



CLC MLST Module

USER MANUAL

User manual for *CLC Multilocus Sequence Typing Module* 24.0

Windows, macOS and Linux

January 5, 2024

This software is for research purposes only.

QIAGEN Aarhus
Silkeborgvej 2
Prismet
DK-8000 Aarhus C
Denmark



Contents

1	Introduction	5
1.1	What is MLST?	5
1.2	The concept of the Multilocus Sequence Typing Module	5
1.3	Acknowledgements	6
2	MLST quick-start	7
3	Handling MLST schemes	8
3.1	Downloading schemes	8
3.1.1	Download custom schemes	9
3.2	Creating new schemes	9
3.3	Merging schemes	11
3.4	Viewing schemes	11
3.4.1	Profile table	12
	Extracting alignments	12
	Extracting sequences	13
	Adding and removing sequence information	14
3.4.2	Allele table	14
	Displaying the sequences	14
	Creating an alignment comparing the allelic types	15
	Removing allele sequences	15
4	Isolates	16
4.1	Assemble and create isolate	16
4.2	Working with isolates	18
4.2.1	Isolate-scheme dependencies and the sequence type	19

4.2.2	Opening contigs	20
4.2.3	Deleting contigs	20
4.2.4	Assigning new contigs	20
4.2.5	Align to nearest allele	21
4.2.6	Assemble to an existing isolate	21
4.3	View a report of an isolate	21
5	Extending schemes	23
5.1	Add isolates to scheme	23
5.2	Add sequences to scheme	23
6	Install and uninstall modules	25
6.1	Installation of modules	25
6.2	Uninstalling modules	27
7	Licensing requirements for the Multilocus Sequence Typing Module	29
7.1	Licensing modules on a Workbench	29
7.1.1	Request an evaluation license	29
	Direct download	30
	Go to license download web page	30
7.1.2	Download a license using a license order ID	31
	Direct download	32
	Go to license download web page	32
7.1.3	Import a license from a file	32
7.1.4	Configure License Server connection	33
	Borrowing a license	35
	Common issues when using a network license	36
7.1.5	Download a static license on a non-networked computer	38
	Bibliography	39

Chapter 1

Introduction

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

The Multilocus Sequence Typing Module provides tools to download MLST schemes from PubMLST (<http://pubmlst.org>)

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 (PubMLST)** (📁)

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 (📁) | **Assembly 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 directly from the databases at PubMLST.org.

To select schemes for download:

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

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.

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

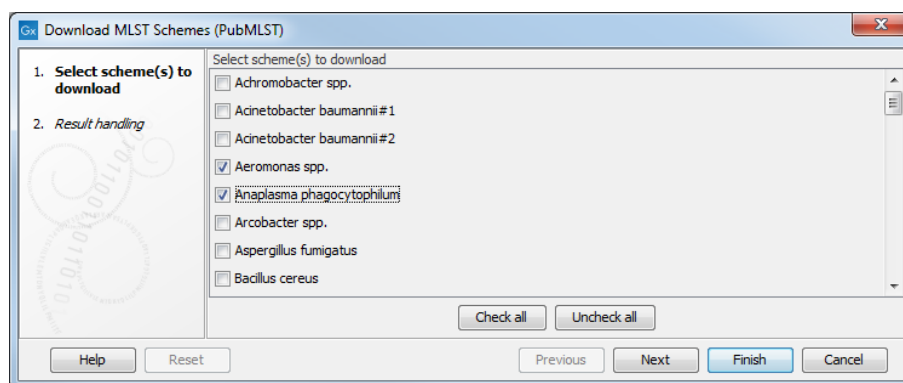


Figure 3.1: Selecting schemes for download.

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.

