CLC MLST Module

USER MANUAL

User manual for

CLC Multilocus Sequence Typing Module 1.7

Windows, Mac OS X and Linux

February 15, 2017

This software is for research purposes only.

QIAGEN Aarhus
Silkeborgvej 2
Prismet
DK-8000 Aarhus C
Denmark

Contents

Chapter 1

Introduction to the Multilocus Sequence Typing Module

1.1 What is MLST?

With the use of nucleotide sequencing it is possible to perform a fine-scale typing of microbial variants which can greatly improve our understanding of bacterial species, genera, populations, ecosystems and epidemiology.

In order to compare experiments it is required that DNA sequencing and analysis is performed using a portable standard. One such standard is multilocus sequence typing (MLST) which was first proposed by Maiden et al. [Maiden et al., 1998]. Currently, MLST is used to type isolates of bacterial and fungal species for epidemiological and evolutionary studies. For some recent reviews of MLST technology and its applications, see [Taylor and Fisher, 2003, Urwin and Maiden, 2003, Sullivan et al., 2005, Maiden, 2006].

Briefly described, an MLST scheme for a given organism defines a number of internal fragments to be sequenced. These fragments have a length of approximately 450- to 500-bp and are usually chosen to lie in well conserved regions such as housekeeping genes to ensure that general primers can be designed for PCR amplification and sequencing of all species members.

For each fragment, the scheme also contains a dynamic list of the different known alleles. Each allele is assigned a number and by using this numbering system each new isolate can be assigned a complete allelic profile based on all the defined fragments. Each distinct allelic profile in the scheme is called a sequence type (ST) and given a unique number.

1.2 The concept of the Multilocus Sequence Typing Module

The idea behind the module is to eliminate much of the manual work previously required to do MLST. Once a reference scheme has been downloaded, a standard typing of an isolate is done with a few clicks.

The data structure of the Multilocus Sequence Typing Module can be described with two concepts: *scheme* and *isolate*. The MLST scheme contains sequence information and allelic profiles of known sequence types. An isolate is a collection of sequences of one bacteria isolate.

The central point of interaction between the scheme and the isolate is the typing of the isolate.

When an isolate has been created, it is compared with the MLST scheme to determine the sequence type. This is done dynamically every time the isolate is opened, providing the opportunity to extend the scheme with new types and immediately have an updated sequence type of the isolate.

The MLST schemes are described in further detail in section 3, and isolates are described in further detail in section 4

1.3 Acknowledgements

Multilocus Sequence Typing Module uses the databases of:

- PubMLST http://pubmlst.org
- MLST.net http://www.mlst.net
- University of Warwick (developed and previously hosted at University College Cork) http://mlst.warwick.ac.uk/mlst/

The databases are accessed through PubMLST.

Chapter 2

MLST quick-start

This section briefly describes how to get started using the main functionality of the Multilocus Sequence Typing Module.

First, download an MLST Scheme:

> **Multilocus Sequence Typing () | MLST Schemes () | Download MLST Schemes ()**

In this dialog, select the correct scheme to use as reference for typing. Save the scheme ().

Next, to find a sequence type for your sequencing data, you will have to *create an isolate*:

> **Multilocus Sequence Typing () | Assemble and Create Isolate ()**

In this dialog, select all the sequencing data for the isolate that you wish to type. (Note that you will have to **Import** () the data first.)

Click **Next** and select the scheme () that you downloaded as reference scheme. Click **Next** through a series of steps where sequences are assigned to each locus in the scheme. If the sequence is not automatically assigned correctly, use the **Add** () and **Remove** () buttons to make sure the right sequences are assigned.

When you click **Finish**, an isolate is created and a sequence type is assigned. Double-click a row in the upper part of the view to inspect and edit the contig.

Chapter 3

Handling MLST schemes

An MLST scheme () consists of the following:

- Names of the loci to be sequenced.
- For each of these loci a list of the known alleles is maintained.
- A list of the known allelic profiles or sequence types described by the scheme.

Given this information, sequence data from a new isolate can be used to assign a sequence type.

In the Multilocus Sequence Typing Module, an MLST scheme can be created in two ways:

- It can be downloaded from a database, or
- Created from sequence data imported into the workbench.

The schemes downloaded from a database will already have information about many sequence types, whereas this information needs to be accumulated by the user for self-made schemes.

3.1 Downloading schemes

MLST schemes can be downloaded from the databases at PubMLST.org, MLST.net and University of Warwick.

To select schemes for download:

> **Multilocus Sequence Typing () | MLST Schemes () | Download MLST Schemes ()**

This will bring up a dialog as shown in figure 3.1.

The list of schemes is dynamically updated with the available schemes. This means that it takes a short while before the dialog is shown (depending on your internet connection). You can select one or more schemes for download.

Click **Next** if you wish to save the scheme(s) without opening. If not, click **Finish**, and the download will begin.

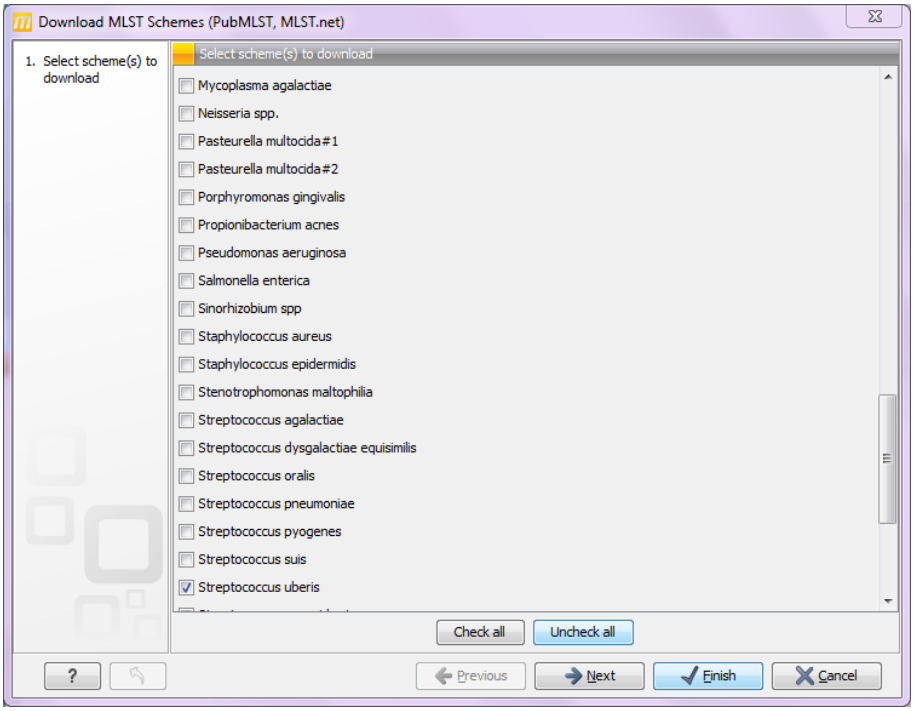


Figure 3.1: *Selecting schemes for download.*

See section 3.4 for information about the content of the downloaded scheme.

3.1.1 Download custom schemes

In addition to the MLST schemes that the Multilocus Sequence Typing Module can automatically download (see section 3, you can also provide a URL to an MLST scheme on another web server. The scheme has to be in the same format as the schemes on PubMLST.

Figure 3.2 shows the dialog where you can enter a URL.

