

BioNumerics Tutorial:

Band matching and polymorphism analysis

1 Aim

Fingerprint patterns do not have well-defined characters. Band positions vary continuously, although they do tend to fall into categories, or *band classes*. BioNumerics allows you to define band position classes, thereby creating a band matching table. A number of statistical techniques can be applied on these tables, including minimum spanning trees, dimensioning techniques such as principal components analysis, etc. In this tutorial you will learn how to create a band matching table and how to search for discriminative bands between a selection of entries.

2 Principle of band matching

Band matching be executed on any selection of entries from the database. In a first step, BioNumerics divides all the bands found among the selected patterns into *classes of common bands* (1 to 8 in Figure 1). As such, every band of a given pattern belongs to a class, and conversely, every band class is represented by a band on one or more patterns. The result is shown in Figure 1.

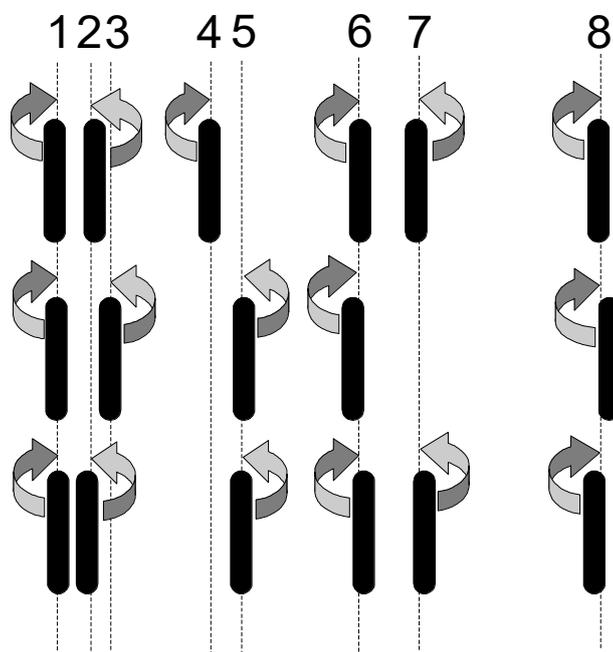


Figure 1: Comparative quantification: bands are assigned to classes.

For each pattern, a particular band class can have two states: present or absent. The binary table for the above example (Figure 1) is shown in Figure 2.

Instead of using binary (+/-) data, the same tables can be generated using band intensities obtained from the curves (band heights or surfaces) or from the two-dimensional pattern contours (volumes or concentrations).

	1	2	3	4	5	6	7	8
Pattern 1	+	+	-	+	-	+	+	+
Pattern 2	+	-	+	-	+	+	-	+
Pattern 3	+	+	-	-	+	+	+	+

Figure 2: Binary presence/absence table of banding patterns.

3 Preparing the database

The **DemoBase Connected** will be used in this tutorial and can be downloaded directly from the *BioNumerics Startup* window or restored from the back-up file available on our website:

1. To download the database directly from the *BioNumerics Startup* window, click the **Download example databases** link, located in the lower right corner of the *BioNumerics Startup* window. Select **DemoBase Connected** from the list and select **Database > Download**. Confirm the download action.
2. To restore the database from the back-up file, first download the file `DemoBase_Connected.bnbk` from <http://www.applied-maths.com/download/sample-data>, under 'DemoBase Connected'.

In the *BioNumerics Startup* window, press the  button, select **Restore database**, browse for the downloaded file and select **Create copy**. Specify a name and click **<OK>**.



In contrast to other browsers, some versions of Internet Explorer rename the `DemoBase_Connected.bnbk` database backup file into `DemoBase_Connected.zip`. If this happens, you should manually remove the `.zip` file extension and replace with `.bnbk`. A warning will appear ("If you change a file name extension, the file might become unusable."), but you can safely confirm this action. Keep in mind that Windows might not display the `.zip` file extension if the option "Hide extensions for known file types" is checked in your Windows folder options.

4 Creating a composite data set

Before a band matching table can be displayed in *BioNumerics*, a *composite data set* needs to be defined first, containing the fingerprint type(s) as input.

1. In the *BioNumerics Startup* window, double-click on the **DemoBase Connected** database to open it.
2. In the *Main* window, click on  in the toolbar of the *Experiment types* panel and select **Composite data set** from the list.
3. Press **<OK>**, enter a name, for example "RFLP-combined" and press **<OK>**.