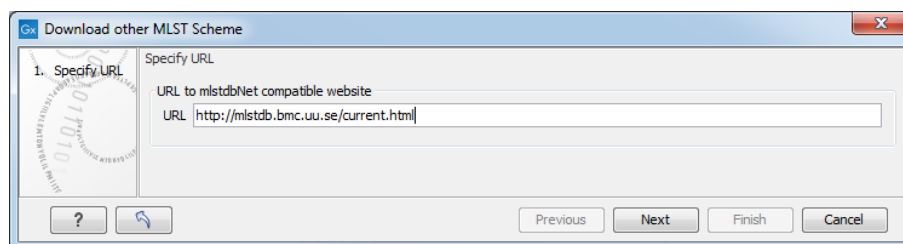


Figure 3.2: Providing a URL for downloading schemes.

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.

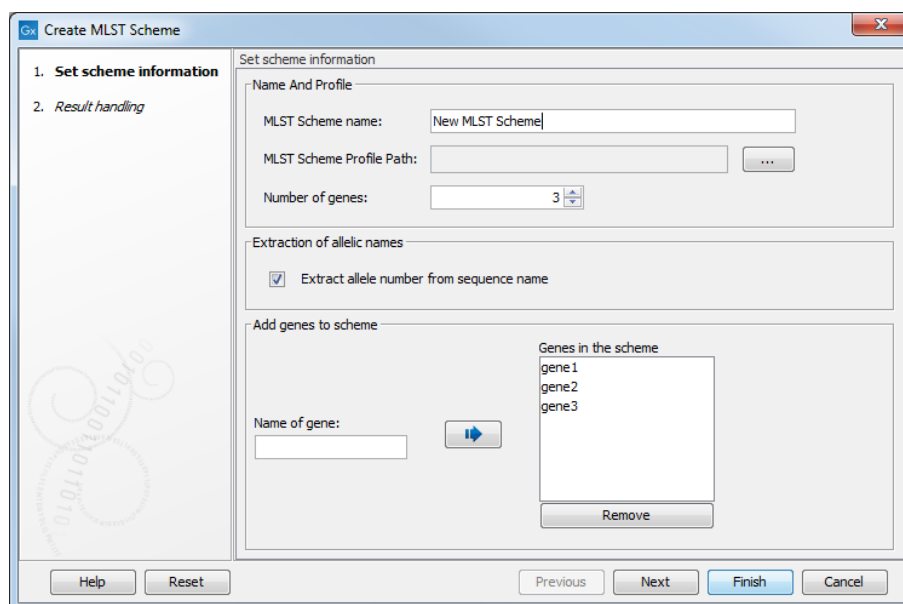
To create a new scheme:

Multilocus Sequence Typing (📁) | **MLST Schemes** (📁) | **Create MLST Scheme** (🛠️)

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

In the top field you can enter a name for the scheme - typically the name of the species in question.

Below, you can specify the location for a text file where all the loci in the scheme are listed. The file must contain a row with the column names 'ST', 'gene_1', 'gene_2', etc. Each subsequent row must contain exactly one loci/sequence type. Note that the file must be in plain text format and the columns must be tab-separated. It is important to update the parameter "Number of genes" to match the number of genes in the selected text file. If such a file does not exist, you can just type, in the "Add genes to scheme" section of the dialog, the name of the locus in the

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

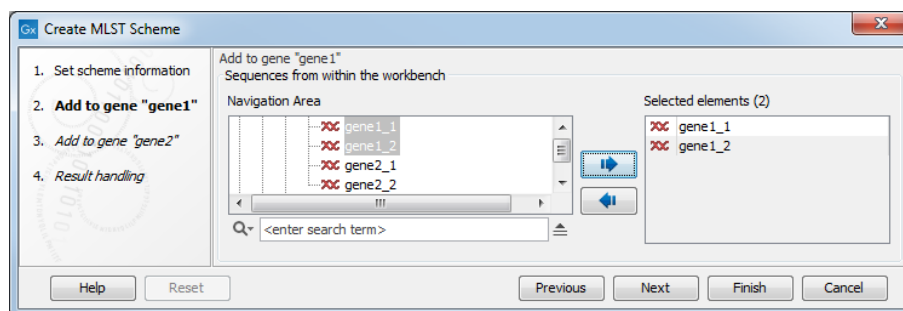
Name of gene field, and click the **Add** (➡) button.

