

MIRU-VNTR plugin

PLUGINS
VERSION 7.6



Contents

1	Starting and setting up BioNumerics	3
1.1	Introduction	3
1.2	Startup program	3
1.3	Creating a new database	3
1.4	Installing the MIRU-VNTR plugin	5
1.5	The BioNumerics main window after installation of the plugin	7
2	Importing and analyzing sequencer data	9
2.1	Importing sequencer data	9
2.1.1	Data format	9
2.1.2	Import example	9
2.2	Calculating and assigning MIRU-VNTRs	13
2.2.1	Introduction	13
2.2.2	Defining MIRU-VNTRs	13
2.2.3	Determining copy numbers	15
2.2.4	Visualizing VNTR data for a single entry	15
3	Importing repeat numbers from external files	17
3.1	Introduction	17
3.2	Import routines	17
4	MIRU-VNTR typing	19
4.1	Assigning MIRU-VNTR types	19
4.2	Submitting new MIRU-VNTR types	20

NOTES

SUPPORT BY APPLIED MATHS

While the best efforts have been made in preparing this manuscript, no liability is assumed by the authors with respect to the use of the information provided.

Applied Maths will provide support to research laboratories in developing new and highly specialized applications, as well as to diagnostic laboratories where speed, efficiency and continuity are of primary importance. Our software thanks its current status for a part to the response of many customers worldwide. Please contact us if you have any problems or questions concerning the use of BioNumerics[®], or suggestions for improvement, refinement or extension of the software to your specific applications:

Applied Maths NV

Keistraat 120
9830 Sint-Martens-Latem
Belgium
PHONE: +32 9 2222 100
FAX: +32 9 2222 102
E-MAIL: info@applied-maths.com
URL: <http://www.applied-maths.com>

Applied Maths, Inc.

11940 Jollyville Road, Suite 115N
Austin, Texas 78759
U.S.A.
PHONE: +1 512-482-9700
FAX: +1 512-482-9708
E-MAIL: info-US@applied-maths.com

LIMITATIONS ON USE

The BioNumerics[®] software, its plugin tools and their accompanying guides are subject to the terms and conditions outlined in the License Agreement. The support, entitlement to upgrades and the right to use the software automatically terminate if the user fails to comply with any of the statements of the License Agreement. No part of this guide may be reproduced by any means without prior written permission of the authors.

Copyright ©1998, 2018, Applied Maths NV. All rights reserved.

BioNumerics[®] is a registered trademark of Applied Maths NV. All other product names or trademarks are the property of their respective owners.

BioNumerics[®] uses following third-party software tools and libraries:

- The Python[®] 2.7.4 release from the Python Software Foundation (<http://www.python.org/>).
- A library for XML input and output from the Apache Software Foundation (<http://www.apache.org>).
- NCBI toolkit version 2.2.10 (<http://www.ncbi.nlm.nih.gov/BLAST/>).
- The Boost c++ libraries (<http://www.boost.org/>).
- Samtools for interacting with SAM / BAM files (<http://www.htslib.org/download/>)
- The 7-Zip command line version (7za.exe) from 7-Zip, copyright 1999-2010 Igor Pavlov. <http://www.7-zip.org/>
- Velvet for Windows, source code can be downloaded from <http://www.applied-maths.com/download/open-source>.
- Ray for Windows, source code can be downloaded from <http://www.applied-maths.com/download/open-source>.
- Mothur for Windows, source code can be downloaded from <http://www.applied-maths.com/download/open-source>.
- Cairo 2D graphics library version 1.12.14 (<http://cairographics.org/>).
- Crypto++ Library version 5.5.2 (<http://www.cryptopp.com/>).
- libSVM library for Support Vector Machines (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>).
- SQLite version 3.7.17 (<http://www.sqlite.org/>).
- Gecko engine version 21 (<https://developer.mozilla.org/en-US/docs/Mozilla/Gecko>).
- pymzML Python[®] module for high throughput bioinformatics on mass spectrometry data (<https://github.com/pymzml/pymzML>).
- Numpy Python[®] library version 1.8.1 (<http://www.numpy.org/>).
- BioPython Python[®] library version 1.64 (<http://www.biopython.org/>).
- PIL Python library[®] version 1.1.7 (<http://www.pythonware.com/products/pil/>).
- The SPAdes genome assembler version 3.7.1 (<http://bioinf.spbau.ru/spades>).

Chapter 1

Starting and setting up BioNumerics

1.1 Introduction

This guide is designed as a tutorial for the *MIRU-VNTR plugin* (Mycobacterial Interspersed Repetitive Units, Variable Number of Tandem Repeats) of BioNumerics. Upon installation, the *MIRU-VNTR plugin* defines the 24 VNTRs used by MIRU-VNTRplus [1], and creates an interface with the MIRU-VNTRplus database (<http://www.miru-vntrplus.org>) to assign known types and submit new types. The *MIRU-VNTR plugin* works in combination with the *MLVA plugin*.

The minimal configuration for the installation of these plugins includes the Fingerprint data module (import of sequence trace files or peak tables) and the Character data module (storage of repeat numbers). For clustering and population genetics, the Tree and Network inference module is also required.

1.2 Startup program

When BioNumerics is launched from the Windows start panel or when the BioNumerics shortcut () on your computer's desktop is double-clicked, the **Startup program** is run. This program shows the *BioNumerics Startup* window (see Figure 1.1).

A new BioNumerics database is created from the Startup program by pressing the  button.

An existing database is opened in BioNumerics with  or by simply double-clicking on a database name in the list.

1.3 Creating a new database

3.1 Press the  button in the BioNumerics *BioNumerics Startup* window to enter the *New database wizard*.

3.2 Enter a name for the database, and press <Next>.

A new dialog box pops up, prompting for the type of database (see Figure 1.2).

3.3 Since we want to create a new database to demonstrate the features of the plugin, leave the default option selected and press <Next>.