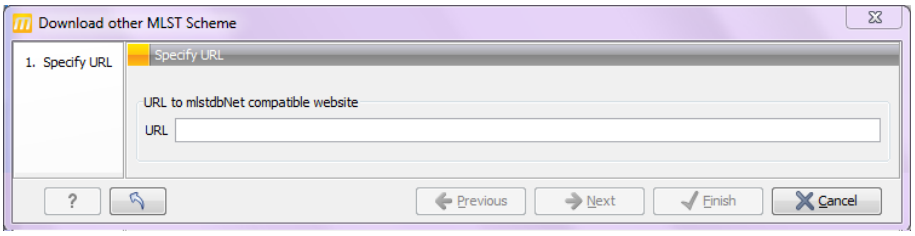


Figure 3.2: *Providing a URL for downloading schemes.*

An example of a URL that can be used to download a scheme is `http://www.pasteur.fr/recherche/genopole/PF8/mlst/Abaumannii.html`.

3.2 Creating new schemes

A scheme can be created from scratch by defining a number of loci and accumulating sequence information about each loci. The schemes have the same format as the schemes that are downloaded.

To create a new scheme:

> **Multilocus Sequence Typing** ( ) | **MLST Schemes** ( ) | **Create MLST Scheme** ( )

This will bring up a dialog as shown in figure 3.3.

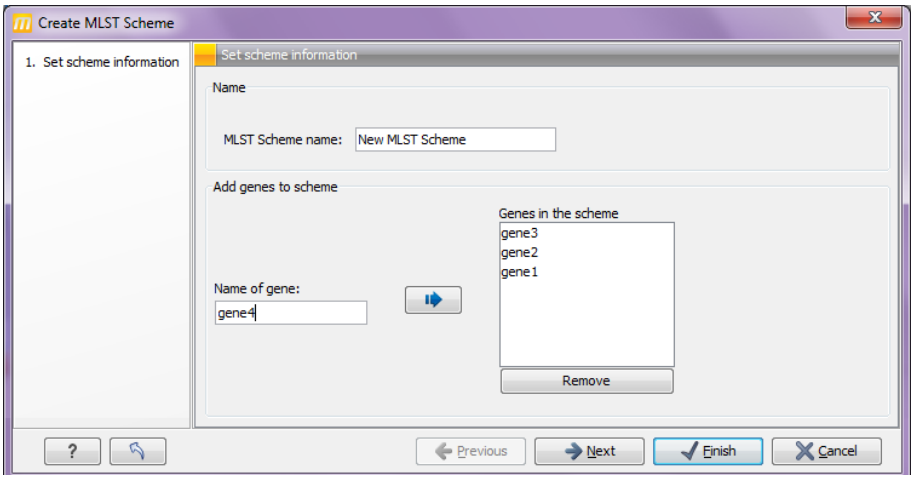


Figure 3.3: *Entering basic information for the scheme.*

In the top field you can enter a name for the scheme - typically the name of the organism in question. Below, you can define a number of loci to be included in the scheme: Type the name of the locus in the **Name of gene** field, and click the **Add** ( ) button.

If you accidentally add a wrong name to the list, select it and click the **Remove** button.

Click **Next** to start assigning sequences to the scheme.

Now follows a number of steps - one for each locus in the scheme. In each of these steps, you can assign a number of allelic sequences to each locus.

In figure 3.4 you can see that the locus called *gene1* from figure 3.3 has its own step where sequences can be added. Clicking **Next** will display a similar dialog for *gene2*, and similar for all the loci that were entered in the first dialog.

It is not mandatory to add sequences to the loci. You can also choose to proceed and create the scheme before you add sequences. If you just add individual sequences as shown in figure 3.4, they will not be combined in an allelic profile and the scheme will not yet contain any sequence types. See section section 5 about how to add **Isolates** to a scheme.

Note! If you wish to use automatic assignment of sequences when creating isolates (explained in section 4.1), there has to be at least one sequence per locus in the scheme.

3.3 Merging schemes

The Multilocus Sequence Typing Module is designed to make it easy to extend a scheme with your own isolates (see chapter 5). This is done be adding sequence information to your local copy of a scheme - either one you created yourself from scratch or a scheme downloaded from a database.

When you extend your own copy of a downloaded scheme, you may find yourself in a situation

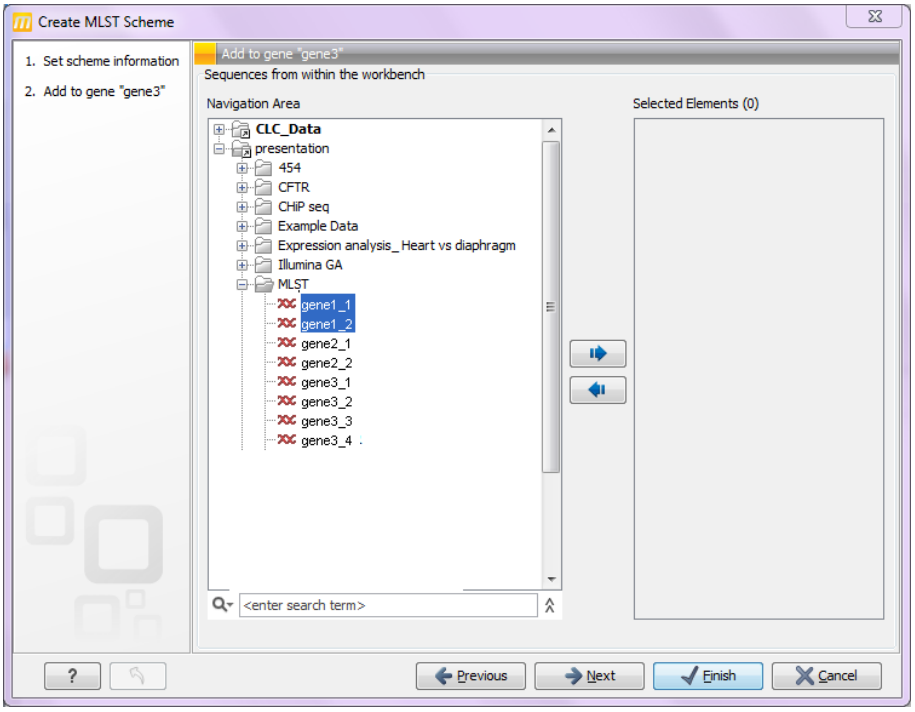


Figure 3.4: *Selecting sequences for gene1.*

where there has been updates to the scheme at the database including new alleles and/or sequence types, and you now have two schemes that you wish to combine into one. Just downloading and using the new scheme from the database would mean discarding your own contributions to your local copy of the scheme.

For this reason, the Multilocus Sequence Typing Module provides a merge function that makes it easy to combine two schemes:

Multilocus Sequence Typing () | MLST Schemes () | Merge MLST Schemes ()

This will bring up a dialog where you select the two schemes that you wish to merge. The schemes have to have the same loci definitions in order to be merged (i.e. the same number of loci with identical names).

Clicking **Next** displays the dialog shown in figure 3.5.

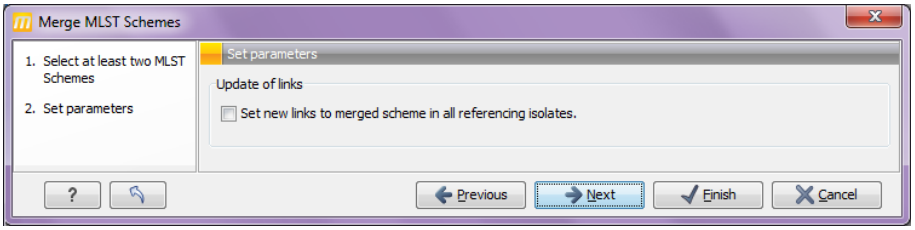


Figure 3.5: *This option makes it possible to update all references in isolates previously referring to one of the two old schemes so that they refer to the new scheme.*

If you check this option, all the isolates that previously referred to one of the two old schemes in order to determine their sequence type can now be updated to refer to the new scheme instead (see section 4.2.1 for more information). Checking this option means that all isolates are typed

according to the new scheme.

