BioNumerics®

MICROBIAL DATA ANALYSIS SOFTWARE

HIV drug resistance plugin

PLUGINS
VERSION 7.6

Contents

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$^{\circledR}$, 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$^{\circledR}$ 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.

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

1.1.1 About the BioNumerics HIV Drug Resistance plugin

The *HIV Drug Resistance plugin* is a plugin tool of the BioNumerics software. BioNumerics is a very comprehensive software suite, designed for generic analysis of virtually all types of biological data. The software has a rich scripting environment that allows for highly specific applications to be developed on top of the powerful databasing and analysis platform that is offered by BioNumerics. A number of script-based applications are provided and officially supported by Applied Maths. These applications are compiled into plugins and are available from the BioNumerics installation or the Applied Maths website.

The minimal configuration of BioNumerics modules for the installation of the *HIV Drug Resistance plugin* includes the Sequence data module, the Character data module and the Classifiers and Identification module.

1.1.2 Principles

The functionality of the *HIV Drug Resistance plugin* is based upon the viral HIV-1 POL gene. This gene codes for the POL poly-protein, which is cleaved into the enzymatic proteins: protease (PR), reverse transcriptase (RT), RNaseH, and integrase (IN). PR and RT are currently the main targets for HIV drugs.

The software can import either processed nucleotide sequences or translated amino acid sequences (see [2] on Figure 1.1). It is also possible to import raw trace files from automated sequencers and assemble them into contigs using BioNumerics' powerful Assembler application (step [1]). Supported sequencers include Applied Biosystems, Amersham and Beckman. The BioNumerics batch import routine has been designed for automatic assembly of large batches of trace files into consensus sequences, thereby generating reports for errors and unresolved positions. Nucleic acid sequences are automatically translated by the plugin ([4]).

The plugin then compares the proteins with the HIV-1 wild type reference sequence ([5]). Most algorithms use the "Consensus B" from the Los Alamos HIV database. The "HXB2" sequence can also be chosen, or the user can store any other sequence. Amino acids that deviate from the reference sequence are stored in mutation lists. These lists can also directly be imported into the database ([3]).

For the interpretation of drug resistance in function of AA mutations, the plugin uses publicly available *algorithms*, including **HIVDB** (Stanford University, CA, USA), **REGA** (REGA Institute, Leuven, Belgium) and **ANRS** (Agence Nationale de Recherches sur le Sida, France). The algorithms, provided as XML files containing *rules* ([6]), are parsed by the software into *Decision Networks*: one Decision Network is generated per algorithm and per drug ([7]). A Decision Network is an operational network that carries out a series of [logical] operations on data, and based on the outcome of these operations, leads to an answer (decision). Decision networks are discussed in detail in the Reference manual, Chapter Decision networks.
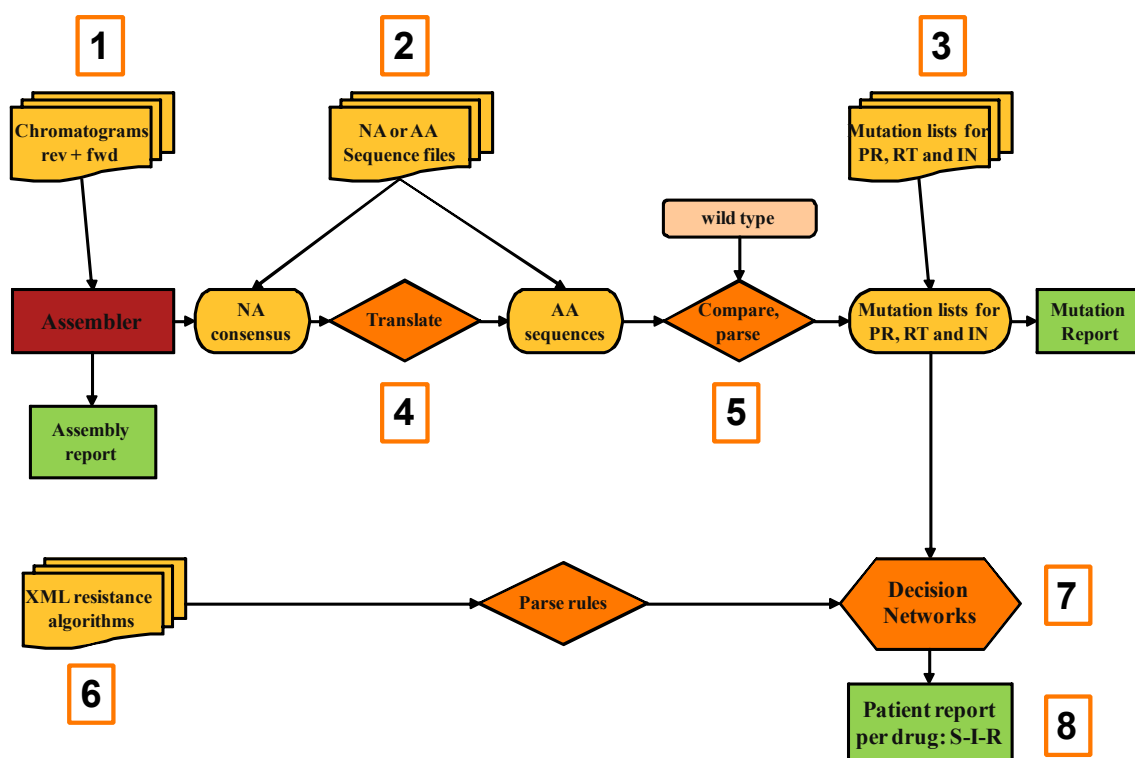
Figure 1.1: Schematic overview of the work flow from sequence import to reporting in the *HIV Drug Resistance plugin*.

For each analyzed POL sequence, a report is generated based upon one or more algorithms and the selected drugs ([8]). If multiple algorithms are installed, comparative reports can be generated as well.

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 1.2).

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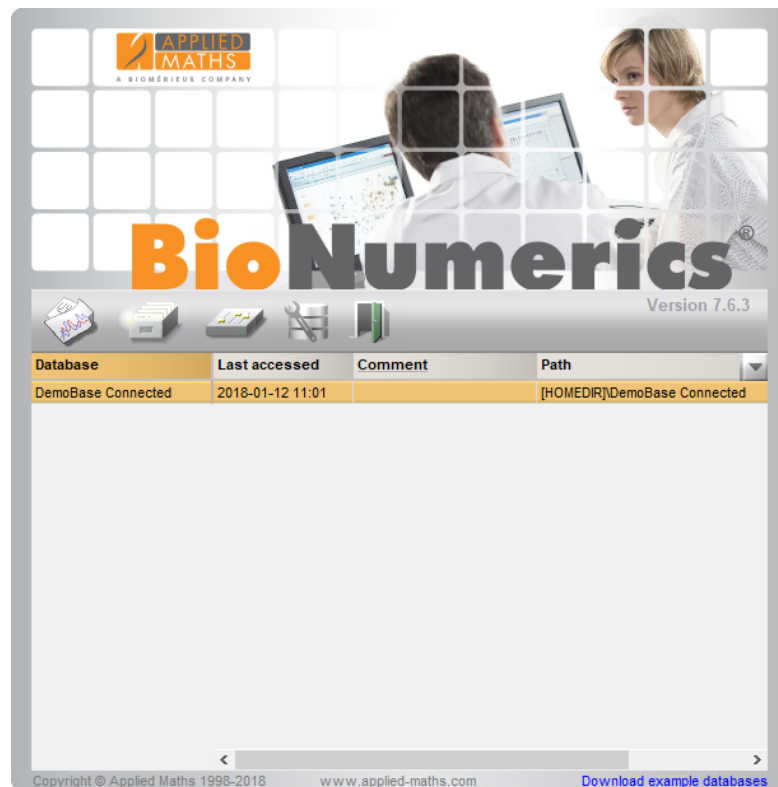
3.2 Enter a name for the database, and press <***Next***>.

Figure 1.2: The *BioNumerics Startup* window.

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

    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>*.
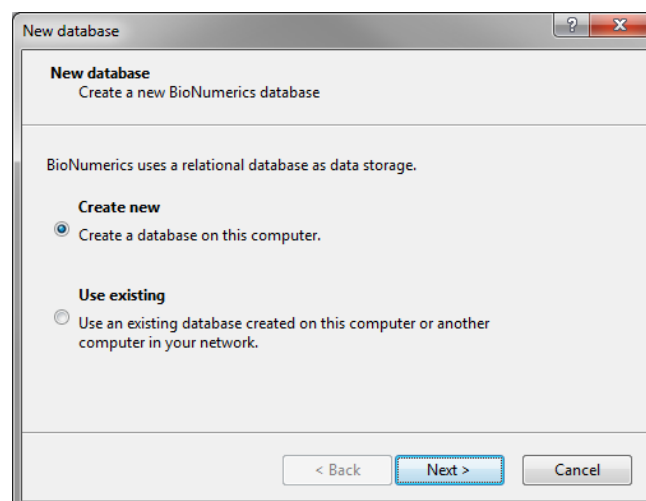

Figure 1.3: The *New database* wizard page.

