Multivariate comparison of variance in R

Anne Le Maître1,2 | Philipp Mitteroecker1

Abstract

1. Comparisons of pattern and magnitude of phenotypic variation are central to many studies in evolution and ecology, but a meaningful comparison of multivariate variance patterns can be challenging. Here, we review an effective exploratory strategy, relative principal component analysis (relative PCA), for the comparison of variance–covariance matrices based on their relative eigenvalues and eigenvectors.

2. Relative PCA allows for the identification of multivariate traits (linear combinations of variables) with maximal or minimal variance ratios between two groups. It can be used to explore the generation and canalization of phenotypic variance throughout ontogeny and phylogeny. Relative PCA also gives rise to a natural metric for the ordination of a sample of covariance matrices. We present a novel biometric justification of these approaches and discuss numerical difficulties as well as strategies for statistical inference, along with a new implementation of these methods in R.

3. In an application of relative PCA to geometric morphometric data on cichlid body shape, we found that the phenotypic variance–covariance structure differs between males and females as well as between allopatric and sympatric populations of *Tropheus moorii*. Divergent selection in these populations mainly affected shape features related to swimming ability, whereas tropic morphology appears to be under stabilizing selection.

4. In biology and biomedicine, individual variation is a key signal. Standard ordination methods as well as scalar summary statistics of multivariate variation often pool different biological factors with heterogeneous variational dynamics and thus conceal differences in variance–covariance pattern among groups. Relative PCA complements existing tests for the proportionality of covariance matrices and is an effective exploratory method to identify multivariate traits that differ in variance between populations, age groups, or treatment groups. By comparing within- and between-population covariance matrices, relative PCA can reveal traits under divergent or stabilizing selection.

KEYWORDS
canalization, multivariate biostatistics, phenotypic variation, relative eigenanalysis, relative principal component analysis, selection, *Tropheus*, variance–covariance matrix
1 | INTRODUCTION

Multivariate studies of phenotypic variation are central to evolutionary biology and ecology. Phenotypic variation – the substrate of natural selection – is affected by genetic and environmental heterogeneity as well as by the organisms’ ability to canalize development in the face of environmental and genetic perturbations (Gibson & Wagner, 2000; Hallgrímsson & Hall, 2005; Hermisson & Wagner, 2004; Mitteroecker, 2009; Pélabon, Hansen, Carter & Houle, 2010). However, ecological studies often measure the overall amount of variation (‘disparity’) (Charitier et al., 2014; Foote, 1993; Zelditch, Sheets & Fink, 2003), many evolutionary studies contrast multivariate patterns of within- and between-population variation in order to infer scenarios of natural selection (Lande, 1979; Marroig & Cheverud, 2001). At the interface of evolutionary and developmental biology, it has been studied how phenotypic variation emerges and canalizes throughout development, and how the pattern of covariances among traits – often referred to as morphological integration – affects evolutionary change (e.g. Armbruster, Pélabon, Bolstad & Hansen, 2014; Cheverud, 1982; Klingenberg, Badyaev, Sovry & Beckwith, 2001; Mitteroecker, Gunz, Neubauer & Müller, 2012; Walter, Aguirre, Blows & Ortiz-Barrientos, 2018; Zelditch, Lundrigan & Garland, 2004).

A biologically meaningful analysis of multivariate variance patterns is much more challenging than the analysis of averages. Standard multivariate tools, such as principal component analysis, do not necessarily permit biological insights into the phenomena producing or canalizing phenotypic variation (see below and Figure 1). Scalar summary statistics of the magnitude of multivariate variation within a group (or of the difference in multivariate variation between groups) lump different factors with different biological dynamics; they rarely are biologically interpretable (Bookstein & Mitteroecker, 2014). For example, two groups, A, B, may have the same amount of total variance (sum of the variances for all variables), but the variance of the first variable may be higher in group A than in group B, and vice versa for the second variable. The finding of equal total variance would thus be misleading and conceal the different variational properties of the variables.

Here, we review an effective exploratory strategy, relative principal component analysis (relPCA), for the comparison of variance–covariance patterns based on their relative eigenvalues and eigenvectors (Bookstein & Mitteroecker, 2014; Flury, 1985). We present a new implementation of these methods in R, with an application to geometric morphometric data on cichlid body shape.

2 | MATERIALS AND METHODS

For a single continuous quantity x, the variation – or spread of the distribution – is commonly quantified by the sample variance

$$\text{Var}(x) = (n - 1)^{-1} \sum (x_i - \bar{x})^2,$$

the average squared deviation from the sample mean. Of course, two distributions can also differ in other properties than the mean and the variance: a distribution can be asymmetric (skewed), multimodal, leptokurtic, platykurtic (fat or thin tails), can have outliers, etc. Meaningful comparison of variation across two or more samples in terms of the variance (Equation 1) requires that the distributions differ only (or at least primarily) in mean and variance, while other properties of the distributions are the same.