Click **Next** if you wish to save the new scheme without opening. If not, click **Finish**, and the new scheme will be opened.

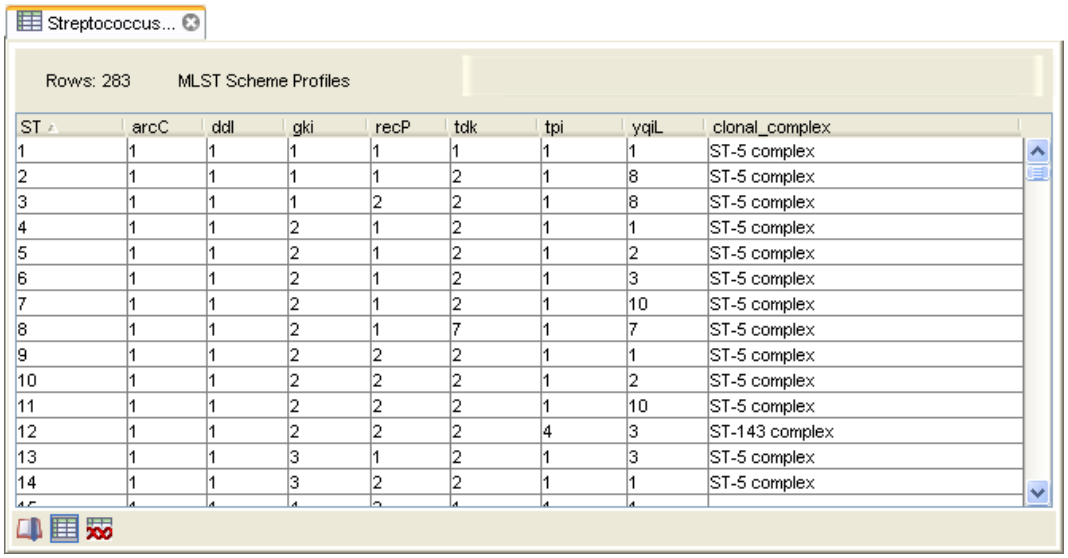
If there are duplicates among the sequence types in the scheme, only one of the types will be included in the new scheme.

3.4 Viewing schemes

There are two ways of viewing the MLST schemes (besides the history view) as described below:

3.4.1 Profile table

The profile table shows all the sequence types in one table (see an example in figure 3.6).



ST	arcC	ddl	gki	recP	tdk	tpi	yqiL	clonal_complex
1	1	1	1	1	1	1	1	ST-5 complex
2	1	1	1	1	2	1	8	ST-5 complex
3	1	1	1	2	2	1	8	ST-5 complex
4	1	1	2	1	2	1	1	ST-5 complex
5	1	1	2	1	2	1	2	ST-5 complex
6	1	1	2	1	2	1	3	ST-5 complex
7	1	1	2	1	2	1	10	ST-5 complex
8	1	1	2	1	7	1	7	ST-5 complex
9	1	1	2	2	2	1	1	ST-5 complex
10	1	1	2	2	2	1	2	ST-5 complex
11	1	1	2	2	2	1	10	ST-5 complex
12	1	1	2	2	2	4	3	ST-143 complex
13	1	1	3	1	2	1	3	ST-5 complex
14	1	1	3	2	2	1	1	ST-5 complex

Figure 3.6: *The profile table.*

Each allelic profile is assigned a number - the **Sequence type (ST)**. The following columns displays the allele type for each locus. It is the combination of these allelic types which defines the sequence type (see section 1.1).

The last column displays information about the **Clonal complex** of this sequence type.

The table can be sorted by clicking the column headers. Pressing Ctrl (⌘ on Mac) while clicking the column makes a secondary sorting within the primary sorting.

You can search the table by using the **Filter** at the upper right part of the view. E.g. typing *ST-143* will show sequence types within the ST-143 clonal complex.

Extracting alignments

If you wish to compare the sequences of two or more sequences types in a multilocus sequence analysis, you can easily extract concatenated alignments from the profile table. To extract an alignment displaying the sequences of one or more loci:

Select the sequence types you wish to align | **right-click** | **Extract Alignment** | **Sequence Types** or **Alleles** (see below)

The two options in the alignment right-click menu provide two ways of composing the alignment:

- **Sequence Types (possibly redundant)**. An alignment is made with alleles from all the selected sequence types, even if the same allele occurs more than once.
- **Alleles (non-redundant)**. Redundant alleles are excluded from the alignment where possible. If you e.g. select three rows in the table and two of the sequence types have the same allele sequence in all the selected loci, the alignment will only include two sequences (the first option would have included all three sequences in the alignment).

Selecting either option will display a dialog similar to the one shown in figure 3.7 displaying the loci in the scheme.

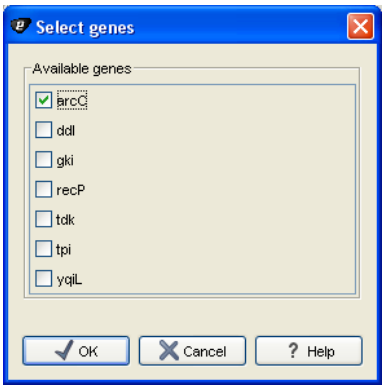


Figure 3.7: *Selecting genes for alignment.*

Before calculating the alignment, the allele sequences for each sequence type will be joined into a "supergene" sequence which is used for the alignment. The checkboxes shown in figure 3.7 let you specify which of the loci that should be used in the joined sequence. Selecting only one locus will create a "non-supergene" alignment with allele sequences for only this particular locus.

When you click **OK**, a dialog with alignment parameters will be shown. To learn more about these parameters, click the dialog's **Help** (?) button.

Extracting sequences

If you wish to extract unaligned concatenated sequences for other analyzes or if you wish to inspect the sequence graphically, you can easily extract the sequences of one or more sequence types to a sequence list:

Select the sequence types you wish to extract | **right-click** | **Extract Sequences** | **Sequence Types** or **Alleles** (see below)

The two options in the right-click menu provide two ways of composing the list (equivalent to extracting alignments (see above)):

- **Sequence Types (possibly redundant)**. A sequence list is made with alleles from all the selected sequence types, even if the same allele occurs more than once.

- **Alleles (non-redundant)**. Redundant alleles are excluded from the list where possible. If you e.g. select three rows in the table and two of the sequence types have the same allele sequence in all the selected loci, the list will only include two sequences (the first option would have included all three sequences in the list).

Selecting either option will display a dialog similar to the one shown in figure 3.7 displaying the loci in the scheme.

The sequences in the sequence list will be created by joining the sequence for each locus that you select in this dialog.

Click **OK** to create the sequence list.

Adding and removing sequence information

For information about *adding* sequence information to a scheme, see chapter 5. To *remove* sequence types from the scheme:

> **Select the sequence types you wish to remove** | **right-click** | **Remove Sequence Type**

This will remove the sequence type from the scheme. The allele sequences will not be removed, since they can be shared among several sequence types. To delete the allele sequences, see chapter 5.

3.4.2 Allele table

The profile table described above provides an overview of the sequence types in a given scheme. The **Allele table** is another way of viewing an MLST scheme which provides an overview of the alleles that are included for each locus (see figure 3.8).

Switching between the two table views of the scheme is done by clicking the buttons at the lower left corner of the view () / ().

At the top of the view is a master table including all loci of the scheme (in the example shown in figure 3.8 there are seven loci). Selecting a row in the master table updates the allele table below which displays all the alleles of the selected locus. The columns in the allele table provide information about allele name, allele number, fragment length, creation data and a list of the sequence types in which a given allele is found.

Displaying the sequences

If you wish to open the sequence of one of the allelic types:

> **Select one or more allelic types** | **right-click** | **Extract as Sequence List**

This will open a sequence list of the allelic types that were selected.