A new dialog box pops up, prompting for the database engine (see Figure 1.4).

    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.
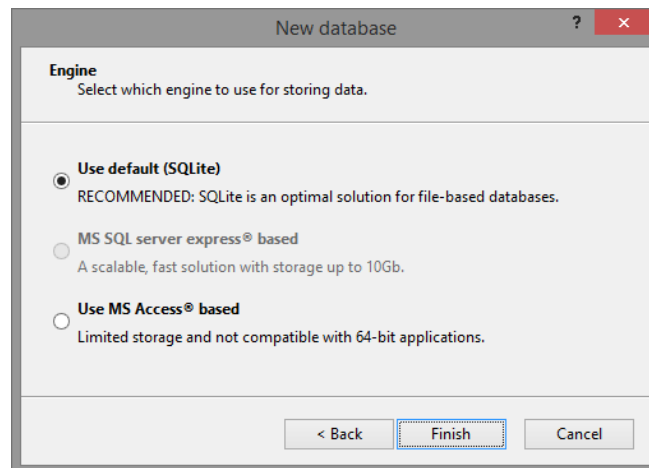
Figure 1.4: The *Database engine* wizard page.

Chapter 2

Installing the HIV Drug Resistance plugin

If a database is opened for the first time, the *Plugins* dialog box will appear by default (see Figure 2.1).

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...* (￼).
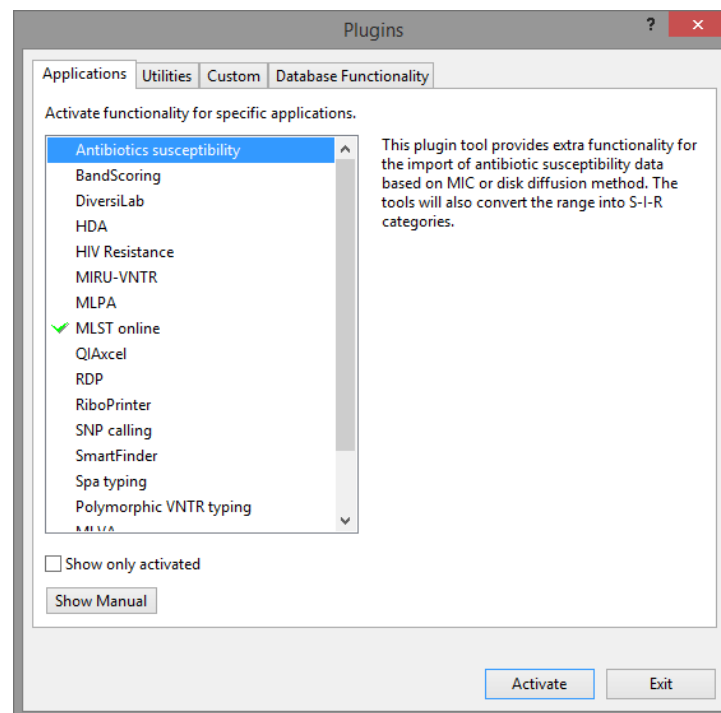

Figure 2.1: The *Plugins* dialog box.

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.

    0.1 Select the *HIV Resistance plugin* from the list in the *Applications tab* and press the <***Activate***> button.

    0.2 The program will ask to confirm the installation of the plugin. Confirm the installation.

The *HIV plugin sequence settings* dialog box pops up, prompting you for a number of options (see Figure 2.2).
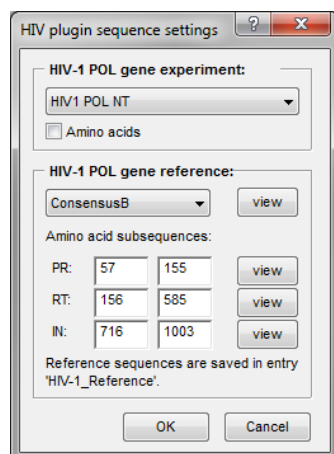

Figure 2.2: The *HIV plugin sequence settings* dialog box.

HIV-1 POL gene experiment: The name for the experiment type that will hold the HIV-1 POL nucleotide sequences (NT) or amino acid sequences (AA) in case the ***Amino acids*** check box is checked.

HIV-1 POL gene reference: The reference sequence that will be used for aligning the sample sequences and determining the mutations/SNPs. Currently two reference sequences are available: ***ConsensusB*** and ***HXB2***. The selected sequence can be viewed using the <***View***> button.
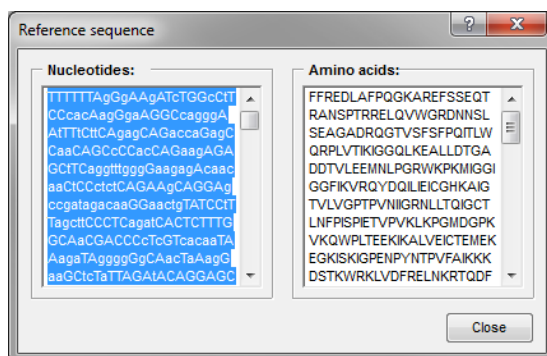

Figure 2.3: The *Reference sequence* dialog box.

The *Reference sequence* dialog box displays the nucleotide (left) and amino acid sequence (right).

The user can also specify an own-defined reference sequence by selecting ***User***. Using the <***View***> button, a nucleic acid and amino acid sequence can be pasted in the editor.

For the selected reference sequence, it is possible to change the start and stop positions for the protease (***PR***), reverse transcriptase (***RT***) and integrase (***IN***) proteins, respectively. The default values are PR: 57-155, RT: 156-585, and IN: 716-1003. The sequences can be viewed using the <***View***> button.

The *Reference amino acid sequence* dialog box displays the amino acid sequence.

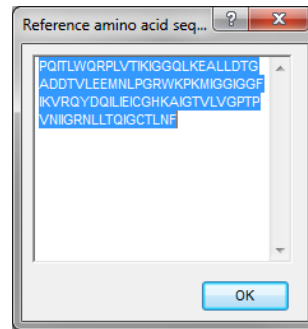   0.3 Press <***OK***> when finished setting the options.

Figure 2.4: The *Reference amino acid sequence* dialog box.

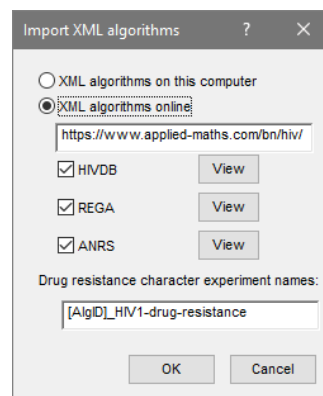The *Import XML algorithms* dialog box prompts for choosing which algorithms to install (see Figure 2.5).

Figure 2.5: The *Import XML algorithms* dialog box.

XML algorithms can be automatically downloaded from the Applied Maths website (which acts as a mirror for the Stanford HIVDB website) or can be imported as files from your local computer.

With ***XML algorithms online*** enabled, the text box will show the URL from which three classes of algorithms are available:

- ***HIVDB*** (Stanford University, CA, USA)

- ***REGA*** (REGA Institute, Leuven, Belgium)

- ***ANRS*** (Agence Nationale de Recherches sur le Sida, France).

With the <***View***> button, the algorithms in XML format can be visualized in the default browser, e.g. to check version dates.

When the option ***XML algorithms on this computer*** is checked, the XML file is prompted for in the next step:

In a first step, the plugin checks the XML file for the presence of the ALGORITHM element's start (<**ALGORITHM**>) and end (</**ALGORITHM**>) tags. If these tags are not present in the selected XML file, the algorithm cannot be parsed by the plugin and the software warns for this.

If the ALGORITHM element tags are detected in the selected XML file, the plugin checks for the presence of the "<**!DOCTYPE ALGORITHM...**> " tag. If no "<**!DOCTYPE ALGORITHM...**> " tag is present in the XML file, the plugin warns you for this. If the selected XML file is a valid algorithm XML document, pressing <***Yes***> imports the algorithm.

For each algorithm that is checked in the *Import XML algorithms* dialog box, a *Character type* experiment will be created, which will contain the final results from analyses with the algorithms. The name of each character experiment should at least include the name of the algorithm (**[AlgID]**). Optionally, one can choose to add a prefix and/or edit the default suggested suffix.

> When working in an Oracle database, make sure the length of the character names does not exceed 24 characters.

0.4 Press <**OK**>.

The XML files of the selected algorithms are downloaded (default option), thus ensuring that always the latest version is used. The downloaded XML files are stored in a sub-folder `HIV XML algorithms` of the database folder. The algorithms are parsed by the software into *Decision Networks*: one Decision Network is generated per algorithm and per drug.

> Additional algorithms can be installed at any time using ***Import decision networks*** (see 3.5).

When the *HIV Resistance plugin* is successfully installed, a confirmation message pops up.