The *Experiment types* panel now lists the composite data type **RFLP-combined**.

4. Double-click on **RFLP-combined** in the *Experiment types* panel.

The *Composite data type* window is shown for **RFLP-combined**.

5. Click on **RFLP1** and select **Experiment > Include experiments** (.
6. Click on **RFLP2** and select **Experiment > Include experiments** (.

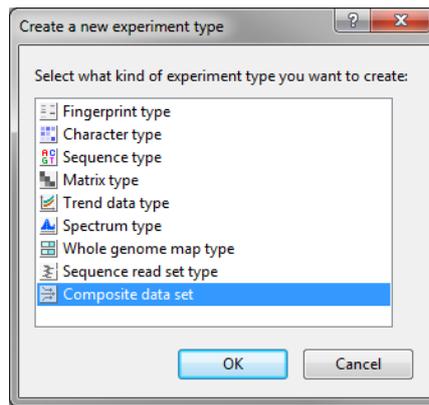


Figure 3: Create a new composite data set.

Both RFLP1 and RFLP2 are selected in the composite data set, and are marked with a green V-sign (see Figure 4).

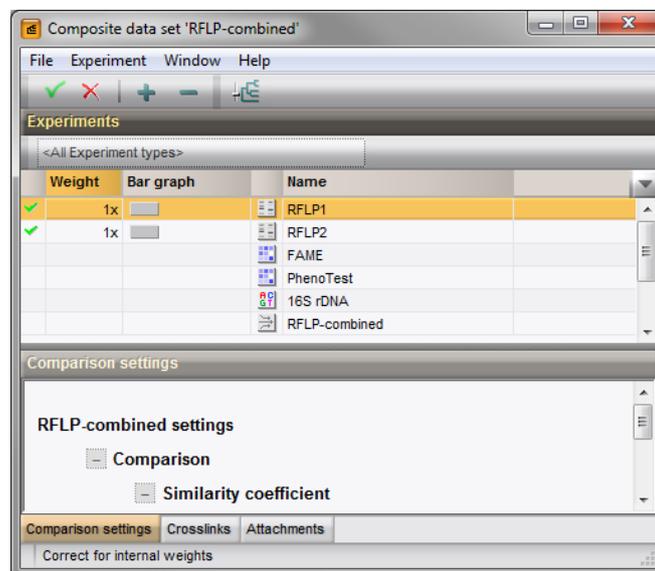


Figure 4: The Composite data set.

7. Close the *Composite data type* window with **File** > **Exit**.

5 Creating a band matching

1. In the *Database entries* panel of the *Main* window, select all entries except those defined as STANDARD: select **Ctrl+A** to select all entries and unselect the entries defined as STANDARD by unchecking the check boxes next to the entries.
2. Highlight the *Comparisons* panel in the *Main* window and select **Edit** > **Create new object...** (➕) to create a new comparison for the selected entries.
3. Click on the (◀) next to the experiment name **RFLP1** in the *Experiments* panel to display the **RFLP1** patterns.
4. Select **Fingerprints** > **Perform band matching...** (📊).

5. Select *Find classes on all entries* and press <Tolerance settings> to enter the band position tolerance settings.

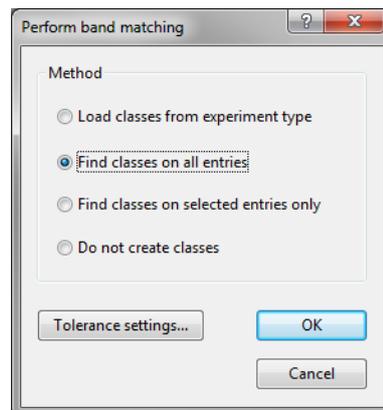


Figure 5: The *Perform band matching* dialog box.

6. Enter an optimization of 0.5%, a position tolerance of 1%, a change of 0%, and a minimum height and minimum surface of 0% (see Figure 6).

The *Position tolerance* is the maximal shift allowed (in percentage of the pattern length) between two bands allowed to consider them as matching.

The *Optimization* is a shift that you allow between any two patterns and within which the program will look for the best possible matching.