Creating an alignment comparing the allelic types

If you wish to create an alignment of the sequences of two or more allelic types:

> **Select the allelic types** | **right-click** | **Align Sequences**

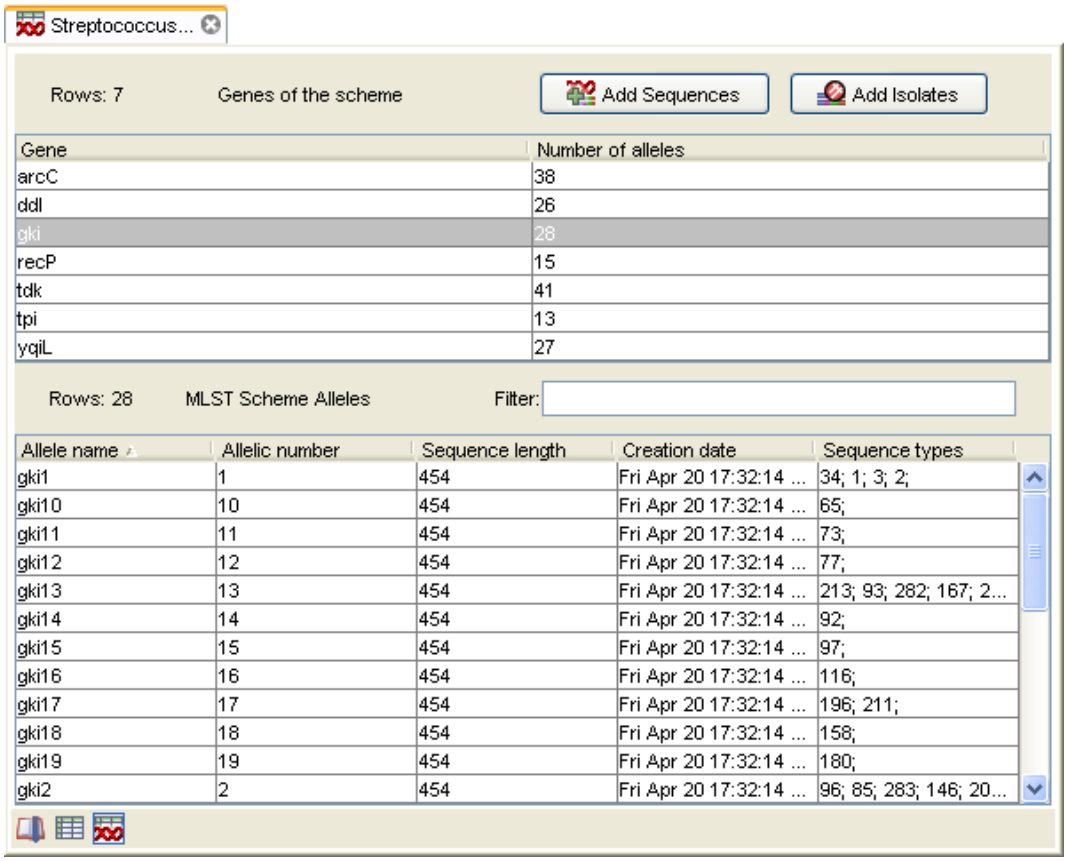


Figure 3.8: *The allele table.*

This will show an alignment parameters dialog. To learn more about these parameters, click the dialog's **Help** (?) button.

Removing allele sequences

To remove one of the allelic types:

Select the allelic type | **right-click** | **Remove**

Note! This will also remove all the sequence types containing this allelic type from the scheme. If you remove an allelic type by accident, you can always **Undo** ( ).

Chapter 4

Isolates

The **Isolate** ( ) is the fundamental data object in the Multilocus Sequence Typing Module. An isolate in this context constitutes a collection of sequencing results - one for each locus. A sequencing result is represented as a contig which has been assembled from one or more sequencing reads (usually a forward and a reverse read) and a reference sequence defined by the scheme. The creation of an isolate is always based on a given scheme and consists of the assembly of sequencing data for each loci, assignment of allele type for each loci given the alleles in the scheme, and final assignment of sequence type given the sequence types in the scheme. In the Multilocus Sequence Typing Module] all these steps are performed automatically and rapidly.

4.1 Assemble and create isolate

An isolate is created by selecting a number of sequences/sequencing data and an MLST scheme.

To create an isolate:

Multilocus Sequence Typing ( ) | Assemble and Create Isolate ( )

This will bring up a dialog as shown in figure 4.1.

In this example, 14 sequences have been selected: a forward and a reverse read for each locus. Even if you do not have sequences for all loci, you can create the isolate and add more sequence later (see section 4.2.6).

Clicking **Next** displays the dialog shown in figure 4.2.

At the top of the dialog, you set the MLST scheme to be used for this isolate by clicking the small button ( ). You have to either create a new scheme or download a scheme and save it in the **Navigation Area** before you can create the isolate (see chapter 3 on how to create and download schemes).

Below there are the options regarding assembling the sequences into contigs:

- **Minimum aligned read length.** The minimum number of nucleotides in a read which must be successfully aligned to the contig. If this criteria is not met by a read, the read is excluded from the assembly.

