**Multiple traits: processes influencing phenotypic correlations in repeatedly expressed traits.**

***Goal:***to develop understanding of how phenotypic correlations between two repeatedly expressed traits are affected by the amount of variation, and the magnitude of correlations occurring at each underlying hierarchical level.

***Step 1. Phenotypic correlations between two repeatedly expressed traits that are not consistently different between individuals.***

**Sub-goal: Introducing within-individual correlations and variance-covariance matrices**

**Introduction:** As we have seen in other modules (Basic Lessons about Variance, Random Regressions), repeatedly expressed traits can vary both within and among individuals. In this module, we apply the same logic to correlations between traits. We do this by introducing the notion that correlations also exist at multiple hierarchical levels. Importantly, we explain how the correlations existing at each level influence the overall phenotypic correlation in the data, and how their influence is mediated by the amount of variation occurring at each level. We will introduce these ideas one step at the time; we will start with focussing on two repeatedly expressed traits that do not vary among individuals. In other words, those two traits only harbour within-individual variation.

As a worked example, we consider two traits that we expect to be correlated because of trade-offs. For example, trait *y* may represent egg size while trait *z* represents egg number (i.e., clutch size). We expect trade-offs in the investment in these two traits because females producing large eggs should typically have fewer resources available for producing a large clutch. This means that we expect a negative correlation within the same individual parent between egg size (*y*) and egg number (*z*).

**Exercise:** We will generate multi-level data for two repeatedly expressed traits. We will consider a scenario where the two traits are assayed simultaneously for a set of individuals. All individuals are sampled repeatedly but at the same time. We are assuming that you will sample 10 individuals with 10 samples per individual throughout. In this exercise, we use the following models:

Set the amount of within-individual variance in each of the two traits (and ):

|  |  |
| --- | --- |
| Within-individual variance Trait Y | Within-individual variance Trait Z |
| [Enter] | [Enter] |

Now set the within-individual correlation. Remember, for this example this correlation should be negative as we are considering trade-offs between egg size (y) and egg number (z).

[Enter]

**Results:**

Above, you have defined the elements of the within-individual “variance-covariance matrix”. This matrix holds the variances of the two traits on the diagonals and their covariance on the off-diagonals:

Here, is the within-individual variance in trait *y*, the within-individual variance in trait *z*, and the within-individual covariance between these two traits.

The values that you have entered above result in the following variance covariance matrix:

Notably, you did not enter the covariance () between the two traits but rather their correlation. As you will learn later on in this module, covariances are important for various calculations (see below), which is why we introduce them here. The correlation is the covariance between the two traits when the associated variances are expressed in standard deviation units. This therefore results in a standardized metric that can be compared across samples differing in variance.

The relationship between the within-individual correlation () and covariance () is:

Here is a plot of the simulated data showing the within-individual correlation, where each individual is given a different colour:

[PLOT 1]

Note that the relationship between the two traits should be negative if you did enter a negative correlation above, and that this relationship exists within individuals. That is, when a female (colour) increases egg size from one instance (e.g. year) to the next, she simultaneously decreases her egg number. In other words, the changes in one trait across repeated expressions of the same individual are associated with changes in another trait.

**Conclusion:** Variance-covariance matrices hold information on the amount of variance existing in each trait as well as the covariances between them. The correlations between two traits are calculated using this information on variances and covariances as derived metrics.

***Step 2. Phenotypic correlations between two repeatedly expressed traits that also differ consistently between individuals.***

**Sub-goal: Understanding the role of repeatability in shaping phenotypic correlations in repeatedly expressed traits**

**Introduction:** In step one, we focussed on within-individual correlations caused by trade-offs between two costly traits that were repeatedly expressed within the same individual. We assumed for simplicity that both traits harboured no among-individual variance. This meant that the individual repeatability was zero for both traits (see module on Basic Lessons for an introduction to repeatability). We will now consider a more complex but realistic scenario where these two traits (egg size and egg number) varied both within- and among-individuals, as we know is the case in natural populations. The models for this step will be:

