BIOMÉRIEUX

BIONUMERICS Tutorial:

Importing and processing a DGGE gel image

1 Aim

Comprehensive tools for the processing of electrophoresis fingerprints, both from slab gels and capillary sequencers are incorporated into BIONUMERICS. In this tutorial we will focus on the import and processing of gel images. Processing tools include lane finding, normalization of the patterns, band finding, etc.

2 Sample data

As an exercise, we will import 8-bit TIFF gel images containing DGGE profiles of bacterial 16S rDNA gene fragments that were amplified using Proteobacteria specific primers.

We will also import information from an external text file containing information about the DGGE profiles (see Figure 1).

Figure 1: Sample information.

The gel images and the text file containing the sample information can be downloaded from the Applied Maths website: go to https://www.applied-maths.com/download/sample-data and click on "DGGE sample data".

3 Preparing the database

1. Create a new database (see tutorial "Creating a new database") or open an existing database.

First we will import the sample information from the external text file `Sample information.txt` (see Figure 1) in our new, empty database.

2. Select **File** > **Import...** (, **Ctrl+I**) to call the *Import* dialog box.

3. In the Import tree, expand **Entry information data**, highlight **Import fields (text file)** and press <**Import**> (see Figure 2).
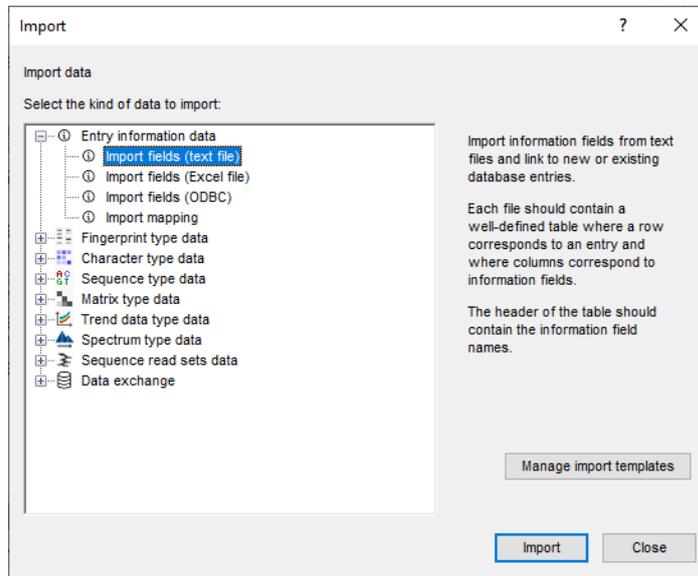
Figure 2: Import tree.

4. Browse for the `Sample information.txt` file in the `DGGE sample data` folder and press <**Next**> (see Figure 3).
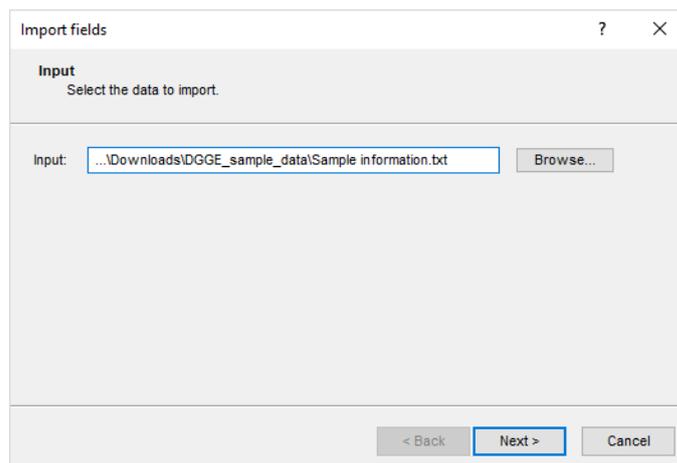
Figure 3: Select the file to import.

This brings up a new dialog box (see Figure 4). Each column in the selected file corresponds to a row in the grid (column 1 in the file corresponds to row 1 in the grid, column 2 corresponds to row 2, etc.).

5. Select the first row in the grid, press the <**Edit destination**> button and select the **Key** field from the list. Press <**OK**>.

The grid is updated (see Figure 4).

6. Highlight the three other external fields in the grid panel (Pond, Sampling site, Sampling date) using the **Ctrl**- or **Shift**-keys. Press the <**Edit destination**> button and select the **Entry info field** option from the list. Press <**OK**>.

7. Press <**OK**> once more to accept the default suggested names and press <**Yes**>.

The grid is updated (see Figure 4).
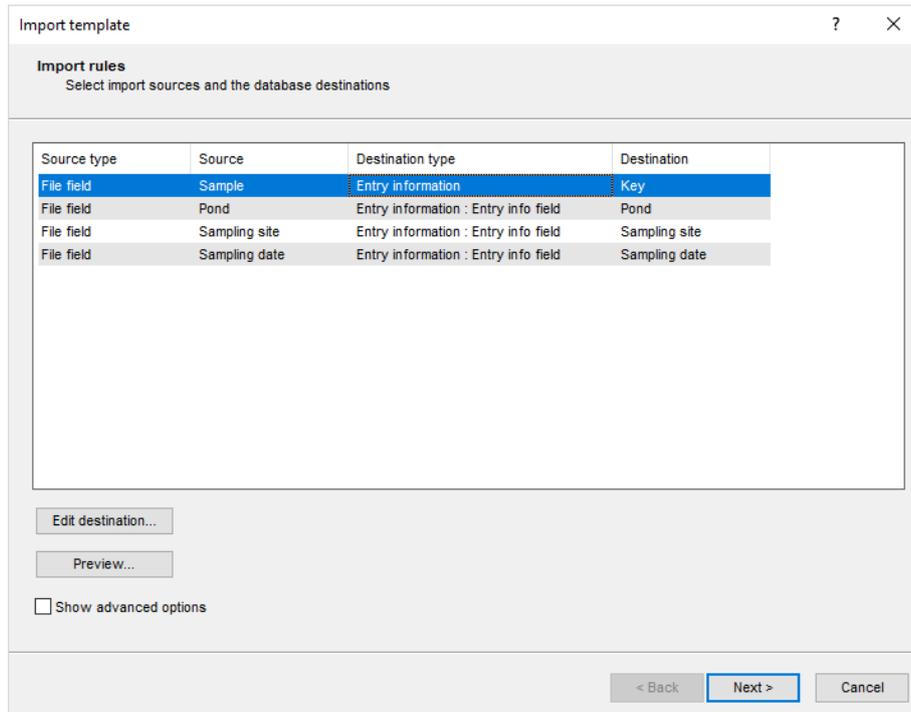
Figure 4: Create a new import template.