A comparison of variance across groups also requires that the variance can – in principle – change independently of the mean (or that the groups do not differ in mean values). In biology, however, the standard deviation (square root of the variance) of many size variables scales with the mean of that variable (large structures vary more than small ones), so that the variance needs to be interpreted relative to the mean. Such variables are thus often log-transformed because the resulting variance is unaffected by linear scaling. Other researchers prefer to divide the variable through its mean, which leads to the coefficient of variation (standard deviation expressed as a multiple of the mean) or its square (Houle, Pélabon, Wagner & Hansen, 2011). For small variations relative to the mean value, the mean-standardized variable is approximately proportional to the natural logarithm of this variable and lead to the same variance estimate. For larger ranges of variation, however, nonlinear transformations, such as the logarithm, severely distort the data space and affect most statistics (Huttegger & Mitteroecker, 2011). Ultimately, it depends on the biological and biometric context whether an association of mean and variance is considered an empirical result or an artefact that requires standardization of the variable.

Furthermore, not every measured variable has a biologically meaningful variance, i.e. it may not estimate an interpretable population property. For instance, consider a study of swimming speed in differently sized fish. If the fish were sampled randomly from a population, the sample variance in the length of the fish is an estimate of the population variance in length, which may have a biological interpretation. If the study design was based on an equal number of fish in, say, five different predefined size classes, the variance in fish length cannot be interpreted; there is no corresponding population property. Of course, the variance can still be computed, and a regression slope of swimming speed on fish length (which equals their covariance divided through the variance in fish length) may be interpretable, but the variance of fish length per se is not. Even if individuals are sampled randomly, variances may not necessarily be directly interpretable. For instance, if we sample fish of different ages, the variance of body length at a given age (a conditional variance) may be interpretable, but the variance of the full sample is not (because age may not have an interpretable variance). The same applies to samples of fossil specimens of different geological age.

Clearly, for a variance to be interpretable, the variable must also have at least an interval scale, enabling the meaningful computation of differences (Houle et al., 2011). On the other hand, a ratio scale, comprising a meaningful origin (zero value) of the variable, is not necessary for computing the variance. But in some situations the
average deviation of individuals from the origin can be more insightful than the deviation from their mean. For instance, leg length has a ratio scale (no leg can be shorter than zero), but this origin is far outside of the actual sample distribution. Variation in leg length thus is measured by the variation around the sample mean (Equation 1). But the difference in length between left and right leg, as a measure of asymmetry, has a meaningful origin, namely perfect symmetry, which is close to or within the sample distribution. Individual variation of leg asymmetry is measured best just by the mean sum of squares (i.e. the variance around the origin); the sample variance of leg asymmetry, often referred to as fluctuating asymmetry, neglects the average magnitude of asymmetry in the sample (directional asymmetry).

Assuming that the variation in two different groups A and B can be meaningfully described by the two sample variances, it is useful to relate them by the ratio of variances, \( \frac{\text{Var} \left( x_A \right)}{\text{Var} \left( x_B \right)} \), as opposed to the difference of variances, \( \text{Var} \left( x_A \right) - \text{Var} \left( x_B \right) \). One property of the ratio of variances is its invariance to linear scaling of the variable, such as a change of unit. For instance, let us multiply \( x \) by the scalar \( a \), then

\[
\frac{\text{Var} \left( ax_A \right)}{\text{Var} \left( ax_B \right)} = \frac{a^2 \text{Var} \left( x_A \right)}{a^2 \text{Var} \left( x_B \right)} = \frac{\text{Var} \left( x_A \right)}{\text{Var} \left( x_B \right)}
\]

This invariance has convenient mathematical properties, which are exploited, for instance, by the F-test. Furthermore, the logarithm (to any base) of a variance ratio equals the difference of the log variances. Exchanging the two groups in the numerator and denominator of the log variance ratio thus only changes the sign of the result:

\[
\log \frac{\text{Var} \left( x_A \right)}{\text{Var} \left( x_B \right)} = \log \text{Var} \left( x_A \right) - \log \text{Var} \left( x_B \right)
\]

\[
= - \left( \log \text{Var} \left( x_B \right) - \log \text{Var} \left( x_A \right) \right)
\]

More important in multivariate biometrics, variance ratios can be compared across different variables even if they differ in unit. Consider, for instance, male and female variance of body height and body mass. The sex differences in variance depend on the unit of measurement, which cannot be meaningfully compared between body height and mass. Yet the ratios of variance are unit-free and scale-invariant; a statement such as “males are 10% more variable than females for body height and 20% more variable for body mass” does not depend on the choice of unit and can be biologically interpreted (Bookstein, 2014; Houle et al., 2011; Huttegger & Mitteroecker, 2011). Even if all variables share the same unit, they
may not be considered equally ‘important’ in a specific biological context, especially if different signals in the data (e.g. different functional anatomical units) are covered by different numbers of variables. For a meaningful interpretation of a multivariate summary statistic, one may thus wish to weight these variables differently. However, variance ratios are unaffected by any such linear weighting, differences of variances imply an (perhaps unwarranted) equal weighting of all variables (Huttiger & Mitteroecker, 2011).

The multivariate biometrics of variance should thus be based on variance ratios rather than differences between variances. However, many geometric concepts in multivariate statistics employ the notion of a ‘distance’ that gives rise to a metric data space or parameter space (Mitteroecker & Huttiger, 2009; Stadler, Stadler, Wagner & Fontana, 2001). The most common metric employed is the Euclidean distance, the square root of the summed squared differences for all measured variables. As the log variance ratio equals the difference of the logged variances, we may thus compare multiple groups on the basis of their log variances. In a space of log variances, the Euclidean distances between groups are scale-invariant, giving rise to a natural multivariate metric for the multivariate biometrics of variance ratios.