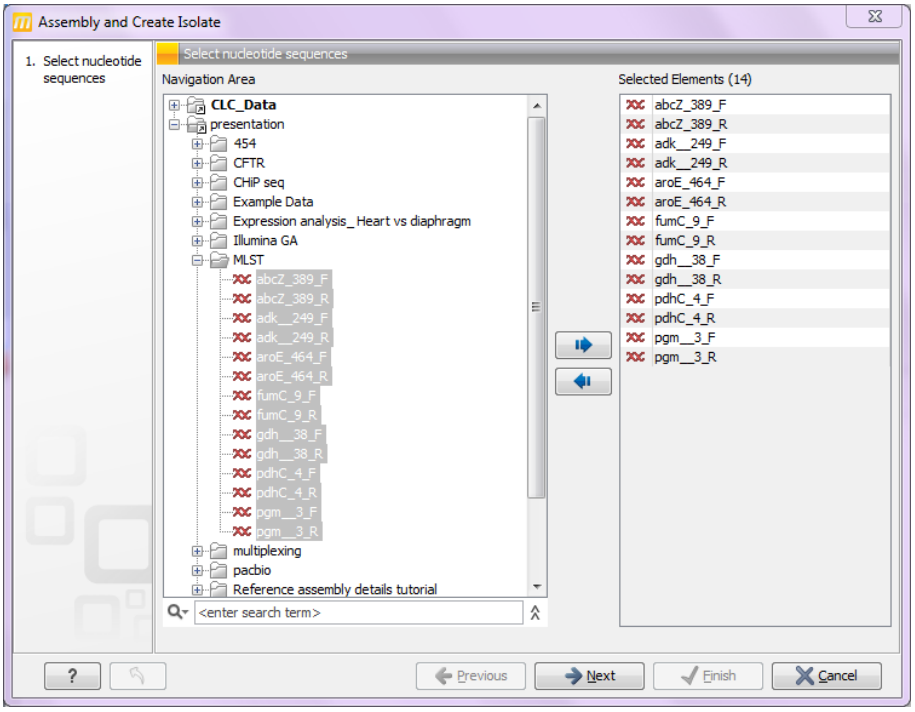


Figure 4.1: *Selecting isolate sequences.*

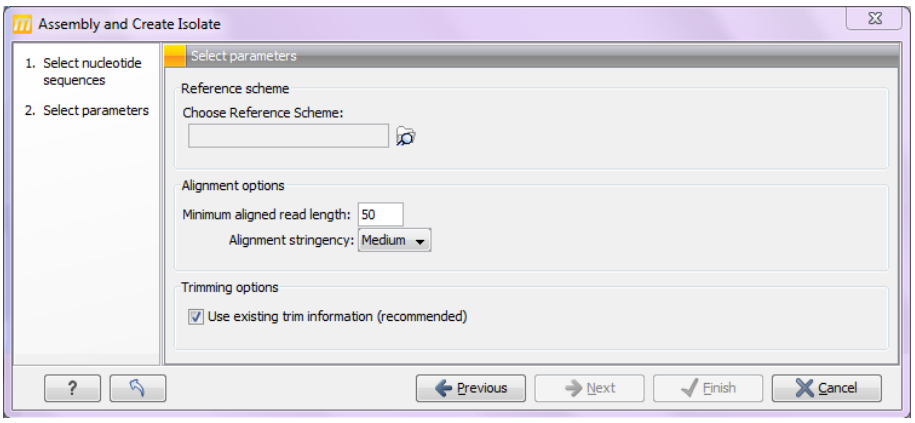


Figure 4.2: *Setting a reference scheme for typing the isolate.*

- **Alignment stringency.** Specifies the stringency of the scoring function used by the alignment step in the contig assembly algorithm. A higher stringency level will tend to produce contigs with less ambiguities but will also tend to omit more sequencing reads and to generate more and shorter contigs. Three stringency levels can be set:
  - **Low.**
  - **Medium.**
  - **High.**
- Use existing trim information. If you have previously trimmed ( ) the sequences, you can select this option to use the trim information. This is generally not necessary since the sequences are assembled to a reference sequence from the scheme. If the contigs in the isolate are not calculated correctly, you can try to trim the sequences first.

When a scheme has been selected, it provides reference sequences for each of the loci. When

you click **Next**, the Multilocus Sequence Typing Module automatically compares each of your sequences to these reference sequences and if they are homologous, your sequence is assigned to this locus. This eliminates the need for manually selecting which sequences belong to which loci.

The following steps in the dialog traverse through the loci, showing which sequences have been assigned to each locus. An example is shown in figure 4.3 where two sequences are assigned to the *abcZ* gene (listed to the right).

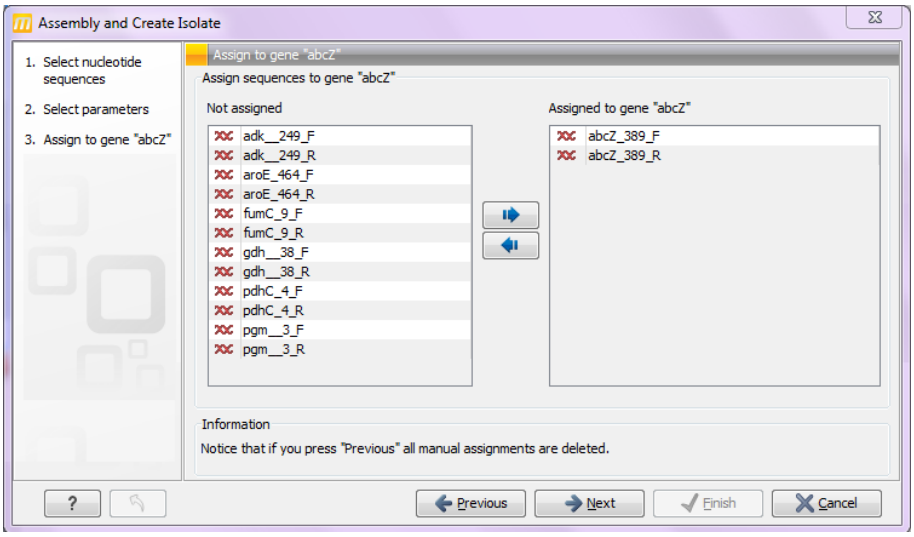


Figure 4.3: *Assigning sequences to the abcZ gene.*

You can of course manually change which sequences should be assigned to this locus by using the **Add** () and **Remove** () buttons.

To the left you see the rest of the sequences which have not yet been assigned to a locus. When clicking **Next**, the Multilocus Sequence Typing Module will try to find sequences matching the reference sequence for the next locus (shown in figure 4.4).

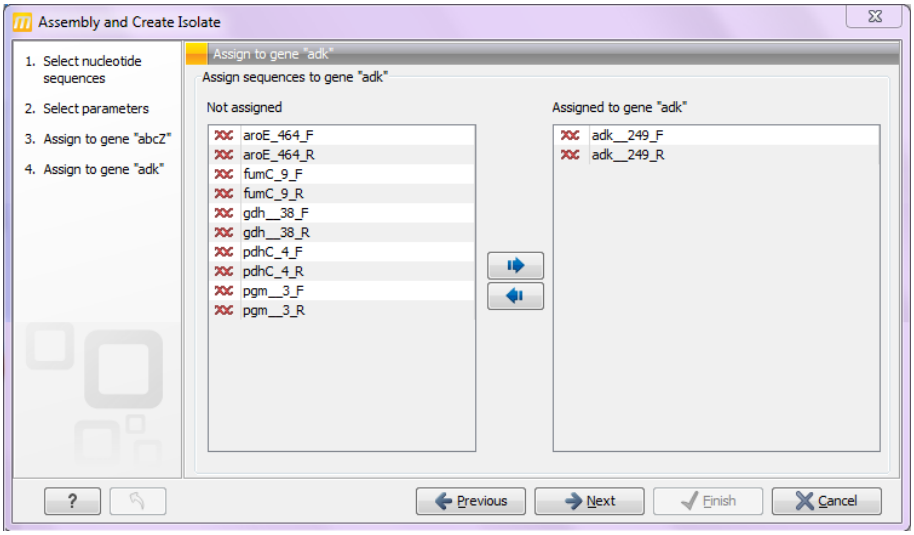


Figure 4.4: *Assigning sequences to the adk gene.*

Clicking **Next** shows the assignment of sequences to the third locus and so on for all the loci.

When clicking **Next** at each step, only the sequences listed in the left side of the dialog are considered for the automatic assignment.

When you press **Finish**, the isolate will be created and opened .

Note! If you press **Finish** while there are sequences listed to the left, they will not be assigned to a locus. Therefore, you should either click **Finish** after the second step when you have selected the scheme, or you should wait until you have stepped through all the assignment steps.

4.2 Working with isolates

A view of an isolate is shown in figure 4.5.

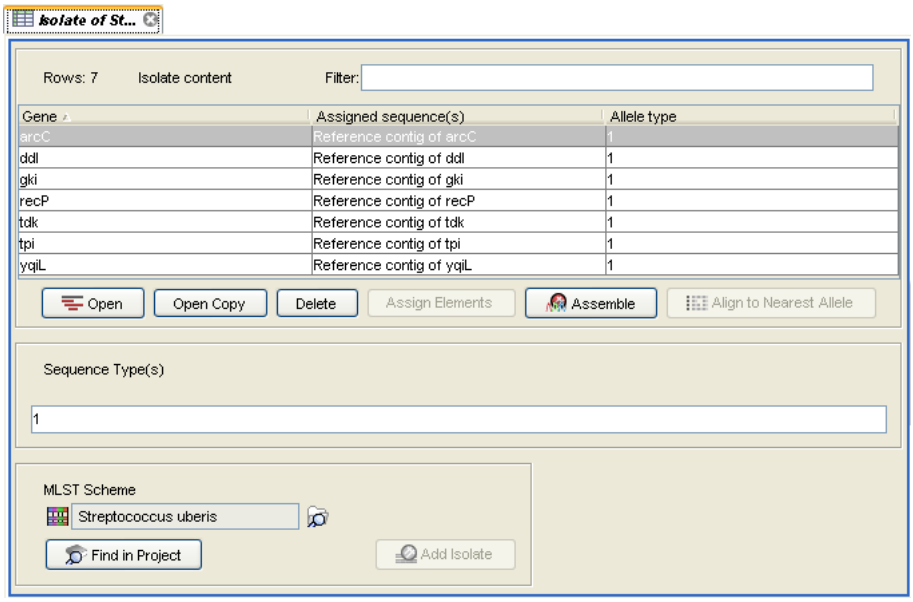


Figure 4.5: *An isolate of Streptococcus Uberis typed as number one.*

At the top there is a row for each locus. Each row represents a contig which is based on the sequences that have been assigned (either automatically or manually) to this given locus and a reference sequence.

