y_i}} + \boldsymbol{\beta_{2r}} \cdot r_{E_{x,y_i}} + \epsilon_{2i}$$

where the parameters $\boldsymbol{\beta_o}$ is the intercept, $\boldsymbol{\beta_1}$ is the slope of the regression of phenotypic correlation $r_{P_{x,y}}$ on the genetic factor, namely, $h_{x,y}$ (model 1) or $r_{A_{x,y}}$ (model 2), and $\boldsymbol{\beta_2}$ is the regression coefficient corresponding to the environmental factor, $e_{x,y}$ (model 1), or $r_{E_{x,y}}$ (model 2). Model 1 assumes that given the data set $\{ r_{P_{x,y_i}}, h_{x,y_i}, e_{x,y_i} \}_{i=1}^{n}$ of n observations, a linear relationship exists between the dependent variable $r_{P_{x,y}}$ and the bivariate vector of regressors $\{h_{x,y}, e_{x,y}\}$. The ϵ term was assumed to have a normal distribution with zero mean and constant variance σ_ϵ^2 (a path analytical representation of model 1 is presented in Supplementary Information section 3.2.2). A similar rationale applies to model 2. The results of the regression analyses using Model 1 were assessed based upon the expectations derived from equation [3] (i.e. zero intercept, and each has a slope parameter equal to positive one).

359 Model 2 was also expected to yield a zero intercept, and , if the genetic and environmental
360 correlations are zero then the phenotypic correlations must also be zero. Otherwise, there are
361 no theoretical relationship that relates $r_{P_{x,y}}$ to $r_{A_{x,y}}$ and $r_{E_{x,y}}$ directly. The least-squares
362 estimates of $\boldsymbol{\beta_o}$, $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ were estimated from the data using the PROC REG procedure of
363 SAS, applied to the data set as a whole as well as by partition. The statistics collected from the
364 results were the intercept (β_o), slope (β_1), and the R-square of each model.

365

366 ## Disparity Index Analyses

367 The disparity index (D) is defined as the absolute value of the difference between the
368 phenotypic correlation and the genetic correlation, $D = |r_{P_{x,y}} - r_{A_{x,y}}|$ (Willis et al. 1991). This
369 study used D for two purposes. One was as a means of validation (Supplementary Information
370 Appendix 2). The second use was to evaluate the closeness of the magnitudes of $r_{P_{x,y}}$ and
371 $r_{A_{x,y}}$. With the aim to further investigate the relationship between $r_{P_{x,y}}$ and $r_{A_{x,y}}$, the disparity
372 index was treated as a random variable, and for this purpose this work derived its probability
373 density function, which used as a basis the transformation of the absolute difference of two
374 uniform random variates, is formulated as $f_D(D) = 2 - 2d$ for $0 \leq D \leq 1$ (Supplementary
375 Information 7.5.2). A method to generate simulated disparity data is presented in
376 Supplementary Information section 7.5.3.

377

Illustrative Examples

Data from selected studies pertaining a range of topics of particular relevance to modern biology were used to help illustrate how the decomposition of $r_{P_{x,y}}$ into $h_{x,y}$ and $e_{x,y}$ becomes an appropriate and pertinent approach to postulate hypotheses concerning the degree of influence that shared genetics and common environmental factors have on shaping an observable association between traits.

Data availability

Supplementary Material is deposited and available in FigShare https:// . It contains File D1 Data consisting of the raw numerical values obtained from the literature review and examples, File SI Supplementary Information including text explaining mathematical and inferential statistical aspects, and additional text and figures; and the File C1 Computing containing code and resources used in statistical and mathematical calculations (SAS, MS Excel).

Results

This work analyzed more than 7700 observations comprising compiled data for the distribution, and examples, plus more than 16000 data points of the transcriptomics example . The data set compiled from journal articles amounted to $n = 6287$ observations, and involved 140 studies in the areas of human genetics, agronomy, forestry, fisheries, animal husbandry, ecology , and life history research. Data was collected from humans ($n = 1069$), animals ($n = 3535$, 39

399 genera, 40 species), plants ($n = 1683$, 26 genera, 33 species). Observations belonged to
400 morphological (4288), physiological (1721), fitness-life history (20), and behavioral (258) traits.
401 Around 60% of the studies were carried out under field conditions, and the rest in lab. The
402 data collected consisted of heritabilities and correlations, and from them the coheritability and
403 coenvironmentability corresponding to each observation, was derived including determination
404 of the partition the datum belong to (see Theoretical Background section). The data set
405 underwent an stringent validation procedure including detection and exclusion of outliers
406 (Supplementary Information Appendix 2). Basic statistics for each variable ($r_{P_{x,y}}, r_{A_{x,y}}, r_{E_{x,y}}, h_{x,y}$
407 and $e_{x,y}$) are presented in Supplementary Information Appendix 3A.

408

409 ## Distribution of the $h_{x,y}$, $e_{x,y}$, $r_{Px,y}$ data

410 Given the heterogeneity of the data in terms of objects of study, traits, sources, methods,
411 precision, it was not the intention to perform a formal metanalysis, nor to combine data of the
412 studies to create weighed averages. Rather, the objective here was to observed occurrence,
413 tendency and variability of the data. Inspection of the coheritability and coenvironmentability
414 data plotted on the 3DHER-plane (Figure 3A) or 2DHER-field (Figure 4A) shows it scattered on a
415 two-dimensional object. The 3D plot of the phenotypic correlation as a function of the genetic
416 and environmental correlations occupied a more variable volume (Figure 3B, 4B). The mean
417 values of coheritability (0.075, SE 0.0039) and coenvironmentability (0.083, SE 0.0022) matched
418 very well with the calculated overall mean phenotypic correlation (0.158, SE 0.0025). The
419 variability, as measured by the standard deviation, was less for the coheritability (0.177) than

420 for the coenvironmentability (0.198) and phenotypic correlation (0.309). The standard error of
421 the estimates was the lowest for coheritability (mostly found below 0.16) (for formulas see
422 Supplementary Information section 3.3). The dispersion was not uniform, all the variables
423 presented abundance around the origin, but became more infrequent towards the borders
424 (Figure 4A) particularly at the neighborhood of the lines $h_{x,y} - e_{x,y} = 1$ and $h_{x,y} - e_{x,y} = -1$.
425 Overall, the tendency of the relationship of $r_{P_{x,y}}$ with $h_{x,y}$, $e_{x,y}$, $r_{A_{x,y}}$ and $r_{E_{x,y}}$ was positive.

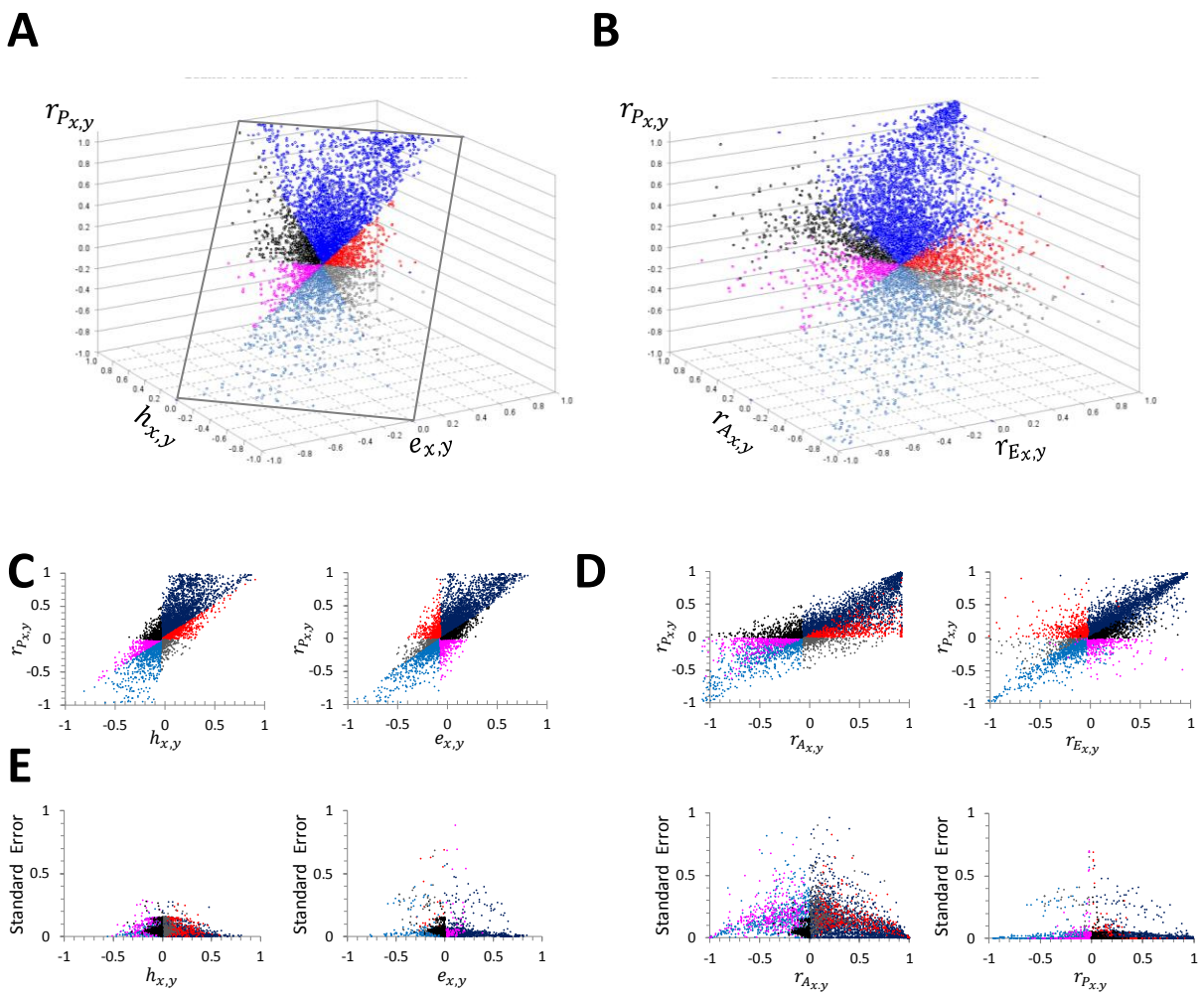


Figure 3. Three-dimensional scatter plot of the phenotypic correlation ($r_{P_{x,y}}$) and the (A) coheritability $h_{x,y}$, coenvironmentability $e_{x,y}$ on the 3DHER-plane; and (B) the genetic correlation $r_{A_{x,y}}$ and the environmental correlation $r_{E_{x,y}}$. The bivariate plot of the phenotypic correlation against (C) the coheritability, and coenvironmentability, and against (D) the genetic and environmental correlations. (E) Standard errors of the parameters estimators. Total sample size $n = 6287$, color refers to each partition.

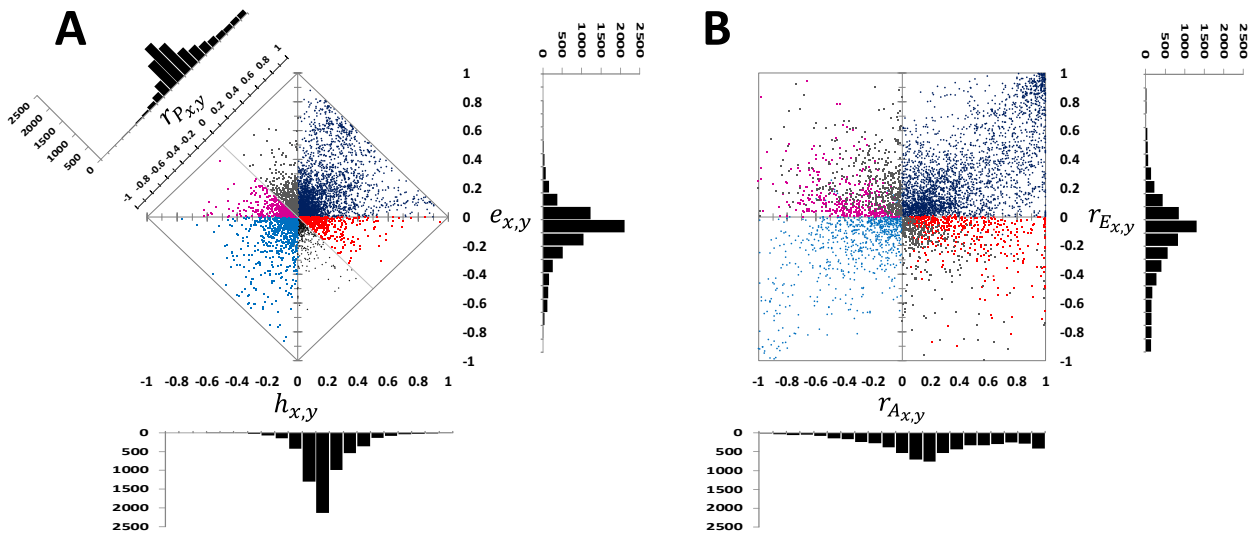


Figure 4: Two-dimensional scatter plot ($n = 6287$) of (A) Coheritability, coenvironmentability, and phenotypic correlations and histograms of the compiled data displayed on the $h_{x,y}$. $e_{x,y}$. $r_{P_{x,y}}$(2DHER)-field. (B) Plot of the genetic and environmental correlations, and histograms. All histograms exhibit a positively skew.

Inspection of the 2DHER-field showed, as expected, well demarcated division between the partitions (4A), a feature not apparent in the plot of the correlations which rather revealed a blurred distinction between partitions (4B). In the plot of the correlations, the data extends throughout its domain occupying most of the area. The histograms approximate a normal distribution, though the shape of all of them appeared somehow skewed towards the positive side.