It is a fundamental theorem that every set of variables can be transformed linearly into another set of variables that are all uncorrelated (Bookstein, 1971; Bibby, 1979; Mitteroecker & Bookstein, 2011). On the basis of these principal components, the variance pattern can be described sufficiently by the p variances for these components; all covariances between them are zero. The orthogonality of the PC axes implies that, geometrically, principal component analysis is a mere rotation of the data space and thus leaves the variance pattern unchanged.

Although less familiar than principal component analysis, it is also possible to transform a set of variables into a set of linear combinations that are uncorrelated separately within two groups, the so-called relative principal components (relative PCs; Bookstein & Mitteroecker, 2014; Flury, 1985). Hence, on the basis of these relative PCs, the difference in multivariate variance–covariance pattern between two samples can be described sufficiently by the corresponding p variance ratios. In contrast to ordinary PCA, however, relative principal component analysis is a non-orthogonal transformation that alters each variance but leaves the variance ratios unchanged (as they are invariant to linear transformation; see Equation 2).

Let \( X_A, X_B \) be the mean-centred data matrices of groups A and B with dimensions \( n_A \times p \) and \( n_B \times p \), respectively. The \( p \times p \) covariance matrix of group A, \( S_A = (n-1)^{-1}X_A'X_A \), can be decomposed into a diagonal matrix of eigenvalues and an orthonormal matrix of eigenvectors:

\[
S_A = E \Lambda E^T.
\]

The eigenvectors, \( e_i \), are the principal component axes, and the principal component scores are given by the orthogonal projection of the data onto these vectors, \( P = XE \), with covariance matrix \((n - 1)^{-1}P'P = \Lambda \). The eigenvalues, \( \lambda_i \), are in descending order, that is, the first PC has the maximal variance, the second PC the second highest, and so on. For instance,

\[
\lambda_1 = \max_a \text{Var}(X_Aa) = \text{Var}(X_Ae_1).
\]

under the constraint that \( a' a = 1 \).

The principal component axes of group A relative to group B are given by the eigenvectors of the \( p \times p \) matrix \( S_B^{-1}S_A \), which equals the inverse of the eigenvalues.

\[
S_B^{-1}S_A = U \Phi U^T.
\]

Because the matrix product \( S_B^{-1}S_A \) usually is not symmetric, the eigenvectors, \( u_i \), are not orthogonal, but the relative principal component scores, \( R_A = X_AU \) and \( R_B = X_BU \), are uncorrelated within both groups:

\[
\text{Cor}(r_{Aa}, r_{Bb}) = \text{Cor}(r_{Aa}, r_{Bb}) = 0 \quad \text{for all } i, j
\]

(Example 2). The relative eigenvalues, \( \phi_i \), are the variance ratios between groups A and B for the ith relative PC, again in descending order. In other words, the first relative principal component is the linear combination for which group A has the maximal variance relative to group B:

\[
\phi_1 = \max_a \frac{\text{Var}(X_Aa)}{\text{Var}(X_Bu_1)} = \frac{\text{Var}(X_Au_1)}{\text{Var}(X_Bu_1)}.
\]

(No constraint on the length of \( a \) is necessary here, but by convention the eigenvectors are of unit length.) The last relative PC is the one with the minimal variance ratio, and those in-between have intermediate variance ratios. A relative eigenvalue close to 1 indicates a corresponding linear combination with similar variance in both groups.

Note that the eigenvalues of \( S_B^{-1}S_A \) are the inverse of the eigenvalues of \( S_A^{-1}S_B \) in reversed order (for instance, the first relative eigenvalue of \( S_B^{-1}S_A \) characterizing the maximal variance ratio of group A relative to group B, equals the inverse of the last relative eigenvalue of \( S_A^{-1}S_B \) characterizing the minimal variance ratio of group B relative to group A). The squared log relative eigenvalues
thus are symmetric (i.e. they remain the same if exchanging group A with group B, cf. Equation 3) and give rise to a fully multivariate metric for variance-covariance matrices:

\[ d_{cov}(S_A, S_B) = \sqrt{\sum_{i=1}^{p} \log^2 \phi_i} \]  

(5)

where \( \phi_i \) are the eigenvalues of either \( S_B^{-1} S_A \) or \( S_A^{-1} S_B \).

This metric is equivalent to the Fisher information metric for multivariate normal distributions with common mean vectors. It is also the Riemannian metric on the space of square symmetric positive definite matrices, a manifold of dimension \( np(np + 1)/2 \) with the form of a convex cone in the vector space of symmetric matrices (Dryden, Koloydenko & Zhou, 2009; Förstner & Moonen, 1999; Lenglet, Rousson, Deriche & Faugeras, 2006; Mitteroecker & Bookstein, 2009).

Relative PCA and the metric in Equation 5 apply only to two groups with distinct variance-covariance matrices. For a set of \( m \) covariance matrices, the \( m(m-1)/2 \) pairwise Riemannian distances (Equation 5) can be used to quantify the heterogeneity of all the variance-covariance patterns. Because these distances span a high-dimensional non-Euclidean space, they cannot be directly plotted, but they can be represented by a principal coordinate analysis (also referred to as metric multidimensional scaling), which approximates (in a least squares sense) the Riemannian distances by the Euclidean distances in a low-dimensional diagram (Mitteroecker & Bookstein, 2009).

