Coriandis: Correlation analysis based on distances
Version 1.12 (compiled 03/11/14)

© Eladio Marquez 2012-2014
Mammals Division
University of Michigan Museum of Zoology
http://www-personal.umich.edu/~emarquez/morph/

Citation:
Márquez, E. J. & Knowles L.L. 2007. Correlated evolution of multivariate traits: detecting co-divergence across multiple dimensions. Journal of Evolutionary Biology, *in press*. doi: 10.1111/j.1420-9101.2007.01415.x

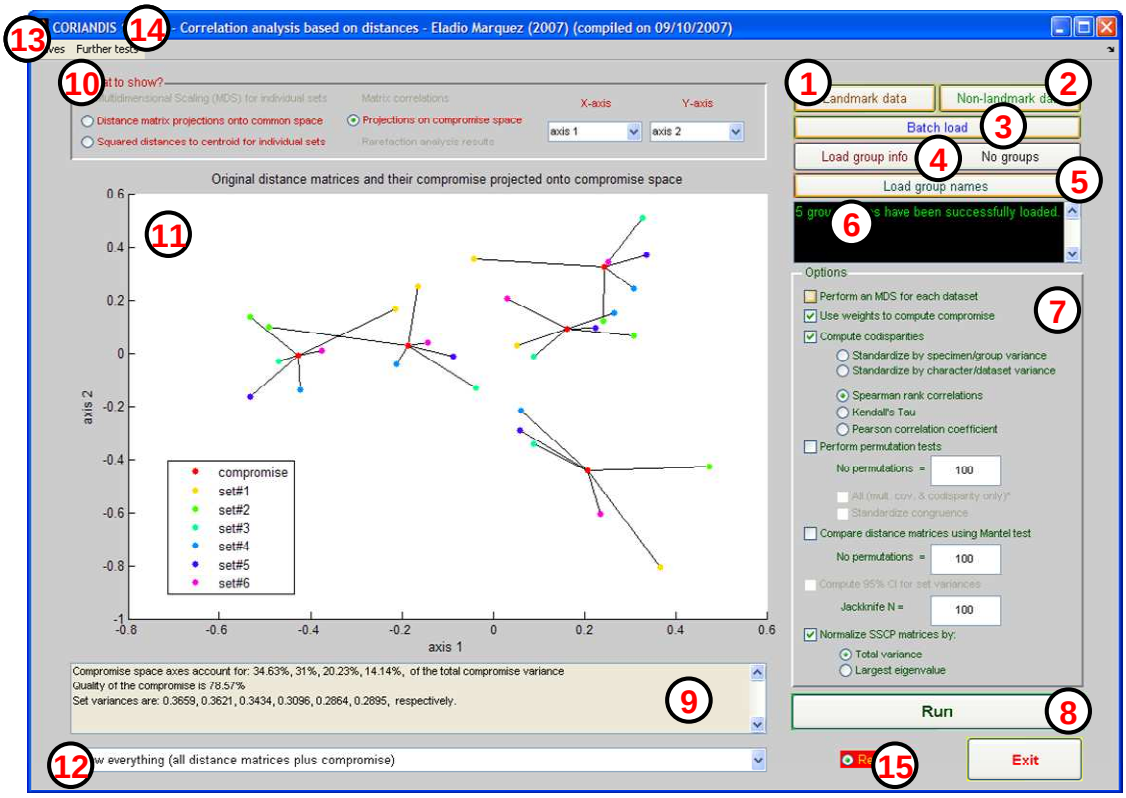
Reference Guide

The following guide presents detailed instructions to operate the program CORIANDIS, whose general purpose is to offer a number of tools to study associations among multivariate datasets. Besides some additional support routines, this software implements most of the methods described in Márquez and Knowles (*in press*) for a broad spectrum of data types, including 2-D landmark and distance data. The main core of these methods, namely the comparison via correlation analyses of matrices of distances among specimens or groups, has been adapted from a technique described by Abdi et al. (2005, 2007), which in turn is an adaptation to distance matrices of the family of techniques known as STATIS (Escoufier 1973). For the mathematical details behind these methods, please refer to either of these articles.

The first portion of this guide describes the methods that CORIANDIS uses to input/output data, including the graphical outputs, and the meaning of analysis options. Refer to the labels in the following page's screenshot to locate particular topics of interest. The second part will show how the program actually works by performing an brief annotated analysis of some of the example datasets provided with this CORIANDIS release. An index of the program functions mentioned in this guide is included at the end.

CORIANDIS uses the main interface (shown in the following page) to both determine what options to use in an analysis' execution, and what kinds of results to show. ***Different outputs will be available depending on what options are chosen***. All outputs contain results obtained ***after the most recent run of any analysis***, so that changing options will not produce by itself new results. The **Run** button must be clicked first.