Multinomial Test

The abundance of the data varied significantly by partition. Partition S_{+1} was the most populated ($n_{S+1} = 3312$) amounting to 53% of the data, which was followed by S_{-1} ($n_{S-1} = 880$, or 14.2%). The occupancy in partitions S_{+2} ($n_{S+2} = 617$, 0.099%) and S_{-3} ($n_{s-3} = 562$, 0.09%) were approximately similar. The lowest count were found in partitions S_{-2} ($n_{S-2} = 473$, 0.076%) and S_{+3} ($n_{S+3} = 367$, 0.059%). The count in partition S_0 was 67 and was not considered in further calculations. Both the chi-square and the chi-LRT gave qualitatively similar results so here I present only the results obtained using the Chi-Square test. A goodness-of-fit test for multinomial distribution under the null hypothesis $Ho: p_{S+1} = 0.53$, $p_{S+2} = 0.1$, $p_{S+3} = 0.06$, $p_{S-1} = 0.14$, $p_{S-2} = 0.07$, $p_{S-3} = 0.1$, showed that the data fitted reasonably well the multinomial model ($\chi^2_{6-1} = 9.348, p = 0.096$) (Supplementary Information Appendix 6B).

Multiple Regression Analysis

The purpose of these analyses was to evaluate the consistency of model equations and parameters, generated from the use of the whole data set and by partition. The phenotypic correlation was modeled as a function of coheritability-coenvironmentability (Model 1), and genetic- environmental correlations (Model 2). Figure 5 presents the regression parameter estimates corresponding to the genetic factor (β_1), the environmental factor (β_2), and the R-square (R^2) for each model. Model 1 showed complete consistency and uniformity ($\beta_1 =$

$\beta_2 = 1;\ R^2 = 1)$, as was clearly seen in producing the same linear regression equation regardless if the analysis used the whole data set or of each partitions. The ratio of slopes to in all cases maintained a 1-to-1 relationship. By Equation [3], the intercept was expected to be zero and the regression coefficients (slopes) of $h_{x,y}$ and $e_{x,y}$ each equaled to unity (Supplementary Information Appendix 3G), and the results from Model 1 satisfied these requirements. In Model 2, the relationship between the slopes of the genetic and environmental correlation varied widely (the test of heterogeneity of slopes shows that they are effectively different among each other, Supplementary Information Appendix 3G). The relationship $\beta_1 r_{A_{x,y}} : \beta_2 r_{E_{x,y}}$, was 1:1.45 (overall) whereas by partition, it oscillated from 1:-0.17 to 1:3.64. Partition S_{+2} and S_{-2} produced models with the lowest R-squares (0.299 for S_{+2}, 0.235 for S_{-2}). Analyses of residuals and the R-squares of Model 1 captured all of the variability of $r_{P_{x,y}}$ $(R^2 = 1)$, in contrast Model 2 exhibited variable degrees in explaining the variability of $r_{P_{x,y}}$, from low to high R-squares (0.23 to 0.94), and each partition yielded a different linear equation.

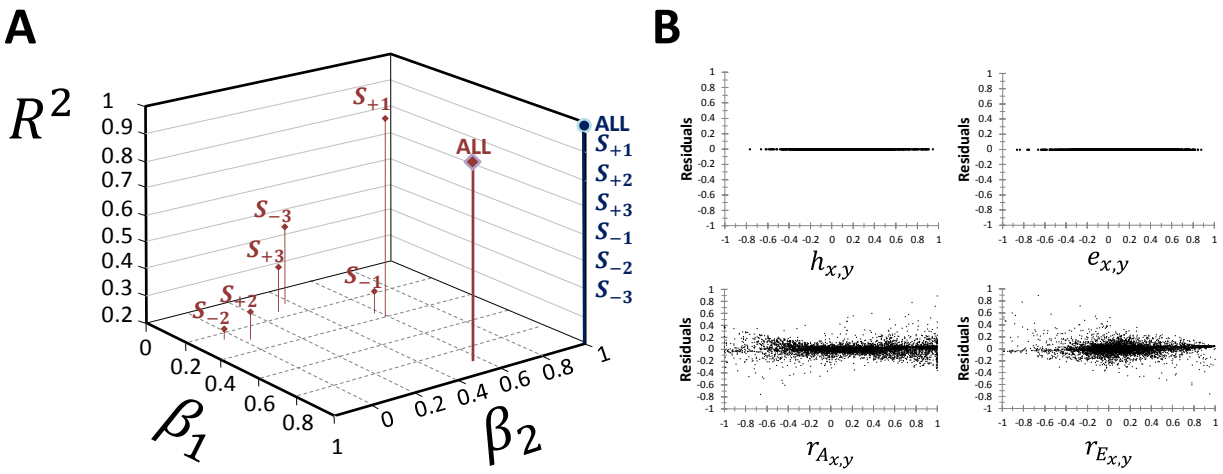


Figure 5. (A) Relationship of the multiple regression parameters (β_1), slope (β_1) for model 1 (blue needle with top circle) and Model 2 (brown needle with rhomboid top). Labels indicate the whole data set (ALL) or the partition subsets. (B) Plots of residual against coheritability and coenvironmentability (Model 1) display low variability dispersed around the horizontal axis indicating that the linear regression model is appropriate for the data. Compare this to the residual plots involving factors $r_{A_{x,y}}$ and $r_{E_{x,y}}$ of Model 2.

The results point out to the dependability of the coheritability and coenvironmentability as the appropriate factors directly related to $r_{P_{x,y}}$, and the consistency of parameter estimates obtained by Model 1 whether using the overall or partition data.

Analyses on $r_{P_{x,y}}$ and $r_{A_{x,y}}$

To address whether the phenotypic and genetic correlation exhibit some degree of similarity that would justify $r_{P_{x,y}}$ to be an appropriate proxy for $r_{A_{x,y}}$, here I now compare the phenotypic and genetic correlation in terms of the Dissimilarity Index (D), namely, the degree of departure of their numerical values. A correlation analysis of the overall data set showed that these two variables maintained a high statistical association $(r = 0.821,\ p < 0.0001,\ n = 6288)$. The regression of $r_{A_{x,y}}$ on $r_{P_{x,y}}$ had a slope of 1.14 (standard error 0.01, R-square 0.67, Supplementary Information Appendix 3F). Approximately 2% of the data had a dissimilarity equal to zero, indicating that $r_{A_{x,y}} = r_{P_{x,y}}$, and if a $D \leq 0.03$ represents an acceptable level of similarity between $r_{A_{x,y}}$ and $r_{P_{x,y}}$, it occurred in 13% of all comparisons. However, there was a considerable departure of their values at the individual level. The mean of the absolute disparity D index was 0.181 (standard deviation 0.175, $n = 6288$) across all comparisons. In 15% of the data, $r_{P_{x,y}}$ and $r_{A_{x,y}}$ differened in sign (partitions S_{-3} and S_{+3}). When both $r_{A_{x,y}}$ and $r_{P_{x,y}}$ were positive $(n = 3929)$, $r_{P_{x,y}}$ was, on average, 130% the magnitude of $r_{A_{x,y}}$. When both $r_{A_{x,y}}$ and $r_{P_{x,y}}$ were negative $(n = 1353)$, $r_{P_{x,y}}$ was 129% the magnitude of $r_{A_{x,y}}$. The relationship $r_{A_{x,y}} > r_{P_{x,y}}$ held in 58% $(n = 3621)$ of all pairwise comparisons, 70% of the which were in partition S_{+1}, and 100% in partitions S_{+2} and S_{+3}. Around 40% $(n = 2526)$ of the overall comparisons involved $r_{A_{x,y}} < r_{P_{x,y}}$, which was found in 67% of comparison in partition S_{-1}, and 100% in partitions S_{-2} and S_{-3}. If the disparity between $r_{P_{x,y}}$ and $r_{A_{x,y}}$ were largely due to measurement error of the genetic correlation, then it would be expected that the

squared disparity (D^2) would approach the sampling variance of $r_{A_{x,y}}$. D^2 and the sampling variance of $r_{A_{x,y}}$ had a correlation not different than zero ($r = 0.17, p = 0.001$) indicating they were not strongly associated, and in 59% ($n = 3353$) of the cases that reported standard errors of the genetic correlation, D^2 displayed a value larger than the estimated sampling variance of $r_{A_{x,y}}$. Under the null hypothesis Ho: the difference between the median of D^2 and the median of $var(r_{A_{x,y}}) = 0$, versus H1: median difference $\neq 0$, the Wilcoxon Rank Sum Test had a critical value 3138479, test statistic 11.65, $p < 0.0001$, shows that there is enough evidence to reject H_0 that the median of D^2 and $var\left(r_{A_{x,y}}\right)$ are quite similar. In addition, the Kruskal-Wallis test rejected the null hypothesis that there is no difference among partitions in mean rank of either D^2 or the Var($r_{A_{x,y}}$). (Supplementary Information Appendix 3I) . Therefore, differences between genetic and phenotypic correlations cannot be entirely explained by sampling error alone. The Disparity Index displayed its largest values when $r_{P_{x,y}}$ was around zero, and progressively decreased as the $r_{P_{x,y}}$ approached the limits of its domain (Figure 6A). The frequency distribution of the D values fitted a triangular model (Figure 6B, Supplementary Information 7.6). The relationship between $r_{A_{x,y}}$ and $r_{P_{x,y}}$ differed greatly by partition (Figure 6). With the goal to evaluate the capacity of the phenotypic correlation to predict the value of the genetic correlation, the simple regression of $r_{A_{x,y}}$ on $r_{P_{x,y}}$ (Figure 6C) resulted in a model where $r_{P_{x,y}}$ a low capacity to explain the variability of $r_{A_{x,y}}$ (R-square 0.674) and a slope equal to 1.141 (SD 0.01, p<0.0001, 95%CI [1.121,1.605], Supplementary Information Appendix 3F). Analyses by partition showed that the equations of the regression lines and the R-squares

varied widely. In 4 out of the 6 partitions the R-squares were below 0.3. The largest R-square values were 0.73 (S_{+1}) and 0.63 (S_{-1}) which correspond to partitions where $r_{A_{x,y}}$ on $r_{P_{x,y}}$ shared the same sign. The lowest R-square were 0.005 $(S_{+3})$and 0.001 (S_{-3}), corresponding to partitions were both phenotypic and genetic correlations differed in sign.

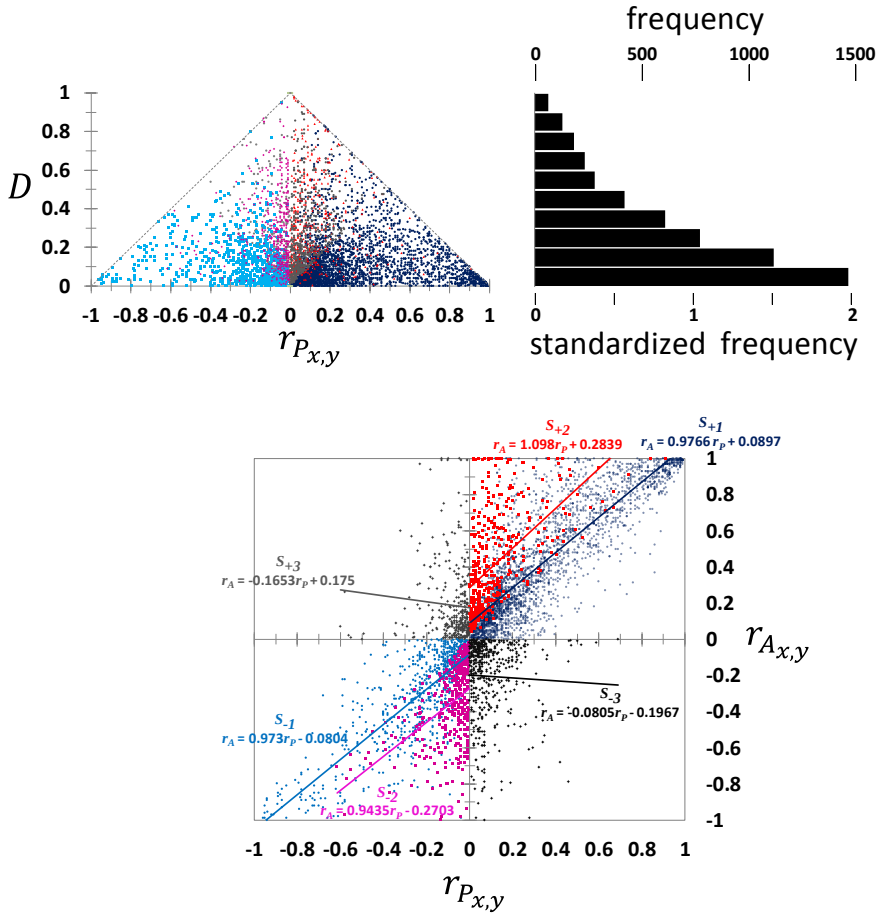


Figure 6. (A) Disparity Index D as a function of the phenotypic correlation value. Each dot represents the absolute value of the difference between $r_{A_{x,y}}$ and $r_{P_{x,y}}$ $(n = 6288)$. (B) Histograms reflecting the frequency of occurrences of comparisons with Disparity Index values within specified ranges. (C) Summary graph of the simple regression between $r_{A_{x,y}}$ (scalar

550 dependent variable) and $r_{P_{x,y}}$ (explanatory variable) showing the model equation and the R-
551 square.

552

553 ## Illustrative Examples

554 To illustrate the application of the decomposition of the phenotypic correlation into the
555 coheritability and coenvironmentability, data from published articles were used to gain insight
556 into relevant topics of modern biology. Further examples can be found in the Supplementary
557 Information Appendix 3.

558

559 **Example 1. Changes due to selection over generations**

560 Ramniwas et al. (2013) tested the hypothesis that abdominal melanisation in *Drosophila*
561 *melanogaster* enhances desiccation resistance. They performed a series of upward and
562 downward selections in the lab and reported the direct response in terms of abdominal
563 coloration change ($R_{melanisation}$) after five generations. In addition, correlated responses in
564 several physiological traits related to water stress were also measured, including cuticular
565 water loss ($CR_{cuticular\_water\_loss}$). Figure 7 presents the results obtained from upward selection in
566 female flies and the direct ($R_{melanization}$) and correlated response ($CR_{cuticular\_water\_loss}$) plotted
567 against the selection differential (S_C) for direct selection on melanisation. It can be seen that
568 as the individuals were becoming darker, the cuticular water loss decreased concomitantly.
569 Employing the regression slope of cumulative response against selection differential (S_C) for

570 darker individuals, the realized heritability for cuticular melanisation was $h_r^2 = 0.46\ (SE\ 0.03)$.