0.5 Press <**OK**>.

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

0.7 Close and reopen the database to activate the features of the *HIV Resistance plugin*.

The *Main* window now contains a menu **HIV** with specific menu items for the *HIV Resistance plugin* (see Figure 2.6).
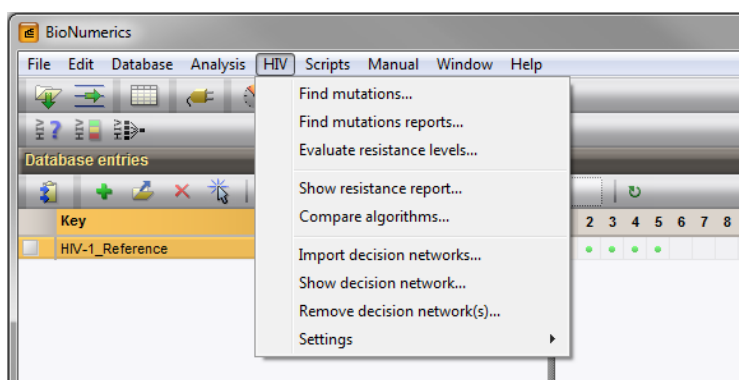

Figure 2.6: The **HIV** menu inserted by the *HIV Drug Resistance plugin*.

In addition, a number of *Experiments* have been automatically created to hold the HIV sequence data and results (see Figure 2.7).

The experiments marked with are *Sequence types*.

- **HIV1 POL NT** will hold the full nucleotide sequences for the POL gene.

- **HIV1 POL NT_TRANSL** will hold the translated sequences derived from the nucleotide sequences in **HIV1 POL NT**.

- **HIV1 PR AA mutations** will store the mutated amino acid positions for the processed protease data.

- **HIV1 RT AA mutations** will store the mutated amino acid positions for the processed reverse transcriptase data.
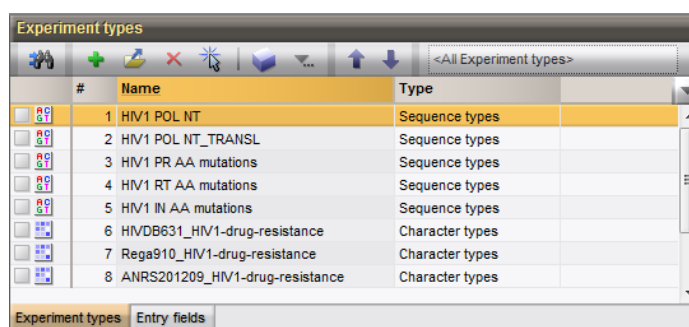
Figure 2.7: The *Experiment types* panel in the *Main* window, containing experiment types created by the *HIV Resistance plugin*.

- **HIV1 IN AA mutations** will store the mutated amino acid positions for the processed integrase data.

In case the ***Amino acids*** check box was checked in the *HIV plugin sequence settings* dialog box (see Figure 2.2), the sequence experiment **HIV1 POL AA** is created and listed in the *Experiment types* panel instead of **HIV1POL NT** and **HIV1 POL NT_TRANSL**.

The experiments marked with 🔲 are *Character type* experiments, which will contain the final results from analyses with the algorithms.

- **Rega***_HIV1-drug-resistance** will contain results from analysis using the REGA algorithm.

- **HIVDB***_HIV1-drug-resistance** will contain results from analyses using the HIVDB algorithm.

- **ANRS******_HIV1-drug-resistance** will contain results from analyses using the ANRS algorithm.

0.8 Double-click on a character type experiment (e.g. **HIVDB***_HIV1-drug-resistance**) to open the *Character type* window (see Figure 2.8).

All drugs used by the algorithm are listed in the *Characters* panel. A default color scale is given to each drug, going from green (susceptible), over yellow (intermediate), to red (resistant).

For the HIVDB algorithm (see Figure 2.8), the degree of resistance is subdivided in 5 levels, ranging from *Susceptible* (1, dark green), over *Potential low-level resistance* (2, light green), *Low-level resistance* (3, yellow) and *Intermediate resistance* (4, orange) to *High-level resistance* (5, red). Note that these levels are determined by the algorithm, not by the plugin. Other algorithms, for example REGA, use a slightly different resistance scale.

0.9 Click on the *Mapping* panel in the *Character type* window (see Figure 2.9).

The *Mapping* panel shows the mapping of the resistance levels to the "SIR" (Susceptible / Intermediate / Resistant) scale.

For the HIVDB algorithm, level 1 and level 2 are mapped to **S** (susceptible). Level 3 and level 4 are mapped to **I** (intermediate), and level 5 is mapped to **R** (resistant). Other algorithms use a slightly different mapping, depending on the number of resistance levels.

0.10 Close the *Character type* window with ***File > Exit***.

The reference is saved in the database entry **HIV-1_Reference**. The protein sequences of protease, reverse transcriptase and integrase are stored in the **HIV1 PR AA mutations**, **HIV1 RT AA mutations**, and **HIV1 IN AA mutations** sequence types respectively, allowing easy comparison with sample sequences.

0.11 Click on the tab *Decision networks* in the *Main* window.

The *Decision networks* panel lists a decision network for each algorithm/drug combination (see Figure 2.10).
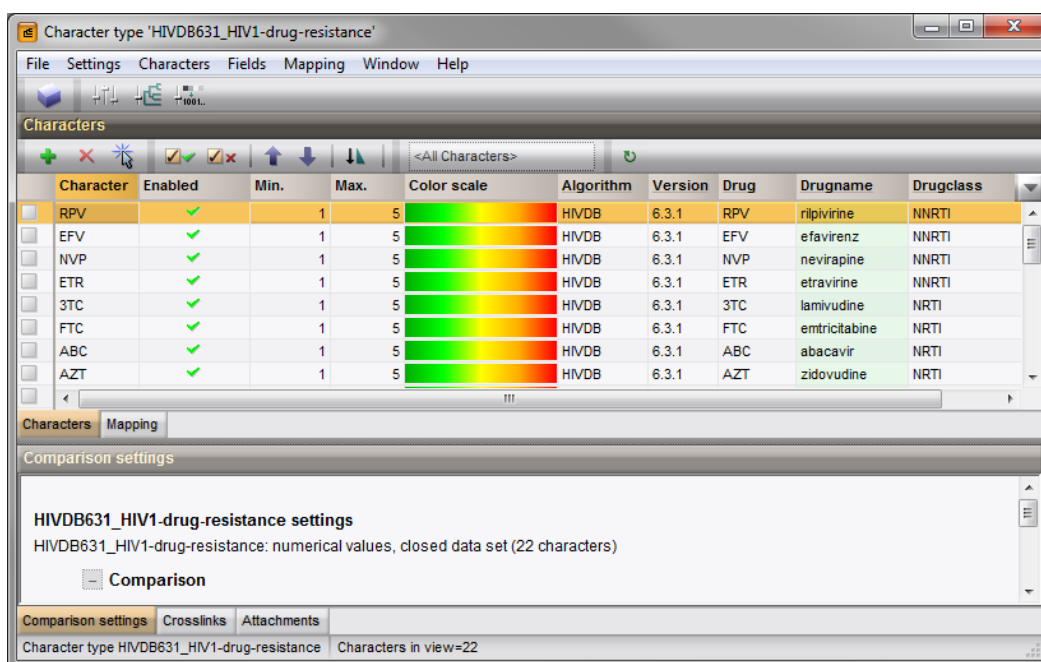
Figure 2.8: The *Character type* window.
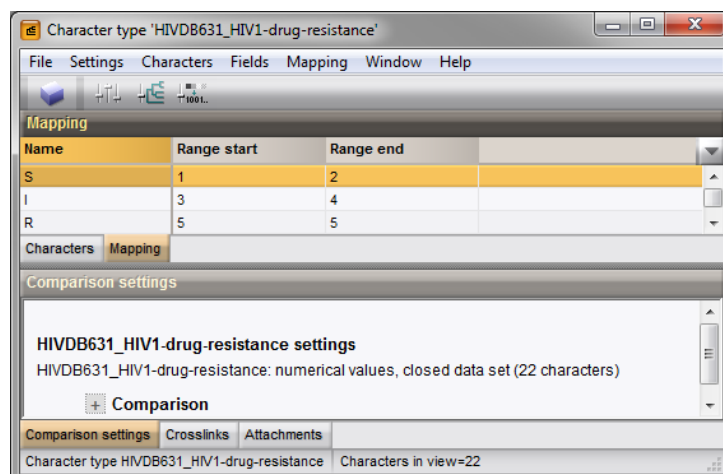
Figure 2.9: The *Character type* window: *Mapping* panel.
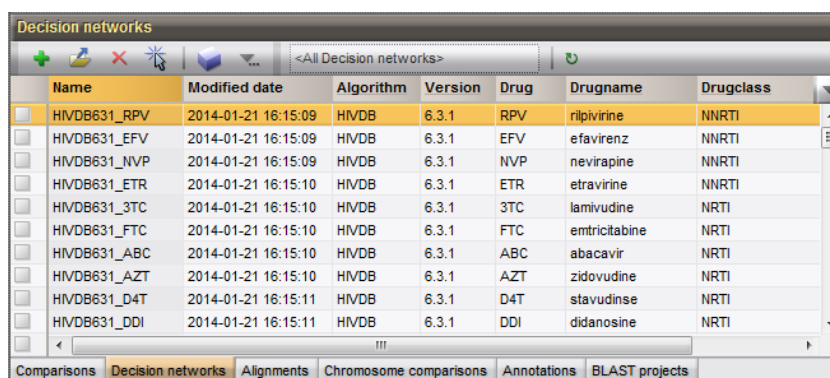
Figure 2.10: The *Decision networks* panel with decision networks created for each algorithm/drug combination.