The information in each row is:

- **Gene**. The name of the locus.
- **Assigned sequence(s)**. The name of the sequences that have been assigned to this locus. In most cases, this will be a reference contig of the selected sequences.
- **Allele type**. This is the allele type based on comparing the consensus sequence of the contig with the allele sequences in the scheme (this is explained further below).

4.2.1 Isolate-scheme dependencies and the sequence type

An isolate (as shown in figure 4.5) is created by defining a scheme to use as a reference. The scheme is used to:

- Automatically assign sequences to loci (see section 4.1).
- Determine the allele type for the consensus sequence for each locus (at the top of the isolate view).
- Determine the sequence type of the isolate.

For each locus, the Multilocus Sequence Typing Module compares the reference sequence of the isolate's contig with all the allelic types in the scheme. If there is a perfect match, an allelic type is assigned. If not, the allele type will be "New allele" (in this case, the new allele can be compared to the nearest allele in the scheme, see section 4.2.5).

The sequence type is shown in the field at the middle of the view and determined by the Multilocus Sequence Typing Module by comparing the allelic profile (all the allele types) of the isolate with the allelic profiles in the scheme (see section 3.4.1).

If none of the allelic profiles are identical to the isolate's allelic profile, the text field will say: "New type". If there are one or more loci where no sequences have been assigned, there could be a number of possible sequence types because one or more allelic type is unknown. All the possible sequence types will then be listed.

The isolate is dynamically linked to the scheme, meaning that if the scheme is updated, the allelic types and the sequence type of the isolate is calculated again. In this case, a dialog is presented to inform you that the information in the isolate view is about to change (see figure 4.6).

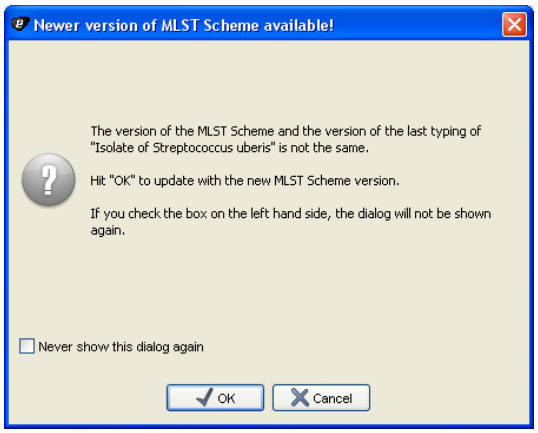


Figure 4.6: *A dialog informing that the information in the scheme has changed.*

If you wish to use another scheme as basis for typing the isolate, click the **Browse** button ( ) next to the scheme name at the bottom of the view.

Clicking **Find in Navigation Area** ( ) will select the scheme that is currently used for typing the isolate.

4.2.2 Opening contigs

The contig of each locus can be opened either by double-clicking the table entry or clicking the **open** ( ) button. This lets you inspect and edit the contig for variations and sequencing errors in a separate view (see figure 4.7).

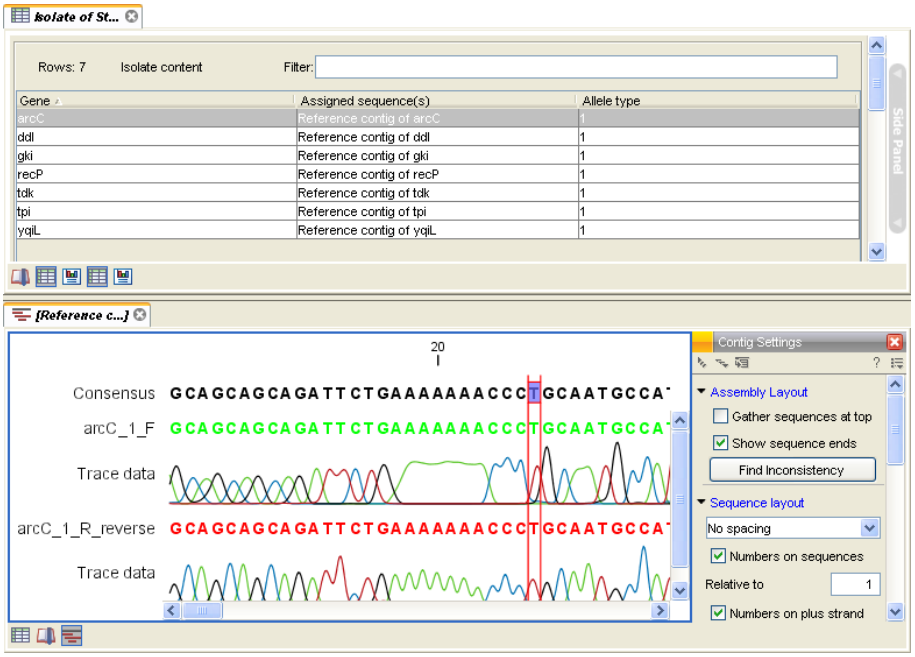


Figure 4.7: *A contig opened from an isolate.*

If you edit the consensus sequence in the contig, the isolate will automatically update the allelic type and subsequently the sequence type. This means that e.g. correcting a sequencing error will immediately reveal the right allelic type.

All changes to the contig are recorded in the contig's history ( ).

4.2.3 Deleting contigs

It may be desirable to delete a contig because of poor quality data or wrong assignment of sequences. To delete a contig:

select the row in the table to delete | Delete

This will remove the contig, and the allelic type will now be: "Unassigned".

4.2.4 Assigning new contigs

If there are no sequences assigned to a locus, you can assign either a sequences or a contig of assembled sequences:

select the relevant row in the table | Assign Elements

This will allow you to assign one sequence or one contig to this locus. Note that if the sequence is not trimmed, you have to do this manually first. Alternatively, use the **Assemble** ( ) button instead. This also allows you to assign several sequences by assembly to the reference sequence in the scheme (see section 4.2.6).

4.2.5 Align to nearest allele

If you have found a new allele, you can compare this with the allele in the scheme which has the highest degree of homology:

select the relevant row in the table | **Align to Nearest Allele** ( )

Multilocus Sequence Typing Module then searches through all the alleles in the scheme and compares them to the new allele. Then an alignment is created with the nearest allele to make it easy to inspect conservation.

4.2.6 Assemble to an existing isolate

Multilocus Sequence Typing Module makes it easy to perform sequence typing in several iterations. If you e.g. find errors in the sequencing data of some of the loci, you can create a preliminary version of the isolate with only sequence data for some of the loci. Then you can re-sequence the missing loci and update the isolate with new sequences.

This is done by assembling to an existing isolate:

Toolbox | **Multilocus Sequence Typing** ( )| **Assemble to Existing Isolate** ( )

In this dialog, select one isolate and a number of sequences. The rest of the dialogs are identical to the dialogs when creating an isolate *de novo* (see section 4.1).

4.3 View a report of an isolate

An isolate can also be viewed in a graphically more pleasant format. Click the **Show Report** ( ) to show the isolate report (see figure 4.8).

The view contains the same information as the standard isolate view, but the layout is improved for printing and graphics export.