8. Press <**Next**> and press <**Finish**>.

9. Specify a template **Name** (e.g. "Import DGGE sample information") and press <**OK**>.

The import template is added to the list and is automatically selected.

10. Press <**Next**> and press <**Finish**>.

24 new entries are imported, each with four database information fields filled in (see Figure 6).

Before we will import a gel image file in this database, we will first create a fingerprint type experiment:

11. In the *Main* window, click on + in the toolbar of the *Experiment types* panel and select **Fingerprint type** from the list (see Figure 5).

12. Press <**OK**>, enter a name, for example **DGGE proteobacteria primers** and press <**Next**>.

13. In the next window, make sure that **Two-dimensional TIFF files** and **8-bit (256 gray values)** are selected and press <**Next**> twice and press <**Finish**> to complete the creation of the new fingerprint type.

The *Experiment types* panel now lists the fingerprint type **DGGE proteobacteria primers** (see Figure 6).
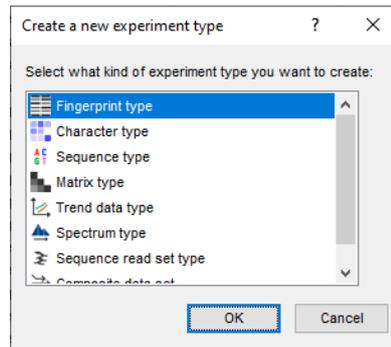
Figure 5: The *Create a new experiment type* dialog box.

You can still change the resolution, inversion, and background subtraction settings later when processing a gel.
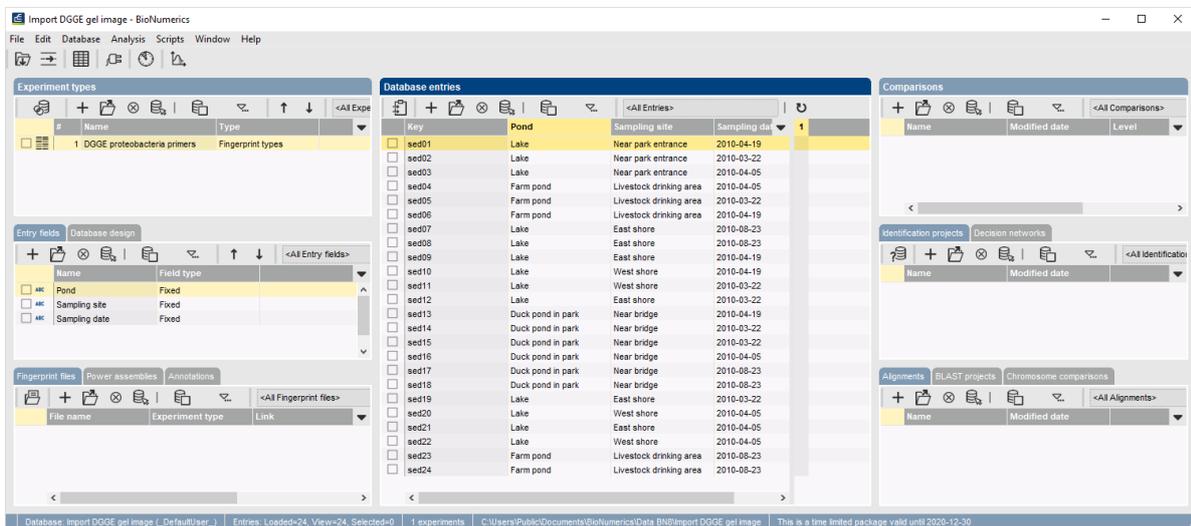
Figure 6: The *Main* window after import of the data.

4 Importing a fingerprint gel file

1. To add a new fingerprint file to the database, select **File** > **Import...** (⬚, **Ctrl+I**) and select **Import gel file** under **Fingerprint type data** (see Figure 7).

2. Press <**Import**> to call the *Select fingerprint file* dialog box.

3. Select the file `Proteobacteria_gel1.tif` in the `DGGE sample data` folder (see Figure 8).

The dialog also asks if you want to edit the image by opening the image editor. Uncheck the **Open in image editor** option if you are sure the file is an uncompressed gray scale TIFF image. For the conversion to an uncompressed gray scale TIFF file make sure the option is checked.

4. Since the example file is an uncompressed gray scale TIFF file, uncheck the option **Open in image editor** and press <**OK**>.

The gel image is now available in the *Fingerprint files* panel (see Figure 9). The file is marked with a red "N" (⬚) indicating that the image has not been processed yet.
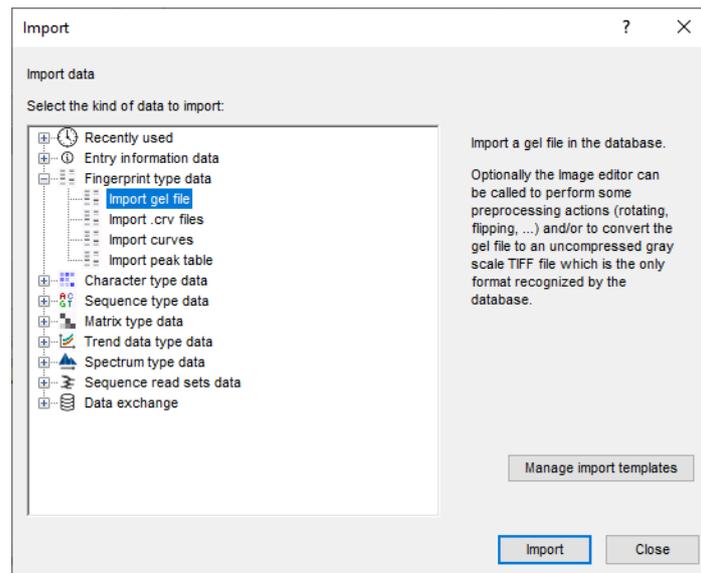
Figure 7: Import tree.
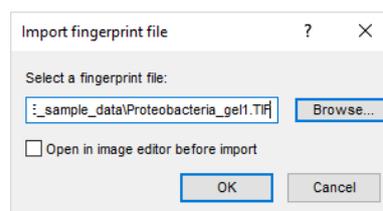
Figure 8: Select the fingerprint image.
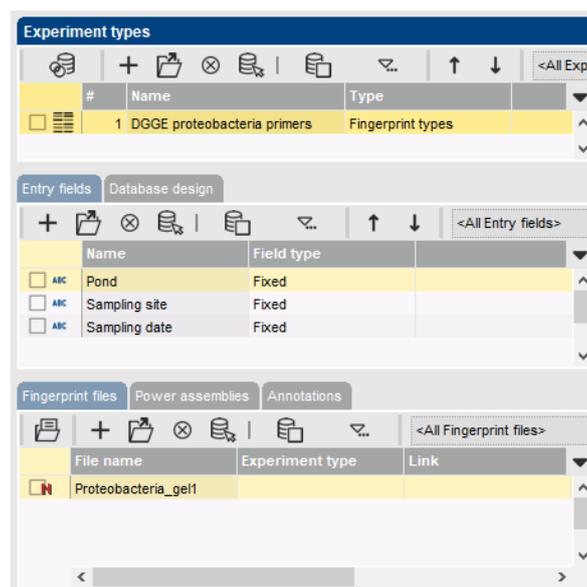
Figure 9: Unprocessed gel image added to the database.