Chapter 3

Importing and analyzing sequences

3.1 Importing sequences

Sequence data can be entered in several ways in BioNumerics using the sequence import tools in the *Import* dialog box. Choose **File** > **Import...** (, **Ctrl+I**) to call the *Import* dialog box.

- Sequences in FASTA format and EMBL/GenBank format can be imported from text formatted files and linked to new or existing entries in the database using the import options **Import FASTA sequences from text files** and **Import EMBL/GenBank sequences from text files**, respectively.

- With the import routine **Download sequences from internet**, sequences can be fetched from online repositories and linked to new or existing entries in the database.

- Sequences can be imported and assembled with the options **Import and assemble trace files** and **Import and assemble traces from FASTA text files**. Binary chromatogram files from Applied Biosystems, Beckman, and Amersham automated sequencers are accepted, and FASTA formatted sequences stored in text files.

A sample data set `HIV_Dataset.txt` of HIV *POL* nucleotide sequences in FASTA format can be downloaded from the download page on the Applied Maths website (http://www.applied-maths.com/download/sample-data, click on "HIV POL sequences data file"). The descriptions below are based on this data set.

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

    1.2 Select **Import FASTA sequences from text files** under **Sequence type data** and press <**Import**>.

This brings up the first step of the wizard (see Figure 3.1).

    1.3 Press the <**Browse**> button.

    1.4 Browse for the `HIV_Dataset.txt` datafile and select the file.

    1.5 Make sure the option **Preview sequences** is checked and press <**Next**>.

The *Preview* wizard page displays all sequences found in the selected file (see Figure 3.2). The **File name** column holds the name of the selected file; the **Length** column displays the size of the sequences; and the **Header** column holds the information that is present in the description line.
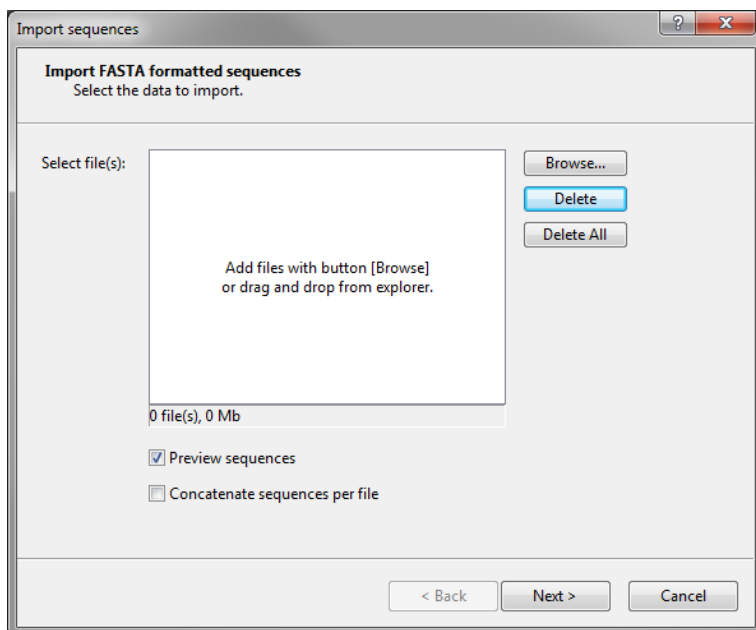
    1.6 Press <**Next**>.

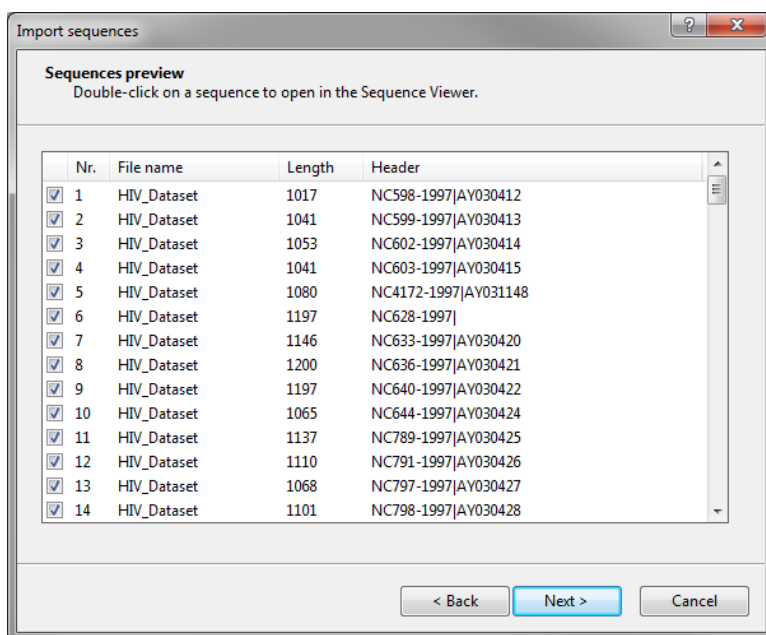**Figure 3.1:** The *Input* wizard page.

Figure 3.2: The *Preview* wizard page.

The way the sequence information should be imported from the selected file into the database needs to be specified with an import template.

> 1.7 Press the <***Create new***> button to create a new import template.

This brings up the *Import rules* dialog box (see Figure 3.3). When sequences are stored in FASTA format, each sequence begins with a single-line description, followed by lines of sequence data. The description line is distinguished from the sequence data by a greater than (">") symbol. The description line contains the *FASTA tags*, separated by a pipe ("|") symbol. Each *FASTA tag* corresponds to a row in the grid (maximum 20 FASTA tags can be parsed from the description line). The text ***FASTA field*** is specified in the ***Source type*** column and the position of the tags in the description line is displayed in the ***Source*** column. The last
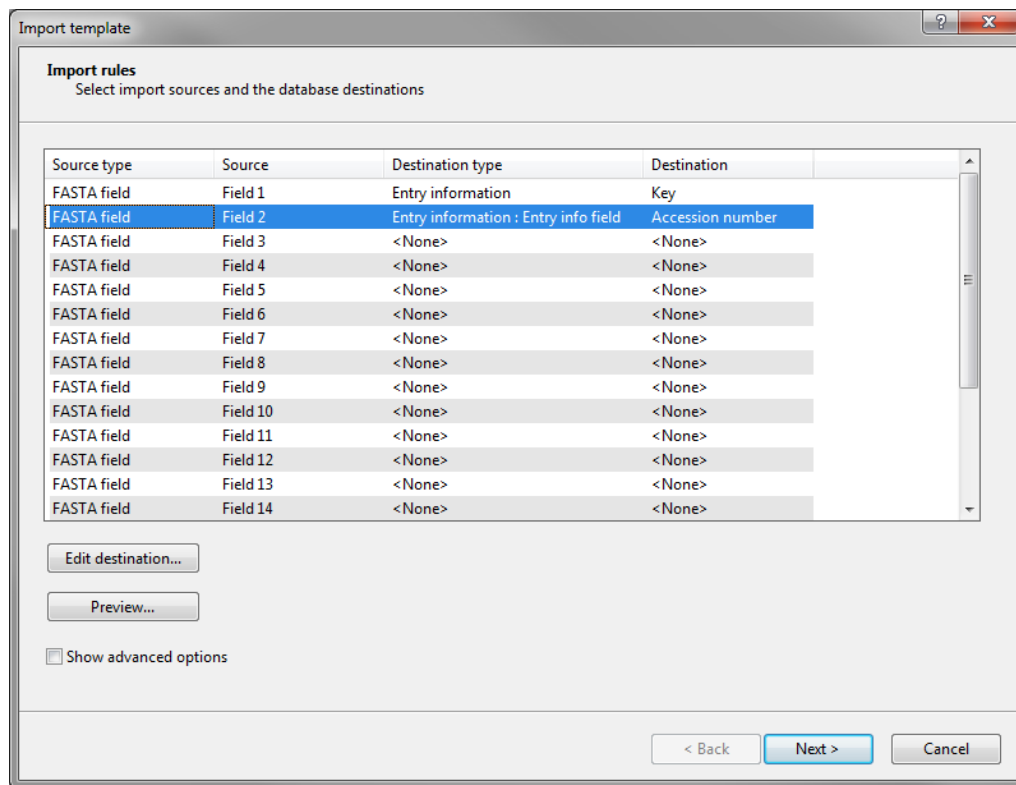
row in the grid holds the name of the file.

Figure 3.3: The *Import rules* dialog box.

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

1.9 Highlight the second row entry in the grid and press the <***Edit destination***> button once more. Select the option <***Create new***> under the topic ***Entry info field*** and press <***OK***>. Provide a name for the new entry information field (e.g. ***Accession number***), press <***OK***> and confirm the creation of the new field with <***Yes***>.