1. Multilocus sequence typing

1.1 Typing results

This report uses the MLST Scheme Streptococcus uberis

1.1.1 Allelic Profile and Sequence Type

Gene	Allele sequence	Allelic profile
arcC	GCAGCAGCAGATTCTGAAAAAAAC CCTGCAATGCCATTAGATACTTGT GTGGCAATGACTGAAGGATCTATT GGTTTCTGGCTCGTTAATGCTCTT GATAACGAACTTAAAGAACAAGGC ATTGAAAAAGAAGTGGCTGCAGTT GTGACACAAGTTATTGTTGATAAA AACGACCAAGCATTGACAAATCCA ACAAAACCAATTGGCCCATTCTTA TCTGAAGAAGAAGCTAAAAAACAA ATGGAAGAAACTGGAGCAAGCTTT AAAGAAGATTCAGGTCGTGGATG GCGTAAAGTGGTGCCATCTCCAAA ACCAGTTGGCATCAAAGAAGCTAA TGTTATTCGTAACTTAGTTGATTGA GGTGTTGTTGTTGTCAGCGCTGGT GGAGGCGGTGTCCCAGTTATTGA AGACCCAACAAC	1
ddl	ACAAACCACATGCTCCAATAGCTT TAAAAGCTTGTTCTGCAAAATATCT CATTTGTGATGTAATGCCTTCATC AATCGTCGCGGGAATAGCCATACT TATCTTATTATCGATGTATTTGGCT TGATAGTCATAGAAGGCAACATCT TTGATGACTTCACCAGGTAGAGTT GATGCTACTTTGGTATTTCCTAATA AGCCAACTTCAATCTCACGTGCTA CGACACCTTGTTCAATTAAAATAC GACTATCATATTTTAAAGCTAATGC	1

Figure 4.8: *An isolate shown as a report.*

Chapter 5

Extending schemes

MLST schemes can be extended in two ways: by adding isolates or by adding sequences:

5.1 Add isolates to scheme

If you have an isolate which represents a new sequence type, you can add this isolate to a scheme and hereby expand the scheme with one more sequence type:

Toolbox | **Multilocus Sequence Typing** ( )| **Add Isolates to MLST Scheme** ( )

or **Open the isolate you wish to add** | **Scroll down to the bottom**| **Add Isolate** ( )

You do not have to specify a scheme since an isolate is always based on a scheme when it is created. The isolate will be added to the scheme that is defined at the bottom of the isolate view (see section 4.2.1).

5.2 Add sequences to scheme

You can also add individual sequences to the alleles in the scheme:

Toolbox | **Multilocus Sequence Typing** ( ) | **MLST Schemes** ( ) | **Add Sequences to MLST Scheme** ( )

This will bring up a dialog as shown in figure 5.1.

Clicking **Next** takes you through a series of step - one for each locus - where you can assign sequences (see figure 5.2).

The sequences added in this way cannot be combined to an allelic profile and a sequence type, and will thus only contribute to expanding the number of allelic types.

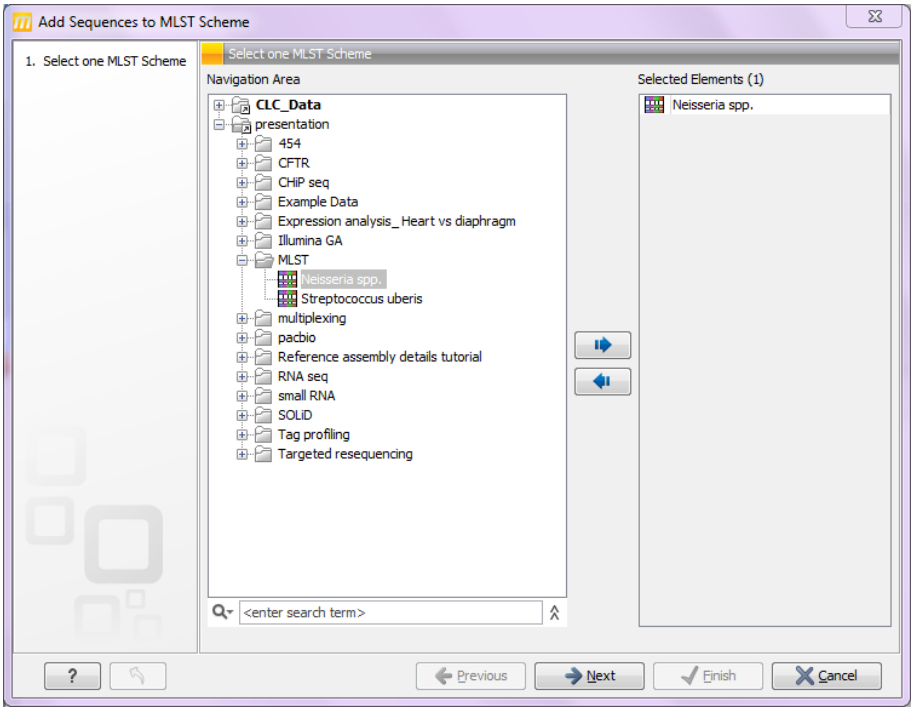


Figure 5.1: *Define a scheme to add sequences to.*

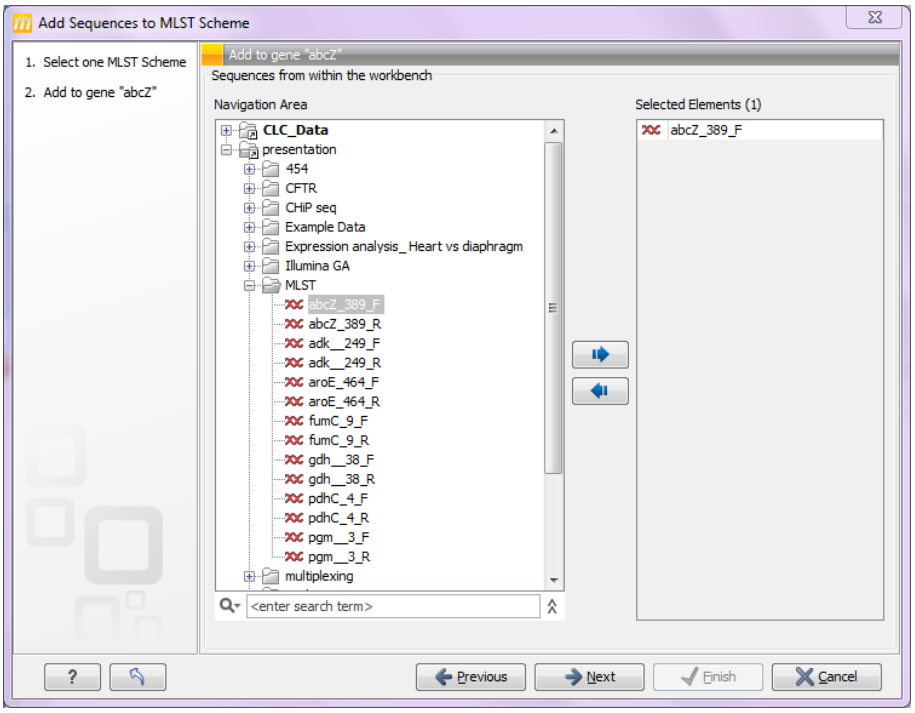


Figure 5.2: *Adding a sequence to gene abcZ.*

Chapter 6

Installation of the Multilocus Sequence Typing Module

Multilocus Sequence Typing Module is installed as a plugin. Plugins are installed using the plugin manager. In order to install plugins on Windows, the Workbench must be run in administrator mode: Right-click the program shortcut and choose "Run as Administrator". Then follow the procedure described below.

Help in the Menu Bar | Plugins... ( )

or **Plugins ( ) in the Toolbar**

The plugin manager has two tabs at the top:

- **Manage Plugins.** This is an overview of plugins that are installed.
- **Download Plugins.** This is an overview of available plugins on CLC bio's server.