5 Processing a fingerprint gel file

5.1 Introduction

Before linking our fingerprint lanes to entries in our database, we must define and process the lanes in the *Fingerprint processing* window.

1. In the *Fingerprint files* panel, double-click on the filename **Proteobacteria_gel1** to open the *Fingerprint* window.

2. In the *Fingerprint* window, select **File** > **Edit fingerprint data...** (▦) to open the *Fingerprint processing* window.

3. In the next dialog box, select **DGGE proteobacteria primers** and press <**OK**>.

The *Fingerprint processing* window opens. In this window we are going to process our TIFF file in 4 steps:

1. **Strips** (defining lanes)

2. **Curves** (extracting densitometric curves)

3. **Normalization**

4. **Bands** (defining bands and quantification)

The steps are shown in the tabs in the bottom left corner of the window. The ◄ and ► buttons in the toolbar are to move back and forth, respectively. The tabs can also be used for the navigation between the different steps.

5.2 Step 1: Define strips

The imported gel image is surrounded by a green rectangle. Since BIONUMERICS recognizes the darkness as the intensity of a band, make sure the bands appear as dark bands on a white background.

✎ In case the bands appear as white bands on a black background, invert the values: press ⇅ to open the *Fingerprint processing settings* dialog box, check/uncheck the **Inverted values** check box, and press <**OK**> to apply the changes.

The first step in processing a gel is to crop the image to remove empty space, and to define the lanes.

4. Delineate the area of the gel lanes by clicking and dragging the nodes of the rectangle to adjust it (see Figure 10).

5. In the toolbar, press ⊞ to let the software search for the individual lanes.

6. Enter "15" as the approximate number of lanes and press <**OK**>.

7. In the toolbar, press ⇅ to open the *Fingerprint processing settings* dialog box.

When opened in the *Strips panel* the *Raw data tab* is selected by default.

8. Adjust the **Thickness** of the image strips so the splines border the edges of the bands (e.g. **45** points).
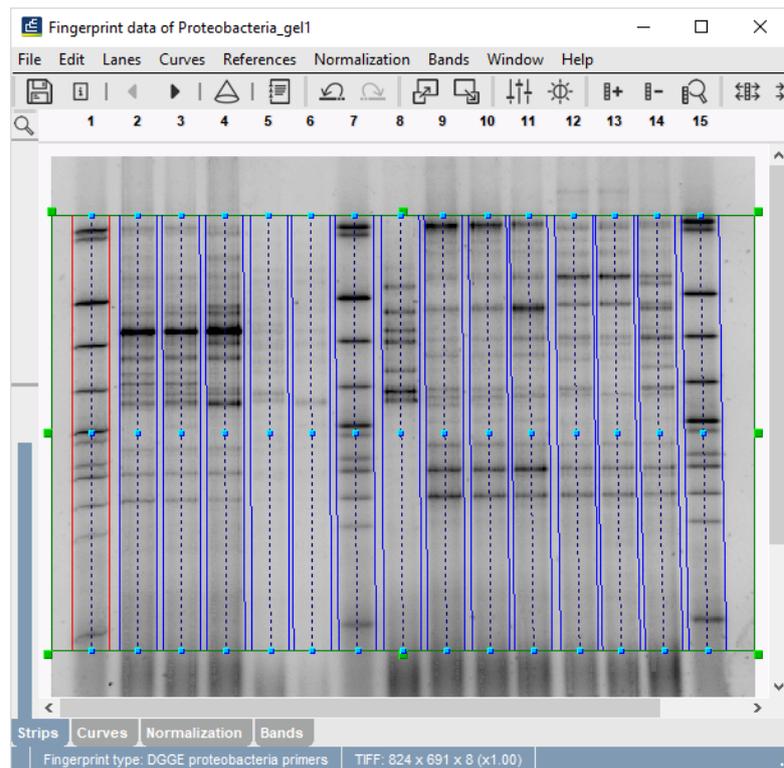
Figure 10: Defining strips.

9. Press <**OK**>.

10. Adjust the horizontal position of each spline as necessary by clicking and dragging a blue node. Hold down the **Shift**-key while dragging a node to adjust the spline locally.

Next, we will edit the tone curve to improve the band visibility.

11. From the pull down menu, select **Edit** > **Edit tone curve...**.

12. Press <**Linear**>.

The visible gray scale interval now ranges between the minimum and maximum values in the image (see Figure 11).

13. Press <**OK**> to save changes to the tone curve settings.

The gel tone curve does not change the image's densitometric values, only the way they appear in the *Fingerprint processing* window.

14. Press 🖫 to save the work already done.

15. Press ▶ to proceed to the **Curves** step or press the *Curves tab*.

5.3   Step 2: Define curves

Now that the lanes have been defined, BIONUMERICS can generate densitometric curves describing the optical density across the spline along each lane. The left panel shows the strips extracted from the image file, the right panel shows the densitometric curve of the selected pat-
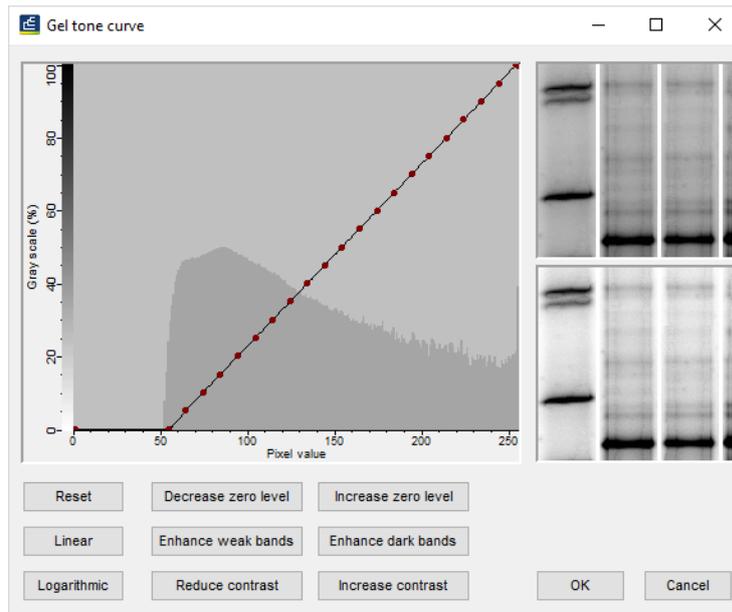
Figure 11: The *Gel tone curve* window.

tern. The area between the blue lines of each lane will be used to calculate the densitometric curve.