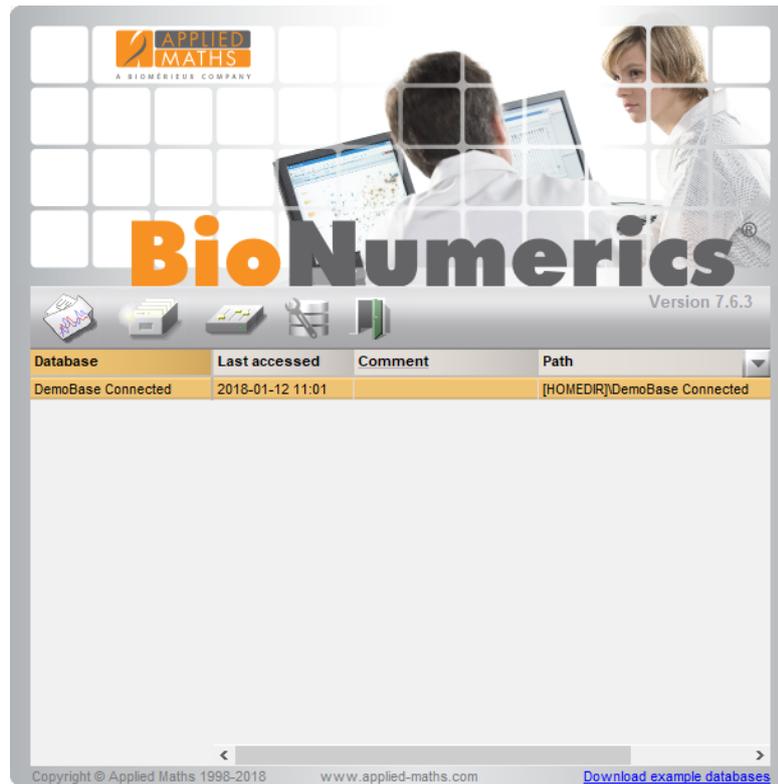


Figure 1.1: The *BioNumerics* Startup window.

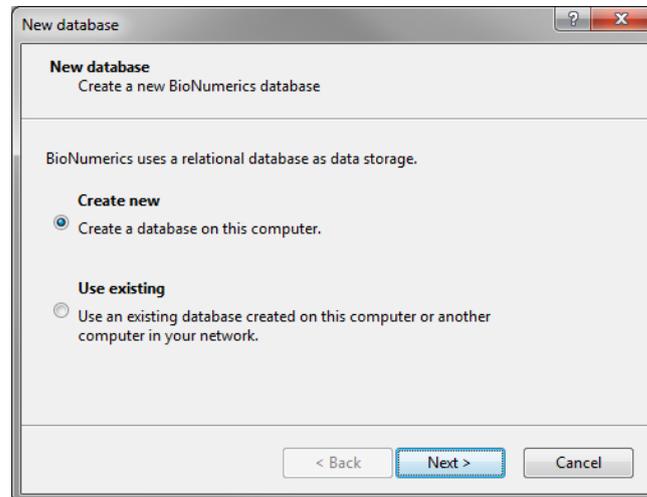


Figure 1.2: The *New database* wizard page.

A new dialog box pops up, prompting for the database engine (see Figure 1.3).

3.4 Leave the default option selected and press *<Next>*.

3.5 Press *<Finish>* to complete the setup of the new database.

The *Plugins* dialog box appears.

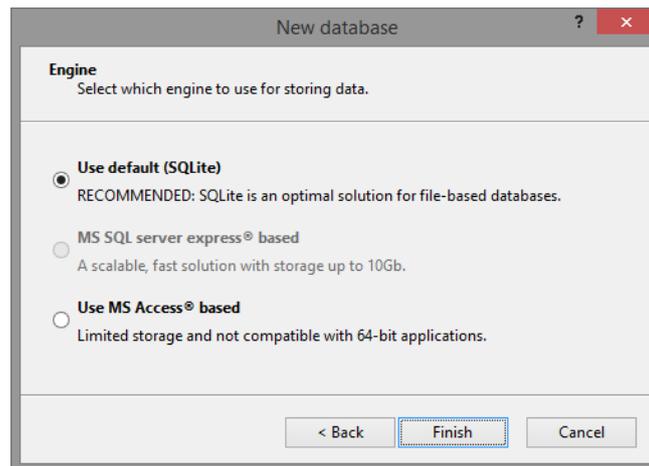


Figure 1.3: The *Database engine* wizard page.

1.4 Installing the MIRU-VNTR plugin

If a database is opened for the first time, the *Plugins* dialog box will appear by default (see Figure 1.4).

If the database has already been opened previously, the *Plugins* dialog box can be called from the *Main* window by selecting *File > Install / remove plugins...* (🔧).

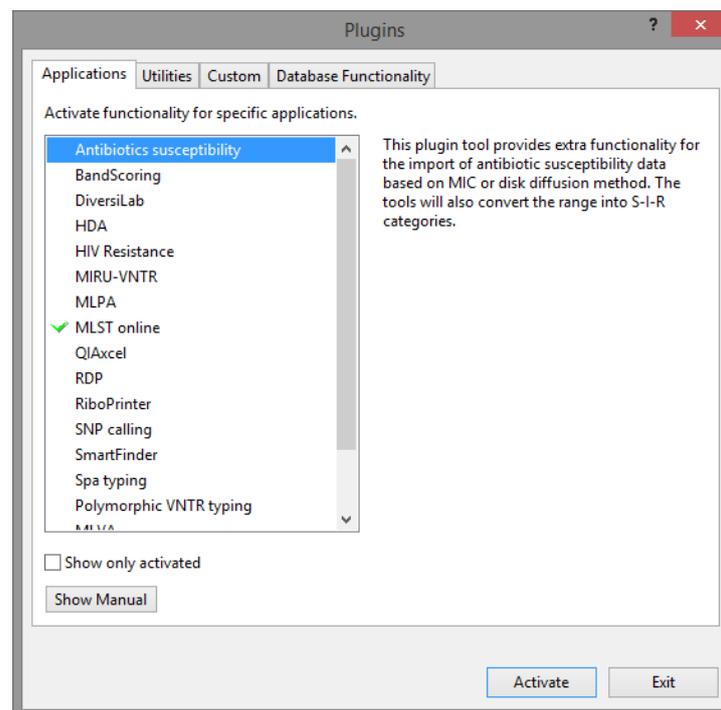


Figure 1.4: The *Plugins* dialog box.

When a particular plugin is selected from the list of plugins, a short description appears in the right panel.

A selected plugin can be installed with the *<Activate>* button. The software will ask for confirmation before installation. Some plugins depend on functionality offered by specific BioNumerics modules. If a required module is missing, the plugin cannot be installed and an error message will be generated.

Once a plugin is installed, it is marked with a green V-sign. It can be removed again with the *<Deactivate>*

button.

If the selected plugin is documented, pressing **<Show Manual>** will open its manual in the *Help* window.

4.1 Select the *MIRU-VNTR plugin* from the list in the *Applications tab* and press the **<Activate>** button.

4.2 The program will ask to confirm the installation of the plugin. Press **<Yes>**.

BioNumerics checks if the *MLVA plugin* is already installed in the database. If this is not the case, the software will install the *MLVA plugin* automatically.

A message appears, stating that the MLVA database components need to be updated.

4.3 Press **<Yes>** to continue.



Updating the MLVA database components will involve creating additional tables and columns in the database to hold MLVA-specific information. Therefore, the current user requires CREATE TABLE and ALTER TABLE privileges on the relational database. This will always be the case for a MS Access, MS SQL Server Express or SQLite database created by BioNumerics. For a shared database on a central database server, please check with your database administrator before proceeding.