571 To determine the realized coheritability between melanisation (trait x) and cuticular water loss

572 (trait y), a regression line was fitted having the correlated response of cuticular water loss as

573 dependent variable against S_C as independent variable (Figure 7). The slope,

574 $b = -0.2673\ (SE\ 0.012, CI_{0.95}\ [-0.27, -0.22])$, represents half the coheritability of the traits

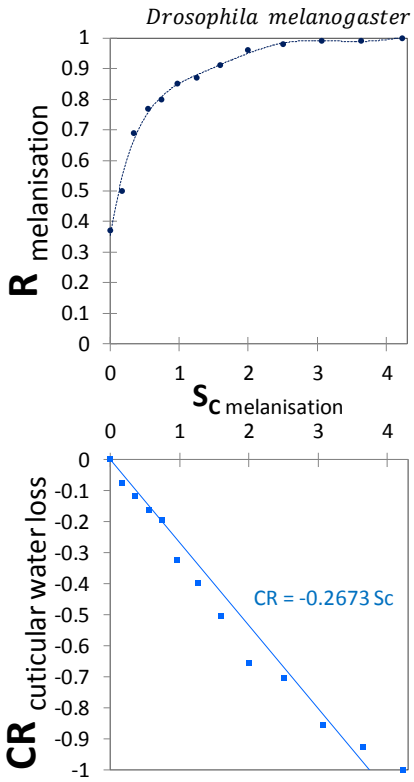
575 $(b = \frac{1}{2} h_{x,y})$ because only measurements obtained from female individuals were used.

576 Therefore, the coheritability between melanisation and the rate of water loss was $h_{x,y\,r} =$

577 $-0.5346(SE\ 0.024)$. This negative coheritability is in agreement with melanisation-

578 dessiccation hypothesis: darker individuals retain more water and are more resistant to

579 desiccation.



580

581 Figure 7. Direct response to melanization and the correlated response of the rate of cuticular
582 water loss among female flies of *Drosophila melanogaster* (Source Ramniwas et al. 2013).
583
584 It is important to recognize, in a selection experiment, that each generation of selected parents
585 constitutes a sample of the genetic correlation, and its value in successive generations of
586 selection depend upon one another sequentially, therefore changes in the value of the genetic
587 correlation will have some of the properties of a random walk (Gromko 1995), and manifest a
588 large standard error.
589
590 **Example 2. Changes through development**
591 How do the contributions of shared genetics and shared environment affect the phenotypic
592 correlation through development? Are the trajectories of the phenotypic correlation,
593 coheritability and coenvironmentability informative about relationships between traits? Here I
594 present the results derived from two field experiments. Diao et al. (2016) studied growth in
595 Japanese larch (*Pinus kaempferi*) in order to evaluate the optimal age for early selection, an
596 aspect particularly important for a long-term plantations. To this aim, the authors determined
597 phenotypic and genetic parameters of absolute height measurements obtained in different
598 ages and compared them to the one at age 16 (Figure 8A). All correlations (phenotypic, genetic
599 and environmental) increased as the trees grew, the phenotypic correlation reached 10% of its
600 maximum at age 11 and the genetic correlation around age 8. The coheritability and

601 coenvironmentability also gradually increased and reached 10% from its maximum at age 6,
602 which reveals that no significant gain from selection can be achieved after this age, setting an
603 age earlier than the one based on the genetic correlation.

604 Another example involved incremental weight gain rather than absolute measurements of
605 weight of Korean native chickens subject to a common diet (Manjula el al. 2017). The aim of
606 the study was to find the best age to make selections that would predict growth at age 20.
607 Measurements of weight were obtained from a number of individuals and subject to a logistic
608 growth model, whose parameters were also treated as traits. Data presented in Figure 8B uses
609 data of 2-week weight increments and the β parameter of the logistic model that captures the
610 asymptotic mature weight gain. The heritability of weight increment decreased consistently as
611 the animals aged. The genetic correlation between the weight increment and the β parameter
612 from early life up to age 8 were negative, and from age 12 to late in life were positive. The
613 phenotypic correlation followed a trend similar to the trajectory of the coenvironmentability
614 indicating that the association between these traits is mainly influenced by common
615 environmental factors. To elucidate what led to changes in the trait-trait relationship between
616 ages 8 and 12, going from negative early in life to positive values late in life, it would
617 necessitate to inquire from other lines of evidence in the metabolic, physiological, and
618 developmental areas. The coheritability had minimal contribution to the phenotypic
619 correlation, and varied in very narrow range of low values (-0.2 to 0.1). One can conclude that
620 weight gain can be better achieved through poultry rearing practices (e.g. diet, feeding regime)
621 than through genetics.

622 Analyses of longitudinal data present challenges because repeated measures from the same
623 subject are often (auto)correlated, and cannot be assumed to be independent, such is the case
624 of growth traits especially when measurements in certain point in time contain the
625 measurement previously obtained. It would benefit to use suitable and informative covariates
626 to help adjust values. The stochastic trends of observations across age/development/time,
627 shows that any measurement at only a single time point, may over- or underestimate the
628 genetic and environmental contributions to the $r_{P_{x,y}}$.

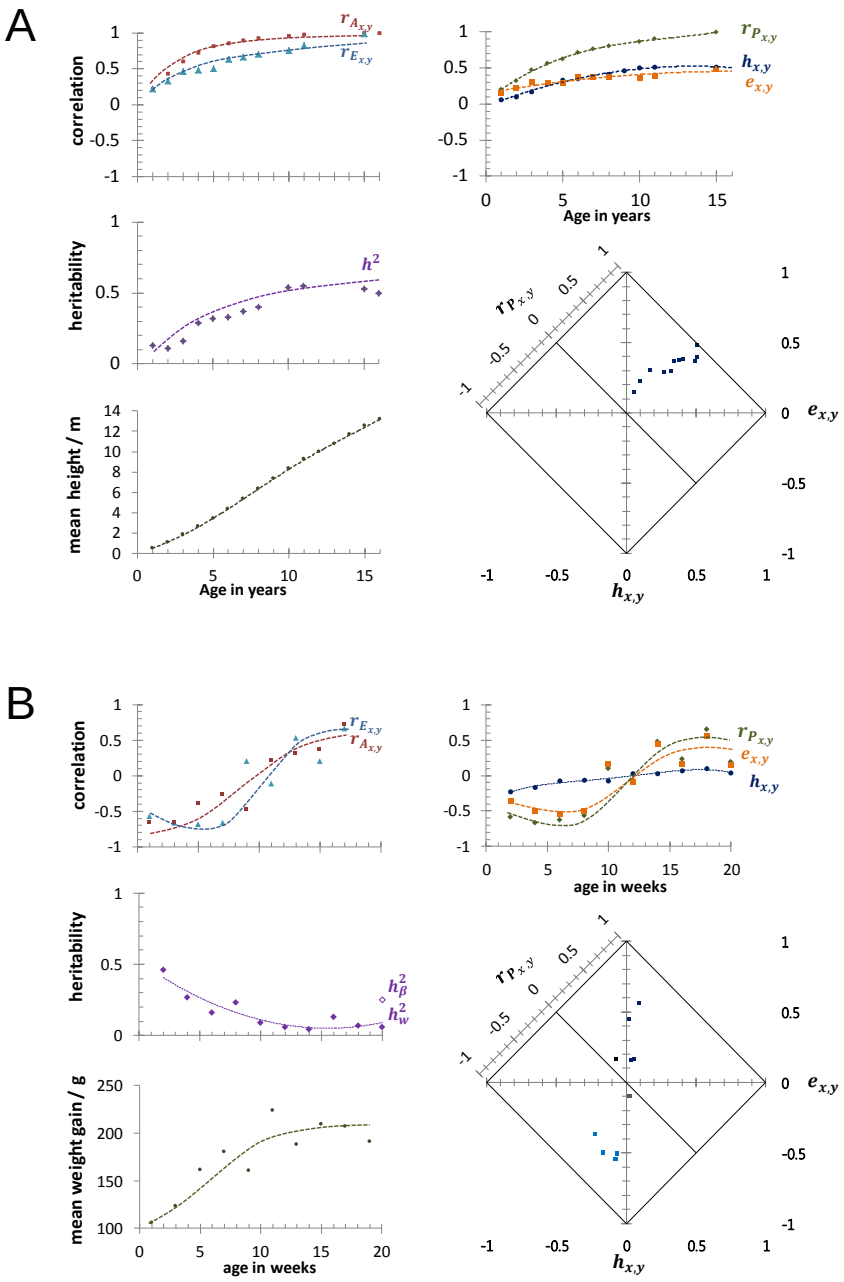
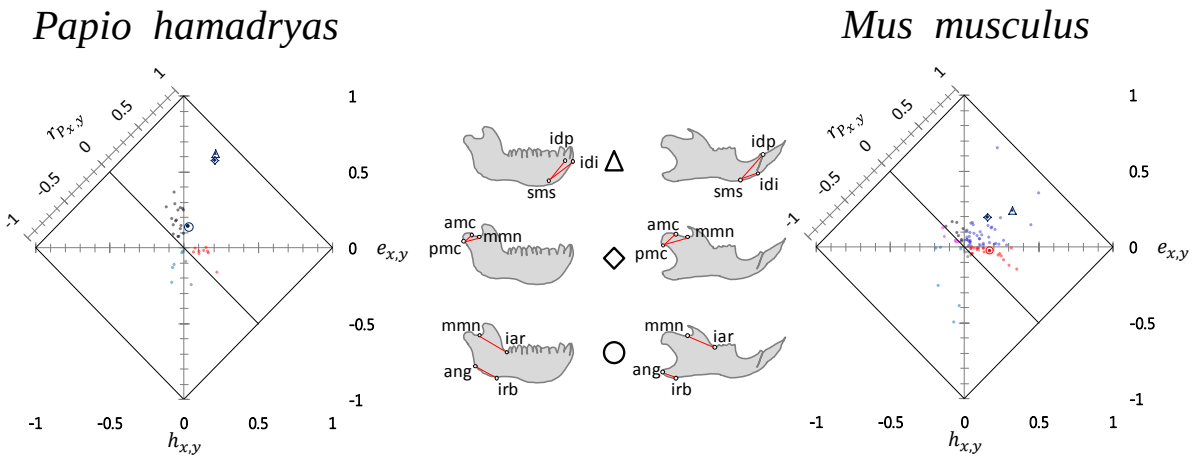


Figure 8. Age-dependent trend in genetic parameters in (A) height growth in japanese larch (*Larix kaempferi*) (Data source Diao et al. 2016), and (B) weight gain (increments) in Korean native chicken (*Gallus gallus domesticus*) (Data source: Manjula et al. 2017).

634 **Example 3. Modularity and Integration**

635 The rationale of morphological integration states that functionally and developmentally related
636 exhibit high levels of phenotypic correlation (Olson and Miller 1958), so that traits belonging to
637 the same functional and/or developmental group are genetically more integrated than traits
638 with different functions or developmental origins (Conner and Via 1993, Waitt and Levin 1998).
639 Willmore et al. (2009) conducted a quantitative genetics study designed to compare patterns of
640 mandibular integration between baboon and mouse. The traits were distances between
641 homologous landmarks in the mandible of each species. Heritabilities of the traits as well as the
642 correlations ($r_{A_{x,y}}$, $r_{P_{x,y}}$) between them were determined in each species separately. Three
643 cases are presented here that reveals the degree of insight provided by the decomposition of
644 the phenotypic correlation. The first case deals with two traits whose genetic correlations are
645 higher than the phenotypic correlation in both species. In the incisive alveolar module, the
646 correlations ($r_{A_{x,y}}$, $r_{P_{x,y}}$) between the distances *sms-idp* and *sms-idi* were $(0.911\,, 0.838)$ in
647 baboon, and $(0.862\,, 0.567)$ in mouse (displayed as a triangle in the 2DHER-field, Figure 9),
648 which by a dimple examination of their magnitudes, it would suggest of a strong genetic
649 influence. The use of the coheritability *and* coenvironmentability expand the inferential space,
650 and point out that most of phenotypic correlation between these traits in baboon was due to
651 the coenvironmentability (75%r_P or 0.624), whereas in the mouse due to the coheritability
652 (60%$r_{P_{x,y}}$ or 0.325). This aspect would otherwise have been overlooked by merely comparing
653 the magnitudes of the correlations.

A second case involves trait-trait associations entirely due to either coheritability or coenvironmentability. In both species the distances *ang-irb* (angular process) and *mmn-iar* (coronoid process) are weakly correlated at the phenotypic level ($r_{P_{x,y}}$: in baboon 0.213, in mouse 0.143; circle)(Figure 9A). However, in baboom the coheritability had a negligible contribution to $r_{P_{x,y}}$ ($h_{x,y} = 0.03$, $e_{x,y} = 0.183$), whereas in mouse the coheritability amounted to most of the phenotypic correlation ($h_{x,y} = 0.170,\ e_{x,y} = -0.027$). Lastly, this case exemplifies that trends in one species cannot be extrapolated to related ones. The phenotypic correlation, coheritability and coenvironmentability in the alveolar *sms-idi* and *sms-idp* (Figure 9A, triangle) and in the condylar *pmc-amc* and condylar-coronoid *pmc-mmn*; (rhomboid) have similar values in baboon and they that plot almost together in the 2DHER-field. In mouse, however, these traits differ in terms phenotypic correlations and the relative amount contributed by the coheritability and coenvironmentability.