The grid is updated (see Figure 3.3).

1.10 Press <***Next***> to go to the next step.

1.11 Make sure the ***Key*** field is checked and press <***Finish***>.

1.12 Specify a template name e.g. "HIV FASTA", optionally add a description and press <***OK***>.

The import template is added to the list and is automatically selected (see Figure 3.4).

1.13 Select the **HIV1 POL NT** sequence type from the ***Experiment type*** list and press <***Next***>.

1.14 In the last step, leave all settings unaltered and press <***Finish***>.

Some 143 sequences are imported and appear as entries in the database.

The entry with key **HIV-1_Reference** was automatically created by the plugin during the installation, and contains the reference sequence ConsensusB (or HXB2).
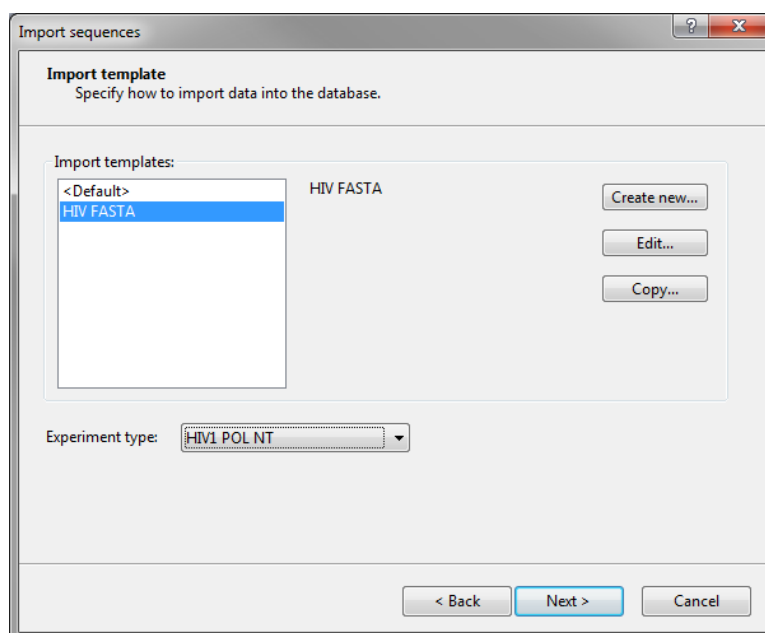
Figure 3.4: Import template added to the list of templates.

3.2   Finding mutations on amino acid sequences

To correctly find the mutated positions on the amino acid sequences, each POL amino acid sequence is aligned against the reference wild type sequence. The plugin uses the sequence types **HIV1 PR AA mutations**, **HIV1 RT AA mutations**, and **HIV1 IN AA mutations**, to store the mutated amino acid positions, insertions and deletions of these proteins.

When using POL nucleotide sequences as starting point in the database, the POL nucleotide sequences are first translated into amino acid sequences. The translation frame is chosen from the 6 possible frames by pairwise alignment with the reference amino acid sequence. The translation is saved in the POL amino acid sequence type experiment (**HIV1 POL NT_TRANSL**).

2.1 Use the **Ctrl-** and **Shift**-keys to select the entries in the database for which you want to perform the mutation search (do *not* select the **HIV-1_Reference**). Check boxes for selected entries are indicated as ☑.

2.2 In the *HIV* menu, select *Find mutations*.

2.3 Press <*Yes*>.

When the calculations are finished, a report pops up (see Figure 3.5).

All mutated positions that were found for the different proteins are listed under *Mutations*. These mutated positions are stored in entry attachment files in the database. The presence of degenerate bases, (possible) stop codons and frame shifts on nucleotide sequences (usually due to sequencing problems and/or base-calling errors) are reported under *Mutations warnings*.

2.4 Close the *Find mutations report window*.

A colored dot is shown in the *Experiment presence* panel for the mutation data of each protein that was present in the sample sequences (see Figure 3.6). In our example data set, only protease (PR) and reverse transcriptase (RT) are present in the sequences, whereas the integrase protein (IN) is not present.

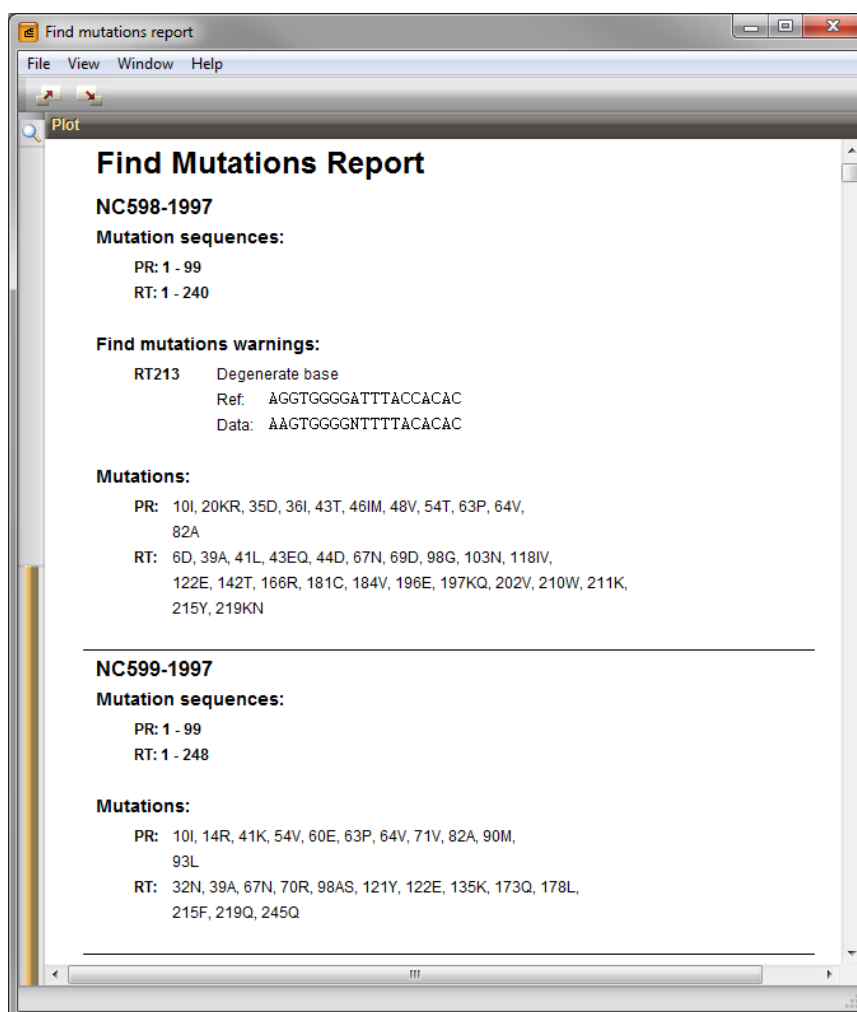2.5 Double-click on one of the entries that was included in the mutation search.

Figure 3.5: The *Find mutations report window*.

2.6 Click on the *Attachments* panel in the *Entry* window.

All mutated positions that were found for the different proteins during the mutation search are stored in entry attachment files. The entry attachment files are listed in the *Attachments* panel (see Figure 3.7).

2.7 Double-click on an attachment file to view all mutations for an individual protein.

2.8 Close the *Entry* window.

2.9 To call the *Find mutations report window* again for a selection of database entries, select **HIV > Find mutations report** in the *Main* window.

3.3   Analyzing resistance levels using algorithms

Using the mutation information as input (see 3.2) resistance levels can be evaluated. More than one algorithm can be evaluated at once.

