BioNumerics Tutorial:

Importing binary character data

1 Aims

This tutorial shows how to import binary data (= 0 and 1 states) in a BioNumerics database and link the data to a character type experiment.

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 ***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.

3 Example data

The example Excel file from which we will import data in this tutorial contains 0 and 1 states for four genes and can be downloaded from http://www.applied-maths.com/download/sample-data.
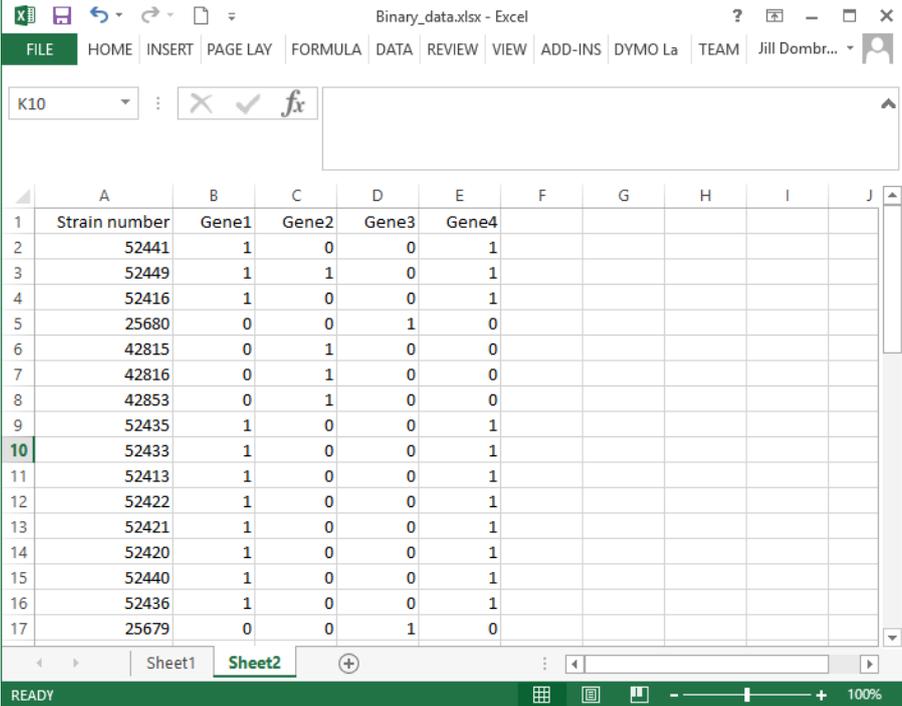
1. Open the file Binary_data.xlsx in Excel to examine the data that will be imported.

Sheet2 contains for 47 isolates following information: a unique identifier (Strain number) and information about the absence ("0") or presence ("1") of four genes (Gene1, Gene2, Gene3 and Gene4) (see Figure 1).

2. Close the Excel file again.

4 Creating a new character type

1. In the *BioNumerics Startup* window, double-click on the **DemoBase Connected** database to open it.

Figure 1: Sheet 2.

Since we will import character data, we will first create a character type to hold this data.

> 2. In the *Main* window, click on ➕ in the toolbar of the *Experiment types* panel and select ***Character type*** from the list. Press <***OK***>.

The *New character type* wizard prompts you to enter a name for the new character type.

> 3. Enter a name, for example "Genes" and press <***Next***>.

In the next step of the wizard, the choice is offered between ***Numerical values*** and ***Binary data***.

> 4. Choose ***Binary data*** since only two possible states are present in our dataset: "0" and "1". Press <***Next***>.

The wizard asks if the character type has an open (***Yes***) or closed (***No***) character set.

> 5. Answer ***No*** and press the <***Finish***> button to complete the setup of the new character type.

The *Experiment types* panel now lists the new character type **Genes**.

5   Import procedure

> 1. Select ***File*** > ***Import...*** (▤, **Ctrl+I**) to open the *Import* dialog box.

> 2. Choose the option ***Import fields and characters (Excel file)*** under the ***Character type data*** item in the tree and press <***Import***>.

> 3. Press <***Browse***> and browse for the downloaded `Binary_data.xlsx` file. Next, press <***Open***>.