667 Figure 9. Modularity and integration. Comparison of mandibular integration between baboon
668 (*Papio hamadryas*), and mouse (*Mus musculus*). At the center are depiction of the mandibles
669 and the distance traits: triangle=*sms-idi* and *sms-idp*; rhomboid=*pmc-amc* and *pmc-mmn*;
670 circle=*ang-irb* and *mmn-iar* (Data source: Willmore et al. 2009).

671

672 Without disregarding the high proportion of negative genetic correlations in the baboon
673 (around 40%), the allometric data clearly indicates that there is a distinct pattern of modular
674 development operating in the mandible of each species, which should be expected given their
675 different ecological niche, feeding adaptations, mastication process, habits, and diet.

676

677

678

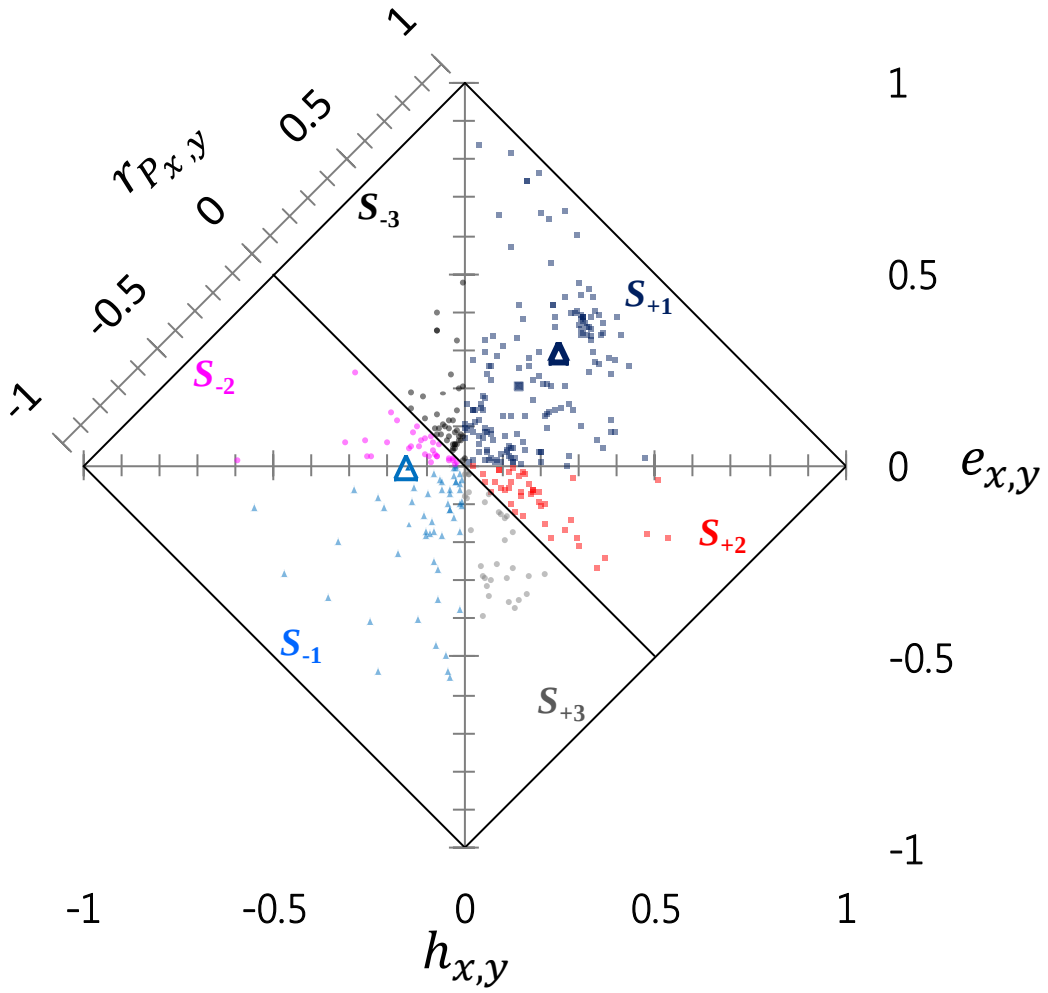
679 **Example 4. Trade-offs of life history traits**

680 A fundamental tenet in the study of fitness in natural populations is that life-history traits limit
681 the individual to attain simultaneously maximal growth, reproduction, and survival (Lande
682 1982). A negative genetic correlations among life-history traits is often adduced to indicate a
683 trade-off that constrain the correlated response to selection in natural populations (Stearns
684 1989, Zera and Harshman 2001). Figure 10 presents results of coheritability and
685 coenvironmentability among life-history traits. Knowing that the sign of the coheritability is

conferred by the genetic correlations, Figure 10 shows sufficient examples of life history traits relationships in both negative and positive directions, congruent with findings in other studies (Roff 1996, Kruuk 2003). Noteworthy is the study of Dutilleul et al. (2015) who analyzed fertility and survival of the nematode *Caenorhabditis elegans* subject to contaminated environments. They found that genotypes that achieve faster sexual maturity (early growth) were more fertile, but had a reduced life span. Their data shows that fertility (x) and early growth (y) were correlated at the phenotypic level $r_{P_{x,y}} = 0.543$, with approximately similar contributions by coheritability ($h_{x,y} = 0.248$) and coenvironmentability ($e_{x,y} = 0.295$). However, with late growth, fertility exhibited a negative phenotypic correlation $r_{P_{x,y}} = -0.154$, which was mostly due to the coheritability ($96\% r_{P_{x,y}}\ or\ -0.148$). The decrease of the phenotypic correlation between fertility from early to late growth could be explained by the antagonistic pleiotropy model which states that if the genes that promote fertility early in life are the same that cause deleterious effects late in life, then fitness is maximized for fertility when the organism is young, at the expense of detrimental performance as the individual age. It would be of great interest to explore whether the described trade-off results from a covariance of both traits to a third unmeasured trait, such as the individual's body size. If individuals reach maturity early at a relatively small adult size (implying shorter growth period) then it would have the added benefit to produce progeny early to compensate for their short lifespan (Charmantier et al. 2006). Generally, traits associated to adaptive response resulting in enhanced fitness (e.g. fertility) will provide an advantage to the individuals that manifest such traits by favoring their reproduction, whose timing and plasticity would depend on the environment. Although

707 selection pressure will always tend towards fitness increase (Fisher 1930, p. 35), it does not
708 imply that fitness necessarily cause the competitive ability of a population to be superior with
709 respect to other populations not interbreeding with it (Lerner and Dempster 1948).

710



711

712 Figure 10. Trade-offs in life history traits in animals. Highlighted are the contrasting results of
713 fertility and growth in *Caenorhabditis elegans* (Dutilleul et al. 2015). Potential trade-offs may

714 be present between fertility and early growth (triangle solid). Fertility and late growth (triangle
715 clear). (Data source: compiled from the literature).

716

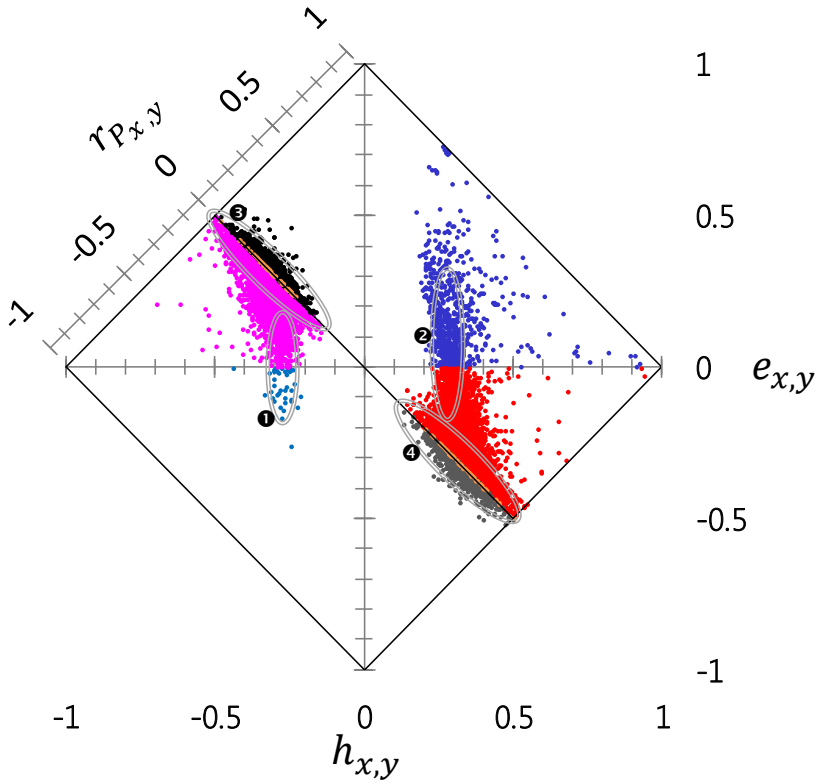
717 The visualization of phenotypic, genetic, *and* environmental information in the 2DHER-field
718 does facilitate the analysis in a single graph. There is an increasing recognition that the
719 environment has a direct influence on the quantitation of genetic parameters underlying life-
720 history traits, therefore it adds value to use the completeness of the data in these analyses to
721 evaluate changes in environmental conditions can influence genetic interactions and
722 covariances among life-history traits (Sgrò and Hoffman 2004).

723

724 **Example 5. Gene co-expression**

725 Gene expression is a major contributor to phenotypic co-variation in human complex trait
726 (Gandal et al. 2018, Lee et al. 2012). Quantitative genetics of gene expression considers the
727 abundance of a particular mRNA transcript as a "trait". Gene expression can be influenced by a
728 variety of biological factors, which are especially susceptible to interact with the environment.
729 To gain insight into the mechanisms that control gene expression, Lukowski et al. (2017)
730 investigated the proportion of co-expression between genes in whole blood samples. Figure 12
731 displays the results of thousands ($n{\sim}14000$) of bivariate comparison in the 2DHER-field, and
732 show that gene expression levels explore all sign configurations among $r_{P_{x,y}}$, $h_{x,y}$, and $e_{x,y}$.
733 Most remarkable is the apparent lack of data points along the transect $-0.1 < h_{x,y} < 0.1$,

734 which distinctly separates and defines two major patterns of coexpression based on the sign of
735 the coheritability. Though classification is not the focus of this work, the visualization in the
736 2DHER-field nevertheless reveals a level of insight and could serve as a guide to further
737 examine functional modes of co-expression.
738



739

740 Figure 12. Shared genetic and shared environmental architecture of transcription in human
741 peripheral blood. Significant genetically correlated probe pairs (study-wide false discovery rate
742 $FDR < 0.05$). Of these 14991 data pairs, 14020 (93%) are located on different chromosomes,
743 7886 have a positive coheritability and 7105 a negative coheritability (Data source: Lukowski
744 et al. 2017).

745 Here special attention is placed on data points circumscribed by the ellipses in Figure 12.
746 Groups 1 and 2 enclose observations with low coheritability, delimited to the interval
747 $-0.3 < h_{x,y} < -0.2$ (Group 1), and $0.2 < h_{x,y} < 0.3$ (Group 2). In these groups, much of the
748 phenotypic correlation is under varying influence of the coenvironmentability. In Group 1 the
749 phenotypic correlations are negative, mostly occupying partition S_{-2}. The positive phenotypic
750 correlations associated to Group 2, includes data with negative (partition S_{+2}) and positive
751 (partition S_{+1}) coenvironmentabilities. Group 3 and 4 are found along a narrow interval around
752 the null phenotypic correlations $(-0.1 < r_{P_{x,y}} < 0.1)$. For the $r_{P_{x,y}}$ to be close to zero, the
753 coheritability and coenvironmentability must have similar magnitudes but with different sign.
754 Numerous instances where the phenotypic and genetic correlation differed in sign were found
755 in partitions S_{+3} and S_{-3}. The values of the genetic correlation averaged -0.019 and oscillated
756 widely throughout all its domain [-1, 1] as can be seen by the large standard deviation (SD)
757 0.883, a fact that is corroborated by the large disparity values (mean $D = 0.84, SD = 0.12$).
758 The coheritability allocated more or less equally to the positive and negative realms of its
759 domain. All this indicated that the genetic correlation when the phenotypic correlation was
760 close to or not different than zero differed greatly in magnitude and direction with respect to
761 $r_{P_{x,y}}$. When sample sizes are large relative to the number of variables, the graphical method is
762 likely to provide the most insight. Gene expression is subject to many factors including
763 methodological bias (see Pereira et al. 2009), and is known to be responsive to environmental
764 clues that alter mRNA abundances accordingly (Grishkevich and Yanai 2013). Therefore, it

765 would be of great interest to conduct assessments of loci known to be up- or down regulated
766 under different environmental conditions (Li and Burmeister 2005).

767 Discussion
768

769 Biological studies rely on the phenotypic correlation between traits as an important
770 measurement of association between the phenotypic values of two traits. There is an added
771 interest to distinguish how much of the correlation can be attributed to additive genetics and
772 how much to the environment. This paper presents a method that captures both the genetic
773 and environmental contributions to the phenotypic correlation, based on quantitative genetics
774 theory, and aims to formalize the decomposition of the phenotypic correlation into the
775 coheritability and coenvironmentability, to further present its mathematical and statistical
776 properties, and to provide a visualization tool capable to handle all the pertinent variables (i.e.,
777 $r_{P_{x,y}}$, $h_{x,y}$, $e_{x,y}$) concurrently.