3.1 Unselect all entries in the database with **Database > Entries > Unselect all entries (all levels)** (🐞, **F4**).

3.2 Select a few entries in the database (e.g. 5). Hold down the **Ctrl**-key while clicking on an entry to select it.
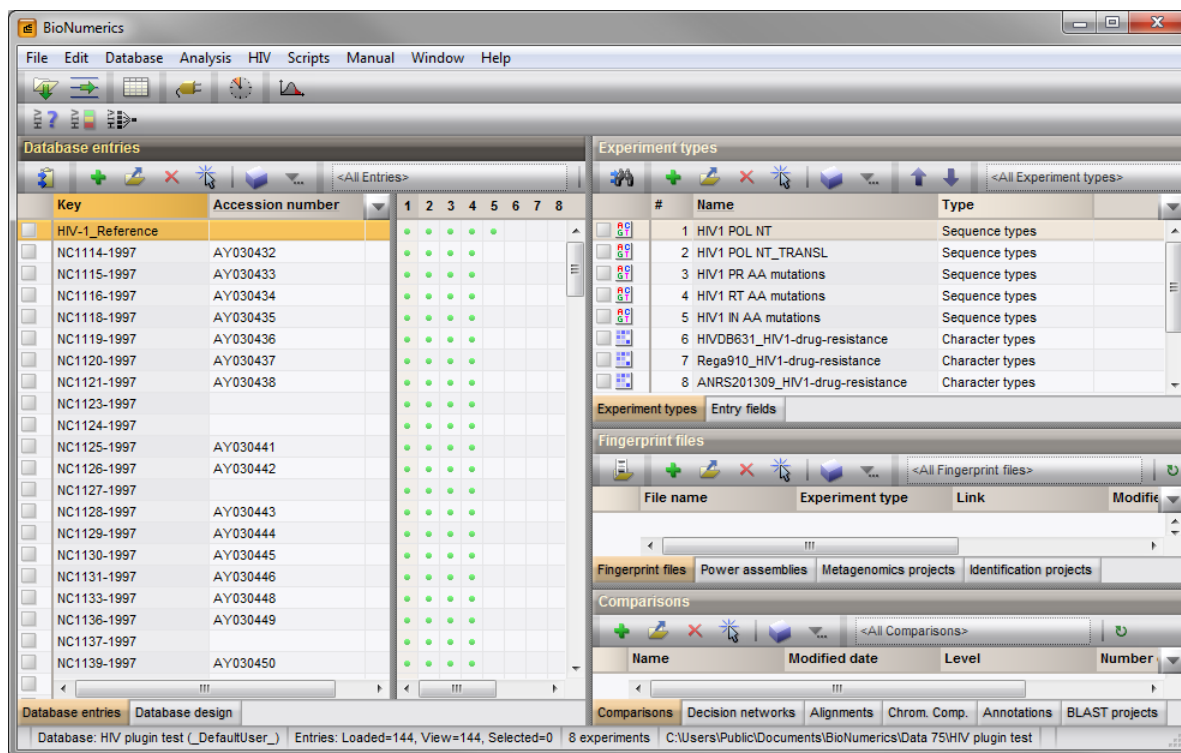
Figure 3.6: The *Main* window after completion of the mutation search.
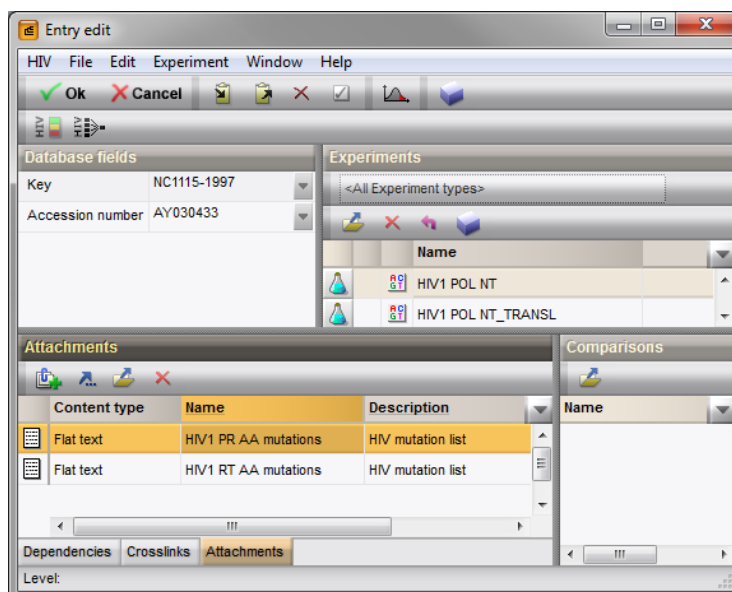

Figure 3.7: The *Entry* window.

      3.3 Select **HIV > *Evaluate resistance levels*** or press the ⊟? button.

This calls the *Select algorithms to run* dialog box (see Figure 3.8).

This dialog prompts to select the algorithms to run.

      3.4 Select all displayed algorithms using the **Shift**-key and press *<OK>*.

The mutations recorded for each selected entry (see 3.2) are fed one by one to the decision networks of the selected algorithm(s) to establish their contribution to the resistance level. In case of ambiguous amino acid
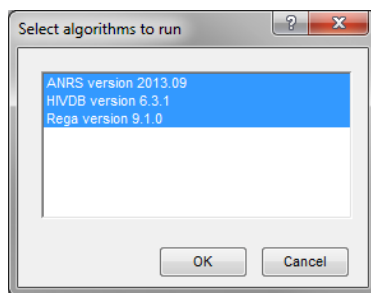
Figure 3.8: Select algorithm(s) to run.

translations, all possible translations are evaluated in the decision network, and the worst case (i.e. inducing highest resistance) is reported.
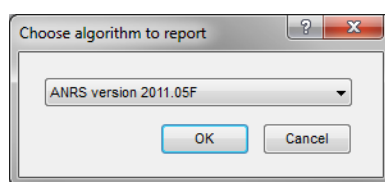

Figure 3.9: The *Select algorithm to report* dialog box.

When the calculations are finished, the plugin asks which algorithm to report in the *Select algorithm to report* dialog box.

> 3.5 From the drop-down list, select e.g. "HIVDB" and press <***OK***>.

If the resistance levels are determined for only one entry, a detailed *Entry report* is shown. When more than one entry was selected, a *Batch report* is shown (see Figure 3.10).

All selected sequence entries are listed in the *Batch report window* (see Figure 3.10). By default, the **Key** information is displayed in the first column, but this can be changed to any other database information field (see 3.6). The available drugs in the selected algorithm (see Figure 2.10) are grouped according to the class (**NNRTI** = non-nucleoside reverse transcriptase inhibitor; **NRTI** = nucleoside reverse transcriptase inhibitor; **PI** = protease inhibitor, **II** = integrase inhibitor).

For the HIVDB algorithm, the degree of resistance is subdivided in 5 levels (see Figure 2.8), ranging from *Susceptible* (1, dark green), over *Potential low-level resistance* (2, light green), *Low-level resistance* (3, yellow) and *Intermediate resistance* (4, orange) to *High-level resistance* (5, red). In the report, both the resistance scale colors and the SIR scale (see Figure 2.9) are shown.

The report can be resized by pressing the zoom buttons and or by dragging the zoom slider with the mouse.

The report can be printed with ***File*** > ***Print*** or copied to the clipboard as a metafile with ***File*** > ***Copy to clipboard***.

The resistance boxes are clickable. When a resistance box is clicked, the *Decision Network* window pops up, with the decision network states for the selected drug. With some experience with the principles of decision networks (see the Reference manual, Chapter Decision networks), this view can be used to determine the mutations or combinations of mutations that have led to the resulting resistance level for the drug.

When clicking on a key, a detailed *Entry report* appears (see Figure 3.11).

In the *Entry report window*, the full information for each drug and the mutations that are responsible for the resistance towards the drug, if any are displayed. By default, no patient information is shown in the report, but this can changed in the *HIV plugin report settings* dialog box (see 3.6).
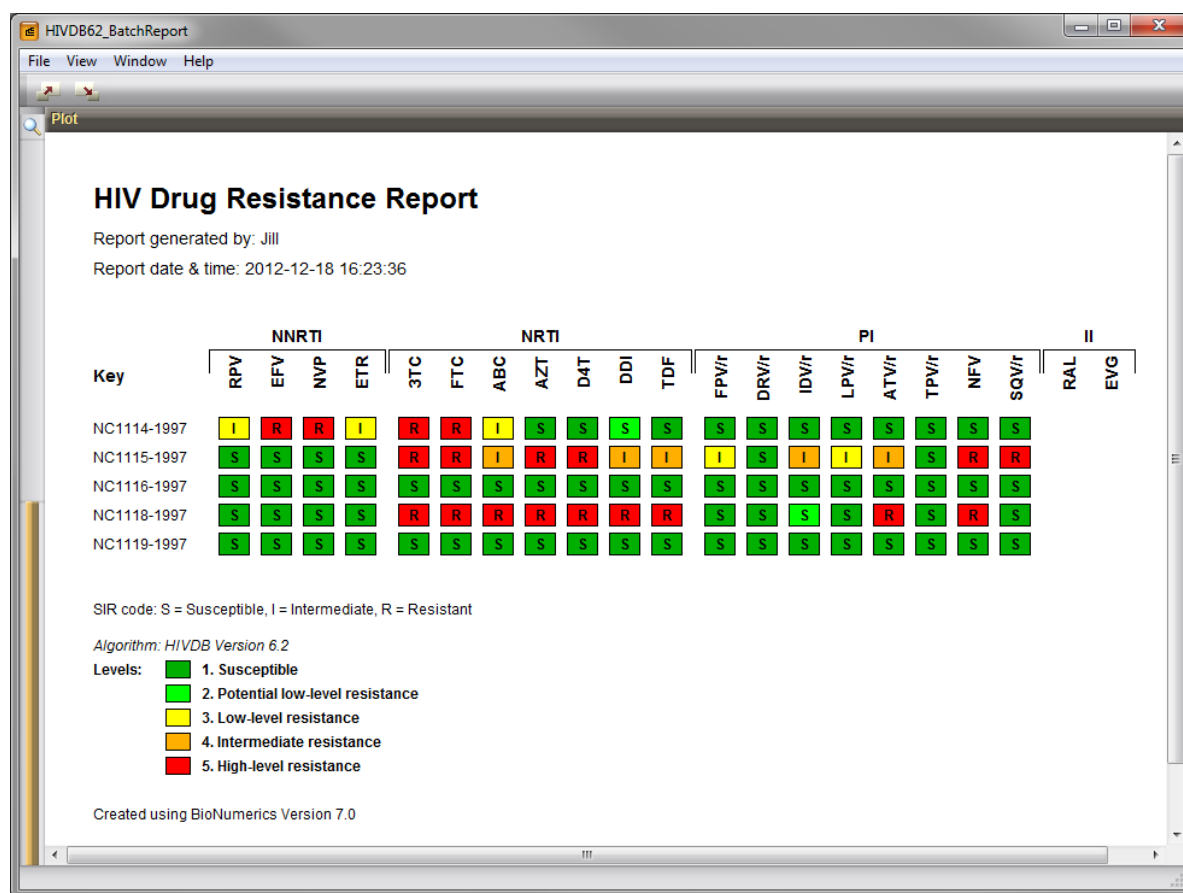
Figure 3.10: The *Batch report window*, with an overview HIV drug resistance report for multiple selected entries.