16. Open the *Fingerprint processing settings* dialog box again (click on the ⫶↕⫶ button).

17. Change the **Averaging thickness** for curve extraction to **15** in the *Densitometric curves tab*.

18. Check **Median filter** and press <**OK**>.

BIONUMERICS can analyze the gel image to determine the optimal settings for removing background noise from the densitometric curves.

19. Select **Curves** > **Spectral analysis...**.

The following settings are shown in the *Spectral analysis* window:

1. **Wiener cutoff scale**: Determines the optimal setting for least square filtering.

2. **Background scale**: Estimation of the disk size for background subtraction.

20. Close the *Spectral analysis* window.

21. Open the *Fingerprint processing settings* dialog box again (click on the ⫶↕⫶ button).

22. Check **Apply least square filtering** and specify a **Least square filtering Cut off** as indicated by the Wiener cutoff scale in the *Spectral analysis* window (use the percentage value, e.g. **1.0**). Least square filtering removes very small peaks from the curves.

23. Check **Apply** in the *Background subtraction panel* and specify a **Background subtraction disk size** as indicated by the background scale in the *Spectral analysis* window (use the percentage value, e.g. **14**). Background subtraction removes large background trends from the curves.

24. After having specified the settings in the *Densitometric curves tab* (see Figure 12) press <**OK**>.

The background noise has been removed from the curves (see Figure 13).
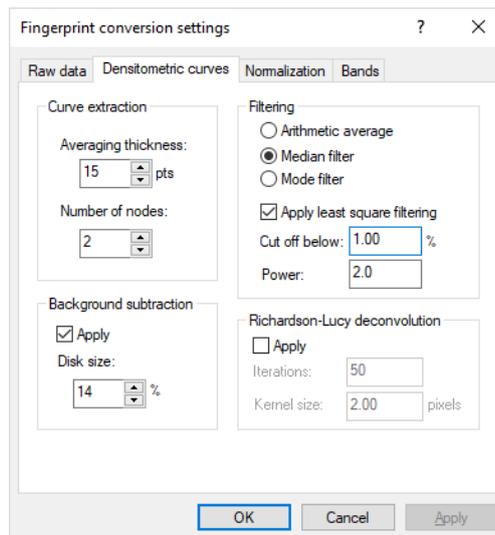
Figure 12: The *Fingerprint processing settings* dialog box: the *Densitometric curves tab*.
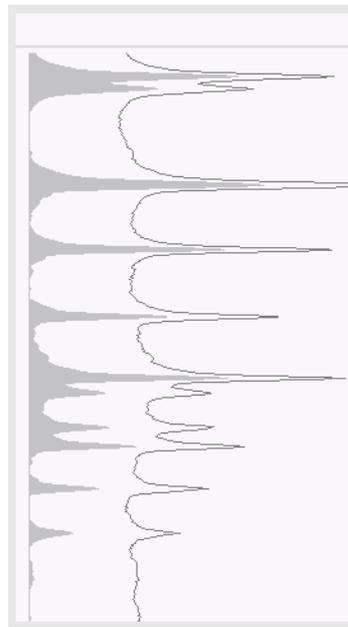

Figure 13: Curves after filtering.

25. Press 🖫 to save the work already done.

26. Press ▶ to proceed to the **Normalization** step or press the *Normalization tab*.

5.4   Step 3: Normalize the gel

Every fingerprint type experiment needs at least one reference system to normalize its gels. Since this is the first DGGE gel we have imported, we need to create a reference system based on the marker pattern(s) contained in this gel. Subsequent DGGE gels, provided they contain the same marker pattern, can be normalized with the same reference system.

27. Press ⫼ to enter the normalized view.

For now, the "normalized view" looks the same as the original view. In a first step we will define the marker lanes and the marker bands.

28. Select marker lane **1** and press (icon) to assign it as a marker lane or select **References** > **Use as reference lane**.

29. Repeat this for marker lane **7** and marker lane **15**.

30. Choose the most suitable marker lane for creating the reference system. For this exercise select lane **7** by clicking anywhere in the lane so that the lane number turns orange.

31. Right-click on the top band in lane **7** and select **Add external reference position**.

Each band in my DGGE marker lane corresponds to a strain. I will name my bands 1 to 12:

32. Enter **1** and press <**OK**>.

(pencil icon) In case you are working with reference patterns for which you know the molecular weights, enter the molecular weight for each reference band but without the unit (e.g. enter **582.6** for a band with size 582.6 kb). This way the calibration curve can be calculated automatically. See tutorial "Importing and processing a PFGE gel image" for more information.

33. Repeat the process for each band in lane **7**.

The reference system for experiment type **DGGE proteobacteria primers** is now defined (see Figure 14).