778

779 Coheritability and genetic correlation
780

781 First, it is important to distinguish between coheritability and genetic correlation. Coheritability
782 quantifies the contribution of additive genetic effects to the phenotypic correlation between
783 two traits, and relates the genetic covariance to the bivariate phenotypic variability (i.e.
784 geometric mean of the phenotypic variances, see Equation [4]). The genetic correlation

785 measures the association between the breeding values of two traits and relates the genetic
786 covariance to the bivariate *genetic* variability (i.e. geometric mean of the genetic variances,
787 equation [6]). Thus, coheritability and genetic correlation do not share the same inferential
788 space. Coheritability, together with the coenvironmentability have an additive relationship with
789 the phenotypic correlation, and is therefore most proximal to the phenotypic correlation, an
790 aspect not shared by the genetic and environmental correlation. In addition, the coheritability
791 has predictive ability in regard to correlated response to selection. The formula of the
792 coheritability (Equation 7) involves the modulation of the genetic correlation by the square root
793 of the trait heritabilities, the latter being independent random variables, coheritability and
794 genetic correlation, notwithstanding sharing the same sign, are not numerically similar nor
795 follow necessarily the same trend nor rank.

796

797 ## The biological and the biometrical concepts of coheritability
798

799 To further conceptualize the meaning of coheritability, it is useful to construe the term in a
800 biological sense as well as in a biometrical sense. Biological coheritability is grounded on a
801 gene-centric approach that rests on the degree of correspondence between specific phenotypic
802 traits and the allelic content of identified, causal genes influencing them. This includes the
803 effect of all loci that affect both traits (i.e. pleiotropic alleles), and the effects of non-randomly-
804 associated, tightly linked loci acting on each trait singly (i.e. loci in linkage disequilibrium) (see
805 Zhang et al. 2018, Tyler et al. 2009), as well as the loci that regulate their expression. Trait

806 correlations caused by pleiotropic alleles are adaptive, different than trait correlations due to
807 linkage disequilibrium, which could erode through recombination (Saltz et al. 2017).

808 The biometric concept of coheritability is based on population-level, statistical principles, and it
809 recognizes that the observed association of traits is the net realization of a multiplicity of
810 shared and not shared genetic factors influencing polygenic traits. It expresses the notion that
811 two traits are inherited together (i.e. maintain the resemblance seen in the parental
812 generation) due to the aggregate additive effect of genetic factors affecting the traits, including
813 those acting on each trait individually, others on both traits. In this sense, the biometrical
814 coheritability derived from multivariate analysis becomes a hypothesis on the degree of
815 contribution of genetics to the phenotypic correlation of traits.

816 Biological coheritability deals with the genetic architecture of quantitative traits based upon
817 identified causal gene variants, quantitative trait loci, additive and dominance effects, mapping,
818 coding and noncoding sequences, and by understanding its variability due to the action of
819 modifier genes adjusting its penetrance and expressivity, the occurrence of recombination
820 events, or by the activity of regulatory elements in other parts of the genome (Short et al. 2018,
821 Mackay et al. 2009). Discerning between pleiotropic and coincident linkage is an area of
822 intensive experimental (Wagner and Zhang 2011, Solovieff et al. 2013, Gardner and Latta
823 2007) effort associated to a diversity statistical and computational (Hackinger and Zeggini
824 2017, Schaid et al. 2016, han and Pouget 2015, Yang et al. 2015, Carter et al. 2007) approaches
825 involving sequencing, fine mapping, and functional characterization of the gene and gene
826 product (Flint and MacKay 2009). The multifunctionality of a pleiotropic gene can be mediated

827 by alternative splicing, by RNA editing, and tissue specific expressions. Pleiotropy can also
828 exhibit multiple phenotypic consequences of a single molecular function, such as those acting
829 on multiple pathways (Singh and Shaw 2012).

830 The biometrical coheritability invokes quantitative genetic concepts such as heritability,
831 correlation (genetic, phenotypic), genetic factors. Its variability depends on population
832 structure, the additive and phenotypic variances of each trait, and the genetic covariance
833 between the traits (Carey 1988). The biometrical concept of coheritability does not carry the
834 assumption that the genetic factors involved in the expression of the traits are related to genes
835 with an additive mode of action, nor the existence of genetic covariance precludes the effects
836 of genes with any degree of dominance or epistasis (cf. Huang and Mackay 2016).

837 Biometrical coheritability is not commensurate with relevance of importance of the trait-trait
838 association. In a recent study on the relationship between amygdala (trait x) and emotion
839 recognition (trait y), Knowles et al. (2015) determined heritabilities $\left(h_x^2 = 0.72, h_y^2 = 0.32\right)$ and
840 genetic correlation $(r_{A_{x,y}} = 0.25)$ of these traits. It allows us to determine a coheritability
841 $h_{x,y} = 0.12$, which appears to be of very low magnitude. Using bivariate linkage and
842 association analyses, they identified the gene $PDE5A$ (locus $4q26$) whose gene product is the
843 phosphodiesterase 5 (PDE5) enzyme expressed in various regions of the brain, and has been
844 linked to deficits in memory recognition. As such the PDE5 enzyme has emerged as a potential
845 drug target for treating cognitive deficits and neurological disease (Teich et al. 2016).

Whereas biological coheritability relates allelic variants and traits, biometrical coheritability focuses on trait-trait associations. At the interface is the extensive use of molecular markers (e.g. SNPs) in genome analyses which has occasionally presented loci harboring markers associated to several, sometimes disparate, traits. These cross-phenotype associations could have as underlying cause pleiotropy, linkage disequilibrium, or be artifactual nature (Solovieff et al. 2013). Besides the 'missing' heritability problem that arises in the estimation of heritability in GWAS (Yang et al. 2017), Gianola et al. (2015) demonstrated that correlation coefficients inferred using markers can provide a distorted picture of the actual genetic correlation between traits due to the lack of knowledge about linkage disequilibrium relationships between QTLs and markers. Thus, speculating about genetic correlations and even more about its causes (e.g. pleiotropy) using genomic data is conjectural (cf. Lee et al. 2012). Pleiotropy is, by definition, a property of a gene or locus, not of a marker. Therefore, caution must be exercised when interpreting (rethorically and conceptually) findings of marker-based genetic parameters, as to distinctly differentiate the causal (i.e., genes) component from the instrument (i.e., the marker). This distinction is critical.

Coheritability under a null phenotypic correlation

A phenotypic correlation equal to or not significantly different than zero, does not necessarily imply that both coheritability and coenvironmentability are also zero or non-significant. The coheritability and coenvironmentability could be similar in magnitude but not in sign.

867 Therefore, an apparent lack of association between traits at the phenotypic level, cannot be
868 dismissed as being of no interest at the genetic level. Disparity analyses revealed that $r_{P_{x,y}}$ and
869 $r_{A_{x,y}}$ differed the most when the phenotypic correlation is around zero.

870 Under a null phenotypic correlation, the definition and metrics of the phenotypes must be
871 reassessed. The ability to define meaningful traits in appropriate continuous or categorical
872 scales would help avoid the confounding of distinct characters as a single phenotype (de
873 Villemereuil 2017). This is of particular important in clinical research where a convergence of
874 multiple traits, symptoms, signs in an individual may be a consequence of a cascade of disease
875 causing defects in a complex network of interacting genes, proteins, and metabolites (Park et
876 al. 2009, Emilsson et al. 2018). Yet given the heterogeneity of sources contributing to
877 comorbidity, it is not obvious whether these traits are properties of a disease at the individual
878 or at the population level. Among several options, two approaches can be used to address this
879 problem. One approach is the use of subphenotype groups. A subphenotype is a subset of
880 individuals drawn from a large cohort who feature a specific set of traits in common (Morris et
881 al. 2010). The intent is to 'concentrate' individuals with similar trait-trait associations, generally
882 using latent class analyses (Famous et al. 2017, Gårdlund et al. 2018), under the premise that, if
883 the traits have a common genetic basis, their detection would be facilitated, otherwise by
884 ascertaining all individuals simultaneously as having the same phenotype, the causal factors
885 would be overlooked.

886 Another approach regards a complex phenotype as a composite of many traits, some of them
887 more closely connected to the genetic cause of the complex phenotype. A trait with this

888 characteristic is referred to as an endophenotype. Endophenotypes are widely employed in
889 psychiatric genetics (Iacono 2018, Doyle et al 2005), livestock research (te Pas et al. 2017),
890 immunology (Gregersen et al. 2015), and clinical research (Benyamin et al. 2007). Glahn (et al.
891 2012) used the absolute value of the coheritability as a criterion to select endophenotypes
892 associated to major depression risk. Similarly, Hammer et al. (2006) proposed to dissect
893 complex phenotypes into traits spanning different levels of biological organization within the
894 individual in order to characterize the full suite of factors that contribute to quantitative
895 variation across cellular, tissue, organ, organismal levels, as well as developmental stages (Cobb
896 et al. 2013). Sun et al (2015) proposed a method to search for a combination of phenotypic
897 characters of multivariate phenotypes and directly maximize the heritability of this combined
898 trait. Though the aim of these models is to enhance the detection of trait-trait associations, it
899 also brings ontogeny as a main factor shaping relationships among traits.

900

901 ## Rules governing trait-trait associations
902

903 Inspection of the scatter plot of $h_{x,y}$, $e_{x,y}$ and r_P on the 3DHER-plane graphically showed that,
904 despite the sizeable sample size, not all places in the plane appeared uniformly populated,
905 suggesting that some areas more inaccessible than others such as close to the border where
906 the relationship $|h_{x,y} - e_{x,y}| = 1$ holds. At this area, data can only be generated by combining
907 heritabilities and genetic correlations both at very high magnitudes. If the construction of traits
908 entails optimization constraints, then heritabilities and genetic correlation reaching

909 simultaneously extreme values in their domains will not be favored. The occupancy question
910 should be better viewed as a consequences of how traits are constructed, and how ontogeny
911 molds their relationship through life. The organized, sequential, and modulating events of
912 development would manifest varying degrees of genetic and environmental control of two
913 correlated traits (Badyaev and Martin 2000, Saltz et al. 2017). Under this condition, not all
914 corners of the 2DHER-field would be accessible to the same degree.

915 The developmental process by which traits are constructed derive from genetic (Kellogg 2004)
916 and epigenetic interactions (Peaston and Whitelaw 2006), as well as its interactions with the
917 environment, all combined to shape the phenotypes as emergent products of such interactions.
918 Certainly, phenotypic structures can be mathematical modeled using simple rules (e.g. egg
919 shape by Stoddard et al. 2017, shell coiling by Raup 1966), or conform to empirical allometric
920 scaling such as volume and surface area (Square-Cube Law; Haldane 1926), metabolic rate and
921 body size (Kleiber's Law; Niklas and Klutchera 2015, Hulbert 2014), speed and body mass
922 (Meyer-Vernet and Rospans 2015). Though a constructionist perspective enhances our ability
923 to account for the relationships among traits, these studies are not meant to simulate the
924 actual biological mechanism of development. Developmental systems are often under strong
925 stabilizing selection to maintain homeostasis, such that patterns of trait-trait covariation
926 through ontogeny or within a particular ontogenic stage generally exhibit conservatism of
927 developmental systems (Badyaev and Martin 2000). Genetic, phenotypic correlations and
928 autocorrelations among traits would reduce independent variation of traits at different ages,
929 limit the number of dimensions in growth trajectories, and overall present a powerful

930 inducement for maintaining certain relationships between traits through ontogeny (Badyaev
931 and Martin 2000).

932

933 ## The biotic environment and its role in trait-trait association
934

935 The manner the coheritability undergoes changes under varying environmental circumstances
936 reveal that the environment can potentially induce changes in the genetic architecture of
937 complex traits (Sikkink et al. 2017), and inferences deduced from studies conducted in one
938 environment cannot be generalized (Aastveit and Aastveit 1993). Stearns et al. (1991)
939 proposed that environmentally induced changes in magnitude and sign of genetic covariances
940 occur when traits are functionally uncoupled at the physiological or developmental levels.

941 Of particular importance is the need to clarify the extent, if any, of the existence of a
942 covariance between the genetic effects of one trait and the environmental effects of the other.
943 Generally, these terms are surmised to be zero. Nevertheless, instances of indirect genetic
944 effects (IGEs, Bijma 2014) exerted by an individual on the trait values of another individual (e.g.
945 maternal effects), admits that there may exist a partly heritable component in the social/biotic
946 environment when IGEs occur (Ørsted et al. 2017). For instance, in isogentic lines of
947 *Caenorhabditis elegans* nematodes, progeny traits such as fecundity and rate of development
948 are due to maternal-dependent provisioning of vitellogenin protein to the embryos (Perez et al.
949 2017), parents' influence juvenile body size by adjusting investment per offspring (Rollinson and

Rowe 2015). Traits related to growth, nutrient assimilation, and nourishment of an organism, which are generally assumed to be an intrinsic property of an individual, is directly affected by the microbiome that colonizes alimentary tracts of animals (Org et al. 2017), or the rhizosphere (Nihorimbere et al. 2011, Wissuwa et al. 2009) and phyllosphere (Kembel et al. 2014, Rosado et al 2018) of plants. The microbial diversity present, in turn, is influenced by host traits (Li et al. 2018). Some researchers regard the commensal microbiome as a virtual organ of the body (Valdes et al 2018) whose genome brings forth emergent, novel capabilities (e.g. processing of drugs, resistance to toxins) to the host that would not otherwise exists (Kho and Lal 2018). Associations among traits manifested in an individual are not only the result from their own adaptations but also could derive from community level relationships. Grab et al. (2019) found that closely related pollinator bee species shared many behavioral and morphological traits including body size, plant fidelity, and visitation rate, but not flower-handling ability. On this basis, they were able to predict characteristics of fruit shape malformation from loss of diversity and abundance of pollinator bees in apple plants, and that phylogenetic diversity and species richness best predicted fruit weight and number of seeds per fruit.