When the database update is finished, the *MLVA plugin* is completely installed.

Next, the *MIRU-VNTR settings* dialog box opens.

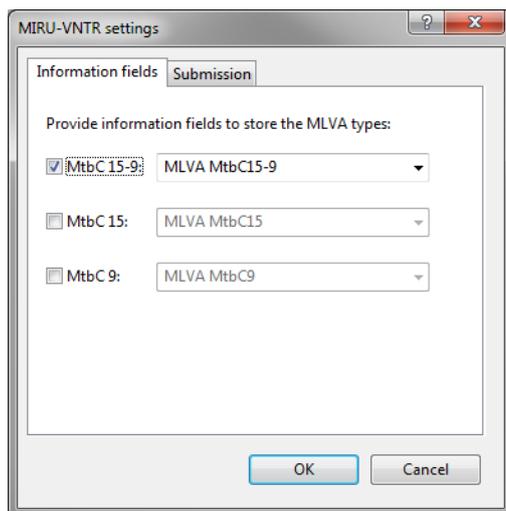


Figure 1.5: The *MIRU-VNTR settings* dialog box: the *Information fields* tab.

In the *Information fields* tab, the user needs to specify the type information that should be assigned to the entries in the database when MIRU-VNTR profiles are screened against the online MIRU-VNTRplus database.

The nomenclature of MIRU-VNTR is split up in two sets of type designations: *MtbC15* and *MtbC 9*, corresponding to a core set of 15 (most discriminatory) VNTRs and 9 additional VNTRs respectively. The *MtbC 15-9* type is defined by a combination of both types.

When enabling an MLVA type check box, the name of the BioNumerics information field that will contain the MLVA type information needs to be specified. One can choose the default suggested name, an existing field or enter a new field name in the *Create new information field* dialog box. When entering new field names manually, please note that information field names cannot start with a space.

In the *Submission* tab, information can be provided for submission of new types. For more information about the submission of new types, see 4.2.



The *MIRU-VNTR settings* dialog box can be called after installation of the plugin using the command *MLVA > MIRU-VNTR settings...*

4.4 Select the MLVA type(s) and specify the information field(s) in the *Information fields tab* and press **<OK>** to continue with the installation of the *MIRU-VNTR plugin*.

If the *MLVA plugin* was installed prior to the installation of the *MIRU-VNTR plugin* and VNTRs have been defined in the database, a dialog box pops up offering the possibility to map the MIRU-VNTRs to existing VNTRs in the database.

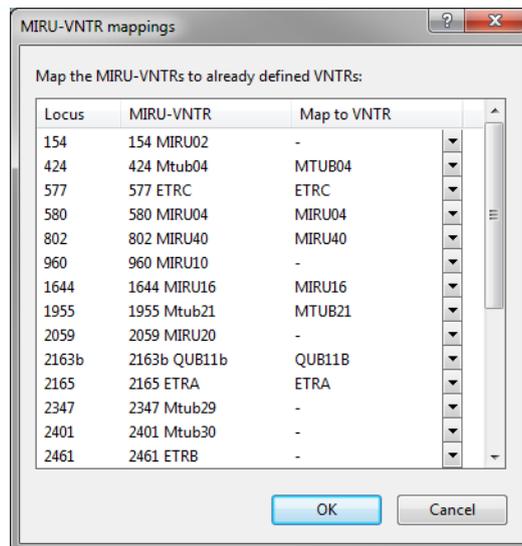


Figure 1.6: The *MIRU-VNTR mappings* dialog box: Mapping MIRU-VNTRs to existing VNTRs.

If existing VNTRs should be mapped to MIRU-VNTRs, specify the correct VNTRs in the 'Map to VNTR' column.



The plugin keeps an internal map of the VNTR names to MtbC loci for online submission (see 4.2) and new type assignment (see 4.1).

4.5 Press **<Proceed>** (or **<Exit>**) to close the *Plugins* dialog box and to continue to the *Main* window.

1.5 The BioNumerics main window after installation of the plugin

Two character type experiments are created by the *MLVA plugin*, i.e. **MIRU-VNTR_frgs** and **MIRU-VNTR_vals**.

MIRU-VNTR_frgs will hold the fragment sizes of the bands, whereas **MIRU-VNTR_vals** will contain the MIRU-VNTR copy numbers (see Figure 1.7). The information fields that will hold the MLVA type information are created upon installation of the *MIRU-VNTR plugin* and are displayed in the *Database entries* panel (see Figure 1.7).

The *MLVA plugin* installs menu items in the main menu of the BioNumerics software under *MLVA*. The three last items that appear in the *MLVA* list (*MLVA > Assign MIRU-VNTR types*, *MLVA > Submit new MIRU-VNTR types*, *MLVA > MIRU-VNTR settings...*) are installed by the *MIRU-VNTR plugin* (see Figure 1.8).

The *MIRU-VNTR plugin* furthermore adds theoretical VNTR definitions for the 24 VNTRs used by MIRU-VNTRplus [1] in the *MLVA management* window (see 2.2.2).

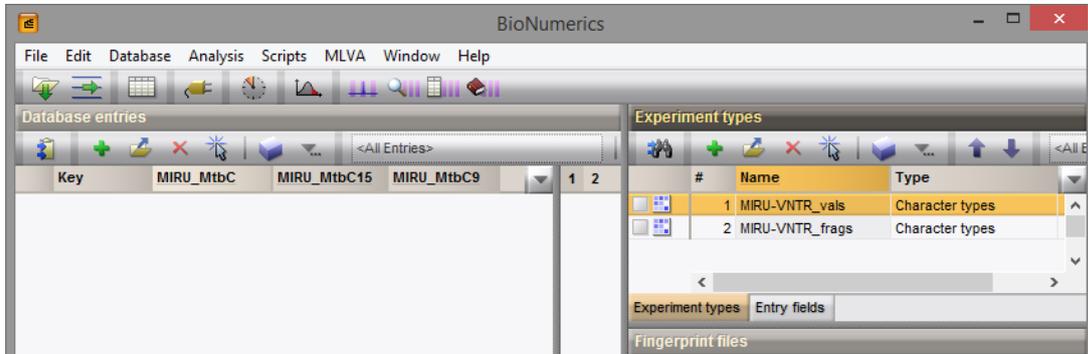


Figure 1.7: The *Main* window after installation of the *MIRU-VNTR* plugin.