34. Select **Normalization** > **Auto assign...** ((icon)).

35. Make sure **Using bands** is selected and press <**OK**>.

36. Carefully inspect the assignments made (see Figure 14).
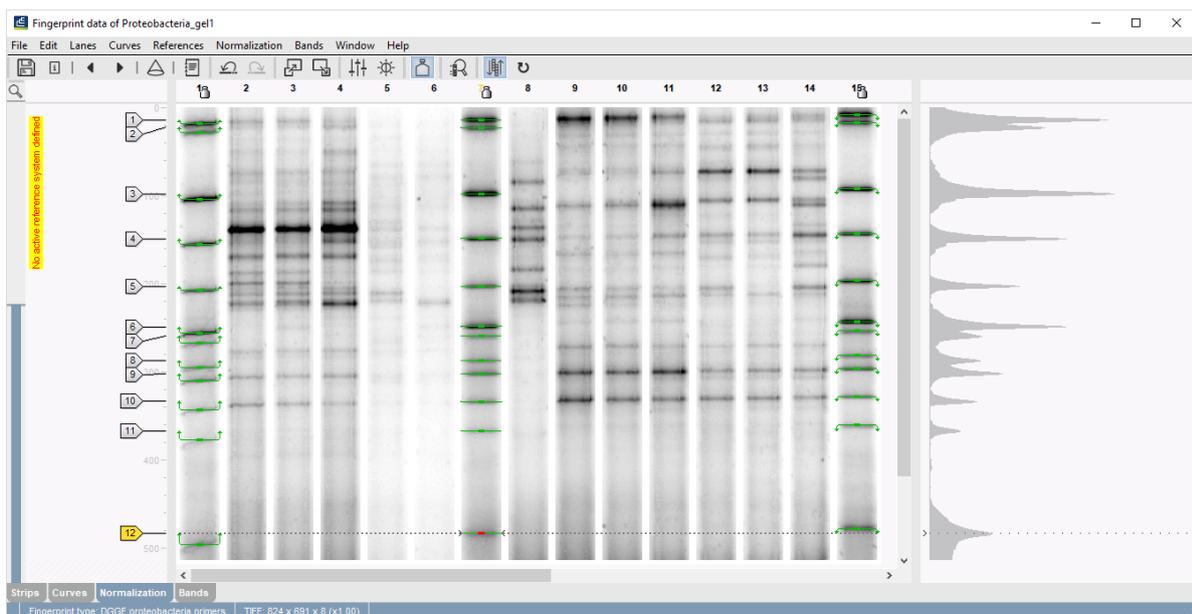
Figure 14: Reference positions assigned.

37. If a band assignment is incorrect, select the band and press the **Del**-key.

38. To assign a reference band manually, first click on the reference position tag, then hold the **Ctrl**-key and click on the reference band to assign it to that reference position.

39. To update the normalization based on the band assignments, select ***Normalization*** > ***Update normalization*** (↻ , **Ctrl+U**).

40. Press ▶ to proceed to the last step. Alternatively press the *Bands tab*.

5.5   Step 4: Define bands

If you want to use the curves to compare the patterns, no bands need to be assigned and this last step can be skipped. If you want to compare the patterns using bands, you will need to assign bands in the sample lanes in this last step. Usually, assigning bands in the sample lanes is done first with the software's automatic band search, followed by manual corrections. Some trial and error might be required to find the best settings for the automatic band search.

41. To automatically search for bands, press ℞ or select ***Bands*** > ***Auto search bands***.

In the *Band search* dialog box, the currently selected lane is shown along the bottom (see Figure 15).
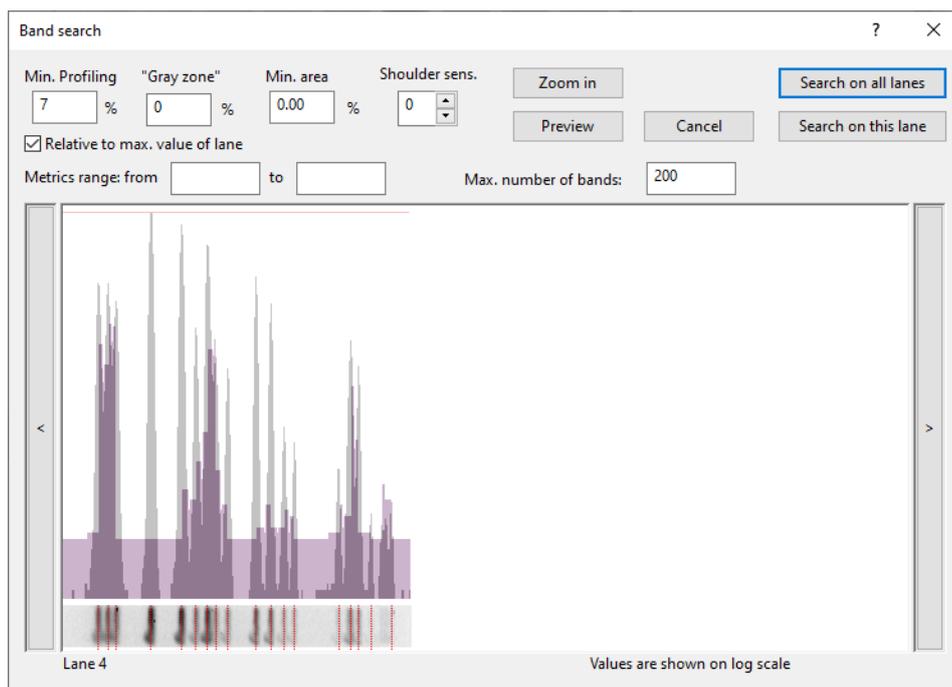
Figure 15: The *Band search* dialog box.

42. To scroll through other lanes, press the < and > buttons on the left and right sides of the curve.

43. Enter **7** for ***Min. Profiling***, **0** for ***Gray zone*** and press <***Preview***>.

The red lines across the lane at the bottom of the window indicate which peaks will be assigned bands with the current settings. The purple area along the curve indicates the minimum profile above which a band is assigned. If a gray zone has been defined (not here), the blue area indicates the zone in which bands are considered "uncertain".

44. Press <***Search on all lanes***> to execute the band search with these settings.

Bands that are found are marked with a green horizontal line. If you see any incorrect band assignments you can edit them manually:

- To add a band, hold down the **Ctrl**-key and click on the spot. The cursor automatically jumps to the closest peak; to prevent this, hold down the **Tab**-key while clicking.

- To select a group of bands, hold down the **Shift**-key and click while dragging the mouse pointer diagonally across the bands.

- To delete one or more selected bands, press the **Del**-key.

- To mark a band as uncertain, click on the band and select **Bands** > **Mark band(s) as uncertain** or press **F5**. To mark a band as certain, click on the band and select **Bands** > **Mark band(s) as certain** or press **F6**.

