

BIONUMERICS Tutorial: Combined analysis of fingerprint data

1 Aim

In BIONUMERICS it is possible to obtain a combined dendrogram from multiple experiments. In this tutorial you will learn how to create a dendrogram based on fingerprint data coming from two different experiments.

2 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 button, located in the toolbar in 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 https://www.applied-maths.com/download/sample-data, under 'DemoBase Con-

nected'. In the *BIONUMERICS Startup* window, press the \bigcirc 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.

3 Creating a composite data set of fingerprint data

A composite data set in BIONUMERICS is a "container" of experiment types, i.e. it holds data coming from one or several experiments. In a first step we need to define a *composite data set*,

containing two fingerprint types 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.

Create a new experiment type	?	×					
Select what kind of experiment type you want to create:							
Character type		^					
Sequence type							
Matrix type							
¹ ∠ Trend data type							
📥 Spectrum type							
≵ Sequence read set type							
➢ Composite data set							
		÷					
ОК	Ca	incel					

Figure 1: Create a new composite data set.

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* (\mathscr{C}).

6. Click on **RFLP2** and select *Experiment* > *Include experiments* (%).

Both RFLP1 and RFLP2 are selected in the composite data set, and are marked with a green V-sign (see Figure 2).

The bar that appears in the 'Bar graph' column displays the weights assigned to each experiment type. Weights can be increased or decreased with *Experiment* > *Increase weight* (\uparrow +) and *Experiment* > *Decrease weight* (\downarrow -) respectively.

In order to treat the two fingerprint types on an objective basis, i.e. the number of bands in the compared patterns, the following is done:

- 7. Select *Experiment* > *Correct for internal weights*. The caption now shows "Correct for internal weights".
- 8. Close the *Composite data type* window with *File* > *Exit*.
- 9. In the Database entries panel of the Main window, select all entries except STANDARD: select Ctrl+A to select all entries and use the Ctrl- key to unselect the entries defined as STANDARD. Alternatively unselect the entries by unchecking the check boxes next to the entries.
- 10. Highlight the *Comparisons* panel in the *Main* window and select *Edit* > *Create new object...* (+) to create a new comparison for the selected entries.

đ	🖆 Composite data set 'RFLP-combined' - 🗆								\times
File Experiment Window Help									
er ⊗ I ↑+ J- ItE									
Experiments									
<all experiment="" types=""></all>									
	Weight	E	Bar graph		Name				•
~	•	1x			RFLP1				^
~	•	1x			RFLP2				
					AFLP				
					PhenoTest				
					FAME				
				Å C	16S rDNA				
				Ъ.	DNA-Hybrid				
				₩.					
Comparison settings									
Г									
	DEL D combined settings								
			incu settings						
	- Comparison								
	- Similarity coefficient								
	Average from experiments								
								\sim	
	- Cluster analysis								
Comparison settings Crosslinks Attachments									
	No internal	wei	ghts						

Figure 2: The Composite data set.

4 Combined analysis of fingerprint data

4.1 Averaging the individual matrices

- 1. Select the experiment **RFLP-combined** in the *Experiments* panel.
- 2. Select Clustering > Calculate > Cluster analysis (similarity matrix)....

When selecting the option **Average from experiments**, the similarity matrices of the individual experiments (in this example **RFLP1** and **RFLP2**) are calculated first using the comparison settings defined for each individual experiment (double-click on the fingerprint type in the *Experiment types* panel and select **Settings** > **Comparison settings...** (HE) to access the settings). From these matrices, a combined matrix is calculated by averaging the values according to the defined weights.

3. Select *Average from experiments* and press <*Next*> and <*Finish*>.

The dendrogram is displayed in the *Dendrogram* panel (see Figure 3).

4. To view the settings used for the calculation of the individual matrices, select *Clustering* > *Show information* (\blacksquare).

4.2 Creating a combined character matrix

In BIONUMERICS it is also possible to merge the characters (i.e. bandclasses in case of fingerprint data) from different experiments, and calculate a similarity matrix based on this merged character table, resulting in a combined dendrogram.