Note that individual mutations may be plotted in green, indicating that the mutation on its own does not induce resistance, but does so in combination with other mutations. Individual mutations shown in red induce resistance on their own. For each drug class, positions that are considered important by the algorithm, but for which the sequence has a different mutation than the one leading to resistance, are listed in the bottom as **Mutations at algorithm positions**. Other mutations, at positions that do not contribute towards resistance according to the version of the algorithm used, but that were detected by the plugin, are listed as **Other mutations**.

The resistance boxes are clickable. When a resistance box is clicked, the *Decision Network* window pops up, with the decision network states for the selected drug (Figure 3.12). With some experience with the principles of decision networks (see the Reference manual, Chapter Decision networks), this view can be used to determine the mutations or combinations of mutations that have led to the resulting resistance level for the drug.

> 3.6 Close the *Decision Network* window and the HIV drug resistance report(s).

> 3.7 Click on a colored dot in the *Experiment presence* panel to open a character card for one of the algorithms.

The output values of the decision networks are listed in the ***Value*** column. The values are automatically mapped to one of the SIR categories (***Mapping*** column) using the mappings specified in the *Character type* window (see Figure 2.9).

> 3.8 Close the character card(s) by clicking in the top left corner of the card.

> 3.9 To call the *Batch report window* again, select ***HIV*** > ***Show resistance report*** in the *Main* window.
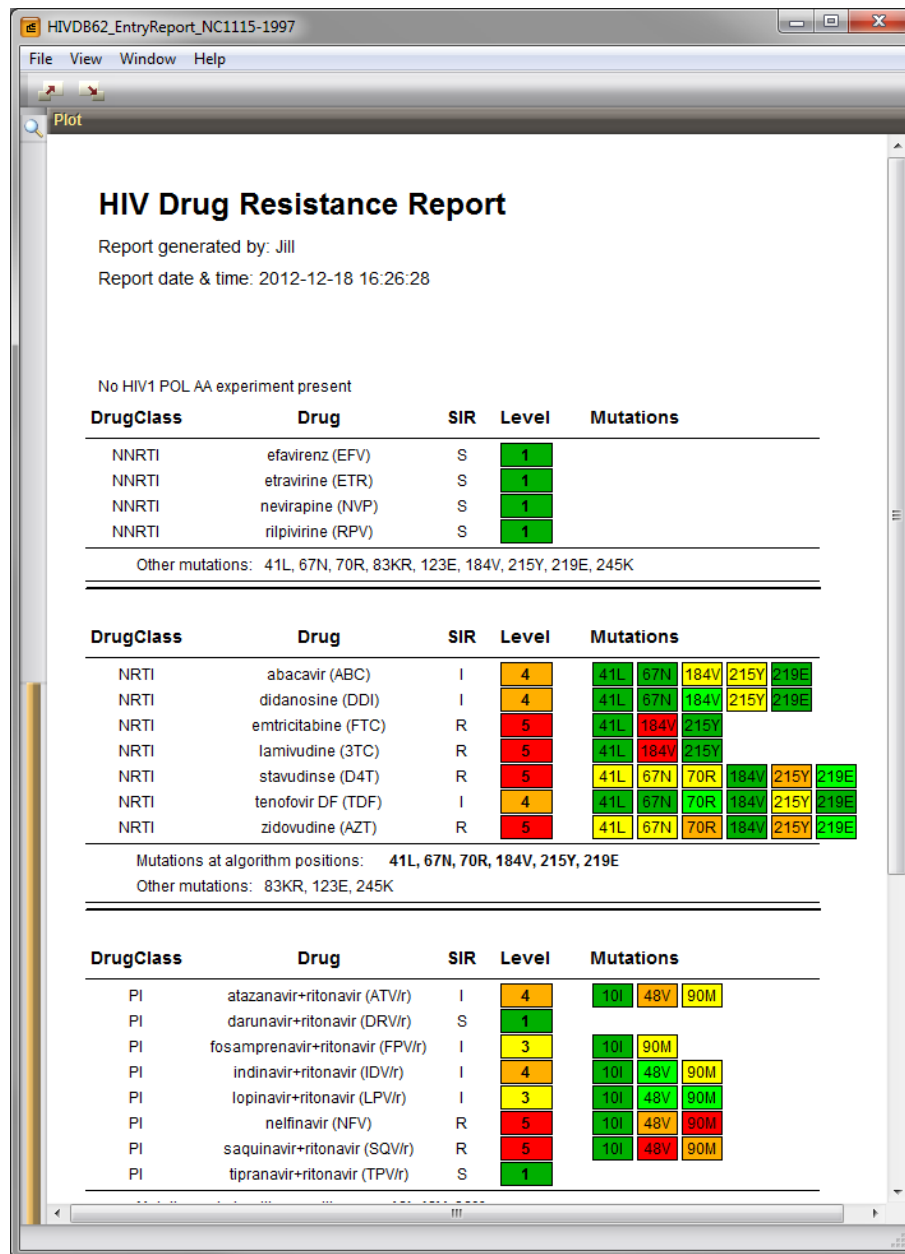
Figure 3.11: The *Entry report window*, with a detailed HIV drug resistance report for a selected entry and algorithm.

The plugin asks which algorithm to report.

   3.10 Select an algorithm from the drop-down list, and press <***OK***>.

If only one entry was selected, the detailed *Entry report window* is shown (see Figure 3.11). When more than one entry was selected, a *Batch report window* is shown (see Figure 3.10).

3.4   Comparing the results of different algorithms

Comparing different algorithms takes a lot of computing time; one report is generated per selected entry.

   4.1 Make sure exactly one entry is selected in the database. This entry needs to have resistance data associated with it.
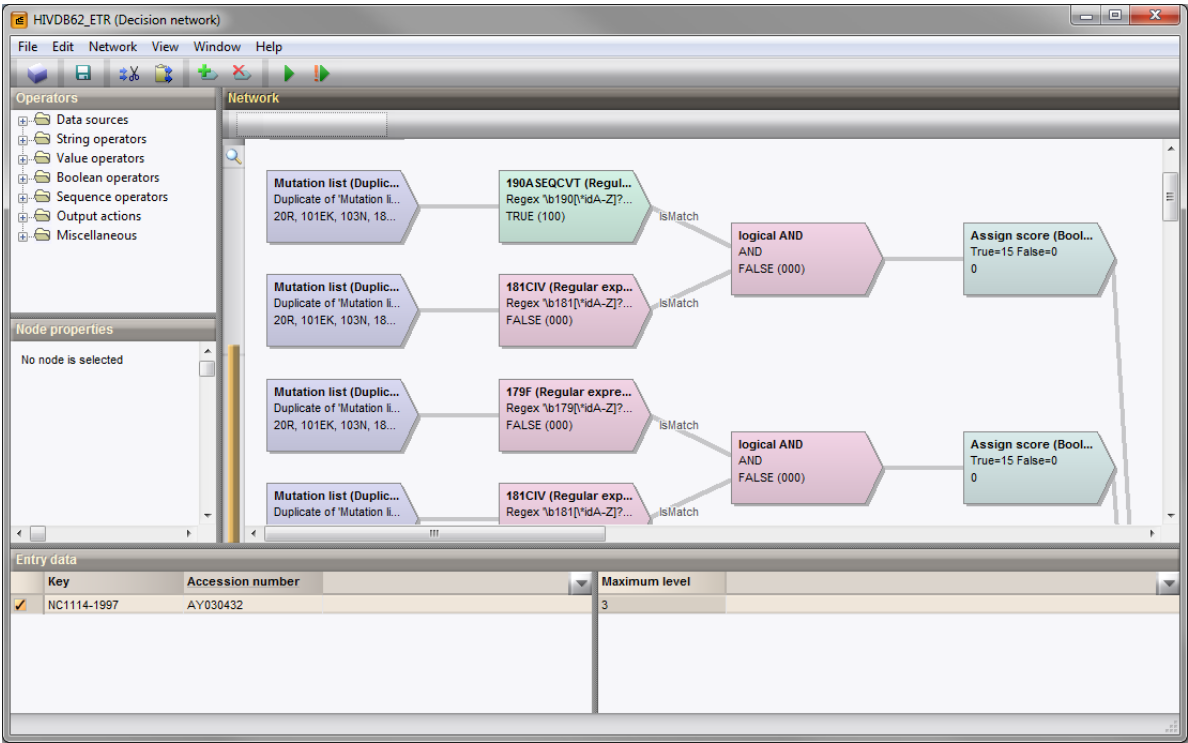
Figure 3.12: Decision network for selected entry and drug.