45. After you are satisfied with the band assignments, press 🖫 to save the file.
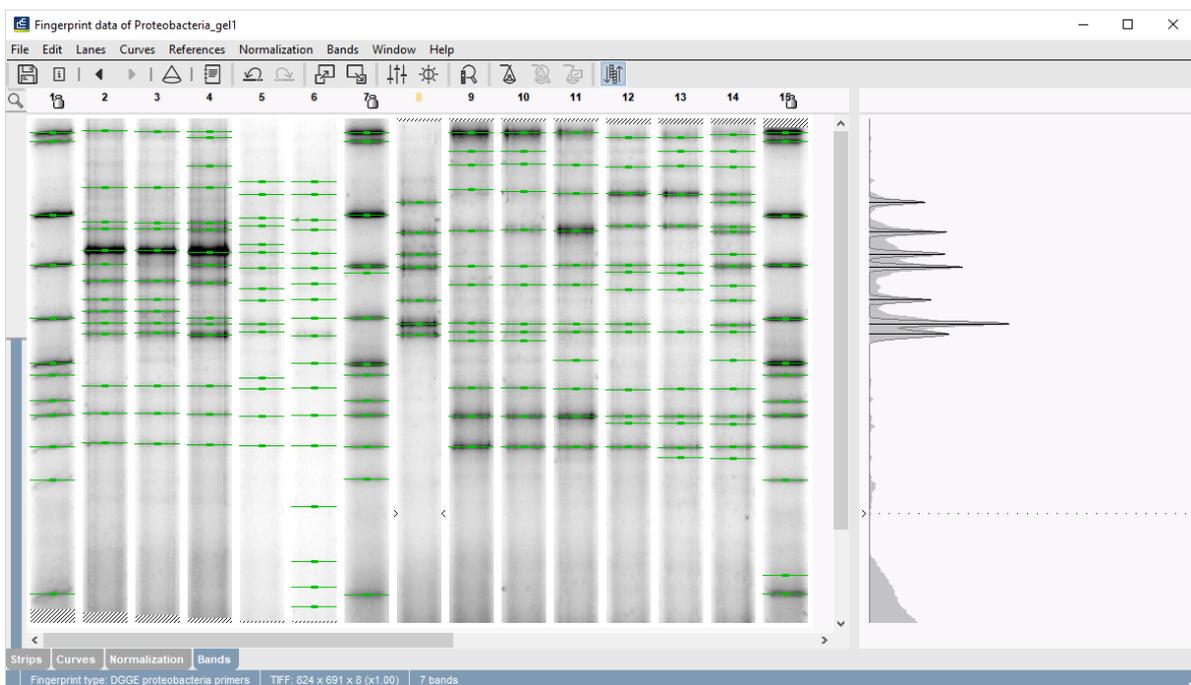
Figure 16: The *Bands tab*.

46. The program may prompt with the following question: "The resolution of this gel differs considerably from the normalized track resolution. Do you wish to update the normalized track resolution?" If the question appears, answer <**Yes**>.

47. Exit the *Fingerprint processing* window by selecting **File** > **Exit**.

48. The software asks: "Settings have been changed. Do you want to use the current settings as new defaults?" Select <**Yes**> so that the settings used for this gel will be saved in the fingerprint type settings.

6   Linking fingerprint data to entries

Although we have created individual fingerprint lanes from our gel image, the software does not know which lanes correspond to which entries in the database.

1. In the *Fingerprint files* panel, double-click on the filename **Proteobacteria_gel1** to open the *Fingerprint* window (see Figure 17).

There are several ways to link the lanes in the database:

- Selecting **Database** > **Add all lanes to database** adds all unlinked lanes as new entries to the database, with the gel lanes linked to these entries. The entries are assigned a unique key.

- Lanes can be linked to new or existing entries one by one using **Database** > **Link lane...** ( ).

- A lane can be linked to an existing database entry by dragging the gray arrow icon to the entry key in the database window. When a lane is linked, the icon becomes purple: .

2. Drag the gray arrow icon of lane **2** to the entry with key **sed01** in the *Database entries* panel and release the mouse.

3. Continue linking the remaining non-reference lanes from **Proteobacteria_gel1** to the database entries as shown in Figure 17.
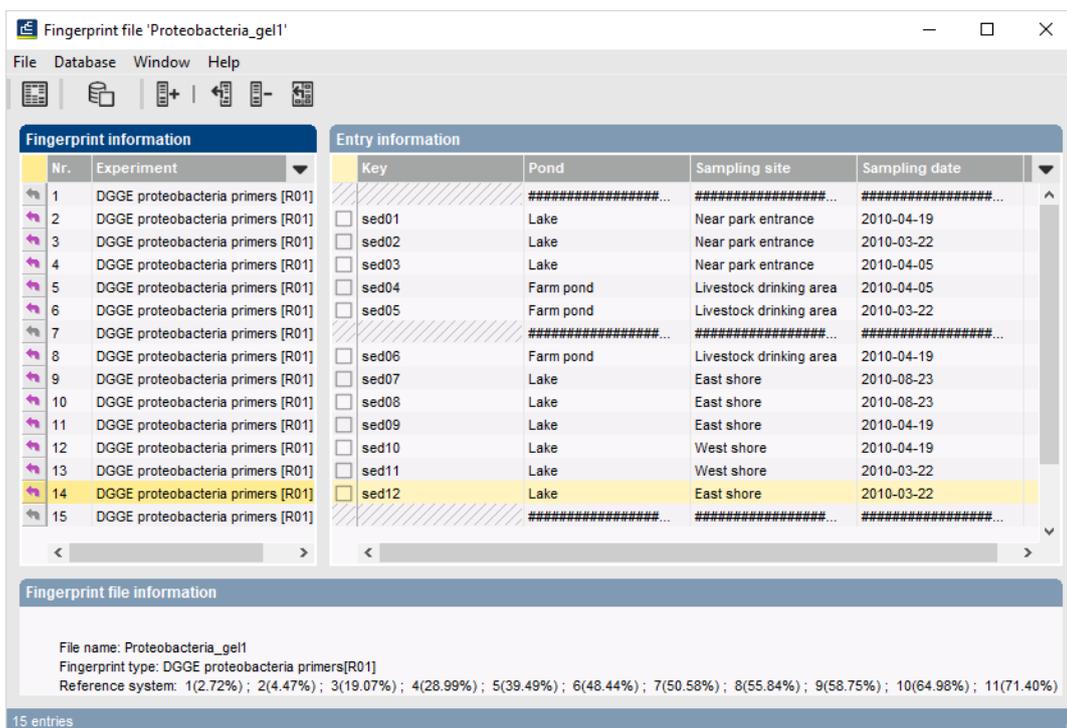
Figure 17: Fingerprint file with linked lanes.

4. After linkage, close the *Fingerprint* window and open the gel strip for one of the entries in the database by clicking on a colored dot in the *Experiment presence* panel.

The card is displayed in raw mode (i.e. not normalized). The band sizes are shown as relative distances from the top (see Figure 18).
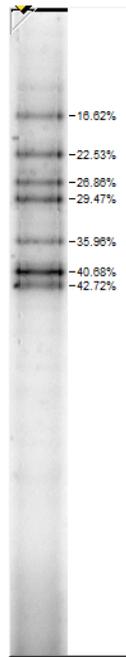
Figure 18: Fingerprint experiment card.