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

When you are done adding loci for your scheme, click **Next** to start assigning sequences to the loci. Note that the option "Extract allele number from sequence name" assumes that the allele number is part of the allele name (e.g. 'atpA_2'). If this is not the case, disable the option.

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.

Figure 3.4: *Selecting sequences for gene1.*

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 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 by 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 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.

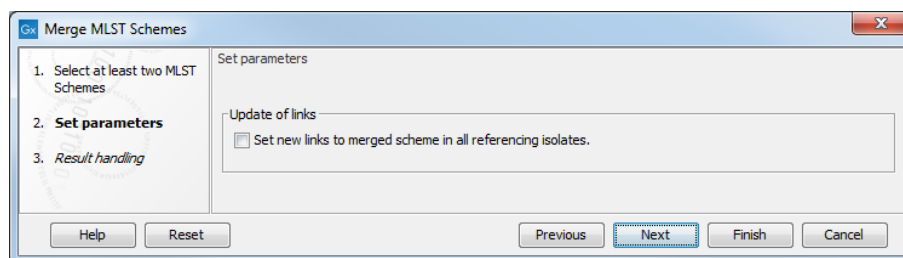


Figure 3.5: This option makes it possible to update all references in isolates previously referring to one of the two old schemes so that they refer to the new scheme.

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	aroC	dnaN	hemD	hisD	purE	sucA	thrA	clonal_complex
1	1	1	1	1	1	1	1	5
2	1	1	1	2	1	1	1	5
3	1	1	1	2	1	1	1	9
4	43	41	16	13	34	13	4	4
5	16	43	45	43	36	39	42	4
6	4	4	4	5	4	4	8	8
7	2	2	3	2	2	2	6	6
8	1	1	2	3	1	1	5	5
9	33	26	30	55	21	87	134	134
10	5	2	3	6	5	5	10	10
11	5	2	3	7	6	6	11	11
12	6	5	6	7	5	7	12	12
13	3	3	7	4	3	3	7	7
14	7	6	8	8	7	8	13	13
15	2	7	9	9	5	9	12	12
16	6	7	10	10	8	10	14	14
17	8	8	11	11	5	11	15	15
18	9	9	6	12	9	12	2	2
19	10	7	12	9	5	9	2	2
20	11	10	13	13	10	13	4	4

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

The two options in the alignment right-click menu provide two ways of composing the alignment (figure 3.7):

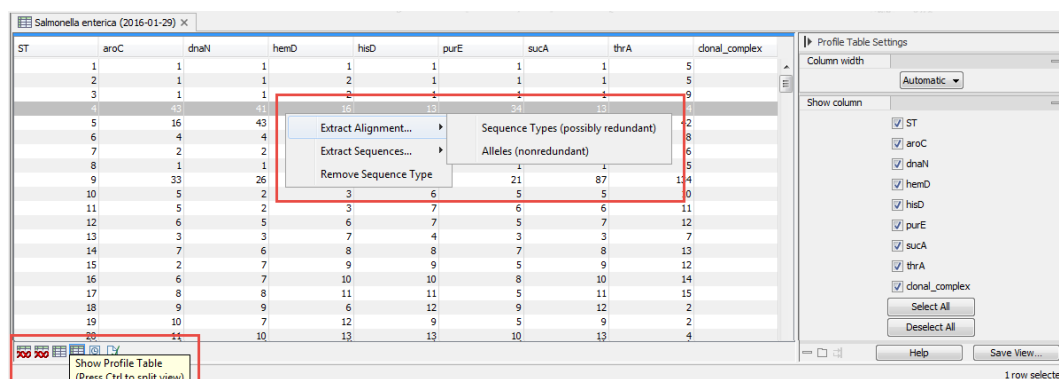


Figure 3.7: Extract alignment from a profile table.