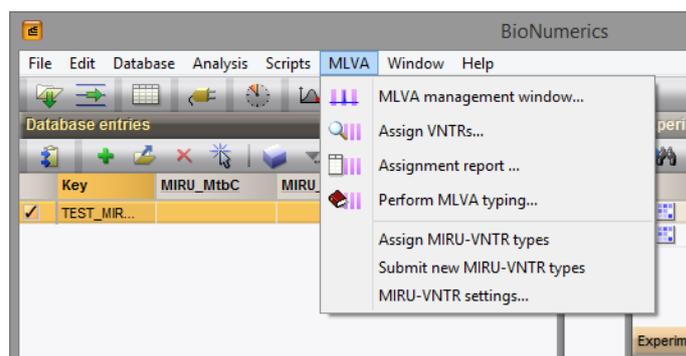


Figure 1.8: MLVA and MIRU-VNTR menu items in the *Main* window.

Chapter 2

Importing and analyzing sequencer data

2.1 Importing sequencer data

2.1.1 Data format

When fingerprints are run on a capillary sequencer, the resulting data can have two fundamentally different formats:

- **Curve files** (also referred to as electropherograms, chromatogram files or trace files): The binary encoded raw data as produced by the capillary electrophoresis equipment. Since it is the raw data, it contains the complete information but some fingerprint preprocessing (e.g. normalization, band assignment, ...) in BioNumerics is required.
- **Peak tables:** Text files containing a listing of peaks from the chromatograms, with their corresponding metrics (sizes in base pairs) and peak height and/or peak area. This type of data has been processed by the software which controls the capillary electrophoresis equipment.



In contrast to the .fsa files generated by Applied Biosystems sequencers, the .scf raw curve files generated by Beckman-Coulter equipment are *not* corrected for spectral overlap of the fluorescent dyes. This is generally not a problem when a single data channel is used. However, in experimental setups where all four dyes are employed (i.e. three data channels and one reference channel), it is advised to import .crv files in BioNumerics. The latter files can be exported from the Beckman-Coulter software and contain curves which are corrected for cross talk.

Both formats can be imported into BioNumerics using the import routines available in the *Import* dialog box. The *Import* dialog box is called with the command **File > Import...** , **Ctrl+I**.

- Curve files (i.e. electropherograms) can be imported in the database with the option **Import curves** listed under **Fingerprint type data** in the import tree.
- Peak files can be imported in the database with the option **Import peak table** listed under **Fingerprint type data** in the import tree.

2.1.2 Import example

An example peak file can be downloaded from the download page on the Applied Maths website (<http://www.applied-maths.com/download/sample-data>, click on "MIRU-VNTR sample data"). This ex-

ample file will be used to illustrate the import and analysis steps in this section.

- 1.1 Select **File > Import...** (📁, **Ctrl+I**) to call the *Import* dialog box, choose **Import peak table** under **Fingerprint type data** and press **<Import>**.
- 1.2 Browse to the downloaded and unzipped example data folder and select the Applied Biosystems GeneMapper MIRU.txt sample file. Press **<Open>**.

The path is displayed in the *Input* wizard page.

- 1.3 Check **Import as fingerprint file**, specify a name (e.g. “MIRU1”) and press **<Next>**.

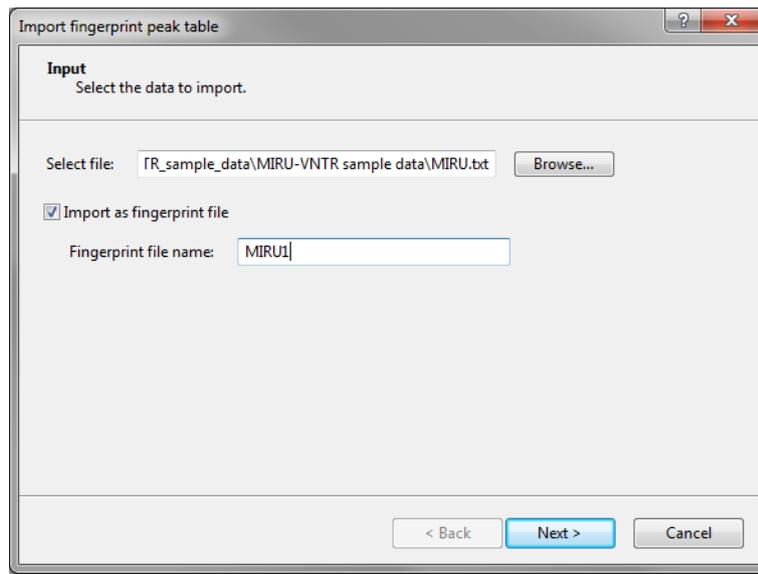


Figure 2.1: Selected peak table file.

The way the information should be imported in the database can be specified with an import template. In the example Applied Biosystems GeneMapper peak file, the dye and sample information is provided. We will create a new import template:

- 1.4 Press the **Create new** button to call the *Import rules* dialog box.

The *Import rules* dialog box lists the information present in the selected file as **Source**, their linked **Source type** and the **Destination** component they are associated with (currently all set to **<None>**).

- 1.5 Select **Sample File** from the list, select **<Edit destination>** and select **Key** as corresponding field. Press **<OK>**.
- 1.6 Select **Dye** from the list, select **<Edit destination>** and select **Fingerprint dye** as corresponding field. Press **<OK>**.
- 1.7 Select **Marker** from the list, select **<Edit destination>** and select **IsMarker** under **Peak info** as corresponding field. Press **<OK>**.
- 1.8 Repeat previous step for the **Size**, **Height** and **Area** rows. Link the rows to the corresponding peak information fields.

The *Import rules* dialog box should now look like Figure 2.2.

- 1.9 Press the **<Preview>** button to check the outcome of all defined import rules (see Figure 2.3).
- 1.10 Close the preview and press **<Next>** twice and **<Finish>**.
- 1.11 Specify a template name (e.g. “My MIRU-VNTR template”) and press **<OK>**.