For example, female passerine birds often adjust their clutch size to changes in breeding density while females also differ consistently in clutch sizes across repeated breeding attempts (Nicolaus et al. 2013). Previously, because the traits did not harbour any among-individual variation, all variation in the data existed at within-individuals. This meant that the phenotypic correlation between the two traits in the dataset was identical to the within-individual correlation. We can see why this is the case by introducing the equation that describes the components affecting the phenotypic correlation () (e.g. Dingemanse & Dochtermann 2013; Dingemanse et al. 2012):

eq1

This monster contains several important elements. The phenotypic correlation () is affected by both the among-individual correlation () and the within-individual correlation (). This former correlation will be addressed in more detail in step 3 of this module. Important for now is to note that the impact of each correlation on the phenotypic correlation is weighted by the following term:

eq 2

This term is called the “geometric mean repeatability” as it represents the square-root of product of the repeatability of trait *y* ()and trait *z* () Note that the term essentially represents one minus this geometric repeatability and that both terms add up to the value of 1. This means that for two traits that are highly repeatable (e.g. sets of morphological traits), the phenotypic correlation will largely reflect the among-individual correlation, whereas for two traits that have a low repeatability (e.g. sets of behavioural traits), the phenotypic correlation instead largely reflects the within-individual correlation (Dingemanse & Dochtermann 2013).

Let us now return to the example dataset that we considered in Step 1, where we focussed on two repeatedly expressed traits that were negatively correlated within-individuals but did not harbour any among-individual variance.

We reprint the equation describing the components shaping the phenotypic correlation:

We have automatically taken forward your inputs from step 1, where you have set the within-individual variances for traits *y* and *z* as well as their covariance but where you assumed that the among-individual variances (, ) were equal to zero:

What you should note is that the geometric mean repeatability (defined above) is equal to zero, while one minus this metric is equal to one, in cases such as these where at least one of the traits harbours no among-individual variance; therefore:

**Exercise:** We will now introduce among-individual variance in both egg size (i.e., >0) and clutch size (i.e., >0). For simplicity, we will assume that an individual’s average egg size over all its repeated observations is not correlated with its average clutch size over all its repeated observations, i.e., that the among-individual correlation is zero.

As in step 1 above, we have set both the number of individuals, and the number of repeated measures per individual, to 10.

Set the amount of within-individual variance in each of the two traits:

|  |  |
| --- | --- |
| Within-individual variance Trait Y | Within-individual variance Trait Z |
| [Enter] | [Enter] |

Set the within-individual correlation, which should be negative in case females trade-off their investment in egg size versus number:

|  |
| --- |
| Within-individual correlation Trait Y, Z |
| [Enter] |

Previously, we assumed that the among-individual variances in *y* and *z* were zero, now set them to some number:

|  |  |
| --- | --- |
| Among-individual variance Trait Y | Among-individual variance Trait Z |
| [Enter] | [Enter] |

Remember, we assume for now that the among-individual correlation is zero.

|  |
| --- |
| Among-individual correlation Trait Y, Z |
| 0 |

**Results:**

As we have seen in step 1, you have defined the elements of the within-individual “variance-covariance matrix”:

Your entries resulted in the following values for this matrix

You have now also defined the values for the among-individual matrix:

Your entries resulted in the following values for this matrix

As you have set both the among- and within-individual variances, you have defined the repeatabilities of both traits. As detailed in module on Basic Lessons, repeatability ranges between 0 and 1 and is calculated as:

The repeatabilities of your two traits are:

|  |  |
| --- | --- |
| Repeatability Trait Y | Repeatability Trait Z |
| [Enter] | [Enter] |

We will now return to the equation that describes the components that shape the phenotypic correlation:

Your entries resulted in the following values:

Note that the phenotypic correlation () no longer matches the within-individual correlation (). This is because both traits now harbour non-zero repeatability; as we have set the among-individual correlation () to zero, in this specific scenario, we can simplify the equation into:

Doing so clarifies that the geometric mean repeatability of the two traits defines how well the phenotypic correlation in the data corresponds to the within-individual correlation occurring because of the presumed trade-off between investments in egg size vs. number. The equation also clarifies that the more repeatable the two traits are, the less apparent is the trade-off if one would focus solely on the overall phenotypic correlation.

This insight can be visualized by presenting your data in two scatter plots.

The first scatter plot shows the correlation between the two traits that your entries (of within- and among-individual variances and covariances) above produced:

[GRAPH]

The second scatter plot also uses your entries but assumes that you entered zero among-individual variances. This second plot therefore depicts the correlation in your data at the within-individual level alone:

[GRAPH]

The first graph should show a weaker overall phenotypic association as explained above.