- **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 3.8 displaying the loci in the scheme.

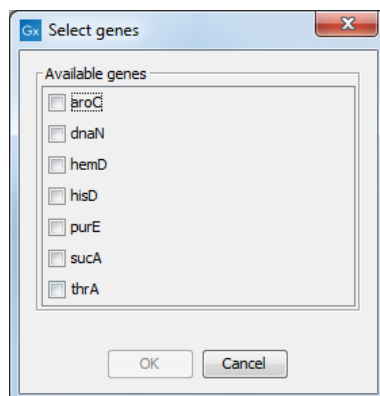


Figure 3.8: Selecting genes for alignment.

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 3.8 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 3.8 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 3.9).

Gene	Number of alleles
aroC	571
dnaN	544
hemD	490
hisD	738
purE	581
sucA	571
thrA	618

Allele name	Allele number	Sequence length	Creation date	Sequence types
aroC1	1	501	Fri Jan 29 11:32:43 CET 2016	1; 2; 3; 8; 815; 890; 911; 1856; 1919; 2138; 2160; 2173; 2209; 2230; 2244; 2254; 2
aroC2	2	501	Fri Jan 29 11:32:43 CET 2016	7; 15; 31; 33; 38; 43; 86; 110; 132; 142; 148; 149; 188; 191; 216; 250; 269; 270; 271; 272
aroC3	3	501	Fri Jan 29 11:32:43 CET 2016	170; 177; 205; 674; 676; 677; 1216; 1238; 1485; 1668; 2117; 2616; 2620

Figure 3.9: The allele table.

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 3.9 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

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 (📁) | Assembly and Create Isolate (🔍)

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

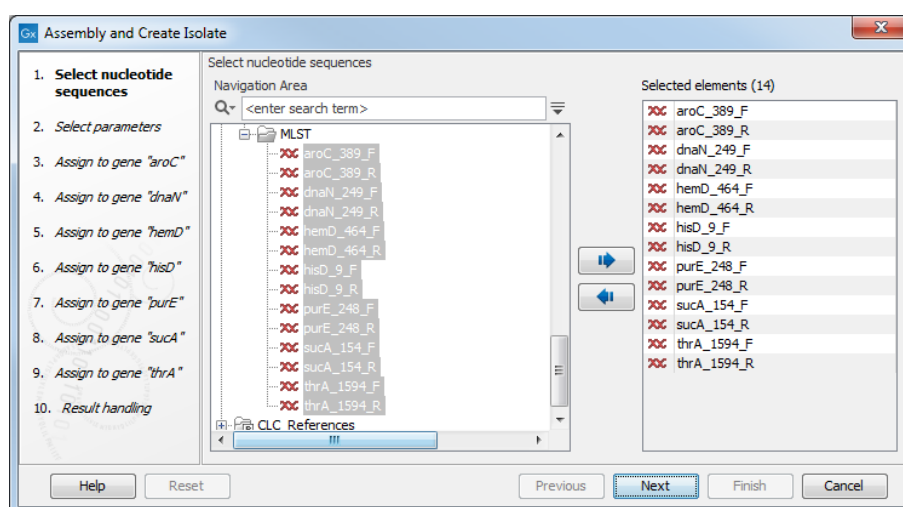


Figure 4.1: Selecting isolate sequences.

In this example, 8 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.