> 4. Specify "Sheet2" as data range and press <***Next***>.
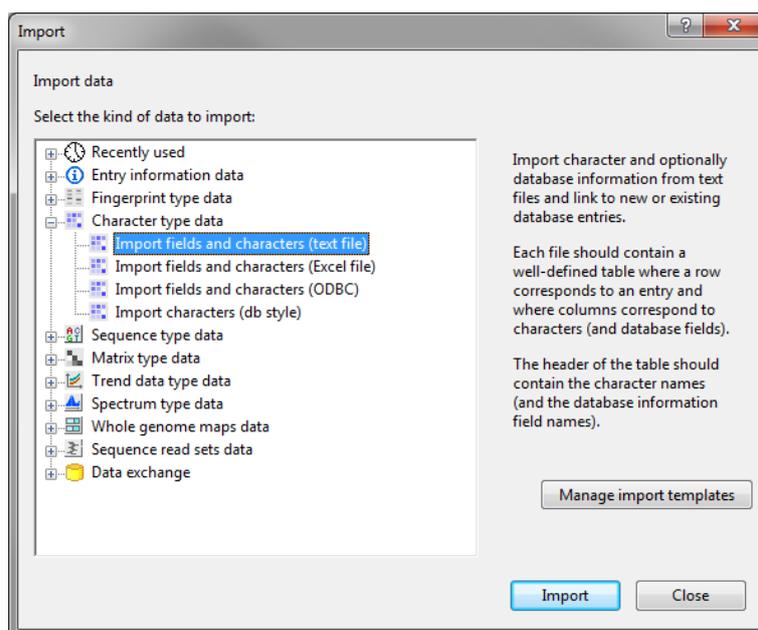
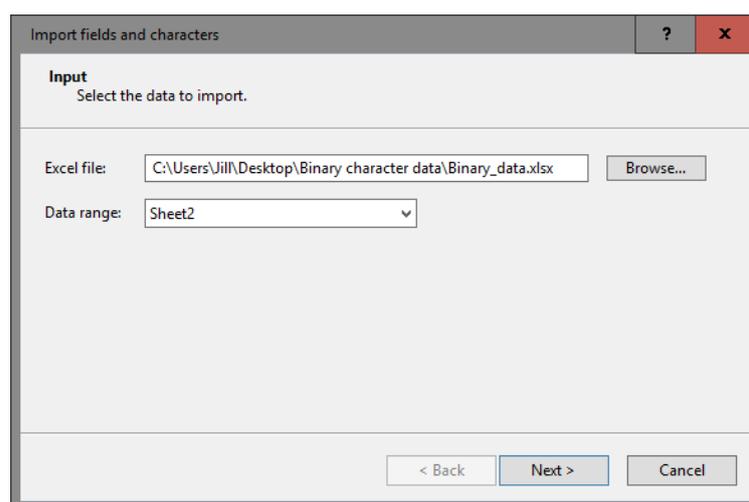**Figure 2:** The *Import* dialog box.

Figure 3: Select data input.

As this is the first time we import character data from Excel into the database, we need to create a new import template by specifying ***Import rules***.

5. Select "Strain number" in the list and click <***Edit destination***> or simply double-click on "Strain number".

6. Select "Strain number" as the BioNumerics destination field in the *Edit data destination* dialog box under ***Entry info field*** (see Figure 4) and press <***OK***>. If this field is not present in the database, choose the ***Create new*** option under ***Entry info field*** and confirm the creation of the new field twice.

7. Select "Gene1", "Gene2", "Gene3" and "Gene4" from the list by holding down the **Ctrl**-key. Click on <***Edit destination***>, select **Genes** under the **Character value** option (see Figure 5) as destination and click <***OK***>.

8. Press <***OK***> and then <***Yes***> to confirm the creation of new characters.
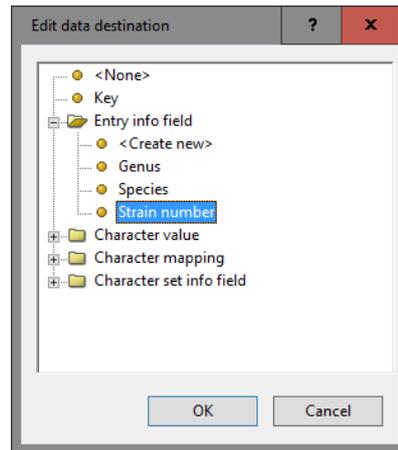
The grid panel is updated (see Figure 6).
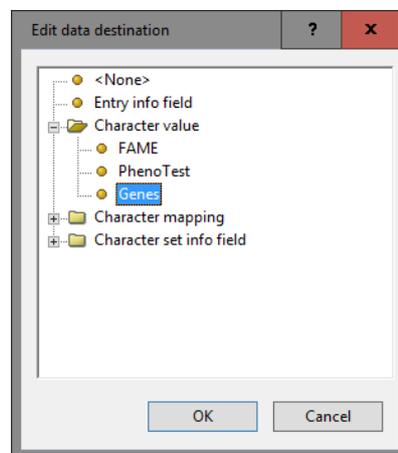
Figure 4: Strain number link.

Figure 5: Link to character experiment.