Relative eigenvalues can also be used to express deviations in the magnitude of multivariate variation (rather than in the full variance-covariance pattern) between two products. The ratio of relative eigenvalues equals the ratio of the generalized variances (determinant of the covariance matrix) of the two groups (Mitteroecker & Bookstein, 2009):

\[ \prod_{i=1}^{p} \frac{\phi_i}{\phi_j} = \frac{\det S_A}{\det S_B}. \]  

(6)

As relative eigenvalues are invariant to linear scaling of the variables, also ratios of generalized variances are unaffected by the scale of variables (Huttegger & Mitteroecker, 2011). Plots of log generalized variances are thus useful to compare the magnitude of multivariate variation across groups.

2.1 | Numerical aspects

The computation of relative eigenvalues requires the inversion of a covariance matrix, which imposes a certain computational burden: only matrices of full-rank can be inverted. For a covariance matrix of full rank, all principal components have non-zero variance, which requires an excess of cases over variables as well as linear independence of the variables. Linear dependences typically arise in the course of standardization of variables. Procrustes shape coordinates, for instance, are standardized for position, scale, orientation. Due to the resulting loss in degrees of freedom, the last four PCs (for 2D landmarks) or seven PCs (for 3D landmarks) have zero variance, and the corresponding covariance matrix cannot be inverted (Dryden & Mardia, 1998; Mitteroecker & Gunz, 2009).

Even for full-rank data, the relative eigenvalues are affected by the number of variables (\( p \)) relative to the number of cases (\( n \)). The larger the ratio \( p/n \), the larger are the first relative eigenvalues and the more unstable are the corresponding eigenvectors (Bookstein, 2017; Mitteroecker & Bookstein, 2011).

There are multiple ways to approach this classic numerical problem, including pseudoinverse and matrix regularization (e.g. Mardia et al., 1979; Mitteroecker, Cheverud & Pavlicev, 2016). The Moore-Penrose pseudoinverse of a matrix \( M \) is \( M^+ = QA^+Q' \), where \( Q \) is the matrix of eigenvectors of \( M \), and \( A^+ \) is a diagonal matrix with the reciprocal of the \( m \) non-zero eigenvalues in the diagonal. The remaining \( p-m \) eigenvalues are set to 0. This is equivalent to reducing the data to the first \( m \) principal components with non-zero variance for the inversion. In practice, especially if \( p/n \) is large, also dimensions with small non-zero variance (often reflecting measurement error and other noise) distort the results as they dominate the matrix inverse. Reliable analysis thus often requires reduction to the first few principal components with a large variance (or another low-dimensional factor structure) prior to relative eigenanalysis. A useful strategy is to start the analysis with the first few leading PCs only, and then adding step-by-step more PCs. Typically, the results of the relative PCA changes with the addition of PCs until all major signals are captured, and then stays roughly constant. Beyond a certain number of PCs, the results will change again due to the increasing noise in the PCs (Bookstein & Mitteroecker, 2014; Mitteroecker & Bookstein, 2009).

2.1.1 | Euclidean approximation

Another way to circumvent numerical challenges is to replace the Riemannian metric by the Euclidean metric. Aguirre, Hine, McGuigan and Blows (2014), for instance, performed standard linear statistics, including principal component analysis, to a set of covariance matrices, which implicitly imposes a Euclidean metric (or Frobenius norm), \( d_{euclidean} \) on the matrices:

\[ d_{euclidean}(S_A, S_B) = \sqrt{\sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij}^2}. \]  

(7)

where \( A = S_A - S_B \). The Euclidean space induced by this metric can be considered the tangent space to the Riemannian space of covariance matrices. For very small differences in variance and covariance, the two metrics will be approximately proportional. A principal coordinate analysis based on the Euclidean metric is equivalent to an ordinary principal component analysis of the vectorized covariance matrices; it requires no matrix inverse and hence no constraint on \( p/n \), but it is not affine invariant and can be difficult to interpret if variance ratios are intermediate to large and if units differ among variables (cf. Figure 1).
2.2 Statistical inference

Like ordinary PCA, relative PCA is primarily an exploratory technique that is applied when no specific hypotheses could have been formulated a priori. As in related multivariate settings, it is useful to start with a test of the proportionality of covariance matrices:

\[ H_0: S_B = k S_A \]  \hspace{1cm} (8)

This test is equivalent to a test of the equality of all the relative eigenvalues of \( S_A \) and \( S_B \):

\[ H_0: \phi_1 = \phi_2 = \ldots = \phi_p . \]  \hspace{1cm} (9)

For two multivariate normal distributions, the corresponding maximum likelihood (ML) test is based on the quantity

\[ \frac{1}{2} np \log \left( \frac{a}{g} \right) \sim \chi^2_{(p - 1)(p + 2)/2} . \]  \hspace{1cm} (10)

The parameters \( a \) and \( g \) are the arithmetic and geometric means of the relative eigenvalues, and \( n \) is the sample size per group. In the case of unequal sample sizes, \( n \) can be estimated by the harmonic mean of the two sample sizes. In the limit of \( n \gg p \), the log likelihood ratio (Equation 10) is distributed approximately as a \( \chi^2 \) with \((p - 1)(p + 2)/2\) degrees of freedom, where \( p \) is the number of eigenvalues being compared (Anderson, 1963; Bookstein & Mitteroecker, 2014).