5. Close the card by clicking in the small triangle-shaped button in the upper left corner.

7 Fingerprint type experiment settings

1. Open the fingerprint experiment type **DGGE proteobacteria primers** by double-clicking on the experiment type in the *Experiment types* panel.

The *Fingerprint type* window shows the defined reference positions in relation to the distance on the pattern in percentage. The reference system is called **R01**. The panel left from the reference system is still blank: the fingerprint still misses a standard pattern. We will link a **Standard** pattern (e.g. lane **7** of **Proteobacteria gel1**, i.e. the one used for defining the reference system) to the **DGGE proteobacteria primers** fingerprint type as follows:

2. Close the *Fingerprint type* window.

3. In the *Fingerprint files* panel, double-click on **Proteobacteria gel1** to open the *Fingerprint* window.

4. In the *Fingerprint* window, add lane **7** to the database by selecting it and pressing ⬅▤ .

5. In the dialog box, enter "marker" and press <***OK***>.

6. Select <***Yes***> to create the new entry in the database.

A new entry **marker** is created with the pattern from lane **7** of **Proteobacteria gel1** linked to it.

7. Close the *Fingerprint* window.

8. Open the fingerprint experiment type **DGGE proteobacteria primers** by double-clicking on the experiment type in the *Experiment types* panel.

9. Drag the ▨ next to **Standard** in the *Settings panel* to the **marker** database entry and release the mouse.

The standard pattern is now displayed in the panel next to the reference positions, and the database entry key **marker** is indicated as the **Standard** (see Figure 19).
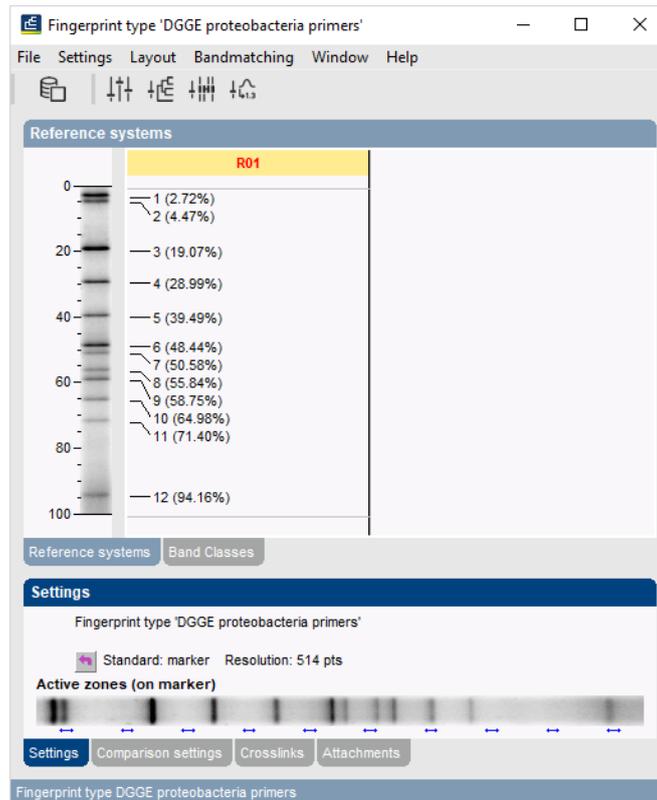
Figure 19: The *Fingerprint type* window.

If a standard pattern is assigned to a fingerprint type, this standard pattern is shown in the *Normalization tab* of the *Fingerprint processing* window to make visual assignment of bands to the reference positions easier. The choice of a standard has no influence on the normalization process since it is only used as visual aid.

8 Importing and processing a second gel file

A second gel file, `Proteobacteria_gel2.tif`, which was generated by DGGE, is present in the `DGGE sample data` folder. For this second DGGE gel, you can use the same fingerprint type, reference system, and conversion settings used for the first gel (see 5).

Because the same reference system is present in the **Proteobacteria_gel2.tif** gel, **you can skip the assignment of the reference positions**: in the third step you just have to assign the reference lanes (lane 1, 8, 15 and 16), execute ***Normalization*** > ***Auto assign...*** ( ) and inspect the assignments made.

After the gel processing, the lanes of the second gel can be linked to existing entries in the *Database entries* panel.

1. Drag the gray arrow icon of lane **2** to the entry with key **sed13** in the *Database entries* panel and release the mouse.