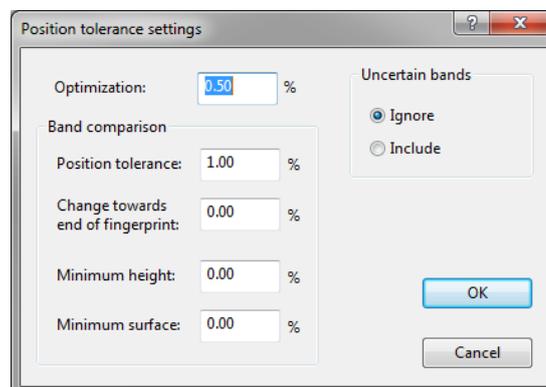


Figure 6: The *Position tolerance settings* dialog box.

7. Press <OK> twice.

The program defines the band classes and associates each band with a class. The band classes are shown as blue lines and the bands are linked to a class in red.

8. You can zoom in on the band matching, e.g. by using the zoom sliders.
9. Band classes can be manually edited to further refine the automatic band matching that was done according to the tolerance settings.

Band class labels are shown in the caption of the *Experiment data* panel. By default, the labels contain the metrics value of the band class.

10. If you double-click on a band class label or click on a band class label and select *Fingerprints > Band class information...* (**Ctrl+I**), a dialog box appears displaying detailed information about the band class (see Figure 8 for an example).

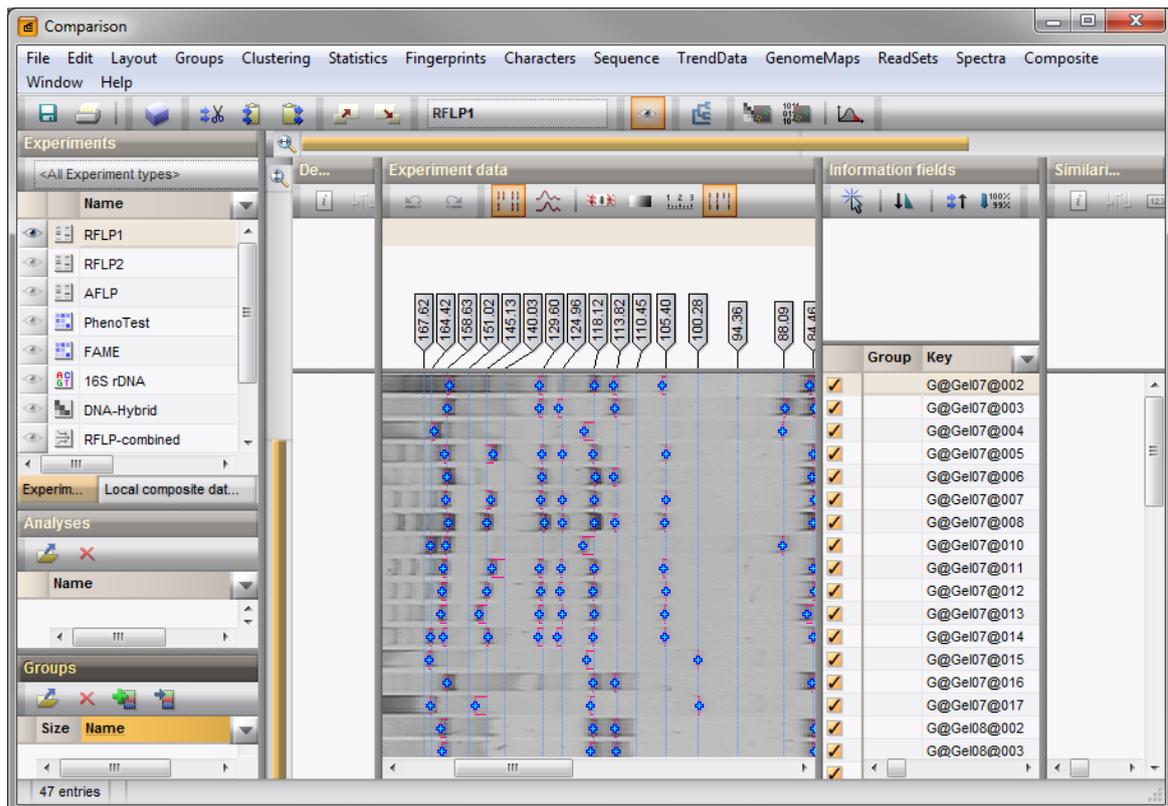


Figure 7: The *Comparison* window with band classes defined.