Main interface



Description of program functions:

1. **Landmark data**: datasets containing 2-D landmarks must be loaded after pressing this button, one at a time. A standard Windows explorer dialog will open; navigate to your folder and open a text-format data file (containing space- or tab-separated columns). A message will show in the secondary program screen (which opened along the main interface) confirming whether a dataset has been loaded, and the explorer dialog will automatically reappear, prompting you to load a second file. After you are done loading files, just click Cancel to return to the main interface. ***Use this button only to load 2-D landmark coordinates***. Currently, CORIANDIS accepts only landmark data formatted in columns X1, Y1, X2, Y2, …, Xn, Yn (for a total of 2n columns). If available, a centroid size column can be added as an additional columns, *after* landmarks. Since the methods implemented in this software are intended to study associations among multivariate traits for a particular set of specimens or taxa, ***all loaded datasets must contain the same number of rows***. Examples of landmark datasets bundled with this software are named `dataset_x.dat`.

2. **Non-landmark data**: non-landmark data may include any measured property, whether uni- or multivariate, including outlines, interlandmark distances, or even ecological inforamtion, as well as inter-specimen (or inter-group) distance

matrices, irrespective of how distances are computed. Non-distance datasets must have the same structure as landmark data (specimens in rows, variables in columns), whereas distance matrices must be entered in full, as squared, symmetrical matrices. As before, ***all loaded datasets, including distance matrices, must contain the same number of rows***, unless group information will be loaded (see below), in which case distance matrices must have as many rows as there are groups. Also, ***if at least one distance matrix is loaded with other datasets, options involving resampling*** (e.g. rarefaction analyses, jackknife confidence intervals) ***will be disabled*** until loaded datasets are removed by pressing **Reset**. Examples of distance matrix datasets bundled with this software are named `distances_x.dat`.

3. **Batch load**: this function allows loading any number of data files simultaneously, by using a *batch file*—a text file with directory address and data type instructions. Each row in a batch file must correspond to a dataset. Fields within each row are separated by commas (so *using commas to name files or folders in the path of loaded files may lead to an error message*). In the current version of this software, only two fields can be defined per dataset: address and type. The *address* field must contain the full directory path where your file is located. The *data type* field must start with the string `TYPE=`, followed by `L` (if the entry correspond to landmark data), `D` (if it is a distance matrix), or `O` (if otherwise). For example, the following batch file would download two files—`dataset1.dat` and `dataset2.dat`—, located in different folders, containing one landmark dataset (type L) and one distance matrix (type D), respectively:

```
C:\landmarks\dataset1.dat,TYPE=L
C:\distances\dataset2.dat,TYPE=D
```

After successfully loading a batch file, this option will become inactivated, and will remain so until the **Reset** function is used. It is possible, however, to load further data files individually using the other two load buttons.

4. **Load group info / No groups**: these buttons allow loading information to categorize loaded specimens (rows in data files) into user-defined groups. If used, means are computed from specimens belonging to the same group and used instead of the actual specimens. Different grouping schemes can be loaded during the course of an analysis, but only one can be used at a time. A *group protocol* file is a text file containing as many rows as specimens in each of the data files. Groups are numbered starting from 1 and increasing continuously. For each row, write the number of the group to which the corresponding specimen belongs to. For example, if 6 specimens have been sampled, the protocol could take any of these three forms:

```
1       1       1
1       1       2
1       2       3
2   ,   2   ,   4
2       3       5
2       3       6
```
(or)

depending on whether we want to define two three-specimen groups, three two-specimen groups, or not define any grouping at all, respectively. The latter case is equivalent to click on the **No groups** button. An example group protocol has been bundled with this release (file's name is `dataset_groupprot.txt`)

5. **Load group names**: a group names protocol is optional, and it consists of a text file containing as many rows as groups are defined in the group protocol (function 4 above). Names can be alphanumeric, and should be written in the same sequential order used in the group protocol. If the **No groups** option has been used, a different name can be given for each row. These names are used to identify groups in certain figures. Examples of group names protocols are bundled with this release (`dataset_groupnames.txt`, which contains names for three groups with five species; and `dataset_specnames.txt`, containing names for a case where the **No groups** options was used, i.e. five names for five species).

6. **Status messages**: CORIANDIS prints messages in this box related to file input, e.g., whether loading a file or batch of files has been successful.