9. Press *<Preview>* to see what you are about to import.

10. Press the *<Close>* button to close the preview.

11. Press *<Next>* and check the ***Strain number*** option.

The character data will be linked to an existing entry in the database if the Strain number in the Excel file and the BioNumerics Strain number correspond.

12. Press *<Finish>*, specify a template ***Name*** (e.g. "Import binary data") and press *<OK>*.

In the *Import template* wizard page, the newly created template is automatically selected.

13. Click *<Next>*.

14. Press *<Finish>* to start the actual import. The progress of the import is shown while the data is added to the BioNumerics database.

The character data is stored in the character type **Genes**.

15. Click on a green colored dot in the *Experiment presence* panel to open the **Genes** experiment card for an entry.

The values are displayed as a "plate" with one column. The "0" values have a white color, the "1" values are colored in black (see Figure 7 for an example).
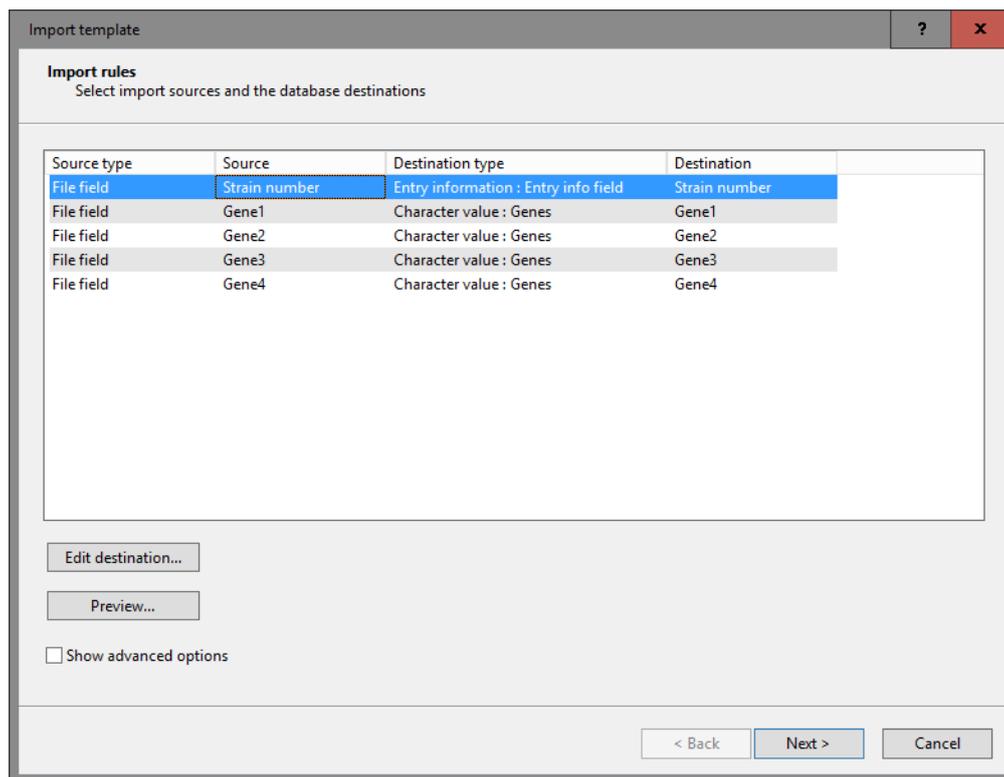
Figure 6: Import rules.

Figure 7: Experiment card: plate representation.

16. Hover over the experiment card with the mouse.

The key of the entry is shown and the name and value of the character being pointed at.

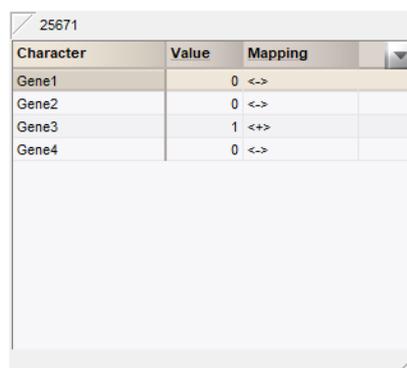17. Click in the upper left corner of the card to close the card.

18. To view the values in a list, double-click on the experiment **Genes** in the *Experiment types* panel, select *Settings* > *General settings...* ( ), select the *Experiment card tab* and change the representation to *List*. Close the two windows.

19. Click on a green colored dot in the *Experiment presence* panel to open the experiment card for an entry (see Figure 8).

The imported values are now displayed in a list.

20. Close the experiment card by clicking in the left upper corner of the card.

Information about the follow-up analysis of binary data sets in BioNumerics can be found in the tutorial "Clustering a binary data set", available on the Applied Maths website.

Figure 8: Experiment card: list representation.