The arithmetic mean of the relative eigenvalues also serves as the maximum likelihood estimate of the scaling factor (the parameter \( k \) in Equation 8) for two proportional covariance matrices (Mardia et al., 1979).

After rejecting the hypothesis of proportionality, relative PCA can be used for identifying the linear combinations that drive this deviation from proportionality. In some situations, it can be interesting to test if the first (or last) two relative eigenvalues differ and thus support a separate interpretation of the corresponding relative eigenvectors. In this case, the test statistic in Equation 10 reduces to

\[ 2n \log \frac{\phi_1 + \phi_2}{2(\phi_1 \phi_2)^{1/2}} \sim \chi^2 . \]  \hspace{1cm} (11)
TABLE 1 Functions of the vcvComp package

| Function name | Description |
|---------------|-------------|
| cov.B         | Computes the between-group covariance matrix (i.e. the covariance matrix of the group means) |
| cov.W         | Computes the pooled within-group covariance matrix |
| cov.group     | Computes the covariance matrix of every group |
| eigen.test    | Performs a ML test of the equality of two successive relative eigenvalues (Equation 11) |
| euclidean.dist| Computes the Euclidean distance between two covariance matrices (Equation 7) |
| mat.sq.dist   | Computes the squared distance matrix of a set of covariance matrices |
| minv          | Computes the inverse or the pseudoinverse of a matrix based on a specified number of PCs or a tolerance threshold for PCs with a small non-zero variance |
| pr.coord      | Performs a principal coordinates analysis of a distance matrix |
| prop.vcv.test | Performs a ML test of proportionality of two covariance matrices (Equation 10) |
| relative.eigen| Computes the Riemannian distance (Equation 5) between two covariance matrices |
| relG V.multi  | Computes the (log-transformed) ratios of the generalized variances of a set of covariance matrices (Equation 6) |
| scaling.BW    | Computes the ML scaling factor between two covariance matrices (k in Equation 8) |

3 | PACKAGE DESCRIPTION AND INSTALLATION

The vcvComp package is written in the R scientific computing language (R Development Core Team, 2018). It provides functions for the comparison of variance–covariance patterns based on relative eigenanalysis (Table 1). It uses the core statistical package stats (R Development Core Team, 2018) for the computation of variance–covariance matrices and suggests other R libraries such as geomorph (Adams, Collyer & Kaliontzopoulou, 2018) for geometric morphometric data and scatterplot3d (Ligges & Mächler, 2003) for the visualization of 3D data.

Furthermore, the package comprises a dataset (Tropheus) of 19 two-dimensional landmarks measured on lateral digital images of the external body of 723 cichlid fishes of the species Tropheus moorii and T. polli, collected from eight locations of Lake Tanganyika (Herler, Kerschbaumer, Mitteroecker, Postl & Sturmbauer, 2010; Kerschbaumer, Mitteroecker & Sturmbauer, 2014). The data are also available via the Dryad repository (Kerschbaumer, Mitteroecker & Sturmbauer, 2013).

The markdown vignette (vcvComp-worked-example.Rmd) provided in the package comprises the step-by-step analysis of the data described in the worked example below. The numerical output and the code chunks used to produce the figures are not presented in the main text of the paper, but they are available in the vignette (see supplementary information).

The package can be installed from the archive file vcvComp_1.0.1.tar.gz (see supplementary information) by typing in the R console:

```r
install.packages("~/vcvComp_1.0.1.tar.gz", repos = NULL, type = "source") # replace ~ by the file path
```

4 | WORKED EXAMPLE: VARIATION OF CICHLID BODY SHAPE

To illustrate the application of relative eigenanalysis by means of the vcvComp package, we studied the variation of body shape within and between different fish populations of the cichlid genus Tropheus of Lake Tanganyika. We used 511 specimens of the sample from Kerschbaumer et al. (2014), consisting of six populations of the Tropheus moorii colour morph ‘Kaiser’. The external body form (body outline, insertion of the fins, position and shape of the orbit, mouth, and operculum) was quantified by 19 two-dimensional landmarks (Herler et al., 2010, see supplementary information). Three of these populations (IKS3, IKS4, IKS5) live in sympatry with the cichlid species T. polli, whereas the three other populations (IKA1, IKA2, IKA3) live alone. As the allopatric and sympatric populations differ in trophic niche and thus presumably also in their selective regime, we investigated if and how they differ in phenotypic variance–covariance structure. We also explored differences in variance pattern between female and male specimens, because these populations show significant sexual dimorphism in mean head shape (cichlids are maternal mouthbrooders; Herler et al., 2010; Kerschbaumer et al., 2014). Finally, we searched for signs of stabilizing and divergent selection among the six Tropheus populations by contrasting within- and between-group covariance matrices.