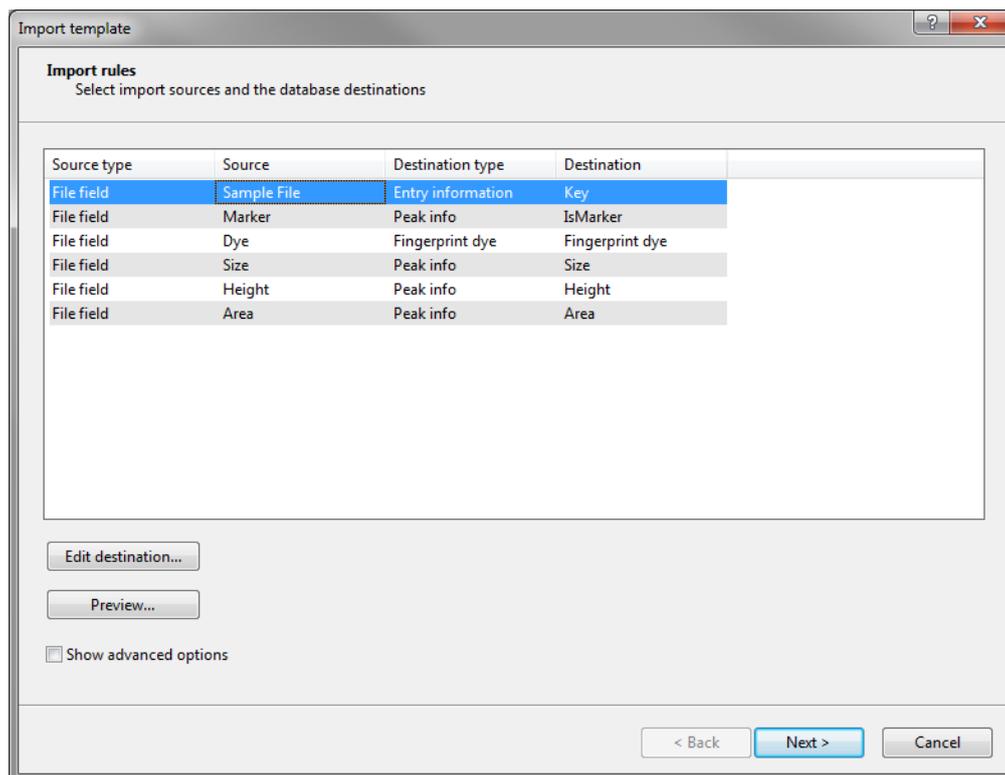


Figure 2.2: Import template with the defined rules.

Nr.	Fingerprint dye	Size	Height	Area	IsMarker	Key
1	B	509.65	1341	21983		1
2	B	503.65	1768	25973		2
3	B	504.96	1147	21440		3
4	B	503.73	1729	22271		4
5	B	503.95	2592	32669		5
6	B	503.64	3434	43946		6
7	B	503.62	2319	29092		7
8	B	502.50	3352	44377		8
9	B	556.93	2312	31726		9
10	B	503.52	2872	36867		10
11	B	503.66	1797	23038		11
12	B	503.82	1480	22157		12
13	B	503.66	1475	22647		13
14	B	504.04	989	15851		14
15	B	504.05	954	16706		15
16	B	517.00	1911	31034		16
17	B	503.64	2462	30290		17
18	B	449.46	3567	48252		18
19	B	503.64	1770	24724		19

Figure 2.3: Preview.

1.12 Make sure *<Create new>* is selected as base fingerprint type experiment, select the new template and press *<Next>*.

1.13 Specify a name for the new base fingerprint type experiment (e.g. "MIRU-VNTR") and press *<OK>*. Confirm the creation of the new experiment in the database.

A fingerprint type needs to be present in the database for each pool (if present) and dye combination. The

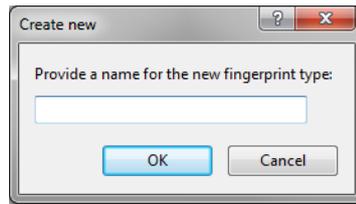


Figure 2.4: New base fingerprint type experiment.

names of these fingerprint types are composed of the base fingerprint type name, followed by the pool name (if present), and the name of the dye. If new fingerprint types need to be created during import, and if no reference system has been specified for the base fingerprint type, a dialog box pops up prompting for some settings.

- 1.14 For this exercise, enter an *Intensity range* of “4095” (= 12-bit), a *Min. fragment length* of “50”, and a *Max. fragment length* of “800”. Press <OK>.

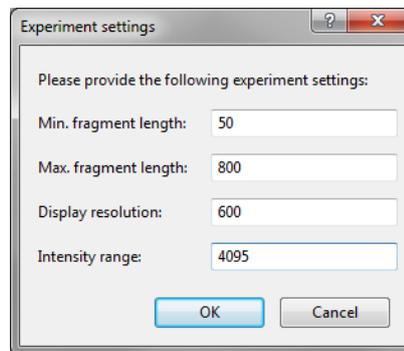


Figure 2.5: Experiment settings.

A new dialog box pops up, listing the three missing fingerprint types.

- 1.15 Confirm the creation of the missing fingerprint type experiments.

- 1.16 Press <*Finish*> to confirm the creation of new entries in the database.

For each import dye, a new fingerprint file is created, composed of the file name specified and the name of the dye (e.g. MIRU1_B). These files are displayed in the *Fingerprint files* panel.

BioNumerics reads the band positions from the mapped “SIZE” column, the peak heights from the mapped “HEIGHT” column, the area information from the mapped “AREA” column and generates densitometric curves using this information. The imported fingerprint lanes are linked to new entries in the database. The lanes are linked to the corresponding fingerprint “dye” type. The names of these fingerprint types are composed of the base fingerprint type name, followed by the pool name (if present), and the name of the dye. The fingerprint type experiments are displayed in the *Experiment types* panel.

If no reference system has been specified for the base fingerprint type, BioNumerics creates for all missing fingerprint types and the base fingerprint type, a linear reference system between the user-defined *Minimum* and *Maximum fragment length* positions, and copies the reference system to a calibration system.

Entries for which fingerprint data was imported are selected in the database.

After data import, the *Main* window looks as in Figure 2.6.