2. Continue linking the remaining non-reference lanes from **Proteobacteria_gel2** to the database entries as shown in Figure 21.
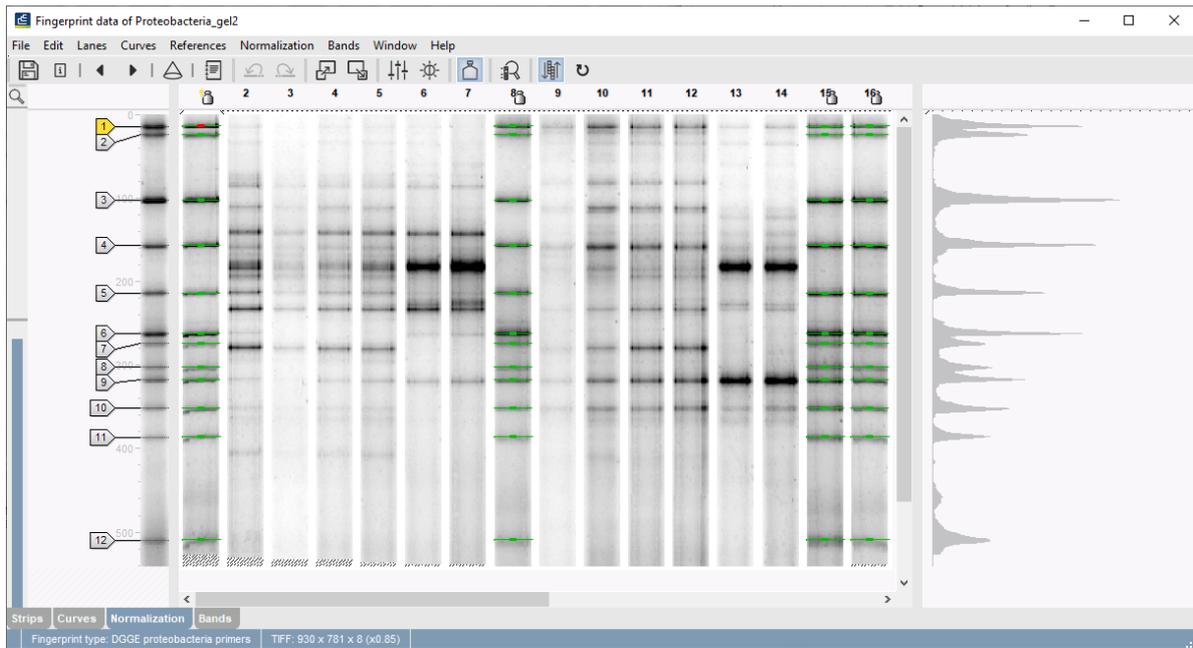
Figure 20: Normalization of the second gel.
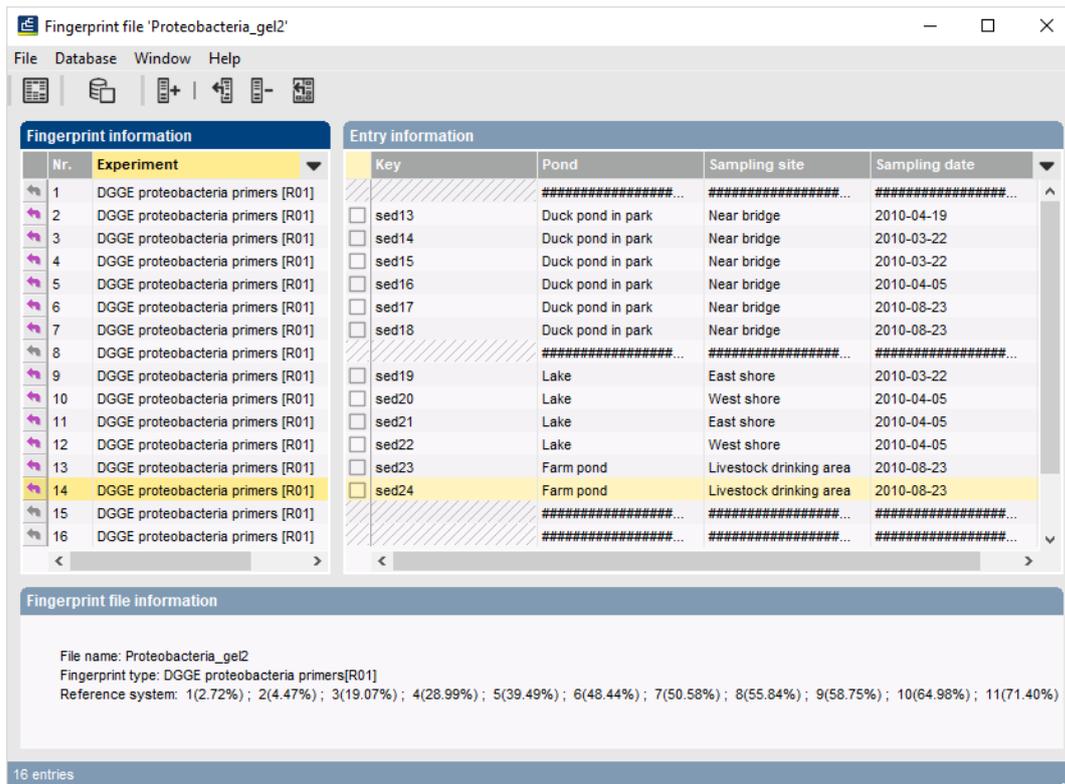
Figure 21: Fingerprint file with linked lanes.

9 Calculating an index of diversity

If one is interested in measuring biodiversity, BioNumerics allows the calculation of various diversity indices (e.g. Simpson's or Shannon-Wiener index of diversity) by means of a script.

First we will add a new field to our database for the storage of the index:

1. Make sure the *Database entries* panel is the active panel, select **Edit** > **Information fields** > **Add information field...**, specify a name (e.g. **Index**) and press <**OK**> to add the field to the *Database entries* panel.

Since the script works on the selected entries, we need to select some entries in the *Database entries* panel:

2. In the *Database entries* panel of the *Main* window, select all entries except 'marker': select **Edit** > **Select all** (**Ctrl+A**) and unselect the 'marker' entry.

3. Select **Scripts** > **Browse internet...**, select the option "Fingerprint related tools" and click on "Calculate diversity indices from banding patterns".

4. Select **DGGE proteobacteria primers** as fingerprint type experiment, choose **Index** as database information field and select a diversity coefficient from the list (e.g. **Diversity index (Simpson)**) (see Figure 22).

5. Press <**OK**> to run the script and confirm the action.
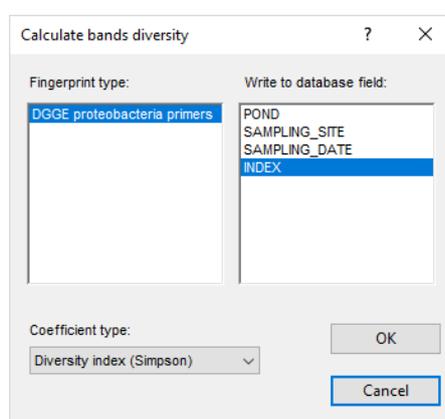
Figure 22: Calculate an index of diversity.

The information in the **Index** field is updated (see Figure 23).

10 Follow-up analysis of DGGE gel images

In this tutorial you have seen how easy it is to import and process gel images in BIONUMERICS. You can now start with the follow-up analysis of the imported data:

- **Cluster analysis:** A cluster analysis can be performed on the DGGE profiles to compare the profiles with each other.

- **Band matching analysis:** In DGGE profiles, one assumes that each individual band/peak refers to a certain microbial taxon, also referred to as a phyplotype or operational taxonomic unit (OTU). Therefore a typical analysis performed DGGE profiles is a band matching analysis. In BIONUMERICS band classes can be named according to the OTU they correspond to and the bands can be quantified via their peak height, peak surface, or 2D band areas, measuring the relative OTU abundances.

- **Ordination analysis:** In BIONUMERICS different ecosystems can be compared with each other based on their fingerprints using Pincipal Coordinates Analysis (PCoA) or Multi-Dimensional Scaling (MDS).

Figure 23: The *Main* window with calculated diversity indices.

- **Search for discriminating bands:** BIONUMERICS can automatically search for discriminating bands, i.e. OTU's that occur under a certain experimental condition, but that are absent or present in much lower abundances under other conditions.

- **Creating charts:** In BIONUMERICS charts can be created that express OTU abundances in function of time or experimental conditions.

- And much more...

More information about these functions can be found in the fingerprint analysis tutorials on our website.