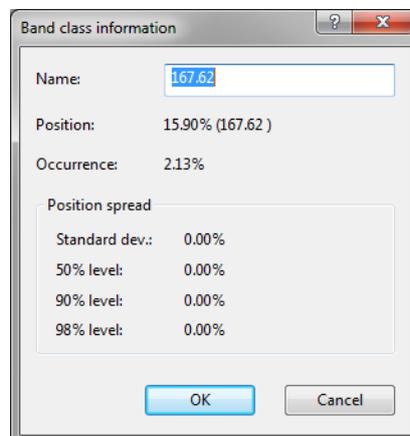


Figure 8: Band class information.

11. If desired, band classes can be renamed, e.g. to indicate a band class of specific interest.
12. A tab-delimited text file of the band matching table can be exported with *Fingerprints* > *Export band matching...*
13. The program asks "Export quantitative band information?". If you answer <Yes>, the band intensities are exported. Answering <No> exports a binary band matching table.

Found band classes can be saved to the experiment type for future use:

14. Select *Fingerprints* > *Save band classes to experiment type...* and confirm.

7 Finding discriminative bands between entries

Discriminative bands can be searched for in a band matching table. This feature works on the selected entries in the comparison and rearranges the characters in the table according to their ability to discriminate the selected entries from the unselected ones.

1. Perform a band matching for **RFLP1** as described in 5.
2. Select *Database > Entries > Unselect all entries (all levels)* (🗑️, F4) to make sure that no entries are selected.
3. In the *Comparison* window, select *Edit > Find item in list...* (🔍, Ctrl+Shift+F) and enter "verc" in the *Search for* field. Press <Select all>.

All *Vercingetorix* entries are now selected in the *Comparison* window.

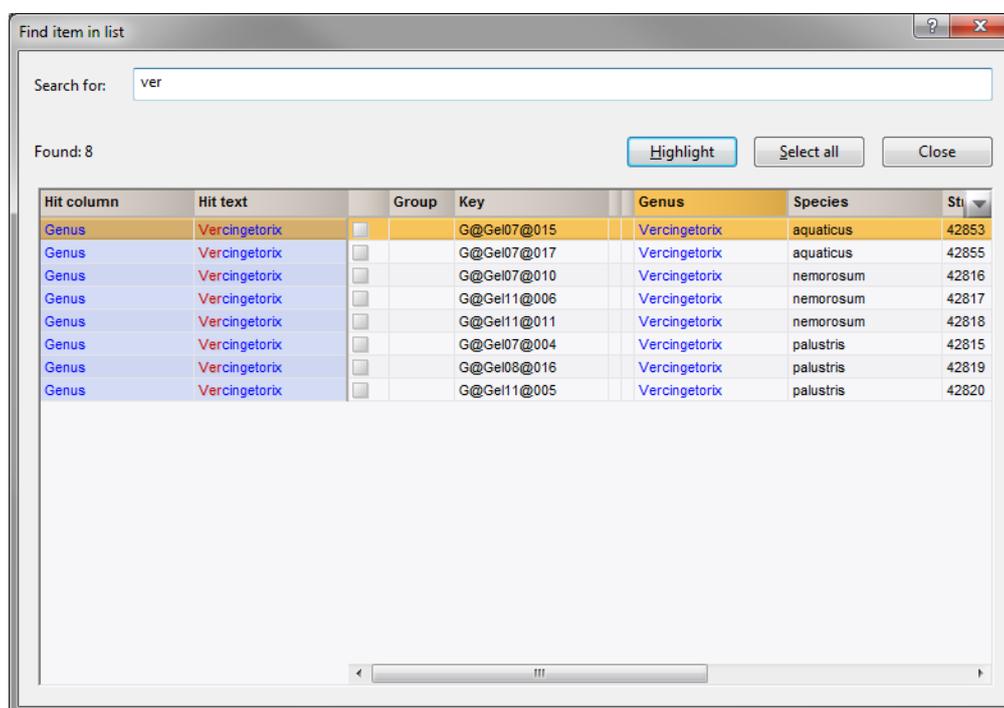


Figure 10: Search for entries.