First, we loaded the vcvComp package and the data.

```r
library("vcvComp")
data("Tropheus")
```

Five specimens are outliers for landmark 2 and were excluded from the sample. After selecting the subsample, we created a new variable combining population and sex.

```r
outliers <- c(18, 56, 155, 351, 624)

Tropheus.IK <- Tropheus[-outliers, ]

# Sample reduced to six populations

Tropheus.IK <- subset(Tropheus.IK, subset = PDP.ID

Tropheus.IK$PDP.ID <- factor(Tropheus.IK$PDP.ID)

# New variable combining population and sex

Tropheus.IK$SexPop <- paste(Tropheus.IK$PDP.ID,
Tropheus.IK$Sex, sep = "_")

Tropheus.IK$SexPop <- as.factor(Tropheus.IK$SexPop)
```
The landmark coordinates were extracted to create a matrix.

```r
phen <- as.matrix(Tropheus.IK which(names(Tropheus.IK) == "Xi") ;
    which(names(Tropheus.IK) == "Yi9")
renames(phen) <- Tropheus.IK$list_TropheusData_ID
```

Then, we performed a generalized Procrustes superimposition (Rohlf & Slice, 1990) of the landmark coordinates using the function `gpagen` of the `geomorph` package.

```r
library("geomorph") # load packages geomorph, rgl and RRPP
# conversion matrix -> array (19 landmarks, 2 dimensions)
PHEN_array <- array(specs(phen, p = 19, k = 2)
# Procrustes superimposition
phen.gpa <- gpagen(PHEN_array, print.progress = FALSE)
# conversion array -> matrix of Procrustes coordinates
proc.coord <- two.d.array(phen.gpa$coords)
colnames(proc.coord) <- colnames(PHEN)
```

We reduced the Procrustes shape coordinates to the first five principal components to avoid collinearities and to guarantee a sufficient excess of cases over variables in further analyses.

4.1 | Population comparison

Because the samples were not balanced regarding sex, we computed the pooled population covariance matrices as unweighted averages of the corresponding male and female covariance matrices.

```r
S.phen.pooled <- cov.group(pc.scores, 
groups = Tropheus.IK$POP.ID, sex = Tropheus.IK$Sex)
```

To explore the heterogeneity of variance–covariance structure in body shape across populations, we performed an ordination analysis of the six pooled within-sex covariance matrices.

```r
eigen.phen <- mat.sqrt(dist(S.phen.pooled, dist = "Riemannian")
proc. <- pr.coordeigen.phen) # ordination
proc$Variance # variance explained
```
The first three principal coordinates together accounted for 88% of total variance (Figure 3a); this also equals the fraction of summed squared Riemannian distances explained by the summed squared Euclidean distances within the first three principal coordinates. The populations living in sympathy (IK3, IK4, IK5) were separated from the allopatric populations (IKA1, IKA2, IKA3) along the third principal coordinate (Figure 3b).

To investigate the actual differences in variance–covariance pattern between sympatric and allopatric populations, we compared the populations IKA1 and IK5, but other pairs of sympatric and allopatric populations led to very similar results. The ML test (Equation 10) indicated that the covariance matrices of IKA1 and IK5 deviate from proportionality at \( p = .02 \).

The generalized variance of IKA1 was only 18% less than that of IK5, but the relative PCA showed that the various shape features deviate strongly in their variational properties across populations (Figure 4a). The first relative PC was roughly twice as variables in IKA1 than in IK5 (first relative eigenvalue was 2.3), whereas the variance of the last relative PC in IKA1 was only half of that in IK5 (last relative eigenvalue was 0.49).

The shape patterns depicted by each relative eigenvector can be visualized by deformations of the average shape along the positive and the negative directions of the corresponding vector (Figure 4b). Note that when the initial variables were reduced to the first principal components, as was the case here, the loadings of the eigenvectors must be multiplied by the loadings of the principal components to get shape patterns.

The shape features captured by relative PC 1 were head shape, relative eye size, and body depth (maximum distance between dorsal and ventral parts); these were the features with maximal excess of variance in allopatric populations relative to sympatric populations (Figure 4b). Or in other words, these were the shape features maximally canaled in the populations living in sympathy.

In Lake Tanganyika, allopatric populations of *Tropheus moorii* live in the whole water column, whereas populations in sympathy with *T. polli* usually are forced to live at greater water depth (Kerschbaumer et al., 2014). The broader trophic niche and the larger environmental heterogeneity in allopatric populations may account for the larger variance in body depth and head shape, but the higher competition and harder living conditions in sympatric populations may also impose a stronger stabilizing selection regime than in allopatric populations.

### 4.2 Comparison between sexes

We separated males and females and performed a principal coordinate analysis of the 12 sex-specific covariance matrices in order to investigate deviations in variance–covariance structure between the sexes.

The first two components together accounted for 62% of total variance (Figure 5) and showed that for all populations, except IKA3, males had higher values than females for the first principal coordinate (Figure 6).

To explore this shared sex difference in variance–covariance pattern, we performed a relative PCA of the females relative to the males of IKA1.

Overall, females were approximately half as variable as males (ratio of generalized eigenvalues was 0.57), but pooling over all dimensions was again misleading here. In fact, the first relative PC was 4.6 times more variable in females than in males (Figure 8).

The first relative PC mainly corresponded to the relative size of the head (Figure 8a), whereas the last three relative PCs were all related to the shape and orientation of the head and mouth (Figure 8b).
Cichlids are mouth brooders and females typically have a larger head and mouth than males. This pattern of sexual dimorphism in body shape was also found for the present *Tropheus moorii* sample (Herler et al., 2010; Kerschbaumer et al., 2014). The increased variance in relative head size likely is a direct consequence of the enlarged head in females, whereas other aspects of head morphology, such as the relative position and orientation of the mouth, seems to be more canalized in females than in males.