To install a plugin, click the **Download Plugins** tab. This will display an overview of the plugins that are available for download and installation (see figure 6.1).

Clicking a plugin will display additional information at the right side of the dialog. This will also display a button: **Download and Install**.

Click the Multilocus Sequence Typing Module and press **Download and Install**. A dialog displaying progress is now shown, and the plugin is downloaded and installed.

If the Multilocus Sequence Typing Module is not shown on the server, and you have it on your computer (for example if you have downloaded it from our website), you can install it by clicking the **Install from File** button at the bottom of the dialog. This will open a dialog where you can browse for the plugin. The plugin file should be a file of the type ".cpa".

When you close the dialog, you will be asked whether you wish to restart the CLC Workbench. The plugin will not be ready for use until you have restarted.

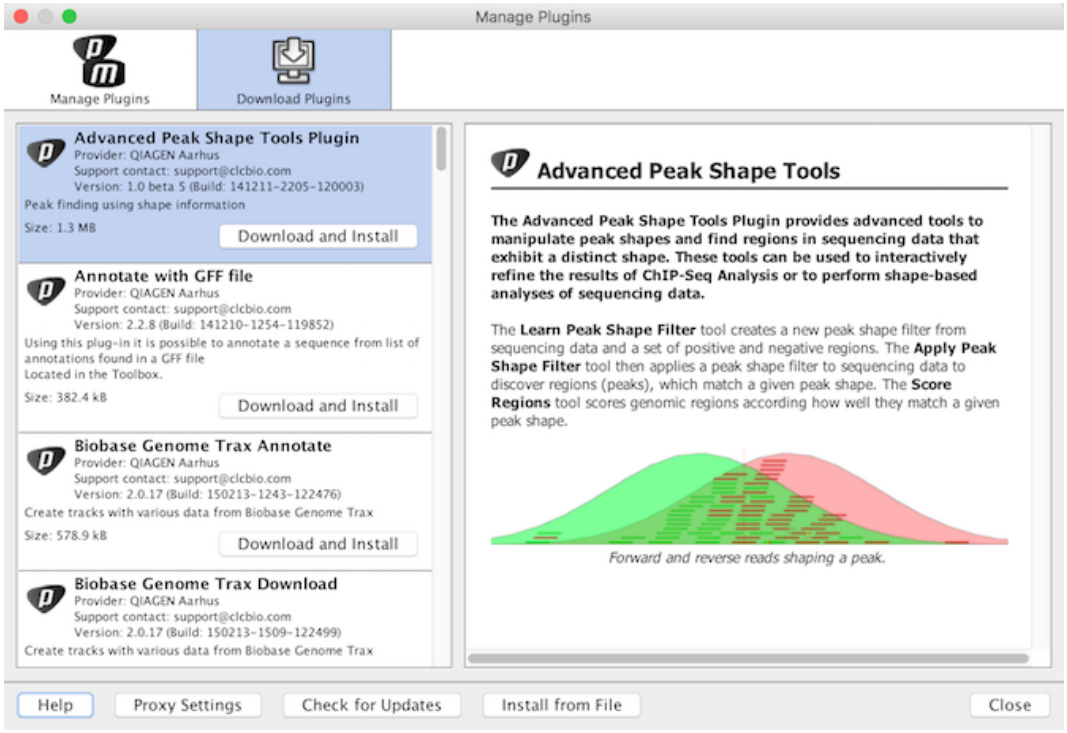


Figure 6.1: *The plugins that are available for download.*

Chapter 7

Uninstall

Plugins are uninstalled using the plugin manager:

Help in the Menu Bar | Plugins... ( )

or **Plugins ( ) in the Toolbar**

This will open the dialog shown in figure 7.1.

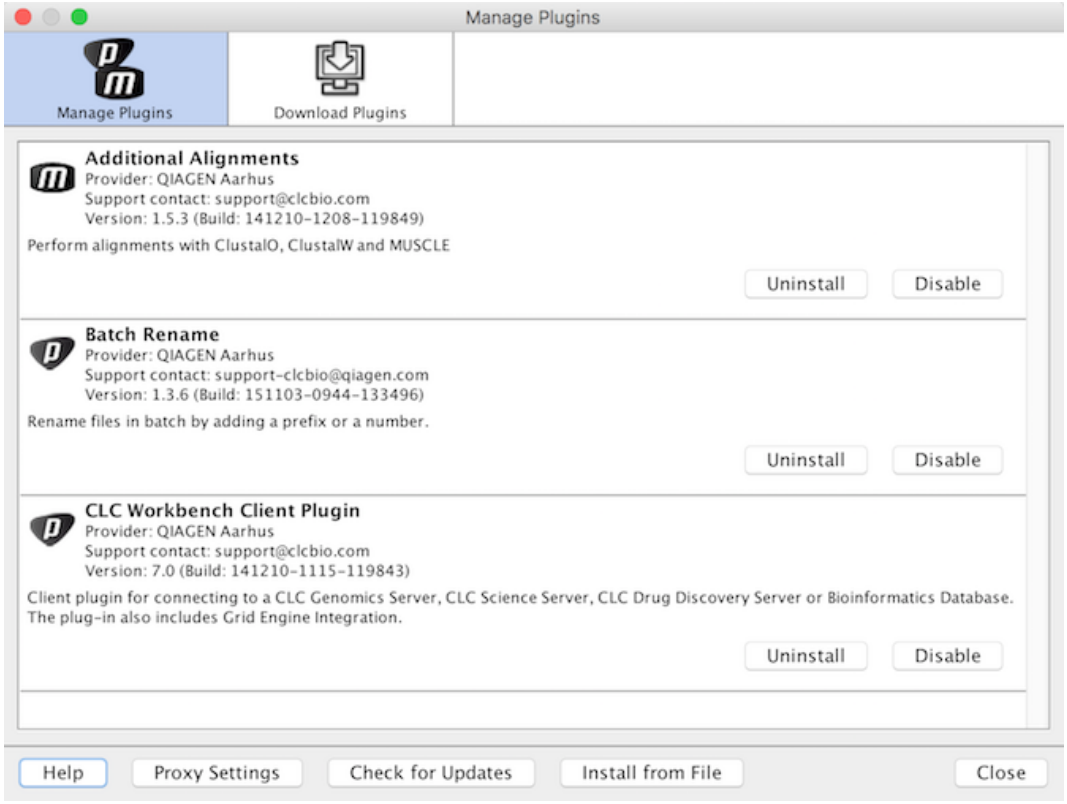


Figure 7.1: *The plugin manager with plugins installed.*

The installed plugins are shown in this dialog. To uninstall:

Click the Multilocus Sequence Typing Module | Uninstall

If you do not wish to completely uninstall the plugin but you don't want it to be used next time you start the Workbench, click the **Disable** button.

When you close the dialog, you will be asked whether you wish to restart the workbench. The plugin will not be uninstalled until the workbench is restarted.

Bibliography

[Maiden et al., 1998] Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., and Spratt, B. G. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A*, 95(6):3140–3145.

[Maiden, 2006] Maiden, M. C. J. (2006). Multilocus sequence typing of bacteria. *Annu Rev Microbiol*, 60:561–588.

[Sullivan et al., 2005] Sullivan, C. B., Diggle, M. A., and Clarke, S. C. (2005). Multilocus sequence typing: Data analysis in clinical microbiology and public health. *Mol Biotechnol*, 29(3):245–254.

[Taylor and Fisher, 2003] Taylor, J. W. and Fisher, M. C. (2003). Fungal multilocus sequence typing–it's not just for bacteria. *Curr Opin Microbiol*, 6(4):351–356.

[Urwin and Maiden, 2003] Urwin, R. and Maiden, M. C. J. (2003). Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol*, 11(10):479–487.

Index