4. In the *Comparison* window choose *Edit > Arrange entries > Bring selected entries to top* (📄, Ctrl+T) to bring the selected entries to the top of the list.
5. Display the band matching table for **RFLP-combined** by clicking on the 🗑️ next to the experiment name **RFLP-combined** in the *Experiments* panel.
6. Display the intensity of the bands as a color scale, with *Composite > Show quantification (colors)* (🎨).
7. Make sure **RFLP-combined** is selected in the *Experiments* panel and select *Composite > Discriminative characters*.

The characters (bands) are reorganized in such a way that the characters positive for the selected entries and negative for the other entries occur left, and those characters negative for the selected entries and positive for the other entries occur right (see Figure 11).

8. Select a band class in the band classes header and *Composite > Sort by character* (📄).

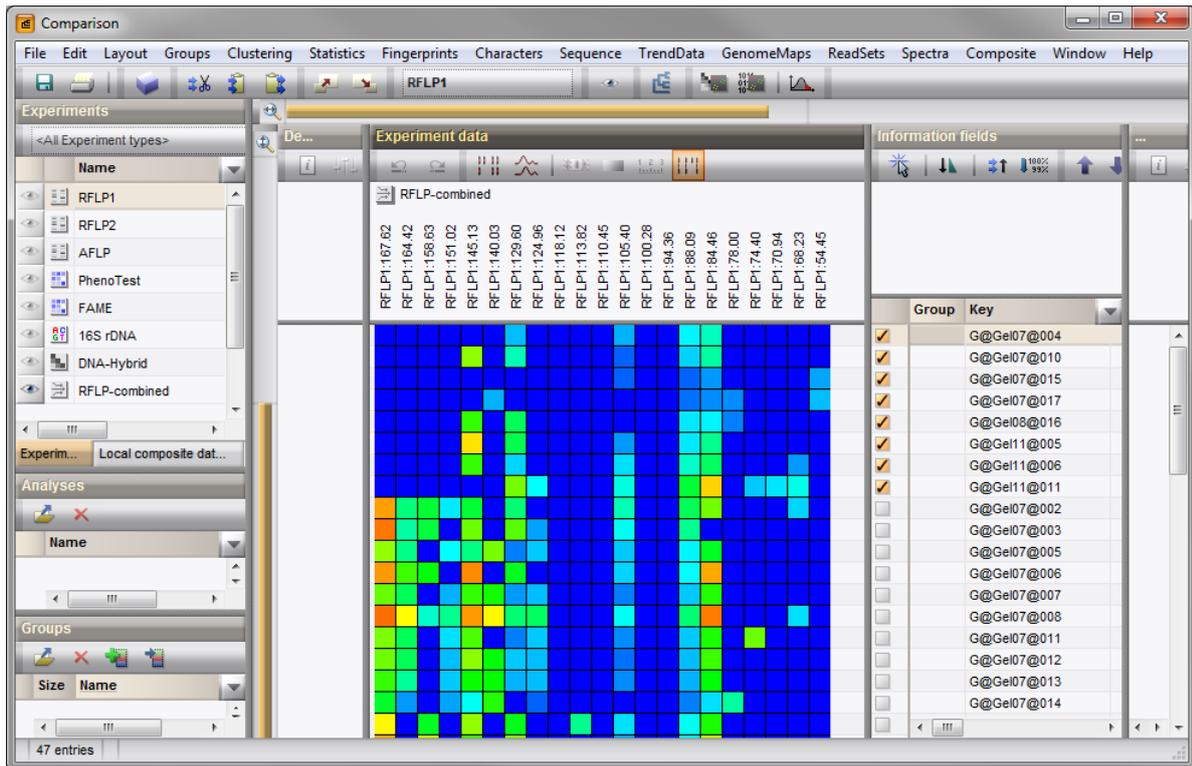


Figure 11: Find discriminative characters.

The entries are now sorted by increasing intensity of the selected band class.

9. Choose *Composite* > *Calculate clustering of characters...* (🔍).

10. In the next dialog box, select *Pearson correlation* and press <OK> to calculate a character dendrogram, which appears horizontally in the caption of the data matrix display of the composite data set.

11. Select *File* > *Save* (💾, **Ctrl+S**) to save the comparison.

12. Enter a name and close the *Comparison* window.

8 Conclusion

In this tutorial you have seen how to create a band matching table and how to look for discriminative bands. Other more sophisticated analytical tools, such as polymorphism analysis and principal components analysis can be applied on band matching tables.