### 4.3 Stabilizing versus divergent selection of cichlid body shape

Under idealized assumptions, the expected amount of phenotypic change due to genetic drift is proportional to the amount of additive genetic variance within the population. Extending this model of neutral evolution to multiple traits leads to the expectation that the between-group covariance matrix for a set of related populations is proportional to the additive genetic covariance matrix of their common ancestral population (Lande, 1979). This rational has inspired statistical tests for natural selection by contrasting the covariance matrix of population means with the pooled phenotypic within-population covariance matrix (as a surrogate of the ancestral genetic covariance matrix; e.g. Ackermann & Cheverud, 2004; Cheverud, 1988; Marroig & Cheverud, 2004; Martin, Chapuis & Goudet, 2008): deviations from proportionality are signs of stabilizing or divergent selection. Most of these approaches, however, only rely on statistical significance tests of proportionality of the between- and within-population covariance matrices ($B$ and $W$, respectively). Relative PCA ideally complements these approaches as an exploratory tool to identify the specific trait combinations that deviate from the null model of neutral evolution (Bookstein & Mitteroecker, 2014). If both divergent and stabilizing selection acted in a set of populations, the first relative PCs of $B$ with respect to $W$ will reveal the trait combinations that were affected by divergent selection (the features with maximal between-population variance relative to within-population variance), and the last relative PCs will show the trait combinations under stabilizing selection (least between-population variance relative to within-population variance).

We computed $B$ and $W$ (pooled by sex) for the *Tropheus* populations based on the first five PCs of the Procrustes coordinates. As we have only six populations in this example, $B$ is estimated with great uncertainty and also the chi-square approximation of equations 10 and 11 is critical; results have to be interpreted with care. The ML test suggested a deviation from proportionality between $B$ and $W$ ($p=0.034$) and thus the action of selective forces.

![Figure 6](image_url) Principal coordinates ordination of the 12 sex-specific covariance matrices. Males in blue, females in red. Populations living in sympatry with *Tropheus polli* in dark colors.

![Figure 7](image_url) Relative eigenvalues (maximal ratios of variance) of females relative to males for the population IKA1.

![Figure 8](image_url) Visualization as thin-plate spline (TPS) deformation grids (Bookstein, 1991) of the shape patterns corresponding to (a) the first relative PC, which has the maximal excess of variance in females relative to males for the population IKA1, and (b) the last relative PC, which has the maximal excess of variance in males relative to females.
However, this test does not specify the magnitude of deviation from proportionality, and especially with the small number of populations in this example, the interpretation of the p-value alone is not sufficient. We thus performed an ordination of the six population covariance matrices (pooled by sex), together with W and B (scaled to fit W using the mean of their relative eigenvalues).

FIGURE 9 Principal coordinates ordination of the six populations (males and females pooled), along with their between-group (B) and their within-group (W) covariance matrices.

FIGURE 10 Relative eigenvalues of the between-group covariance matrix versus the within-group covariance matrix for the six Tropheus populations.

### Code-snippet for Figure 9

```r
Bsc <- B / scalingBW(B, W) # scale B to W
# Create an array of group covariance matrices, B and W
B.bw <- array(c(B,phen.pooled, W, Bsc),
              dim = c(dim(B,phen.pooled)[1],
                      dim(W,phen.pooled)[2],
                      dim(B,phen.pooled)[3] + 2))
dimnames(B.bw) <- list(dimnames(B,phen.pooled)[[1]],
                        dimnames(W,phen.pooled)[[2]],
                        c(dimnames(B,phen.pooled)[[3]],"B","W"))
# Compute the Riemannian distances + ordination
eigen.phen.bw <- mat.sqrt{dist(B.bw, dist = "Riemannian")}
pcoa.bw <- pcoa(eigen.phen.bw)
```

Figure 9 shows that relative to the heterogeneity of population covariance matrices, B clearly deviates from W along the first principal coordinate. The interpretation of the relative PCs of B and W thus seems warranted.

Figure 10 shows the relative eigenvalues of B with respect to W. The first relative eigenvalue is more than 5 times larger than the second one, which is significant at \( p < .05 \), and similarly for the last relative eigenvalue (Equation 11). This supports an interpretation of the first and last relative PC, even though the absolute relative eigenvalues cannot be evolutionarily interpreted without knowing the variance ratio expected under neutral evolution, which depends on the number of generations since divergence and effective population size (Lande, 1979). One way to estimate this expected variance ratio is based on genetic data using the \( F_{ST} \) statistic (Holsinger & Weir, 2009). Under pure genetic drift, 

\[
B = F_{ST}/(1-F_{ST})W
\]

(Lynch & Walsh, 1998; Martin et al., 2008). Kerschbaumer et al. (2014) reported \( F_{ST} \) values for these populations.

However, this test does not specify the magnitude of deviation from proportionality, and especially with the small number of populations in this example, the interpretation of the p-value alone is not sufficient. We thus performed an ordination of the six population covariance matrices (pooled by sex), together with W and B (scaled to fit W using the mean of their relative eigenvalues).