The coheritability and coenvironmentability are the most proximal factors affecting the phenotypic correlation

The coheritability and coenvironmentability are the most proximal factors influencing additively the value of the phenotypic correlation. There is a tendency to surmise that the

971 genetic correlation *per se* will have a preponderant influence on the phenotypic correlation
972 merely because its magnitude is sufficiently large. The theoretical and experimental examples
973 presented in this paper make such proposition untenable. Equation [5] shows that the genetic
974 correlation is embedded within the expression of the phenotypic correlation, where the
975 magnitude of the genetic correlation is modulated by random variables, namely the trait
976 heritabilities. Two traits may display high heritability values, yet be poorly correlated at the
977 genetic level. Highly genetically correlated traits may each display low heritabilities (Drobniak
978 and Cichoń 2016). Therefore, the coheritability is not a linear transformation of the genetic
979 correlation (i.e. it does not preserve rank). A given value of the coheritability may be the result
980 of a numerous set of heritability and genetic correlation values, the coheritability is not a
981 monotonic transformation of the genetic correlation because the variables do not possess a
982 one-to-one relationship. This nuanced interpretation of Equation [5] must be contrasted with
983 other views that treat heritabilities as scalars (cf. Cheverud 1988). Therefore the simple
984 comparison of $r_{P_{x,y}}$ and $r_{A_{x,y}}$ (or their matrices) cannot presume to convey information on the
985 degree on which shared genetics influences $r_{P_{x,y}}$. The argument presented here is persuasive
986 for three reasons. First, the phenotypic correlation and coheritability share the same common
987 denominator, a fact that distinguishes them from the genetic correlation, therefore comparison
988 of $r_{P_{x,y}}$ and $h_{x,y}$ is made on the same basis. Second, the partition of the phenotypic
989 correlation into coheritability and coenvironmentability directly follows, in a standardized form,
990 from the partition of the phenotypic covariance into genetic and environmental covariance ,
991 maintaining additivity of the terms. That is a property that $r_{A_{x,y}}$ and $r_{E_{x,y}}$ do not have with $r_{P_{x,y}}$

since their denominators involve different and independent random variables. Third, results of the regression analyses involving $r_{A_{x,y}}$ and $r_{E_{x,y}}$ as regressors failed to produce a general relationship that can hold to any data set.

Differing from this approach that compares $r_{A_{x,y}}$ and $r_{P_{x,y}}$ values in their own right, Cheverud (1988) postulated that phenotypic correlations could be used as a proxy for genetic correlations. The regression analyses presented in this study have shown that genetic correlation do not map directly to the phenotypic correlation, and overlooking genetic information (i.e. $r_{A_{x,y}}$) can qualitatively affect inferences (evolutionary, statistical, practical breeding outcomes) in ways a purely phenotypic approach cannot predict or explain (Rubin 2016, Kruuk et al. 2008). Therefore, correspondences between phenotypic-based predictions on genetic outcomes are not robust for all plausible assumptions regarding the underlying genetics of traits (Hadfield et al. 2007). By observing equation [5], to consider $r_{P_{x,y}}$ as a suitable predictor of $r_{A_{x,y}}$ assumes that the phenotypic and genetic covariances possess the same sign, that the phenotypic covariance is entirely genetic in nature (i.e. no detectable environmental covariance is present), and that the heritabilities of both traits are unity, all becoming very demanding assumptions (Supplementary Information section 7.4). Thus, phenotypic data is not an adequate predictor of underlying genetics of natural or breeding populations. Since both genetic and environmental correlations combine together to give the phenotypic correlation, therefore, as Falconer and MacKay (1996, page 314) stated, 'this dual nature makes it clear that the magnitude and even the sign of the genetic correlation cannot be determined from the phenotypic correlation alone'.

Uses of the decomposition of the phenotypic correlations

Besides the applications presented in the illustrative examples of this work, there are further uses of the coheritability and coenvironmentability. These parameters could be utilized as objective means to construct common-cause classification of diseases that cluster together diseases with genetic and environmental similarities under the premise that shared genetics and environment would manifest similarities between diseases that have a common-cause nosology (Wang et al 2017). A promising application is enhancing the capability to recognize biomarkers and traits that are readily impacted by a disease challenge (Boulton et al. 2018, de Matos et al. 2018). Clinical diagnostics if construed as a type of indirect selection would benefit by helping detect measurable traits (i.e. phenotypic characters, biomarkers) that are both easily accessible, (positively or negatively) correlated with disease traits, and coheritable with it (see Yin et al. 2017, Mehr et al. 2018, Ngo et al, 2018; cf. Bastaranche et al. 2018).

Also, in the area of genomic prediction, knowledge on the degree genetic effects contribute to the phenotypic correlation can be utilized to simultaneously improve prediction accuracy of parameter estimates of breeding values of multiple traits (Crossa et al. 2017, Montesinos-Lopez et al. 2018). Another area where the decomposition of the phenotypic correlation would be

1032 particularly useful is in high-throughput phenotyping. It would exploit the data in their full
1033 capacity including the genetic interdependencies of traits (Xavier et al. 2017, Sun et al. 2017).

1034 The coheritability of somatic, homologous traits common to both sexes would bring insight in
1035 the study of sexually dimorphic species especially in the manner males and females respond to
1036 selection in different directions given the tendency towards gender-optimal phenotypes (Mank
1037 2009, Poissant et al. 2009, Chippindale et al. 2001). Application of the decomposition of the
1038 phenotypic correlation in morphometrics, a tool for the quantitative description and statistical
1039 analysis of morphology, could allow to model phenotypic changes between traits, emphasizing
1040 those trait-trait correlations that have strong, underlying genetic component, and decrease bias
1041 in certain direction of the morphospace (Polly 2008).

1042 Climate change studies rely on models that attempt to translate meteorological data into a
1043 plausible biological response. Recent progress in modeling tree mortality under drought has led
1044 to the incorporation of plant hydraulic traits in the parameterization of simulation models in
1045 homogeneous forest communities (Choat et al. 2018, Martínez-Vilalta et al. 2009). The
1046 coinheritance of these traits with other morphological, physiological and life-history traits
1047 would add value to statements of climate change on trait diversity in natural habitats.

1048

1049

1050 ## Conclusions
1051

The decomposition of the phenotypic correlation into coheritability and coenvironmentability is consistent with classical statistical genetics theory and therefore placed on a firm statistical footing. The decomposition provides interpretable results that employs the totality of the data (phenotypic, genetic, environmental), utilizes univariate (heritabilities) and bivariate (correlations) statistical measures, accounts for all combinations of magnitude and direction, standardizes covariance terms to a common denominator, all integrated into a simple, intuitive yet robust framework to simultaneously inform on portions of r_P contributed to joint genetic effects and joint environmental deviations. The graphical representations of $h_{x,y}$, $e_{x,y}$, and r_P in the 3DHER-plane or 2DHER-field help visualize all pertinent variables concurrently.

Improving our knowledge of underlying genetic basis of trait-trait associations helps to understand and predict species responses in a more variable environment. Although the coheritability and coenvironmentability, in a biometric sense, *per se* cannot capture subtleties of the causal factors underlying the genotypes or even their complete functional relationships between them, it nevertheless explore an important inferential space that point to relationships that deserve further scrutiny, refine searches, and in general, to supplement results obtained from other experimental lines of evidence.

Literature Cited

[000] Aastveit, A.H. and K. Aastveit, 1993 Effects of genotype-environments interactions on genetic correlations. Theor. Appl. Genet. 86:1007-1013. https://doi.org/10.1007/BF00211054

[000] Badyaev, A., and T.E. Martin, 2000 Individual variation in growth trajectories: phenotypic and genetic correlations in ontogeny of the house finch (Carpodacus mexicanus). J. Evol. Biol. 13: 290-301. https://doi.org/10.1046/j.1420-9101.2000.00172.x

[000] Barabási, A.L. N. Gulbahce, and J. Loscalzo, 2011 Network medicine: a network-based approach to human disease. Nature Review Genetics 12:56-68. http://doi.org/10.1038/nrg2918

[000] Baradat, P., 1976 Use of juvenile-mature relationships and information from relatives in combined multitrait selection. Proc. Joint Meeting on Advanced Generation Breeding. June 14-18, Bordeaux, France. P. 121-138.

[000] Bastarache, L., J.J. Hughey, S. Hebbring et al. 2018 Phenotype risk scores identify patients with unrecognized mendelian disease patterns. Science 359:1233-1239. https://doi.org/10.1126/science.aal4043

[000] Beddard, P.R., C.S. Hsu, L.P.S. Spangelo et al. 1971 Genetic, phenotypic and environmental correlations among 28 fruit and plant characters in the cultivated strawberry. Ca. J. Genet. Cytol. 13:470-479. https://doi.org/10.1139/g71-071

[000] Benyamin, B., T. I. A. Sørensen, K. Schousboe, M. Fenger & P. M. Visscher & K. O. Kyvik 2007. Are there common genetic and environmental factors behind the endophenotypes associated with the metabolic syndrome? .Diabetologia 50:1880–1888. https://doi.org/10.1007/s00125-007-0758-1

[ ] Bijma P. 2014. The quantitative genetics of indirect genetic effects: a selective review of modelling issues. Heredity (2014) 112, 61–69. https://doi.org/10.1038/hdy.2013.15

[000] Boulton, K. M.J. Nolan, Z. Wu et al., 2018. Phenotypic and genetic variation in the response of chickens to *Eimeria tenella* induced coccidiosis. Genet. Sel. Evol. 50:63. https://doi.org/10.1186/s12711-018-0433-7

[14] Carey, G. 1988. Inference about genetic correlations. Behavior Genetics 18(3):329-338. https://doi.org/10.1007/BF01260933

[000] Carter, G.W., S. Prinz, C. Neou et al. 2007 Prediction of phenotype and gene expression for combinations of mutations. Mol. Syst. Biol. 3:96. https://doi.org/10.1038/msb4100137

[000] Charmantier A., C. Perrins, R.H. McCleery, B.C. Sheldon, 2006 Quantitative genetics of age at reproduction in wild swans: support for antagonistic pleiotropy models of senescence. PNAS 103:6587-6592. https://doi.org/10.1073/pnas.0511123103

[000] **Chen, X., B. Hawkins, C.-Y Xie, , C.C. Ying, 2004 Age trends in genetic parameters and early selection of lodgepole pine provenances with particular reference to the Lambeth model. Forest Genetics. 10. 249-258.**

[000] Cheng, R., J. Borevitz, R.W. Doerge, 2013. Selecting informative traits for multivariate quantitative trait locus mapping helps to gain optimal power. Genetics 195:683-691. https://doi.org/10.1534/genetics.113.155937

[000] Cheverud, J.M. 1998. A Comparison of Genetic and Phenotypic Correlations. Evolution 42(5):958-968. https://doi.org/10.1111/j.1558-5646.1988.tb02514.x

[000] Chippindale A.K., J.R. Gibson, W.R. Rice 2001 Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. PNAS98: 1671–1675. https://doi.org/10.1073/pnas.98.4.1671

[000] Choat, B., T.J. Brodribb, C.R. Brodersen et al., 2018 Triggers of tree mortality under drought. Nature 558:531-539. https://doi.org/10.1038/s41586-018-0240-x

[000] Cobb, J.N., G. DeClerk, A. Greenberg et al., 2013 Next generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126:867-887 https://doi.org/10.1016/j.tplants.2017.08.011

[000] Conner, J., and S. Via. 1992. Natural selection on body size in *Tribolium*: possible genetic constraints on adaptive evolution. Heredity 69:73–83.

[000] Conner J.K., K. Karoly, C. Stewart, et al. 2011 Rapid independent trait evolution despite a strong pleiotropic genetic correlation. Am. Nat. 2011 178:429-41. https://doi.org/10.1086/661907

[000] Crossa, J., P. Perez-Rodriguez, J. Cuevas 2017. Genomic selection in plant breeding: methods, models, and perspectives. Trends in Plant Sciences 22: 961-975. https://doi.org/10.1016/j.tplants.2017.08.011

[000] DeStephano, A., Seshadri, S., Beiser, A., Atwood, L.D., Massaro, J.M., Au, R., Wolf, P.A., DeCarli, C. 2009. Bivariate Heritability of Total and Regional Brain Volumes: The Framingham Study. Alzheimer Disease & Associated Disorders. 23(3):218-223. https://doi.org/10.1097/WAD.0b013e31819cadd8

[000] de Matos, R., C. Ferreira, S.K. Herukka, et al., 2018 Quantitative genetics validates previous genetic variants and identifies novel genetic players influencing Alzheimer's disease cerebrospinal fluid biomarkers. J. Alzheimer's Dis. 66: 639-652. https://doi.org/10.3233/JAD-180512

[000] de Reggi, C. 1972 [Generalization of the notion of heritability of several characters: coheritability]. Ann. Genet. 15:41-4.

[000] de Villmereuil, P. 2017 Quantitative genetic methods depending on the nature of the phenotypic trait. The Year in Evolutionary Biology 2017. Annals of the New York Academy of Sciences. https://doi.org/10.1111/nyas.13571

[000] Diao, S., Hou, Y., Xie, Y., and Sun, X. 2016. Age trends of genetic parameters, early selection and family by site interactions for growth traits in *Larix kaempferi* open-

1153 **pollinated families. BMC Genetics 17:104. https://doi.org/ 10.1186/s12863-016-**
1154 **0400-7**

1155 **[000] Doyle, A.E., E.G. Willcut, L.J. Seidman, et al., 2005 Attention-Deficit/Hypesensitivity**
1156 **Disorder endophenotypes. Biol Psychiatry 57:1324-1335.**
1157 **https://doi.org/10.1016/j.biopsych.2005.03.015**

1158 **[000] Drobniak, S. M. and Cichon, M. 2016. No evidence for the positive relationship**
1159 **between genetic correlations and heritabilities. bioRxiv**
1160 **https://doi.org/10.1101/039388**

1161 **[000] Dutilleul M., B. Goussen B, J.-M. Bonzom, S.Galas, D. Réale 2015 Pollution Breaks**
1162 **Down the Genetic Architecture of Life History Traits in *Caenorhabditis elegans*. PLoS**
1163 **ONE 10(2): e0116214.doi:10.1371/journal.pone.0116214**

1164 **[000] Emilsson, V. , and M. Ilkov, J.R. Lamb et al. 2018 Co-regulatory networks of human**
1165 **serum proteins link genetics to disease. Science 361769-773.**
1166 **https://doi.org/a0.1126/science.aaq1327**

1167 **[000] Falconer, D. S. and T. F. C. Mackay 1996. *Introduction to quantitative Genetics* Ed 4.**
1168 **Longmans Green, Harlow, Essex, UK. P. 312-334.**
1169 **[000] Famous, K.R., K. Delucchi, L.B. Ware et al., 2017 Acuate respiratory distress syndrome**
1170 **subphenotypes respond differently to randomized fluid management strategy. Am. J.**
1171 **Crit. Care Med. 195:331-338. https://doi.org/10.1164/rccm.201603-0645OC**

[000] Fisher, R. A., 1915 Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. Biometrika, 10, 507-521. https://doi.org/10.1093/biomet/10.4.507

[000] Fisher, R.A., 1930 Fundamental Theorem of Natural Selection. Trans. R. Soc. Edinburgh 52, 399–433.