In order to combine fingerprints in one character table, we need to perform a bandmatching on

Comparison					- U ×
File Edit Layout Groups Clustering	Statistics Fingerprints Characters Sequence TrendD	ata ReadSets Spectra	Composite Window Help		
B@ 6 ;x 17 C;	RFLP-combined ◎				
Experiments	(Q)				
<all experiment="" types=""></all>	Condrogram	Experiment data	Information fields		Similarities
Name 🔫	이 밖 집 단 년 년	123 ABC	€, ↓ \$↑ ↓₩ ↑ ↓		i 114 mm
RFLP1		권 RFLP			
C RFLP2					RFLP-combined
AFLP	RFLP-combined				0 20 40 60 80 100
PhenoTest	40 80 80 100		Group Key Genus	Species 🔻	
C FAME			G@Gel07@017 Vercingetorix G@Gel07@004 Vercingetorix	aquaticus palustris	
16S rDNA			G@Gel11@005 Vercingetorix G@Gel08@018 Vercingetorix	palustris palustris	
DNA-Hybrid			G@GeI07@010 Vercingetorix G@GeI11@008 Vercingetorix	nemorosum	
			G@Gel07@005 Ambiorix G@Gel03@005 Perdrix	sp. pseudoarchaeus	
< >			G@Gel08@005 Perdrix G@Gel08@015 Perdrix	pseudoarchaeus pseudoarchaeus	
Analyses Local composite datasets			G@Gel08@002 Perdrix G@Gel08@003 Perdrix	pseudoarchaeus	
FÅ @			G@Gel08@007 Perdrix G@Gel08@007 Perdrix	pseudoarchaeus pseudoarchaeus	
			G@Gel09@008 Perdrix G@Gel09@009 Perdrix	pseudoarchaeus pseudoarchaeus	
A DELD combined			G@Gel09@007 Perdrix G@Gel09@002 Perdrix	pseudoarchaeus	
E REFECCIONDINES			G@Gel09@005 Perdrix G@Gel09@005 Perdrix	pseudoarchaeus aberrans	
			G@Gel07@002 Ambiorix G@Gel07@003 Ambiorix	svivestris aberrans	
			G@Gel08@013 Ambiorix G@Gel08@014 Ambiorix	sylvestris sylvestris	
			Gig/Gel11g002 Ambiorix G/g/Gel7/g007 Ambiorix G/g/Gel11g009 Ambiorix	sp. sylvestris	
			GRIGel07@005 Ambiorix GRIGel07@011 Ambiorix	sylvestris sylvestris	
			G@Gel07@012 Ambiorix G@Gel11@010 Ambiorix	sylvestris aberrans	
Groups			G@Gel03@010 Ambiorix G@Gel11@003 Ambiorix	sylvestris sylvestris	
			G@Gel09@010 Ambiorix G@Gel09@011 Ambiorix	sylvestris	
Size Name 🔽			G@Gel09@004 Ambiorix G@Gel07@008 Ambiorix	sylvestris sp.	
			G@Gel11@008 Ambiorix G@Gel07@014 Ambiorix	sylvestris	
			G@Gel08@008 Perdrix G@Gel08@012 Perdrix	50. 50.	
			<	>	< > v
47 entries in comparison 47 entries sele	ected in database				

Figure 3: The Comparison window.

each of the fingerprint type experiments (RFLP1 and RFLP2 in our example):

- 5. Select **RFLP1** in the *Experiments* panel and select *Fingerprints* > *Perform band matching...* (||||).
- 6. Select *Find classes on all entries* and press < OK >.
- 7. Select **RFLP2** in the *Experiments* panel and select *Fingerprints* > *Perform band matching...* (||||).
- 8. Select *Find classes on all entries* and press <*OK*>.
- 9. Click on the onext to the experiment name **RFLP-combined** in the *Experiments* panel to display the binary band matching table of RFLP1 and RFLP2 (see Figure 4).
- 10. Select *Clustering* > *Calculate* > *Cluster analysis (similarity matrix)....*

To calculate a similarity matrix based on the merged band matching table choose a coefficient under *Character data*. Binary coefficients measure the similarity based upon common and different bands; numerical coefficients use the band intensities for the calculation of the similarity.

11. Select a binary coefficient, e.g. *Dice* and press < *Next* > and <*Finish* >.

The resulting dendrogram is displayed in the *Dendrogram* panel (see Figure 4).



Figure 4: The Comparison window.