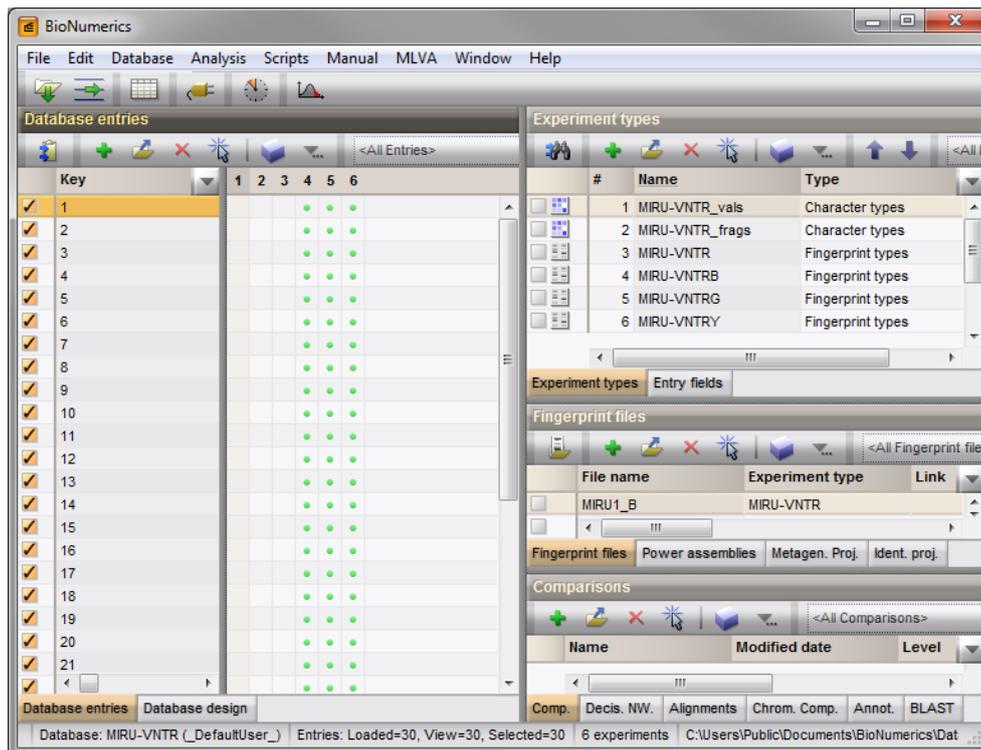


Figure 2.6: The *Main* window after import of MIRU-VNTR data.

2.2 Calculating and assigning MIRU-VNTRs

2.2.1 Introduction

With the *MLVA plugin*, VNTR copy numbers can be calculated for a selection of database entries. In the *MLVA plugin* tutorial, step-by-step instructions are given for the calculation of VNTR copy numbers. The *MLVA plugin* will not be covered in detail in this section.



If the *MLVA plugin* was not installed prior to the installation of the *MIRU-VNTR plugin*, the *MLVA plugin* is installed together with the installation of the *MIRU-VNTR plugin* (see 1.4).

2.2.2 Defining MIRU-VNTRs

Before the *MLVA plugin* can determine copy numbers from the fragment length information, the plugin should have sufficient information about what VNTR alleles can be found in which pool and dye, what the repeat lengths are, etc.. Upon installation of the *MIRU-VNTR plugin*, the 24 VNTRs used by MIRU-VNTRplus [1] to type *Mycobacterium* are automatically added to the VNTR list in the BioNumerics database.

2.1 Select *MLVA > MLVA management window...* () in the *Main* window to display the *MLVA management* window (see Figure 2.7).

The *MLVA management* window (see Figure 2.7), lists the 24 VNTRs used by MIRU-VNTRplus with the theoretical settings for the offset (*Offset*), repeat length (*Repeat size*), minimum (*Min. count*) and maximum (*Max. count*) expected copy numbers, and tolerance (*Tolerance*). The associated fingerprint type (*Fingerprint type* column) still needs to be provided manually (see 2.2) as this depends on the experimental setup of the lab.

2.2 Highlight a VNTR from the list (e.g. 154 MIRU02) and select *Edit > VNTRs > Edit VNTR...* ().

Name	Offset	Repeat size	Min. count	Max. count	Tolerance	Fingerprint type	Start	Stop	Value
154 MIRU02	402	53	0	25	10.0		392.0	412.0	0
424 Mtub04	537	51	0	25	10.0		445.0	465.0	1
577 ETRC	171	58	0	25	10.0		498.0	518.0	2
580 MIRU04	175	77	0	25	10.0		551.0	571.0	3
802 MIRU40	354	54	0	25	10.0		604.0	624.0	4
960 MIRU10	482	53	0	25	10.0		657.0	677.0	5
1644 MIRU16	565	53	0	25	10.0		710.0	730.0	6
1955 Mtub21	116	57	0	25	10.0		763.0	783.0	7
2059 MIRU20	437	77	0	25	10.0		816.0	836.0	8
2163b QUB11b	67	69	0	25	10.0		869.0	889.0	9
2165 ETRA	195	75	0	25	10.0		922.0	942.0	10
2347 Mtub29	335	57	0	25	10.0		975.0	995.0	11
2401 Mtub30	252	58	0	25	10.0		1028.0	1048.0	12
2461 ETRB							1081.0		

Figure 2.7: The *MLVA management* window displaying the theoretical MIRU-VNTR definitions.

This brings up the *Edit VNTR* dialog box, displaying the default settings of the selected VNTR (see Figure 2.8).

Figure 2.8: The *Edit VNTR* dialog box.

- **Offset:** Each fragment consists of a repeat portion and a constant portion due to the fact that the primers do not occur exactly at the start and the end of the repeat region. This parameter specifies the size (in base pairs) of the constant portion.
- **Repeat size:** This parameter specifies the size of a unit repeat block (in base pairs).
- **Copy range:** Specifies the minimum and maximum number of copies that the plugin will consider during the copy number determination. Within the same lane and the same dye, more than one VNTR can be loaded if there is a sufficient difference in fragment size. Imposing a limit on the copy range allows the software to distinguish which band corresponds to which VNTR.

- **Tolerance:** This parameter specifies the maximum difference in base pairs between the expected fragment length and the actual length.
- **Fingerprint type:** This drop-down list contains all fingerprint types defined in the database and is used to indicate in which fingerprint type this VNTR was run. The fingerprint type is determined by the dye and – if several pools are used – by the pool tag.

2.3 Select the correct fingerprint type from the **Fingerprint type** drop-down list (see Figure 2.8) and press **<OK>** to save the changes.

The **Fingerprint type** column is updated in the *MLVA management* window.

2.4 Assign the correct fingerprint types to the other VNTRs (repeat Instruction 2.2 and Instruction 2.3).

Link the three VNTRs in the example database to the correct fingerprint types:

- 154 MIRU02: **MIRU-VNTRB**
- 2531 MIRU23: **MIRU-VNTRG**
- 4348 MIRU39: **MIRU-VNTRY**

2.5 After having entered the information for all VNTRs in the *MLVA management* window, select **File > Exit (Alt+F4)** to close the *MLVA management* window.

2.2.3 Determining copy numbers

VNTR copy numbers can be determined for selected entries in the database.

2.6 Select the entries in the database for which you want to calculate the copy numbers. To select all entries at once, use **Edit > Select all (Ctrl+A)**.

2.7 Select **MLVA > Assign VNTRs...**  to display the *Assign VNTRs* dialog box.

2.8 Leave the default settings unaltered and press **<OK>** to determine VNTR copy numbers in the fingerprints of the selected entries.

After assignment of the VNTRs in the database, the *VNTR reporting* window is displayed listing all selected entries (keys) as rows and the selected VNTRs as columns (see Figure 2.9).

Each cell, corresponding to an entry / VNTR pair, provides information about the VNTR copy number assignment. For more information about this window, we refer to the *MLVA plugin* manual.

2.9 Select **File > Exit (Alt+F4)** to close the *VNTR reporting* window.



The *VNTR reporting* window can be called at any time from the *Main* window with **MLVA > Assignment report ...** .