[000] Flint, J. and T.F.C. MacKay 2009 Genetic architecture of quantitative traits in mice, flies and humans. Genome Research 19:723-733. https://doi.org/10.1101/gr.086660.108

[000] Gandal, M.J., J.R. Haney, N.N. Parikshak et al. 2018 Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science 359:693-697. https://doi.org/10.1126/science.aad6469

[000] Gårdlund, B., N.O. Dmietrieva, C.F. Pieper et al., 2018 Six subphenotypes in septic shock: latent class analysis of the PROWESS Shock Study. J. Crit. Care 47:70-79. https://doi.org/10.1016/j.jcrc.2018.06.012

[000] Gardner, K.M. and R.G. Latta 2007 Shared quantitative trait loci underlying the genetic correlation between continuous traits. Molecular Ecology 16:4195-4209. https://doi.org/10.1111/j.1365-294x.2007.03499.x

[000] Gianola, D., G. Campos, M.A. Toro, M.A. et al. 2015. Do Molecular Markers Inform About Pleiotropy? Genetics 201:23–29. https://doi.org/10.1534/genetics.115.179978

1191 **[000] Glahn D.C., J.E. Curran, A.M. Winkler, M.A. Carless et al. 2012 High Dimensional**
1192 **Endophenotype Ranking in the Search for Major Depression Risk Genes. *Biological***
1193 ***Psychiatry 71*(1)6–14. https://doi.org/10.1016/j.biopsych.2011.08.022**

1194 **[000] Grab, H., M.G. Bramstetter, N. Amon et al. 2019 Agriculturally dominated landscapes**
1195 **reduce bee phylogenetic diversity and pollination services. Science 363:282-284.**
1196 **https://doi.org/10.1126/science.aat6016**

1197 **[000] Gregersen, P.K., G. Klein, M. Keogh et al., 2015 The Genotype and Phenotype**
1198 **Registry: a living biobank for the analysis of quantitative traits. Immunol. Res. 63.107-**
1199 **112. https://doi.org/10.1007/s12026-015-8711-8.**

1200 **[000] Grishkevich V. and I. Yanai 2013 The genomic determinants of genotype ×**
1201 **environment interactions in gene expression. Trends Genet. 29:479-87.**
1202 **https://doi.org/10.1016/j.tig.2013.05.006**

1203 **[000] Gromko, M.H. 1995 Unpredictability of correlated response to selection. Pleiotropy**
1204 **and sampling interact. Evolution. 1995 Aug;49(4):685-693.**
1205 **https://doi.org/10.1111/j.1558-5646.1995.tb02305.x**

1206 **[000] Gui, H., J. S. Kwan, P. C. Sham, S. S. Cherny and M. Li 2017. Sharing of Genes and**
1207 **Pathways Across Complex Phenotypes: A Multilevel Genome-Wide Analysis. Genetics**
1208 **206:1601-1609. https://doi.org/10.1534/genetics.116.198150**

[000] Guo, Z., Wang,W., Cai, T.T., Li, H. 2006. Optimal Estimation of Co-heritability in High-dimensional Linear Models. arXiv:1605.07244

[000] Hackinger, S. and E. Zeggini 2017 Statistical methods to detect pleiotropy in human complex traits. Open Biol. 7:170125. https://doi.org/10.1098/rsob.170125

[000] Hadfield, J.D., Nutall, A., Ososrio, D., Owens, I.P.F. 2006. Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. Journal of Evolutionary Biology 20:549–557. https://doi.org/10.1111/j.1420-9101.2006.01262.x

[000] Haldane, J.B.S. 1926. On being the right size. Harper's Magazine (March 1926).

[000] Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D. 2006. Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci. 2006 Dec;11(12):587-93. Epub 2006 Nov 7. https://doi.org/10.1016/j.tplants.2006.10.006

[000] Han, B., J.G. Pouget, K. Slowikowski et al. 2015. Using genotype data to distinguish pleiotropy from heterogeneity: deciphering coheritability in autoimmune and neuropsychiatric diseases. bioRxiv online Nov. 6, 2015; doi: http://dx.doi.org/10.1101/030783

[000] Hazel, L.N. 1943 The Genetic Basis for Constructing Selection Indexes. Genetics 28: 476–490.

[000] **Höblinger, A., C. Erdmann, C.P. Strassburg, et al. 2009 Coinheritance of hereditary spherocytosis and reversibility of cirrhosis in a young female patient with hereditary hemochromatosis Eur. J. Med. Res. . 14: 182–184. https://doi.org/10.1186/2047-783X-14-4-182**

[000] Hoskens, H., J. Li, K. Indencleef et al. 2018 Spatially dense 3D facial heritability and modules of co-heritability in a father-offspring design. Front. Genet 9:554. https://doi.org/10.3389/fgene.2018.00554

[000] Hotelling, H. (1953) New light on the correlation coefficient and its transforms. J. Roy. Statist. Soc. Ser. B, 15:193- 225.

[000] Huang W. and Mackay T.F.C. 2016 The genetic architecture of quantitative traits cannot be inferred from variance component analysis. PLoS Genet 12(11): e1006421. https://doi.org/10.1371/journal.pgen.1006421

[000] Hulbert, A.J. 2014. A skepti's view: "Kleiber's Law" or the "3/4 Rule" is neither a law nor a rule but rather an empirical approximation. Systems 2:186-202. https://doi.org/10.3390/systems2020186.

[000] Iacono, W.G. 2018. Endophenotypes in psychiatric disease: prospects and challenges. Genome medicine 10:11. https://doi.org/10.1186/s13073-018-0526-5

[000] Janssens, M.J.J. 1979 Co-heritability: its relation to correlated response, linkage, and pleiotropy in cases of polygenic inheritance. Euphytica 28:601-608. https://doi.org/10.1007/BF00038926

[000] Jia, Y., and J.L. Jannick 2012 Multiple-trait genomic selection methods increase genetic value prediction accuracy. Genetics 192:1513-1522. https://doi.org/10.1534/genetics.112.144246

[000] Kellogg, E.A. 2004 Evolution of developmental traits. Curr. Opin. Plant. Biol. 7:92-98. https://doi.org/10.1016/j.pbi.2003.11.004

[000] Kembel, S. W., T. K. O'Connor, H. K. Arnold, et al. 2014 Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. PNAS 111:13715-13720. https://doi.org/10.1073/pnas.1216057111

[000] Kho, Z.Y. and S. K. Lal, 2018 The Human Gut Microbiome – A Potential Controller of Wellness and Disease. Front. Microbiol. 9:1835. https://doi.org/10.3389/fmicb.2018.01835

[000] Knowles, E.E.M., D.R. McKay, J.W. Kent et al. 2015 Pleiotropic locus for emotion recognition and amygdala volume identified using univariate and bivariate linkage. Am. J. Psychiatry 172:190-199. https://doi.org/10.1176/appi.ajp.2014.14030311

[000] Kruuk, L.E.B. 2003 Estimating genetic parameters in natural populations using the animal model. Phil. Trans. R. Soc. Lond. B 359:873-890. https://doi.org/10.1098/rstb.2003.1437

[000] Kruuk, L.E.B., J. Slate, A.J. Wilson 2008 New Answers for Old Questions: The Evolutionary Quantitative Genetics of Wild Animal Populations. Annual Review of Ecology, Evolution, and Systematics 39:525-548. https://doi.org/10.1146/annurev.ecolsys.39.110707.173542

[000] Lande, R. 1982 A quantitative genetic theory of life history evolution. Ecology 63:607-615. https://doi.org/10.2307/1936778

[000] Lander, E.S. and N.J. Schork 1994 Genetic dissection of complex traits. Science 265:2037-2048. https://doi.org/10.1126/science.8091226

[000] Lee, S.H. ,Yang, J.,Goddard, M.E., Visscher, P.M., Wray, N.R. 2012. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics Applications Note 28(19)2540–2542. doi:10.1093/bioinformatics/bts474.

[000] Lerner, I.M. 1950 Population Genetics and Animal Improvement. Cambridge University Press. P. 97, 233.

[000] Lerner, I.M. and E.R. Demspter 1948 Some aspects of evolutionary theory in the light of recent work on animal breeding. Evolution 2:19-28. https://doi.org/10.1111/j.15585646.1948.tb02728.x

[000] Li, J. and Burmeister, M., 2005 Genetical genomics: combining genetics with gene expression analysis. Human Molecular Genetics 14: R163–R169 https://doi.org/10.1093/hmg/ddi267

[000] Li, Y., X. Wu, T. Chen et al., 2018 Plant phenotypic traits eventually shape its microbiota: a common garden test. Front. Microbiol. 9:2479. https://doi.org/10.3389/fmicb.2018.02479

[000] Lukowski, S.W., Lloyd-Jones, L.R., Holloway, A. 2017. Genetic correlations reveal the shared genetic architecture of transcription in human peripheral blood. Nature Communications 8: 483. https://doi.org/10.1038/s41467-017-00473-z

[000] Mackay, T.F.C., Stone, E.A., Ayroles, J.F. 2009. The genetics of quantitative traits: challenges and prospects. Nat. Rev. Genet, 10:565-577. https://doi.org/10.1038/nrg2612

[000] Martinez-Villalta, J., H. Cochard, M. Mencuccini et al. 2009 Hydraulic adjustments of Scots pine across Europe. New Phytol. 184:353-364. https://doi.org/10.1111/j.1469-8137.2009.02954.x

[000] Manjula,P., Park,H.-B.,Seo, D. Choi,N., Jin, S., Ahn,S.J., Heo,K.N., Kang,B.S., Lee,J.H. 2018. Estimation of heritability and genetic correlation of body weight gain and growth curve parameters in Korean native chicken. Asian-Australas J Anim Sci.31(1): 26–31. https://doi.org/10.5713/ajas.17.0179

[000] Mank, J.E. 2009 Sex chromosomes and the evolution of sexual dimorphisms: lessons from the genome. Am. Nat. 173:141-150. https://doi.org/10.1111/j.1558-5646.1984.tb00346.x

[000] Melton, P.E., S. Rutherford, V.S. Voruganti et al. 2010 Bivariate genetic association of KIAA1797 with heart rate in American Indians: the Strong Heart Family Study. Hum Mol Genet. 2010 Sep 15;19(18):3662-71. https://doi.org/10.1093/hmg/ddq274

[000] Mehr, A. P., M.T. Tran, K. M. Ralto et al. 2018 De novo NAD+ biosynthetic impairement in acute kidney injury in humans. Nat. Med. 24:1351-1359. https://doi.org/10.1038/s41591-018-0138-z

[000] Meyer-Vernet, N. and J.P. Rospars 2015 How fast do living organisms move: maximum speeds from bacteria to elephants and whales. Am. J. Phys. 83:719-722. https://doi.org/10.1119/1.4917310

[000] Montesinos-Lopez, O.A., A. Montesinos-Lopez, J. Crossa, J. et al., 2018. Multi-trait, multi-environment deep learning modeling for genomic-enabled prediction of plant traits. G3:Genes|Genomes|Genetics (Bethesda) 8:3829-3840. https://doi.org/10.1534/g3.118.200728

[000] Morris, A.P., C.M. Lindgren, E. Zeggini et al. 2010 A powerful approach to subphenotype analysis in population-based genetic association studies. Genet. Epidemiol. 34:335-343. https://doi.org/10.1002/gepi.20486

[000] Muñoz, M.L., D. Jaju, S. Voruganti, et al. 2018 Heritability and genetic correlations of heart rate variability at rest and during stress in the Oman Family Study. J. Hypertens. 36:1477-1485. https://doi.org/10.1097/HJH.0000000000001715

[000] Nei, M. 1960. Studies on the application of biometrical genetics to plant breeding. Memoirs of the College of Agriculture Kyoto University 82: 1-100.

[000] Ngo, T.T.M., M.N. Moufarrej, M.L. H. Rasmussen et al. 2018 Non-invasive blood tests for fetal development predict gestational age and preterm delivery. Science 360:1133-1136. https://doi.org/10.1126/science.aar3819

[000] Nihorimbere, V., M. Ongena, M. Smargiassi, and P. Thonart 2011 Beneficial effect of the rhizosphere microbial community for plant growth and health. Biotechnol. Agron. Soc. Environ. 15:327-337.

[000] Niklas, K.J. and U. Kutschera 2015. Kleiber's Law: How the Fire of Life ignited debate, fuel theory, and neglected plants as model organisms. Plant Signalling & Behabior 10:7. https://doi.org/10.1080/15592324.2015.1036216

[000] Olson, E.C. and R.L. Miller 1958 Morphological integration. University of Chicago Press. Chicago, IL.