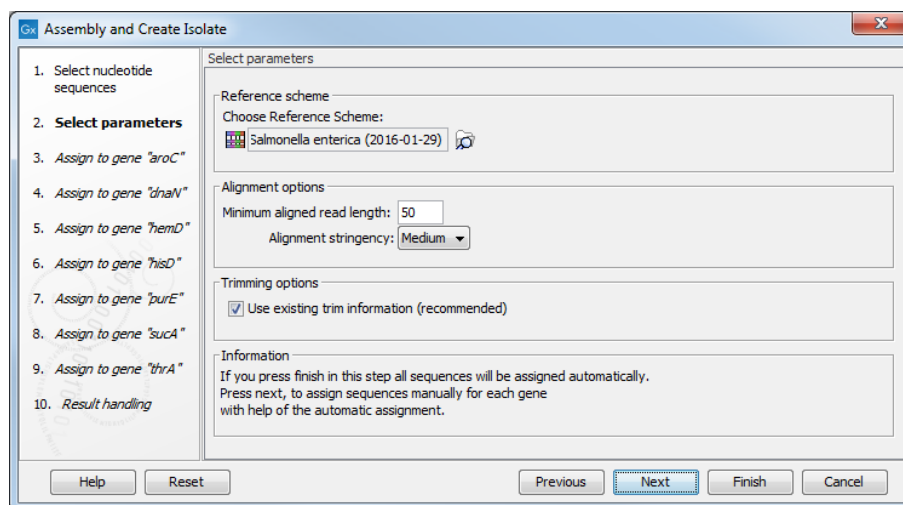




Figure 4.2: Setting a reference scheme for typing the isolate.

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.
- 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 or 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 *dnaN* gene (listed to the right).

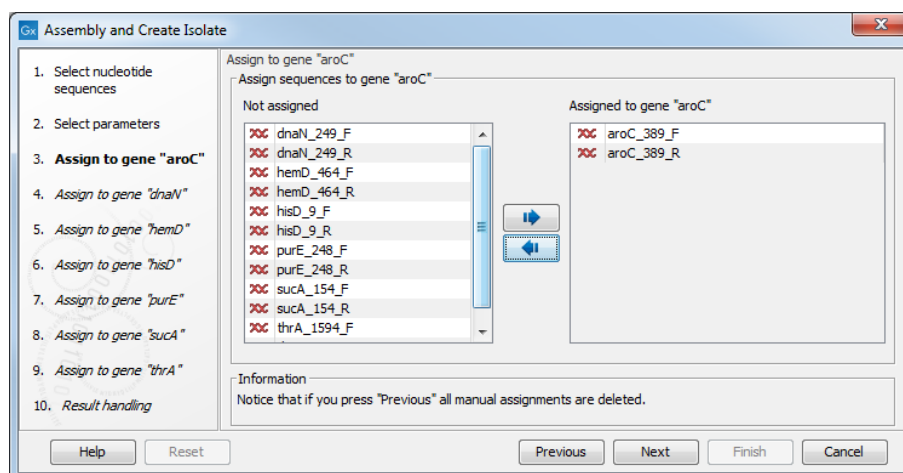


Figure 4.3: Assigning sequences to the *dnaN* gene.

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, 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 4.4.

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).