2.2.4 Visualizing VNTR data for a single entry

After VNTR assignment, the copy numbers for each detected VNTR are stored in the character type **MIRU-VNTR_vals**. The character type **MIRU-VNTR_frgs** holds the observed fragment size (in base pairs) for each VNTR character.

2.10 Click on a colored dot in one of the character type columns (**MIRU-VNTR_vals** or **MIRU-VNTR_frgs**) in the *Experiment presence* panel in the *Main* window to display the character card for an entry.

2.11 Close an experiment card by clicking in the small triangle-shaped button in the top left corner.

Key	154 MIRU02	2531 MIRU23	4348 MIRU39
1	2	5	2
2	2	6	2
3	2	5	2
4	2	3	2
5	2	5	2
6	2	5	2
7	2	5	2
8	2	5	2
9	3	5	2
10	2	5	2
11	2	6	2
12	2	5	3
13	2	6	2

Figure 2.9: The *VNTR reporting* window, here displaying only the VNTRs for which data are available in the example data set.

Character	Value	Mapping
154 MIRU02	509.6	<+>
424 Mtub04		
577 ETRC		
580 MIRU04		
802 MIRU40		
960 MIRU10		
1644 MIRU16		
1955 Mtub21		
2059 MIRU20		
2163b QUB11b		
2165 ETRA		
2347 Mtub29		
2401 Mtub30		
2461 ETRB		
2531 MIRU23	409.3	<+>
2687 MIRU24		
2996 MIRU26		
3007 MIRU27		
3171 Mtub34		
3192 MIRU31		
3690 Mtub39		
4052 QUB26		
4156 QUB4156		
4348 MIRU39	641.8	<+>

(a)

Character	Value	Mapping
154 MIRU02	2.0	<+>
424 Mtub04		
577 ETRC		
580 MIRU04		
802 MIRU40		
960 MIRU10		
1644 MIRU16		
1955 Mtub21		
2059 MIRU20		
2163b QUB11b		
2165 ETRA		
2347 Mtub29		
2401 Mtub30		
2461 ETRB		
2531 MIRU23	5.0	<+>
2687 MIRU24		
2996 MIRU26		
3007 MIRU27		
3171 Mtub34		
3192 MIRU31		
3690 Mtub39		
4052 QUB26		
4156 QUB4156		
4348 MIRU39	2.0	<+>

(b)

Figure 2.10: The *Experiment card* window from a) *MIRU-VNTR_frags*: molecular observed size; b) *MIRU-VNTR_vals*: copy number.

Chapter 3

Importing repeat numbers from external files

3.1 Introduction

When VNTR copy numbers are already stored in external text, Excel, or other ODBC-compatible file, these copy numbers can be imported in BioNumerics and stored in the character type experiment **MIRU-VNTR_vals** using import routines available in the *Import* dialog box.

3.2 Import routines

The *Import* dialog box is called with the command **File > Import...** (, **Ctrl+I**). The import tree options are organized in groups based upon the type of data (see Figure 3.1).

- With the *Import fields and characters (text file)* option, listed under the topic *Character type data* in the Import tree, the copy numbers can be imported from text files in the database and linked to new or existing database entries.
- With the *Import fields and characters (Excel file)* option, listed under the topic *Character type data* in the Import tree, the copy numbers can be imported from an Excel file in the database and linked to new or existing database entries.
- With the *Import fields and characters (ODBC)* option, listed under the topic *Character type data* in the Import tree, the copy numbers can be imported from ODBC-compatible files in the database and linked to new or existing database entries.

The import routines will not be covered in detail in this manual.

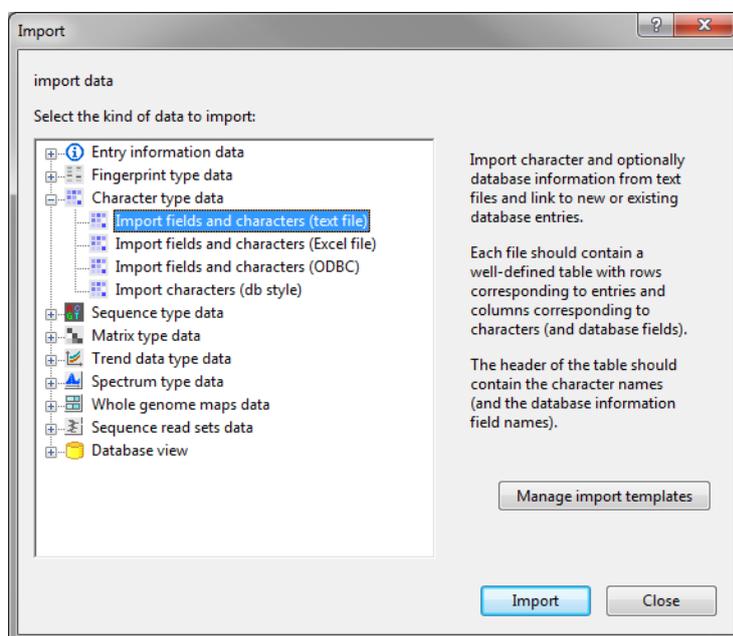


Figure 3.1: The character import routines in the *Import* dialog box.

Chapter 4

MIRU-VNTR typing

4.1 Assigning MIRU-VNTR types

With the installation of the *MIRU-VNTR plugin*, an interface is created with the MIRU-VNTRplus database (<http://www.miru-vntrplus.org>) to assign known MIRU-VNTR types to entries in the BioNumerics database. The MLVA types that can be assigned to entries in the database, depend on the MtbC information fields specified in the *MIRU-VNTR settings* dialog box.

- 1.1 Select *MLVA > MIRU-VNTR settings...* in the *Main* window to call the *MIRU-VNTR settings* dialog box.

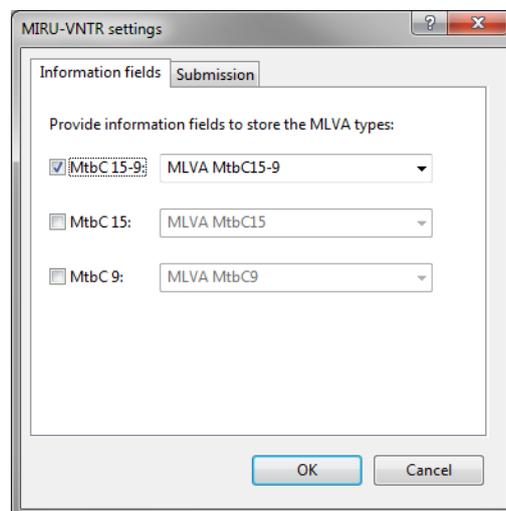


Figure 4.1: The *MIRU-VNTR settings* dialog box: the *Information fields* tab.