[000] Oren, Y., A. Nachshon, A. Frishberg, R. Wilentzik, and I. Gat-Viks 2015 Linking traits based on their shared molecular mechanisms eLife 4: e04346. https://doi.org/10.7554/eLife.04346

[000] Org, E., Y. Blum, S. Kasela et al., 2017 Relationship between gut microbiota, plasma metabolite and metabolic syndrome traits in the METSIM cohort. Genome Biol. 18:70 https://doi.org/10.1186/s13059-017-1194-2

[000] Ørsted, M. P.D. Rohdde, A. A. Hoffmann et al. 2017. Environmental variation partitioned into separable heritable components. Evolution 72:136-152. https://doi.org/10.1111/evo.13391

[000] Park, J., D.S. Lee, N.A. Christakis, A.L. Baraba 2009. The impact of cellular networks on disease comorbidity. Molecular Systems Biology 5:262; https://doi.org/10.1038/msb.2009.16

[000] Peaston AE1, Whitelaw E. 2006 Epigenetics and phenotypic variation in mammals. Mamm. Genome 17:365-74. https://doi.org/10.1007/s00335-005-0180-2

[000] Pereira, V., D. Waxman, A. Eyre-Walker 2009 A problem with the correlation coefficient as a measure of gene expression divergence. Genetics 183:1597-1600. https://doi.org/10.1534/genetics.109.110247

[000] Perez, M.F., M. Francesconi, C. Hidalgo-Cacedo, B. Lehner 2017. Maternal age generates phenotypic variation in *Caenorhabditis elegans*. Nature 552:106-109. https://doi.org/10.1038/nature25012

[000] Pick, J.L., Hutter, P., Tschirren, B.T. 2016. In search of genetic constraints limiting the evolution of egg size – direct and correlated responses to artificial selection on a prenatal maternal effector. Heredity 116:542-549. https://doi.org/10.1038/hdy.2016.16

[000] Plomin, R. and J.C. DeFries 1979. Multivariate Behavioral Genetic Analysis of Twin Data on Scholastic Abilities. Behavior Genetics 9(6):505-517. http://dx.doi.org/10.1007/BF01067347

[000] Poissant, J. A.J. Wilson, D.W. Coltman 2010 Sex-specific genetic variance and the evolution of secual dimorphisms: a systematic review of cross-sex correlations. Evolution 64:97-107. https://doi.org/10.1111/j.1558-5646.2009.00793.x

[000] Polly, P.D. 2008 Developmental dynamics and G-matrices: Can morphometric spaces be used to model phenotypic evolution? Evol. Biol. 35:83-96. https://doi.org/10.1007/s11692-008-9020-0

[000] Posthuma, D. , Baaré, W.F.C. , Hulshoff Pol, H.E. , Kahn, R.S. , Boomsma, D.I., De Geus, E.J.C. 2003. Genetic Correlations Between Brain Volumes and the WAIS-III Dimensions of Verbal Comprehension, Working Memory, Perceptual Organization, and Processing Speed. Twin Research 6(2) 131-139.

[000] Ramniwas S., B. Kajla, K. Dev, R. Parkash 2013 Direct and correlated responses to laboratory selection for body melanisation in *Drosophila melanogaster*: support for the melanisation-desiccation resistance hypothesis. J Exp Biol. 216:1244-54. https://doi.org/10.1242/jeb.076166

[000] Rao, D.C., Rice, T. 2014. Path analysis in genetics. In: Wiley StatsRef: Statistics Reference Online. John Wiley and Sons. p. 1-17. https://doi.org/10.1002/9781118445112

[000] Raup, D.M. 1966 Geometric analysis of shell coiling: general problems. J. Paleontol. 40:1178-1190.

[000] Roff, D.A. 1996. The Evolution of Genetic Correlations: An Analysis of Patterns. Evolution 50(4):1392-1403. https://doi.org/10.1111/j.1558-5646.1996.tb03913.x

[000] Rollinson, N., and L. Rowe 2015 Persistent directional selection on body size and a resolution to the paradox of stasis. Evolution 69:2441-2451. https://doi.org/10.1111/evo.12753

[000] Rosado, B.H.P., L.C. Almeida, L.F. Alves et al. 2018 The importance of phyllosphere on plant functional ecology: a phyllo manifesto. New Phytol. 219:1145-1149. https://doi.org/10.1111/nph.15235

1389 **[000] Rubin, H. 2016 The phenotypic gambit: selective pressure and EES methodology in**
1390 **evolutionary game theory. Biol. Philos. 31:551-569. https://doi.org/10.1007/s10539-**
1391 **016-9524-4**

1392 **[000] Sae-Lim, P., H. Mulder, B. Gjerde et al. 2015 Genetics of Growth Reaction Norms in**
1393 **Farmed Rainbow Trout. PLoS ONE 10(8): e0135133.**
1394 **https://doi.org/10.1371/journal.pone.0135133**

1395 **[000] Saltz, J.B., F.C. Hessel, M.W. Kelly 2017 Trait correlations in the genomic era. Trends**
1396 **in Ecol. Evol. 32:279-290. https://doi.org/10.1016/j.tree.2016.12.008**

1397 **[000] Schaid, D.J., X. Tong, B. Larrabee et al. Statistical methods for testing genetic**
1398 **pleiotropy. Genetics 204:483-497. https://doi.org/10.1534/genetics.116.189308**

1399 **[000] Searle, S.R. 1961. Phenotypic, genetic and environmental correlations. Biometrics**
1400 **17(3):474-480.**

1401 **[000] Sgrò, C.M. and A.A. Hoffman 2004 Genetic correlations, tradeoffs and environmental**
1402 **variation. Heredity 93:241-248. https://doi.org/10.1038/sj.hdy.6800532**

1403 **[000] Short, P.J., J.F. McRae, G. Gallone eta l. 2018. De novo mutations in regulatory**
1404 **elements in neurodevelopmental disorders. Nature 555:611-616.**
1405 **https://doi.org/10.1038/nature25983**

1406 **[000] Sikkink, K.L. R.M. Reynolds, W.A. Cresko et al. 2015 Environmentally induced changes**
1407 **in correlated responses to selection reveal variable pleiotropy across complex genetic**
1408 **network. Evolution 69:1128-1142. https://doi.org/10.1111/evo.12651**

1409 [000] **Singh, N.D. and K.L. Shaw 2012 On the scent of pleiotropy. PNAS 109:5-6.**
1410 **https://doi.org/10.1073/pnas.1118531109**

1411 **[000] Sölkner, J., H. Grausgruber, A.M. Oyeko et al. 2008 Breeding objectives and the**
1412 **relative importance of traits in plant and animal breeding: a comparative review.**
1413 **Euphytica 161:273-282. https://doi.org/10.1007/s10681-007-9507-2**

1414 **[000] Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., Smoller, J.W. 2013. Pleiotropy in**
1415 **complex traits: challenges and strategies. Nature Reviews 14:483-495.**
1416 **https://doi.org/10.1038/nrg3461**

1417 **[000] Stearns, S.C. 1989 Trade-offs in life history evolution. Functional Ecology 3:259-268.**
1418 **https://doi.org/10.2307/2389364**

1419 **[000] Stearns, S. G. de Jong, B. Newman 1991 The effects of phenotypic plasticity on**
1420 **genetic correlations. *TREE* 6, 122–6. https://doi.org/10.1016/0169-5347(91)90090-K**

1421 **[000] Stoddard MC, Yong EH, Akkaynak D, Sheard C, Tobias JA, Mahadevan L 2017. Avian**
1422 **egg shape: Form, function, and evolution. Science 356(6344):1249-1254.**
1423 **https://doi.org/10.1126/science.aaj1945**

[000] **Sun J., H.R. Kranzler, J. Bi 2015 An Effective Method to Identify Heritable Components from Multivariate Phenotypes. PLoS ONE 10(12):e0144418. https://doi.org/10.1371/journal.pone.0144418**

[000] **Sun, J., J.E. Rutkoski, J.A. Poland et al. 2017 Multitrait, random regression, or simple repeateability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield. Plant Genome 10 No 2 https://doi.org/10.3835/plantgenome2016.11.0111**

[000] **te Pas, M.F.W., O. Madsen, M.P.L. Calus et al. 2017 The importance of endophenotypes to evaluate the relationship between genotype and external phenotype. Intl. J. Mol. Sci. 18:472. https://doi.org/10.3390/ijms18020472**

[000] **Teich, A.F., Sakurai, M., Patel, M. et al. 2016 PDE5 Exists in Human Neurons and is a Viable Therapeutic Target for Neurologic Disease. Journal of Alzheimer's Disease 52:295–302. https://doi.org/10.3233/JAD-151104**

[000] **Traglia, M., Bseiso, D., Gusev, A., et al. 2017 Genetic Mechanisms Leading to Sex Differences Across Common Diseases and Anthropometric Traits. Genetics 205:979-992. https://doi.org/10.1534/genetics.116.193623**

[000] **Tyler, A.L., F.W. Asselberg, S.M. Williams et al., 2009 Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. Bioessays 31:220-227. https://doi.org/10.1002/bies.200800022**

[000] **Valdes, A.M. J. Walter, E. Segal, T.D. Spector 2018 Role of the gut microbiota in nutrition and health. BMJ 361:k2179. https://doi.org/10.1136/bmj.k2179**

[000] Waitt, D. E., Levin, D.A. 1998. Genetic and phenotypic correlations in plants: a botanical test of Cheverud's conjecture. Heredity 80:310-310. https://doi.org/10.1046/j.1365-2540.1998.00298.x

[000] Wambua, S., J. Mwacharo, S. Uyoga et al. 2006. Co-inheritance of alpha+-thalassaemia and sickle trait results in specific effects on haematological parameters. Br. J. Haematol. 2006 Apr;133(2):206-9. https://doi.org/10.1111/j.1365-2141.2006.06006.x

[000] **Wang, H. G., H. Poon, N. J. Cox, A. Rzhetsky 2017 Classification of common human diseases derived from shared genetic and environmental determinants. Nat Genet. 49: 1319–1325. https://doi.org/10.1038/ng.3931**

[000] Wang K, Gaitsch H, Poon H, Cox NJ, Rzhetsky AQ 2017 Classification of common human diseases derived from shared genetic and environmental determinants. Nature Genetics 49:1319–1325. https://doi.org/10.1038/ng.3931

[000] **Wellman, A., B. A. Edwards, S. A. Sands, R. L. Owens et al. 2013 A simplified method for determining phenotypic traits in patients with obstructive sleep apnea J Appl Physiol 114: 911–922. https://doi.org/10.1152/japplphysiol.00747.2012**

[000] **Willis, J. H., J. A. Coyne, and M. Kirpatrick 1991 Can one predict the evolution of quantitative characters without genetics? Evolution 45:441–444. https://doi.org/10.1111/j.1558-5646.1991.tb04418.x**

[000] Wicklin, R. 2013. Simulating with SAS®. SAS Institute, Cary, NC.

[000] **Willmore, K.E., C.C. Roseman, J. Rogers, J. M. Cheverud, J. T. Richtsmeier 2009. Comparison of Mandibular Phenotypic and Genetic Integration between Baboon and Mouse. Evol Biol. 2009 March ; 36(1): 19–36. https://doi.org/10.1007/s11692-009-9056-9**

[000] **Wissuwa, M., M. Mazzola, C. Piccard 2009 Novel approaches in plant breeding for rhizosphere-related traits. Plant Soil 321:409-430. https://doi.org/10.1007/s11104-008-9693-2.**

[000] **Wu, T., Boezen,H.M., Postma,D.S., Los,H. , Postmus, P.E., Snieder, H., Boomsmae, D.I. 2010. Genetic and environmental influences on objective intermediate asthma phenotypes in Dutch twins. Eur Respir J 2010; 36: 261–268. https://doi.org/10.1183/09031936.00123909**

[000] **Wu, T., Treiber, F.A., H. Snider, H. 2013. Genetic influence on blood pressure and underlying hemodynamics measured at rest and during stress. Psychosom Med. 75(4): 404–412. https://doi.org/10.1097/PSY.0b013e31828d3cb6**

[000] **Xavier, A., B., Hall, S. Castel et al. 2017 Using unsupervised learning techniques to assess interactions among complex traits in soybeans. Euphytica 213:200. https://doi.org/10.1007/s10681-017-1975-4**

[000] **Zera, A.J. and L.G. Harshman 2001 The physiology of life history trade-offs in animals. Ann. Rev. Ecol. Syst. 32:95-126. https://doi.org/10.1146/annurev.ecolsys.32.081501.114006**

[000] **Yamada, Y. 1968. On the Realized heritability and genetic correlation estimated from double selection experiments when two characters are measured Japan. Poultry Sci. 5(3):148-151.**

[000] Yang, C., C. Li, Q. Wang et al. 2015 Implications of pleiotropy: challenges and opportunities for mining Big Data in biomedicine. Front. Genet. 6:229 . https://doi.org/10.3389/fgene.2015.00229

[000] **Yang, C., X. Wang, J. Liu, M. Ng 2016. Introduction to statistical methods for integrative data analysis in genome-wide association studies. In: K.-C. Wong (ed) Big Data Analytics in Genomics. Springer. http://doi.org/10.1007/978-3-319-41279-5_1**

[000] **Yang, J., Zeng, J., M.E. Goddard, N.R. Wray, P.M. Visscher 2017 Concepts, estimation and interpretation of SNP-based heritability. Nature Genetics 49:1304–1310. http://doi.org/10.1038/ng.3941**

[000] **Yin, P., Anttila, V., Siewert, K.M., Palotie, A., Smith, G.D., Voight, B.F. 2017. Serum calcium and risk of migraine: a Mendelian randomization study. Human Molecular Genetics 26(4)820–828. https://doi.org/10.1093/hmg/ddw416**

[000] **Zhang, Y.Y., Z.Y. Fu, J. Wei et al. 2018 A *LIMA1* variant promotes low plasma LDL cholesterol and decreases intestinal cholesterol absorption. Science 360:1087-1092. https://doi.org/10.1126/science.aao6575**

[000] Zhang, Y., Z. Liang, H.-T. Hsueh, V. A. et al . 2012 Characterization of Streptokinases from Group A Streptococci Reveals a Strong Functional Relationship That Supports the Coinheritance of Plasminogen-binding M Protein and Cluster 2b Streptokinase. J. Biol. Chem. 287: 42093–42103. https://doi.org/10.1074/jbc.M112.417808