**Conclusion**: The phenotypic correlation reflects the within-individual correlation when measured for repeatedly expressed traits; however, their correspondence is a function of how repeatable the two traits are. You may explore this idea more fully by re-running the exercise for different levels of trait repeatabilities.

***Step 3. Level-specific correlations.***

**Sub-goal:** Understanding the role of level-specific correlations in shaping phenotypic correlations in repeatedly expressed traits

**Introduction:** As a final step in our analysis of the within-individual trade-off between egg and clutch size, we will introduce another level of complexity. We will consider the idea that correlations often differ between hierarchical levels because of multiple mechanisms shaping phenotypic correlations. We will continue our example where individuals have to allocate limited resources into multiple costly traits, causing negative within-individual correlations due to trade-offs. We will now introduce the notion that there is a second mechanism causing covariance, namely among-individual variation in the acquisition of resources. This may lead to a situation where individuals with many resources are able to produce both large eggs and large clutches whereas individuals with fewer resources produce small eggs and small clutches instead. Variation in acquisition combined with allocation trade-offs therefore will lead to level-specific associations (van Noordwijk & de Jong 1986). In statistical terms, an individual’s average egg size will correlate positively with its average clutch size over all its expressions. Essentially, what we have now set up is a situation where the correlations between egg size and number are opposite across hierarchical levels.

**Exercise:** Set the within- and among-individual variances for the two traits. We suggest that you create a dataset where the associated repeatabilities are 0.5 as this will result in a situation where the phenotypic correlation is affected equally by correlations at each of the two levels. You will also have to set the within- and among-individual correlation. We suggest that you start with setting them to the same value but with opposite signs (e.g. within: r=-0.5 vs. between r=+0.5). As in step 1, we have already set both the number of individuals, and the number of repeated measures per individual, to 10.

Set the amount of within-individual variance in each of the two traits:

|  |  |
| --- | --- |
| Within-individual variance Trait Y | Within-individual variance Trait Z |
| [Enter] | [Enter] |

Set the within-individual correlation, which should be negative in case female trade-off their investment in egg size versus number:

|  |
| --- |
| Within-individual correlation Trait Y, Z |
| [Enter] |

Set the amount of among-individual variance in each of the two traits:

|  |  |
| --- | --- |
| Among-individual variance Trait Y | Among-individual variance Trait Z |
| [Enter] | [Enter] |

Set the among-individual correlation, which should be positive if females with more resources can produce both large eggs and large clutches.

|  |
| --- |
| Among-individual correlation Trait Y, Z |
| [Enter] |

**Results:**

You have defined the elements of the within-individual “variance-covariance matrix”:

Your entries resulted in the following values for this matrix:

You have now also defined the elements of the among-individual matrix:

Your entries resulted in the following values for this matrix

As you have set both the among- and within-individual variances, you have set the repeatabilities of both traits. Those are:

|  |  |
| --- | --- |
| Repeatability Trait Y | Repeatability Trait Z |
| [Enter] | [Enter] |

We will now again return to the equation describing the components affecting the phenotypic correlation:

Your entries resulted in the following values:

Note that if you have entered different values for the within-individual correlation () versus the among-individual correlation (), the overall phenotypic correlation is not accurately reflecting either one. Only when you would have entered the same value and the same sign for both the within- and among-individual correlation, would the overall phenotypic correlation be the same. So do try different values for the two types of correlations and check how this affects the overall phenotypic correlation in your data.

Another issue that you may investigate by changing the entry values above is the role of repeatability. You may for example choose within- and among-individual correlations of equal strength but opposite sign (e.g. -0.8 vs. 0.8) for a range of different repeatabilities. You could also play with creating a dataset where one trait has a high and another trait has a low repeatability. You will find that when both traits are highly repeatable, the phenotypic correlation will more closely reflect the among-individual correlation. By contrast, if both traits have a low repeatability, the phenotypic correlation will more closely reflect the within-individual correlations.

But how can we visualize the patterns of correlation within vs. among individuals? We can do this by producing three types of scatter plots.