In the *Information fields* tab, the user needs to specify the type information that should be assigned to the entries in the database when MIRU-VNTR profiles are screened against the online MIRU-VNTRplus database.

The nomenclature of MIRU-VNTR is split up in two sets of type designations: *MtbC15* and *MtbC 9*, corresponding to a core set of 15 (most discriminatory) VNTRs and 9 additional VNTRs respectively. The *MtbC 15-9* type is defined by a combination of both types.

When enabling an MLVA type check box, the name of the BioNumerics information field that will contain the MLVA type information needs to be specified. One can choose the default suggested name, an existing field or enter a new field name in the *Create new information field* dialog box. When entering new field

names manually, please note that information field names cannot start with a space.

- 1.2 Select the MLVA type(s) and specify the information field(s) in the *Information fields tab* and press **<OK>**.

Screening can be done for all entries present in the database, or for any selection of entries in database.

- 1.3 Make a selection in the database using the **Ctrl-** and **Shift-**keys. To select all entries at once, use **Edit > Select all (Ctrl+A)**.
- 1.4 In the *Main* window, select **MLVA > Assign MIRU-VNTR types**.
- 1.5 If no selection is present in the database, a message pops up prompting to confirm to run the tool on all entries in the database. Press **<OK>** to start the profile screening for all entries in the database.

The MIRU-VNTR copy numbers – stored in the **MIRU-VNTR_vals** character type experiment – are screened against the MIRU-VNTRplus database. If matching MtbC types are found online, the type information is written in the MLVA information field(s). The MLVA type is denoted as "???" if no match could be found with the online information available.

Key	MLVA MtbC15-9	MLVA MtbC15	MLVA MtbC9	1	2	3	4	5	6
1	1-1	1	1	•	•	•	•	•	•
2	3-4	3	4	•	•	•	•	•	•
3	10-3	10	3	•	•	•	•	•	•
4	6-7	6	7	•	•	•	•	•	•
5	39-2	39	2	•	•	•	•	•	•
6	?-?	?	?	•	•	•	•	•	•

Figure 4.2: Assigning MLVA types to entries in the BioNumerics database.



MIRU-VNTR types cannot be obtained with the example data set provided, since it contains only information for three VNTRs and MIRU-VNTR typing is based on 9 or 15 VNTRs.

4.2 Submitting new MIRU-VNTR types

With the installation of the *MIRU-VNTR plugin*, an interface is created with the MIRU-VNTRplus database (<http://www.miru-vntrplus.org>) to submit new types to this online database. Before new unknown types can be submitted to the MIRU-VNTRplus database, user information needs to be provided in *Submission tab* of the *MIRU-VNTR settings* dialog box.

- 2.1 Select **MLVA > MIRU-VNTR settings...** in the *Main* window to call the *MIRU-VNTR settings* dialog box and click the *Submission tab* (see Figure 4.3).

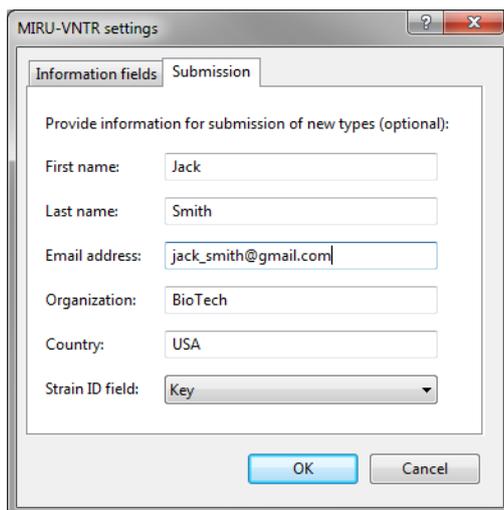
User submission information includes **First name**, **Last name**, **Email address**, **Organization name**, and **Country**. In addition to these user fields, an information field holding strain information (**Strain ID field**) needs to be selected from the list of available information fields. The user information and strain ID will be sent along with the submitted profile to the MIRU-VNTRplus database, and will be stored in the online database if a new type is assigned to the submitted profile. No user or strain information is recorded if existing profiles are submitted to the online database.



Submitting a new MIRU-VNTR type to the MIRU-VNTRplus database is only possible if the strain ID field is not empty.

- 2.2 Specify the submission settings and press **<OK>** to save the settings to the database.

Submitting new types can be done for all entries present in the database, or for any selection of entries in database. Of course it only makes sense to submit MLVA type information if the MtbC15 and/or MtbC9 type



The screenshot shows a dialog box titled "MIRU-VNTR settings" with two tabs: "Information fields" and "Submission". The "Submission" tab is active. The dialog contains the following fields:

- First name: Jack
- Last name: Smith
- Email address: jack_smith@gmail.com
- Organization: BioTech
- Country: USA
- Strain ID field: Key (dropdown menu)

At the bottom of the dialog are "OK" and "Cancel" buttons.

Figure 4.3: The *MIRU-VNTR settings* dialog box: the *Submission* tab.

is assigned a "??", and if the complete copy number information is present for the MtbC15 and (optionally) the MtbC9 type.

2.3 Make a selection in the database using the **Ctrl-** and **Shift-**keys.

2.4 In the *Main* window, select *MLVA > Submit new MIRU-VNTR types*.

2.5 If no selection is present in the database, a message pops up prompting to confirm to run the tool on all entries in the database. Press **<OK>** to start the profile submission for all entries in the database.

The MIRU-VNTR copy numbers (stored in the **MIRU-VNTR.vals** character experiment), the user information and the strain IDs of the selected entries are submitted to the MIRU-VNTRplus database.

- New profiles that are submitted to the MIRU-VNTRplus database automatically acquire a new type number (MtbC15 and/or MtbC9). The user information is recorded together with the profile and the new type. The new type information – returned by the MIRU-VNTRplus database – is written in the corresponding BioNumerics information field(s).
- Re-submissions are not stored in the online database. The MIRU-VNTRplus database simply returns the existing types and does not record any user information.



Copyright 1998-2018, Applied Maths NV. All rights reserved.

Please contact us for any additional information you might require, we will gladly help you!

Headquarters

📍 Keistraat 120 • 9830 Sint-Martens-Latem • Belgium
☎ +32 922 22 100 ✉ info@applied-maths.com

USA and Canada

📍 11940 Jollyville Rd., Suite 115N • Austin, TX 78750 USA
☎ +1 512 482 9700 ✉ info-us@applied-maths.com

Bibliography

- [1] C. Allix-Beguec, D. Harmsen, T. Weniger, P. Supply, and S. Niemann. Evaluation and strategy for use of miru-vnrplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of mycobacterium tuberculosis complex isolates. *Journal of clinical microbiology*, 46(8):2692, 2008.