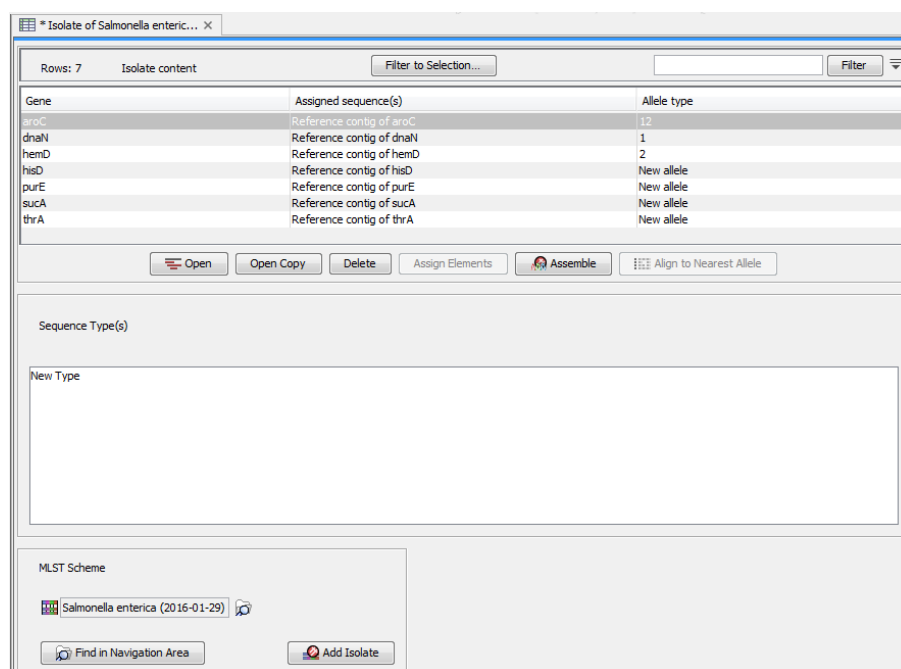


Figure 4.4: An isolate of *Streptococcus Uberis* typed as number one.

4.2.1 Isolate-scheme dependencies and the sequence type

An isolate (as shown in figure 4.4) 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.

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 4.5).

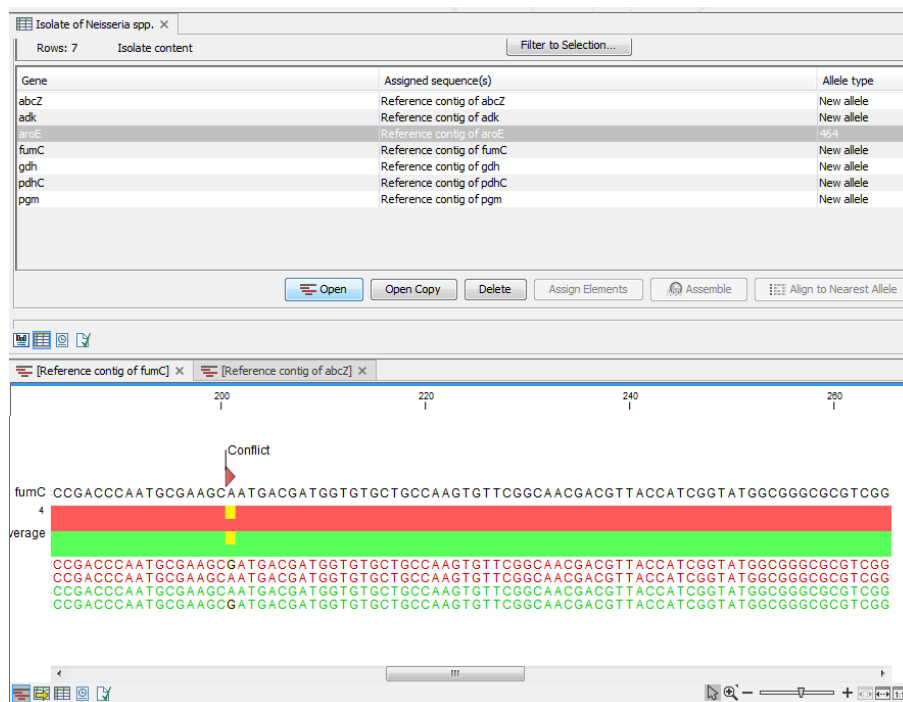


Figure 4.5: A contig opened from an isolate.

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 (🧬)

As input for this tool, select exactly one isolate and at least one nucleotide sequence. It is possible to select a nucleotide sequence list (see figure 4.6) or multiple nucleotide sequences.

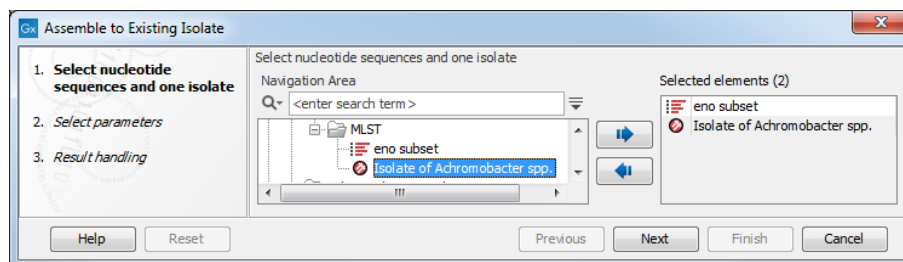


Figure 4.6: Assembling an isolate.