7. **Options**: in this space, you can choose what analysis to perform in a particular run of the program, along with some options that you can change for specific analyses. Depending on what options you choose, different combinations of outputs (Function 10) and saves (Function 13) will become available.

    a. **Perform an MDS for each dataset**: if this option is checked, CORIANDIS will carry out a non-metric multidimensional scaling on the matrix of Euclidean distances obtained from each dataset. In the present implementation, the intergroup or interspecimen distances in the matrix are used to obtain a configuration of points, corresponding to groups/specimens, in two dimensions, so that Euclidean distances among these points approximate distances in the original matrix. The method can be used to assess the overall similarity among groups/specimens as implied by each multivariate dataset. The quality of the "fit" of the approximation is given by a measure of "stress", that is minimized by the procedure. Minimized stresses are reported in the results box (function 9) when MDS plots are chosen (function 10). See Krzanowski (2000) for more details.

    b. **Use weights to compute compromise**: a 'compromise' is built as an average of the SSCP matrices derived from centering each loaded distance

matrix. As such, it can be interpreted as the structure of interspecimen/ intergroup similarilty averaged over multiple multivariate traits. If this option is checked, weights are applied to each trait before computing the compromise. Weights are chosen so that traits that are most congruent with other traits have a larger influence on the computation of the compromise than traits that are too different from the rest. See Abdi et al. (2005, 2007) and Márquez and Knowles (*in press*) for further details.

c. **Compute codisparities**: if this option is checked, the squared distances of each specimen/group to the origin are computed for each dataset, and plotted in the graph area (function 11). Codisparity among traits are obtained as correlations of these squared distances, and receive their name from its original usage in interspecific comparisons (*they should probably not be termed as such if only intraspecific comparisons are been carried out*). Codisparities and their *P*-values (which correspond to the same null hypothesis as an ordinary correlation analysis) can be saved from the **Saves** menu (function 13). By default, reported *P*-values correspond to known distributions, selected according to the type of correlation being computed (see below). If permutations are used, however, then *P*-values are obtained using this approach.

   Additional options:

   i. *Standardize by specimen/group variance*: squared distances are transformed so that added variation at each group equals 1, facilitating comparisons among groups with different degrees of distinctiveness.
   ii. *Standardize by character/dataset variance*: squared distances are transformed so that variation at each dataset equals 1 when added over all specimens/groups.
   iii. *Spearman/Kendall/Pearson correlation metrics*: choose here what statistic to use to obtain co-disparity values.

d. **Perform permutation tests**: the specified number of permutations will be used to tests hypotheses of congruence, codisparity, and multivariate covariance. Resulting *P*-values and permutation runs can be saved from the **Saves** menu.

   Additional options:

   i. *All (mult. cov. & codisparity only)**: when checked, this option instructs the program to enumerate all possible permutations of specimens/group in significance tests for covariance or codisparity. ***This option is not feasible if the number of specimens/groups is too large, or above 10, but could produce error messages even with smaller sample sizes if available computer resources are insufficient***.
   ii. *Standardize congruence*: if checked, this option instructs CORIANDIS to use permutation runs to rescale congruence values, which is necessary if sample sizes are small or if it is

intended to compare results obtained from samples of different sizes. See Márquez and Knowles (*in press*) for further deails.

e. **Compare distances using Mantel test**: this option allows computing and testing (using random permutations) matrix correlations among distance matrices.

f. **Compute 95% CI for set variances**: when checked, CORIANDIS obtains N subsamples of the original datasets by dropping a random number of specimen from each (the same specimens are dropped from all datasets). The actual number dropped is randomly picked for each of the N subsampling rounds bewteen 1 and a third of the sample size. Confidence limits are then computed as the 2.5 and 97.5 percentiles of the resulting distributions of set variances. Variances and, if selected, confidence limits are reported in the results box (function 12) after each run.

g. **Normalize SSCP matrices by***: if this option is checked, SSCP matrices (which are obtained by centering the loaded or computed Euclidean distance matrices) are normalized so that information of differences among trait variances (or disparities) are removed prior to computing congruences. ***This option should be checked when computing congruences, but unchecked when computing set variances (and their CI), since normalization will render the differences between set variances uninterpretable***. See Abdi et al. (2005, 2007) and Márquez and Knowles (*in press*) for further details.

   Additional options:
   i. *Total variance*: standardize SSCP matrices by their trace, so that each trait has approximately the same total variance.
   ii. *Largest eigenvalue*: standardize SSCP matrices so that their largest eigenvalue is 1, and the others are rescaled accordingly.

8. **Run**: clicking this button causes CORIANDIS to perform default analyses plus those specified in the options section (function 7). The time it takes to run an analysis will be somewhat determined by the number of permutations and jackknife subsamples required. When done, the appropriate options will be made available in the output panel (function 10) and **Saves** menu (function 13), selecting **Projections on compromise space** as the default output (explained below).

9. **Results box**: text in this box complements results shown graphically, and will contain different information depending on the chosen output.