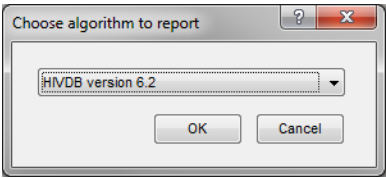**Figure 3.13:** Character card: (a) ANRS, (b) HIVDB.

Figure 3.14: Select algorithms to report.

4.2 Select *HIV > Compare algorithms*.

The *Select algorithms to report* dialog box asks to select the algorithms to compare.

4.3 Select the algorithms you want to compare e.g. **HIVDB** and **Rega** and press *<OK>*.

When the calculations are finished, the *Compare algorithm report window* appears (Figure 3.16).
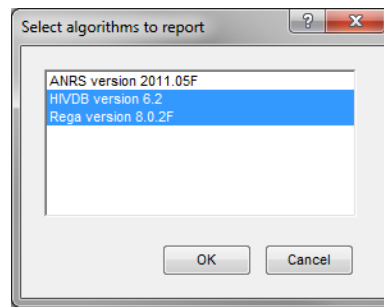
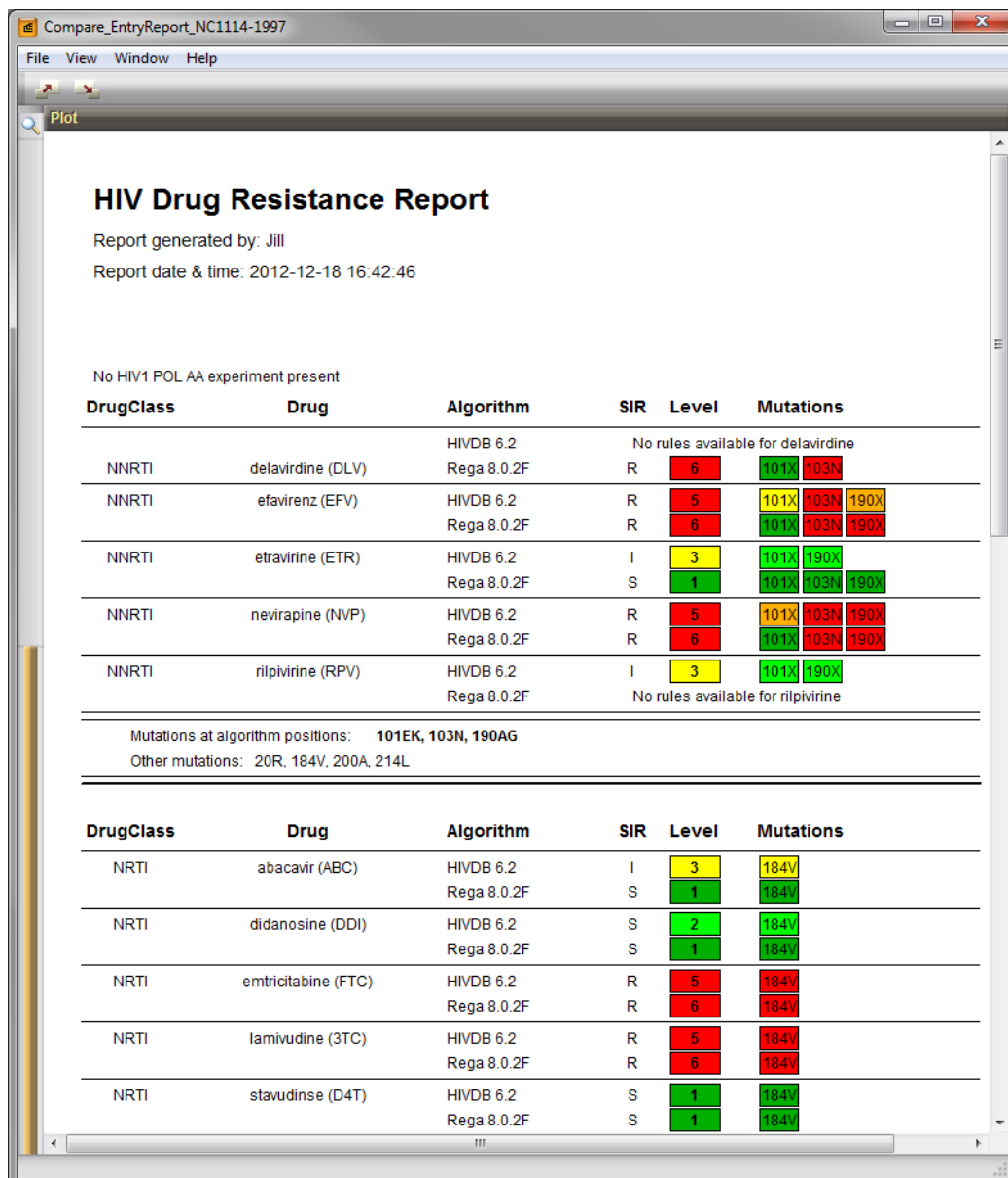**Figure 3.15:** The *Select algorithms to report* dialog box.

Figure 3.16: The *Compare algorithm report window*: a detailed HIV drug resistance report for a single sequence entry and multiple algorithms.

The *Compare algorithm report window* is similar to the *Entry report window*, but it lists for each drug the results for two (or more) algorithms under each other. Similar as for the other report windows, the resistance

indicator blocks are clickable and will cause the decision network for the selected drug and algorithm to be opened.

3.5   Managing drug resistance decision networks

After installation of the *HIV Resistance plugin*, it is possible to update algorithms or install new algorithms at any time.

5.1 With **HIV** > **Import decision networks**, you can add decision networks to the system, based upon new XML algorithms.

This calls the *Import XML algorithms* dialog box, as discussed in 2 (see Figure 2.5).

5.2 Selecting **HIV** > **Show decision network** brings up the *Select Decision Network* dialog box (see Figure 3.17). Alternatively, press the 🔲 button.
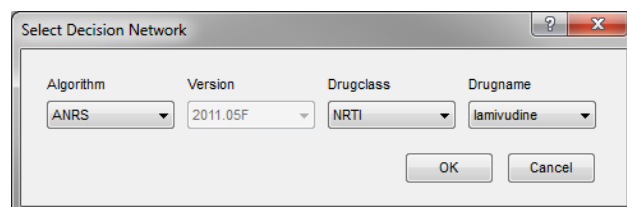

Figure 3.17: The *Select Decision Network* dialog box.

It is possible to search for and display a decision network according to algorithm, version, drug class and drug name.

5.3 Existing decision networks that are no longer of use (e.g. older versions of algorithms) can be removed with **HIV** > **Remove decision network(s)**.

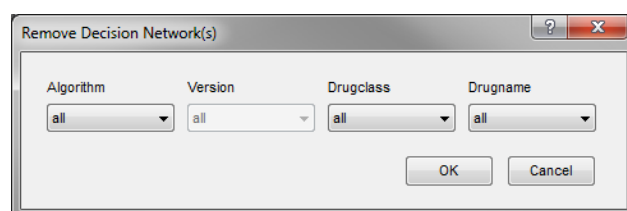The *Remove Decision Network(s)* dialog box appears (see Figure 3.18).

Figure 3.18: The *Remove Decision Network(s)* dialog box.

It is possible to delete a decision network according to algorithm, version, drug class and drug name. **All** can also be chosen for each of the categories.

3.6   Settings of the HIV Drug Resistance plugin

The *HIV plugin sequence settings* dialog box can be opened with **HIV** > **Settings** > **Sequences** (see Figure 2.2).

Information fields shown in the reports can be changed with **HIV** > **Settings** > **Reports**. This calls the *HIV plugin report settings* dialog box (see Figure 3.19).
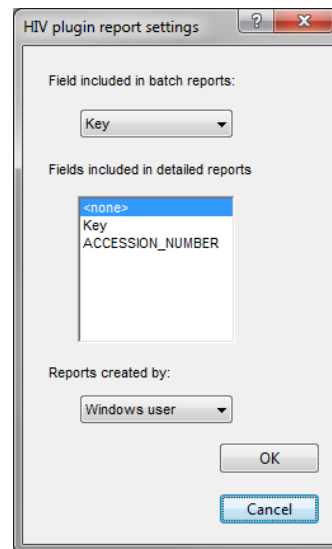
Figure 3.19: The *HIV plugin report settings* dialog box.

A ***Field to be included in batch reports*** can be selected from the list, and one or more ***Fields to be included in the detailed reports***.

The ***Windows user*** or ***Database user*** name can be displayed in the *Report windows*. If the ***Database user*** is an empty string, the user name will not be displayed.

For information on how to create database users, see the Reference manual, Chapter User management.

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