**FIGURE 11** Visualization of the shape patterns corresponding to (a) the first relative PC, which has the maximal excess of variance between the six *Tropheus* populations relative to the variance within populations, and (b) the last relative PC, which has the maximal excess of variance within populations relative to that between populations.
populations ranging from 0.033 to 0.085, which translates into ratios of between- to within-population variance of 0.034–0.093. The first relative eigenvalue (2.06) clearly exceeded this threshold and suggests strong divergent selection. The last relative eigenvalue (0.025) may indicate weak stabilizing selection.

The first relative PC corresponded to overall body depth and the positions of caudal, dorsal and ventral fins (Figure 11a). These features, which determine the hydrodynamics and swimming ability of the fish, were under strong divergent selection in the studied Tropheus populations: their inter-population variance strongly exceeded the variation expected for neutral evolution.

The shape pattern corresponding to the last relative eigenvector mainly involved the shape of the head, especially the position and orientation of the mouth (Figure 11b). These features, which crucially affect feeding performance, likely were under stabilizing selection. Canalization of trophic head morphology is also supported by our finding that females show little variation in mouth position.

5 | CITATION OF VCVCOMP

Scientists using vcvcComp in a published paper should cite this article or the vcvcComp package directly. Citation information can be obtained by typing:

citation("vcvcComp")

ACKNOWLEDGEMENTS

This research was supported by the Austrian Science Fund (FWF grant P29397). The authors declare no conflict of interest.

AUTHORS’ CONTRIBUTIONS

P.M. conceived the ideas and designed the methodology; A.L.M. wrote the R scripts; A.L.M. and P.M. analysed the data and wrote the manuscript.

DATA AVAILABILITY STATEMENT

The example dataset used in this manuscript is packaged with vcvcComp and is publicly available via CRAN (https://CRAN.R-project.org/package=vcvcComp). The data are also available via the Dryad repository (Kerschbaumer et al., 2013) https://doi.org/10.5061/dryad.fc02f.

ORCID

Anne Le Maître https://orcid.org/0000-0003-2690-7367

REFERENCES

Ackermann, R. R., & Cheverud, J. M. (2004). Detecting genetic drift versus selection in human evolution. Proceedings of the National Academy of Sciences, 101, 17946–17951. https://doi.org/10.1073/pnas.0405919102

Adams, D. C., Collyer, M. L., & Kaliontzopoulou, A. (2018). Geomorph: Software for geometric morphometric analyses. R package version 3.0.6.

Aguirre, J. D., Hine, E., McGuigan, K., & Blows, M. W. (2014). Comparing G: multivariate analysis of genetic variation in multiple populations. Heredity, 112, 21–29. https://doi.org/10.1038/hdy.2013.12

Anderson, T. W. (1963). Asymptotic theory for principal component analysis. The Annals of Mathematical Statistics, 34(1), 122–148. https://doi.org/10.1214/aoms/1177704248

Armbuster, W. S., Pélabon, C., Bolstad, G. H., & Hansen, T. F. (2014). Integrated phenotypes: understanding trait covariation in plants and animals. Philosophical Transactions of the Royal Society London B: Biological Sciences, 369(1649), 20130245. https://doi.org/10.1098/rstb.2013.0245

Bookstein, F. (1991). Morphometric tools for landmark data: Geometry and biology. Cambridge, UK: Cambridge University Press.

Bookstein, F. L. (2014). Measuring and reasoning: Numerical inference in the sciences. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139161923

Bookstein, F. L. (2017). A newly noticed formula enforces fundamental limits on geometric morphometric analyses. Evolutionary Biology, 44(4), 522–541. https://doi.org/10.1007/s11692-017-9424-9

Bookstein, F. L., & Mitteroecker, P. (2014). Comparing covariance matrices by relative eigenanalysis, with applications to organismal biology. Evolutionary Biology, 41(2), 336–350. https://doi.org/10.1007/s11692-013-9260-5

Chartier, M., Jabbour, F., Gerber, S., Mitteroecker, P., Sauquet, H., von Balthazar, M., Schönberger, J. (2014). The floral morphospace—A modern comparative approach to study angiosperm evolution. New Phytologist, 204(4), 841–853. https://doi.org/10.1111/nph.12969

Cheverud, J. (1982). Phenotypic, genetic, and environmental morphological integration in the cranial. Evolution, 36, 499–516. https://doi.org/10.1111/j.1558-5646.1982.tb05070.x

Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. Evolution, 42(5), 958–968. https://doi.org/10.1111/j.1558-5646.1988.tb02514.x

Dryden, I. L., Kolodydenko, A., & Zhou, D. (2009). Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. Annals of Applied Statistics, 3(3), 1102–1123. https://doi.org/10.1214/09-AOAS249

Dryden, I. L., & Mardia, K. (1998). Statistical shape analysis. New York, NY: John Wiley and Sons.

Flury, B. N. (1985). Analysis of linear combinations with extreme ratios of variance. Journal of the American Statistical Association, 80(392), 915–922. https://doi.org/10.1080/01621459.1985.10478203

Foote, M. (1993). Contributions of individual taxa to overall morphological disparity. Paleobiology, 19(4), 403–419

Förstner, W., & Moonen, B. (1999). A metric for covariance matrices. In F. Krumm & V. S. Schwarze (Eds.), Quo vadis geodesia ...? Festschrift for