10. **What to show?**: from this output panel, you can control what plots/results to show at a particular time. Only results produced since last run will be enabled. Some results can produce more than one plot, which can be selected using the pull-down menu at the bottom of the interface (function 12). These additional plots are included in the following descriptions:

a. **Multidimensional Scaling (MDS) for individual sets**: plots the two dimensions to which distance matrices are fitted using MDS (see **Options** descriptions above). Pull-down menu can be used to select among datasets.

b. **Matrix correlations**: plots corresponding cells from pairs of distance matrices (datasets), which can be selected using the pull-down menus at the left of the panel. Correlation coefficients and *P*-values from Mantel tests are printed in the results box (function 9).

c. **Distance matrix projections onto common space**: the "common space" is obtained as the eigenvectors of the congruence matrix, and this plot contains the set coefficients in such eigenvectors. Values along 'axis 1' equal the weights used to compute compromise (see **Options** above).

d. **Projections on compromise space**: in this option, which is the default output after each run, all specimens/groups and traits/sets are plotted in the same space, obtained by projecting each dataset plus their weighter average ('compromise') onto the compromise space (see above for details). The compromise itself is plotted as a red circle and connected by wires to other traits, so that the resulting plot consists of a set of 'star' plots centered at the compromise, each star corresponding to one specimen or group. This allows both looking into similarities among specimens/groups, and interpreting such similarities in terms of congruence among traits. The pull-down menu at the bottom allows separate plotting of the compromise and distance matrices, plus a Scree plot for the compromise space.

   Along with these plots, enabling this option prints, in the results box (function 9), the proportion of the variance accounted for by each of the eigenvectors of the compromise space (also charted in the available Scree plot), the 'quality' of the compromise, and the variance of each dataset (as explained above, only meaningful if normalization of SSCP matrices is turned off). Quality equals the proportion of the variance accounted for by the first eigenvector of the common space (see above), and thus is a proportional measure of the amount of information in the data used to compute weights. See Abdi et al. (2005, 2007) for further details.

e. **Squared distances to centroid for individual sets**: this option is available after codisparity values are computed. Clicking on it produces a stacked-bar chart where total height equals the total or standardized sum of squared distances of each trait/dataset to the origin (a measure of variance or disparity), and that allows looking into how much each species differs from the rest and to interpret such differences in terms of individual traits. These squared distances can also be plotted against one another by selecting appropriate options in the bottom and upper pull-down menus.

f. **Rarefaction analysis results**: by clicking this option (which is made available after successfully completing a rarefaction analysis, function 14), two graphs are enabled. The first one plots correlations between rarefied and actual data as a function of (decreasing) sample size at each rarefaction run. This plots contains (1) correlations between the congruences averaged over rarefied samples and original congruence values (in red), correlations between each rarefied sample and the original congruence (in black), and averages of the latter (in blue). The second plot allows examining the average and 95% confidence limits for the congruence between selected pairs of traits/datasets as a function of the sample size of the rarefied sets.

11. **Plot area**: if visible, the legend in plots can be dragged within the plot area by clicking on it. Right-clicking on this area produces a pop-up menu that allows toggling the legend on and off, and also saving the current plot as an image file. Several bitmap and vector formats are available.

12. **Additional plots pull-down menu**: allows switching between graphs when more than one is available for a particular type of results (see description of function 10 for details).

13. **Saves menu**: this menu allows saving most of the outputs produced by a CORIANDIS run, including in some cases individual subsamples from permutation and resampling procedures. When saving results corresponding to multiple sets (e.g., distance matrices), the first column of the output indicates to what set the columns at their right belong.

14. **Further analysis menu**:
    a. **Rarefaction analysis of congruence**: this option allows computing conrgruence at smaller sample sizes by obtaining a number of subsamples from the original dataset, hence aiding to determine whether actual sample size is sufficient to produce consistent results. This option is only enabled if no distance matrices were loaded in CORIANDIS, only after a run has been completed, and when the SSCP normalization option has been checked. When selecting this option, an input dialog will open, prompting to specify some additional options:

        i. *Maximum size of rarefied dataset*: must be equal or smaller than the number of specimens/groups to be rarefied;
        ii. *Minimum size of rarefied dataset*: must be equal or larger than 3;
        iii. *Number of rarefied replicates per sample size*: this is the number of random subsamples to with each of the sample sizes between maximum and minimum specified above;
        iv. *Step size for rarefaction*: number of steps in sample size between the maximum and minimum to be used for subsampling.

        Rarefaction analyses start after confirming these inputs, after which ***two*** file save dialogs will prompt for file names to save analyses

outputs: the first dialog automatically produces three different files, (1) average congruence matrices at each rarefied sample size (indicated in the first column of the output), (2) lower, and (3) upper confidence limits for these congruences; the second dialog will save a single file with three columns: (1) rarefied sample size, (2) matrix correlations between average congruence matrices from rarefaction analyses and the original congruence matrix, and (3) average correlations between the congruence of each rarefied sample and the original congruence matrix. *Either or both file save dialogs can be ignored* by clicking the Esc key or Cancel button.

15. **Reset**: clicking this button will erase all loaded datasets and return the program to startup conditions.