First, we can simply plot the raw data in a scatter plot; this plot represents the overall phenotypic association between the two traits:

[GRAPH]

Second, we can calculate each individual’s mean value for each of the two traits, and plot these two values in a scatter plot. This plot of individual-mean values represents a visual of the among-individual correlation:

[GRAPH]

Third, we can calculate how each observation deviates from an individual’s mean value for each of the two traits, and plot these two values in a scatter plot. This plot of within-individual deviations from individual-means represents a visual of the within-individual correlation:

[GRAPH]

**Conclusion**: Repeatedly expressed traits often vary across multiple hierarchical levels, such as within and among individuals. This means that the phenotypic correlation summarizes the correlations existing at different levels but that their respective influences are weighted by the amount of variance existing at each level. In other words, the correlation between two traits that have low repeatabilities (e.g. physiology, behaviour) will largely represent within-individual correlations, whereas correlations between traits that have high repeatabilities (e.g. morphology) will largely represent among-individual correlations. For this reason it is important to partition the phenotypic correlation into its underlying components when one is interested in level-specific patterns of correlations, such as correlations within individuals due to trade-offs.

**Step 4. The Bivariate mixed-effects model**

**Sub-goal:** Understanding the difference between a univariate and a multivariate mixed-effects model, and how level-specific correlations are modelled.

**Introduction:** In the previous modules (Basic Lessons, Non-stochastic Environments, and Random Regression) we have considered univariate mixed-effects models. As you have seen, the univariate mixed-effects model enabled us to estimate the variance attributable to variation within- and among-individuals in a single trait (*y*) with the following equation:

The variation in intercepts () among individuals was assumed to be normally distributed () with a mean of zero and a variance () and is called the *among-individual variance* (estimated as : the variance across random intercepts of individuals) :

:

A residual error () was also assumed to be normally distributed, with zero mean and a variance () representing the *within-individual variance*:

:

In the bivariate mixed-effects models, we are estimating these parameters simultaneously for two traits. That is, the model can be formulated as a set of two phenotypic equations (one for *y* and one for *z*):

As was the case for univariate models, the random intercepts () and the within-individual contributions () to *y* and *z* are modelled as having means of zero. However, in this bivariate case, neither the random intercepts nor the residual errors are independent. Instead, the random intercepts are now distributed assuming a multivariate normal distribution with a variance-covariance structure () specifying the among-individual variances ( and ) and the among-individual covariance between the two attributes () :

The residual errors () are likewise assumed to be drawn from a multivariate normal distribution, with means of zero, within-individual variances ( and ), and within-individual covariances ():

From these estimated matrices, we can calculate the phenotypic variances for each trait by adding up the variances estimated at each level:

In the same fashion, we can calculate the phenotypic covariance between the two traits by adding up the covariances estimated at each level:

With this information in hand, we can now calculate the overall phenotypic correlation in the data.

**Conclusion**: Bivariate mixed-effects models differ distinctly from univariate mixed-effects models as the former assumes multivariate normality while the latter assumes univariate normality. Bivariate mixed-effects models estimate variances and covariances within and among each specified level from which overall phenotypic variances and covariances, as well as correlation, can be subsequently derived.

Step 5. **Environmentally-induced correlations.**

**Sub-goal: Introducing integration of plasticity as a mechanism causing correlations.**

**Introduction:** In the previous steps we have, considered a bivariate phenotypic equation where two traits correlated among- and/or within-individuals without considering plasticity as a mechanism causing this variation. This is why the double equation that we re-print here harboured no fixed effects apart from the statistical intercepts:

In this final step, we will investigate the role of within-individual plasticity in generating correlations among repeatedly expressed traits. Returning to our example of egg size and egg number, we can imagine that individuals breed repeatedly across a set of years, sometimes encountering “good” and sometimes encountering “bad” years. We might expect females to plastically adjust both their eggs size and their clutch size to food availability, where we expect females to down-regulate both traits when encountering reduced food availability. We can model this situation by including food availability () as a fixed-effect covariate affecting both egg size () and egg number ():

To keep matters simple, we will assume in this example that the two traits are neither correlated within nor correlated among individuals due to other mechanisms than variation in food availability. At the same time, we are assuming that both traits harbour both among- and within-individual variation.