Example of CORIANDIS usage:

The following worked example uses some of the sample datasets bundled with CORIANDIS. It is recommended that you carry out the procedures detailed in this example as you read through, while free experimentation with the supplied datasets is highly encouraged until you become familiar with the different options of the program. To begin, use the **Batch load** function and load the file `dataset_batchAll.txt`. ***Make sure to edit this file so that file paths point to the correct folders*** (otherwise, an error message will be generated). If load was successful, the secondary screen should read:

```
Data in file XXX\dataset_1.dat will be loaded as LANDMARK data
Data in file XXX\dataset_2.dat will be loaded as LANDMARK data
Data in file XXX\dataset_3.dat will be loaded as LANDMARK data
Data in file XXX\dataset_4.dat will be loaded as LANDMARK data
Data in file XXX\distances_5.dat will be loaded as a DISTANCE MATRIX
Data in file XXX\distances_6.dat will be loaded as a DISTANCE MATRIX
```

where `XXX` corresponds to the actual path where your files are. The status box should read:

```
Batch results: 6 data files successfully loaded.
```

In the next step, click **No groups**; the status box should read:

```
No group protocol file has been defined for loaded data.
```

Although CORIANDIS allows changing group information during an analysis (i.e., without re-loading data), keep in mind that grouping applies only to non-distance data, so that a single distance matrix with dimensions $m \times m$ can only be used in combination with a dataset that has m specimens or groups. An error message will be generated if dimensions of loaded distance matrices are not consistent with those of loaded non-distance data.

Next, click **Load group names** and select file `dataset_specnames.txt`. The following message will be printed:

```
5 group names have been successfully loaded.
```

Alternatively, these analyses can be run using only the four landmark-based `dataset_x.dat` files supplied, which can be batch-loaded usiong `dataset_batchLM.txt`; or, for a case with three groups, group information can be loaded and named with `dataset_groupprot.txt` and `dataset_groupnames.txt`, respectively. In this case, use `distances_x_3g.txt` for the correct dimensions of distance matrices.

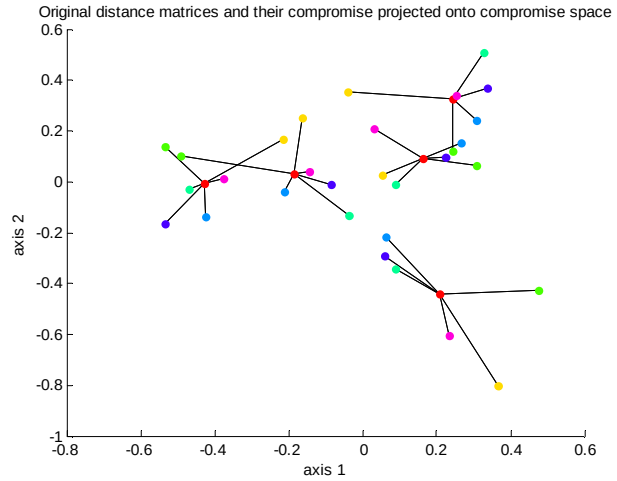
In summary, the program now has loaded 6 files, each corresponding to a (multivariate) "dataset", "set", "trait", or "character", each of which contains measurements for 5 individual specimens ("species" 1 through 5 according to labels in `dataset_specnames.txt`).

Let us then examine results obtained by changing program options. First, uncheck **Compute codisparities** and leave checked only **Use weights**… and **Normalize**…, then click **Run**. The plot in the following page is produced (legend has been removed by right-clicking in the plot area).

Before interpreting this graph, it is useful to experiment by re-running the analysis after unchecking **Use weights**… and/or normalizing by **Largest eigenvalue** instead of **Total variance**. In these runs, pay particular attention to differences in the result box below the plot area. For default conditions, CORIANDIS will print the following text in the results box:

```
Compromise space axes account for: 34.63%, 31%, 20.23%, 14.14%,  of the
total compromise variance
Quality of the compromise is 78.57%
Set variances are: 0.3659, 0.3621, 0.3434, 0.3096, 0.2864, 0.2895,
respectively.
```

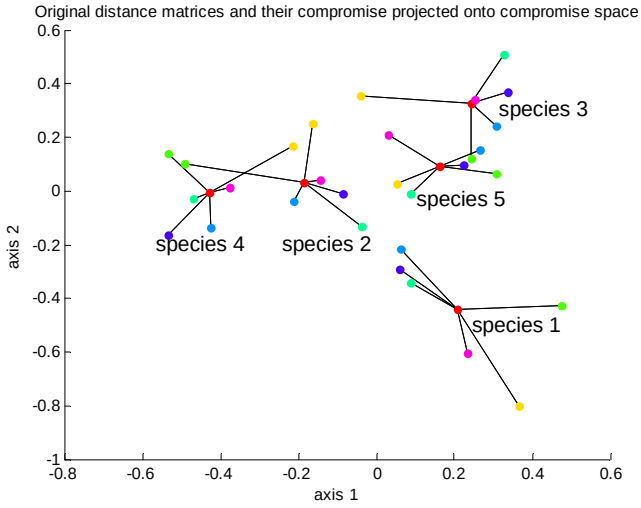
The percent values at the top are obtained from a PCA of the "compromise" SSCP matrix (see above and refs. for details). In this case, there are only 4 PCs because there are only 5 specimens. We can see that the shown axes (1 and 2) account for just over 65% of the variation of this compromise. In this output, the "quality of the compromise" is the percentage of variation of inter-trait congruence that is accounted for by the first eigenvector of the congruence matrix—in other words, it is a measure of how much of the information in these datasets is used in defining the weights used to compute the compromise. At the same time, it is a measure of how appropriate it is to use a unidimensional scheme as the one implemented in CORIANDIS to weight datasets. Finally, set variances (i.e. variances of the 6 loaded datasets) are very similar in this case due to the normalization step used in this run, and therefore not interpretable until this option is unselected.



We see that this graph contains plots for 5 clusters (stars), each corresponding to a different species. In order to match clusters to species, it is necessary to save the data used to make this plot and match the sequence of the specimens/groups in the resulting file to the sequence of specimens/groups in the input. For this, we can use **Save specimen/group scores on compromise space** from the **Saves** menu, which produces the following output:

```
 0.20807978  -0.44118623  -0.03923854  -0.07471933
-0.18610591   0.03135741   0.36998607  -0.05339945
 0.24341745   0.32512548  -0.09417023  -0.18577659
-0.42755268  -0.00807325  -0.23453353   0.00012008
 0.16216137   0.09277659  -0.00204377   0.31377529
```

which contains the scores of 5 specimens (rows) in the 4 PCs of the compromise space (columns). Matching these scores and the plot they produce (compromise, or red, points in the plot area), we get:

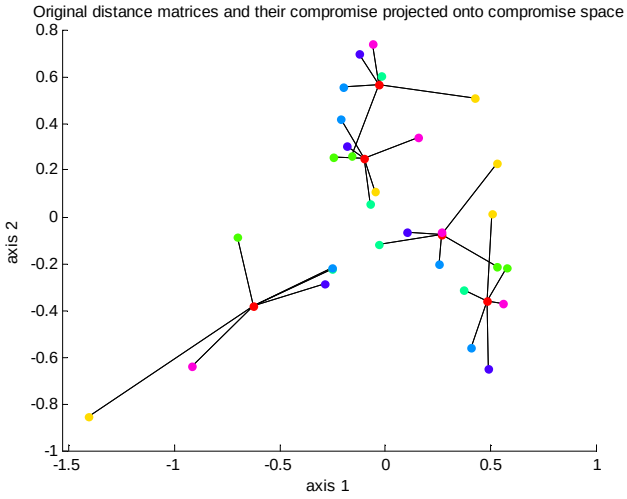


Congruence and multivariate covariance measure how similar the interspecific locations of traits/datasets (represented as colored points) are in this space. If two traits tend to be consistently different or similar between pairs of species, they are said to be (positively) congruent, and will show in this plot as a general tendency to cluster together within species. This is the case of traits #4 and #5 in this example (represented as points in blue tones). If two traits are inconsistent in how similar they are between species (as traits #1 and #3/#4 are in this example), then low congruence and covariance values are expected.

Trait variance or disparity is also visible in this graph, as it is proportional to the area occupied by datasets (all points of the same color). In order to properly interpret variances, though, it is necessary to re-run the anlysis after unchecking the normalization option. The new results box will change to:

```
Compromise space axes account for: 33.82%, 31.79%, 19.98%, 14.41%,  of
the total compromise variance
Quality of the compromise is 78.57%
Set variances are: 1.277, 0.4483, 0.3612, 0.7041, 0.5777, 0.8691,
respectively.
```

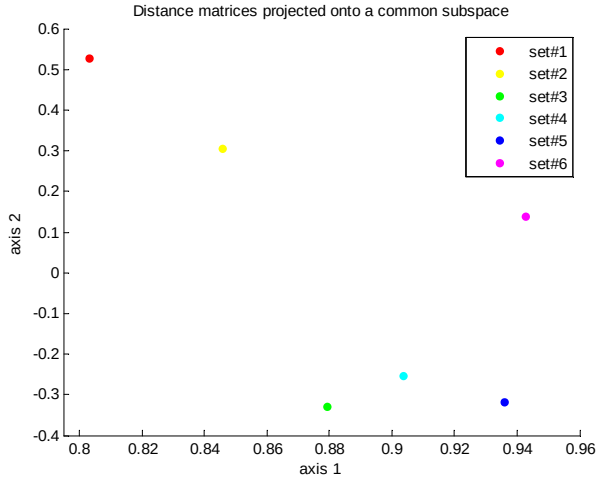
Note that the main distinction with respect to the previous result is the remarkable difference among variances. In this case, trait #1 has over three times more variance than trait #3, and this is reflected in the resulting projection on compromise space:



where trait #1 (yellow points) clearly occupies a larger area of the space. Variances can be exported from the **Saves** menu by using **Save compromise space statistics**. A single column will be saved containing the variance of each dataset/trait, plus a last row containing the quality the compromise, as a proportion. If the option **Compute 95% CI for set variances** is checked, these confidence intervals will be printed alongside variances in the results box, and as two additional columns if **Save compromise space statistics** is used (for low 2.5% and high 97.5% limits, respectively.