**Exercise:** We will generate a dataset for 100 individuals each sampled 10 times. Set the within- and among-individual variances for the two traits. We suggest that you create a dataset where the associated repeatabilities are 0.5. You will also have to set the within- and among-individual correlations. For the purpose of this example, we suggest that you set their values to zero. We will assume that the environment is changing stochastically between the 10 time steps, and that the average phenotype in the population has a value of zero.

Set the amount of within-individual variance in each of the two traits:

|  |  |
| --- | --- |
| Within-individual variance Trait Y | Within-individual variance Trait Z |
| [Enter] | [Enter] |

Set the within-individual correlation:

|  |
| --- |
| Within-individual correlation Trait Y, Z |
| [Enter] |

Set the amount of among-individual variance in each of the two traits:

|  |  |
| --- | --- |
| Among-individual variance Trait Y | Among-individual variance Trait Z |
| [Enter] | [Enter] |

Set the among-individual correlation:

|  |
| --- |
| Among-individual correlation Trait Y, Z |
| [Enter] |

Set population-level slope (i.e., level of plasticity to food availability) for each trait:

|  |  |
| --- | --- |
| Slope () | Slope () |
| [Enter] | [Enter] |

You have set the following bivariate phenotypic equation:

Your entries resulted in the following values:

You have also defined the elements of the among-individual and residual within-individual variance-covariance matrices:

Your entries resulted in the following values for these matrices:

As we have seen in previous steps, we can visualize the patterns at the overall phenotypic level, within individuals, and among individuals.

First, we plot the raw data in a scatter plot; this plot represents the overall phenotypic association between the two traits:

[GRAPH]

Second, we calculated each individual’s mean value for each of the two traits, and plotted these two values in a scatter plot. This plot of individual-mean values represents a visual of the among-individual correlation:

[GRAPH]

Third, we calculated how each observation deviates from an individual’s mean value for each of the two traits, and plot these two values in a scatter plot. This plot of within-individual deviations from individual-means represents a visual of the within-individual correlation:

[GRAPH]

Here is also the plot showing how the environment is changing over time:

[GRAPH]

If you chose a within- and among-individual correlation equal to zero and also set non-zero slopes for x for both traits, you will see that the data show a within-individual correlation. The correlation emerges at this level because we set the simulation to produce a situation where all individuals were sampled equally often and at the same time (i.e., there was no variation in the timing of sampling). This meant that the environment (food availability) varied within but not among individuals, and therefore caused correlations between egg size and number within but not among individuals either.

We will now run a bivariate mixed-effects model on these data, where we will fit the following phenotypic equation with the following random effects structures:

Here we print the point estimates derived from a Bayesian-analysis of these data (true values are between brackets):

With these values, we can calculate the point estimates for the level-specific correlations:

|  |  |
| --- | --- |
| Among-individual correlation | Residual within-individual correlation |
| [calculate] | [calculate] |

Importantly, the estimates printed above are from a model where the effect of food availability (x) was not modelled. Given the structure of the data, where food availability affected both traits but only varied within individuals, the residual within-individual correlation will thus be nonzero even if you modelled it to be zero.

We will now run an updated bivariate mixed-effects model on these data, where we will fit same model but now we do model the effect of food availability (x) on egg size and number:

Here we print the point estimates derived from a Bayesian-analysis of these data (true values are between brackets):

With these values, we can calculate the point estimates for the level-specific correlations:

|  |  |
| --- | --- |
| Among-individual correlation | Residual within-individual correlation |
| [calculate] | [calculate] |

This new model should properly recover the effect of x on y and z. You will notice that the residual within-individual covariance () should now better resemble the estimate that you should have expected based on the within-individual correlation that you entered previously. This is because the within-individual covariance is now explained by the fixed effect covariate, and therefore food availability no longer contributes to the residual within-individual variances.

**Conclusion:** Integration of plasticity, where the same environmental factor affects multiple phenotypic traits, leads to patterns of trait correlations in phenotypic data. Provided that such environmental factors varied solely within individuals, environmental variation would cause residual within-individual correlations if not accounted for. Though not further articulated here, we can imagine that such types of environmental effects could also cause among-individual correlations in situations where the sampling design was such that it resulted in among-individual variation in the timing of sampling. In such situations, among- and within-individual correlations would occur in the presence of integration of plasticity, provided that its effects were not modelled in the fixed-effects structure of the statistical model. We will explore this issue and others in another module (Sampling and Bivariate Models).