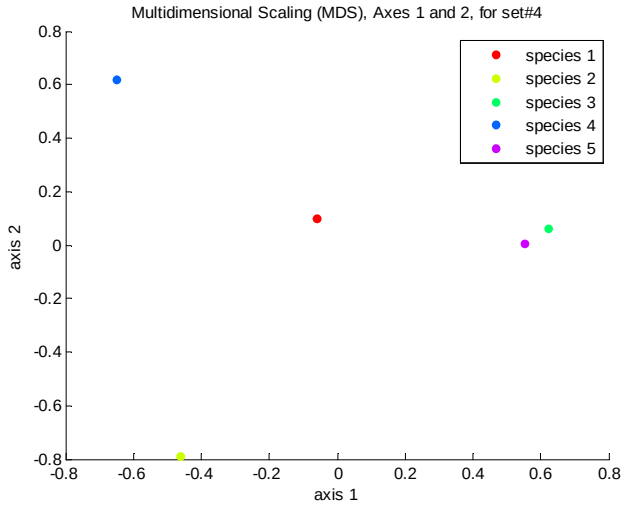
The next graph available, **Distance matrix projections onto common space**, plots one point per dataset, on a space built from PCA of the congruence matrix, so that sets that

are closer to each other over all dimensions are more congruent. Values along the first of these axes are actually the ones used as weights to compute the compromise. From these data, projection of distance matrix on the first two axes of their common space is:



In this case, we see how sets #1 and #2 are incongruent relative to the other sets.

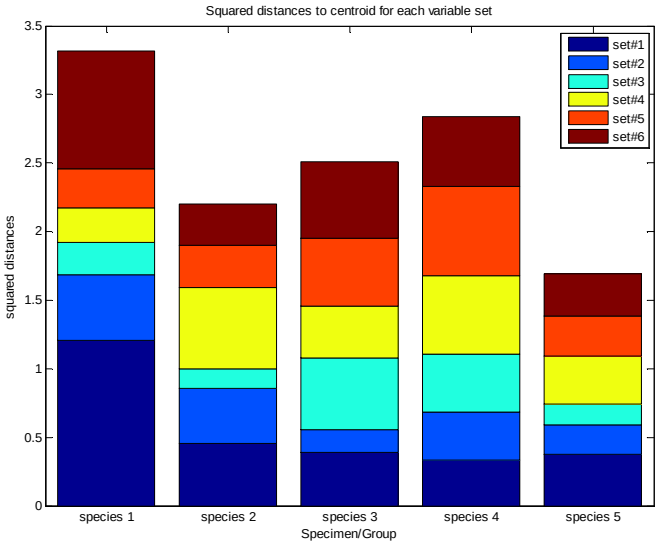
For the next run, let us perform a MDS (non-metric multidimensional scaling) analysis, by checking the appropriate options. The resulting graphs (six in this case, selectable via the pull-down menu at the bottom) depict interspecific distances for each dataset, as two-dimensional approximations. For trait #4, for example, the resulting plot is:



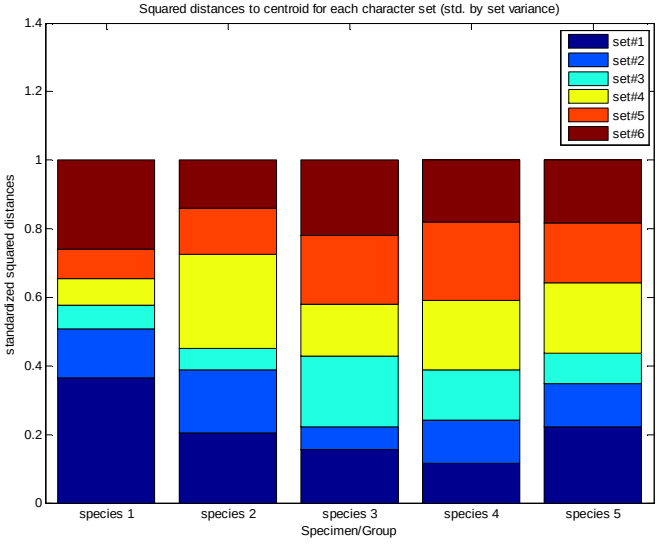
in which species 3 and 6 appear clearly much more similar to each other than to the rest.

Next, check the option **Compute codisparities**, while leaving both standardization options unchecked, and select **Pearson correlation coefficient** as your measure of correlation. For now, do not select permutation tests. After running this analysis, select to

show **Squared distances to centroid for individual sets**, which are the quantities used to compute codisparities. The resulting graph should look like this:



In this chart, the total height of each bar results from the addition of the squared distances of each trait separately, which a measure of trait disparity. This chart can therefore be interpreted as a decomposition of a species distinctiveness from other species in terms of specific traits. Thus, for example, we can see that species 1 departs considerably from other species, although this seems to be largely a function of trait #1 alone. We can compare the relative contributions of different traits to each species' distinctiveness by standardizing the squared distances so that their sum is 1 for all species. For this, we check **Standardize by specimen/group variance**, and re-run the analysis. The resulting chart looks like this:



In this case, we can see how the contribution of different traits to a species' divergence (possibly a function of divergence rates) varies across traits and species. Plots of these

contributions between pairs of traits can be produced using the pull-down menu at the bottom of the interface.

Codisparity measures the degree to which two traits tend to contribute similarly or proportionally to the divergence of several species. In the charts above, we ask: are two traits consistently contributing less or more to species divergence, as if the traits themselves were diverging together?

Actual codisparity values can be obtained from the **Saves** menu, along with their corresponding *P*-values (which correspond to the same null hypothesis as an ordinary correlation coefficient). In the present case, codisparities are:

```
 1.00000000   0.36134053  -0.65259086  -0.60808955  -0.94073027   0.58293425
 0.36134053   1.00000000  -0.89732405   0.40627338  -0.54567235  -0.46294574
-0.65259086  -0.89732405   1.00000000  -0.14913897   0.73061163   0.19716664
-0.60808955   0.40627338  -0.14913897   1.00000000   0.39332036  -0.99683503
-0.94073027  -0.54567235   0.73061163   0.39332036   1.00000000  -0.36689036
 0.58293425  -0.46294574   0.19716664  -0.99683503  -0.36689036   1.00000000
```

and *P*-values are:

```
 0.00000000   0.55014452   0.23257162   0.27654850   0.01716659   0.30227703
 0.55014452   0.00000000   0.03888057   0.49732189   0.34143982   0.43234872
 0.23257162   0.03888057   0.00000000   0.81081666   0.16088803   0.75059577
 0.27654850   0.49732189   0.81081666   0.00000000   0.51243865   0.00021364
 0.01716659   0.34143982   0.16088803   0.51243865   0.00000000   0.54356334
 0.30227703   0.43234872   0.75059577   0.00021364   0.54356334   0.00000000
```

In this case, we can see how the only significant ($P < 0.05$) codisparities are between traits #1 and #5, and between traits #4 and #6. These are, however, *negative* codisparities, meaning that whenever one of the traits is very divergent, the other one is closer to the species average. These particular results, of course, must be confirmed with a sufficiently large dataset in order to be taken seriously. Squared distances corresponding to each dataset can be plotted using the bottom and upper pull-down menus.

Next, the option **Compare distance matrices using Mantel test** can be used to compare pairwise matrix correlations among distance matrices, and test them by randomly permuting the rows and columns of one of the matrix (i.e., Mantel test). Output for this analysis can be accessed through the plot control panel and from the **Saves** menu. The result box will print a message like the following:

```
Matrix correlation for set 1 and set 2 is r = 0.71, with P-value = 0.07
Mantel test was based on 100 permutations.
```

whereas the pair of datasets being reported can be changed using the upper pull-down menus. The plot compares the actual distances extracted from distance matrices for each of the datasets, which can be saved in vector form from the **Saves** menu, selecting **Saves vectorized distances**.

Last, the option of performing permutation tests, if cheched, allows computing *P*-values for congruence and multivariate covariances, while replacing the distribution values

obtained from codisparity analyses. These values can be exported from the **Saves** menu. The option to **Standardize congruence**, which is only available if normalization is checked, allows for computation and testing of a congruence matrix that, although less intuitive to interpret, is comparable across studies based on different numbers of specimens (see Márquez and Knowles, in press for further details).

Cited references

Abdi, H., Valentin, D., O'Toole, A.J. & Edelman, B. 2005. DISTATIS: The analysis of multiple distance matrices. In: Proceedings of the IEEE Computer Society: International Conference on Computer Vision and Pattern Recognition, San Diego, pp. 42–47.

Abdi, H., Valentin, D., Chollet, S., & Chrea, C. 2007. Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. Food Quality and Preference, 18: 627-640.

Escoufier, Y. 1973. Le traitement des variables vectorielles. Biometrics 29: 751–760.

Krzanowski, W.J. 2000. Principles of Multivariate Analysis: A User's Perspective. Oxford University Press, Oxford.

Márquez, E. J. & Knowles L.L. 2007. Correlated evolution of multivariate traits: detecting co-divergence across multiple dimensions. Journal of Evolutionary Biology, *in press*. doi: 10.1111/j.1420-9101.2007.01415.x

Index