The rest of the dialogs are identical to the dialogs displayed 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 4.7).

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	GCAGCAGCAGATTCTGAAAAAAC CCTGCAATGCCATTAGATCTTGT GTGGCAATGACTGAAGGATCTATT GGTTTCTGGCTCGTTAATGCTCTT GATAACGAACTTAAAGAACAAGGC ATTGAAAAAGAAGTGGCTGGAGTT GTGACACAAGTTATTGTTGATAAA AACGACCAAGCATTGACAAATCCA ACAAACCAATTGGCCATTCTTA TCTGAAGAAGAAGGTAAAAACAA ATGGAAGAACTGGAGCAAGCTTT AAAGAAGATTGAGTCTGGGATG GCGTAAAGTGGTGCCATCTCGAAA ACGAGTTGGGATCAAAGAAGCTAA TGTTATTGTAACITAGTTGATTGA GGTGTGTTGTTGTGAGCGCTGGT GGAGGCGGTGTCCTCCAGTTATTGA AGACCAAGAAC	1
ddl	ACAAACCAGATGCTCCAATAGCTT TAAAAGCTTGTTCTGCAAAATATCT CATTGTGATGTAATGCCTTCATC AATCGTCGCGGAATAGCCATACT TATCTTATTATCGATGTAATTTGGCT TGATAGTCATAGAAGGCAACATCT TTGATGAGTTGACCAGGTAGAGTT GATGCTACTTTGGTATTTCCTAATA AGCCAACTTCAATGTGACGTGCTA CGACACCTTGTTCAATTAAATAC GACTATGATATTTTAAAGCTAATGC	1

Figure 4.7: 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.

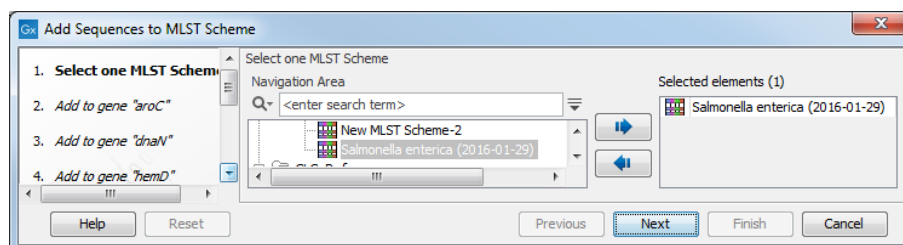


Figure 5.1: Define a scheme to add sequences to.

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

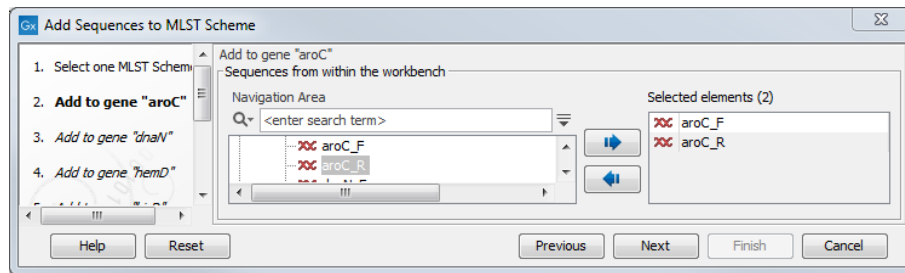


Figure 5.2: Adding a sequence to gene *abcZ*.


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.

Chapter 6

Install and uninstall modules

6.1 Installation of modules

Note: In order to install plugins and modules, the *CLC Workbench* must be run in administrator mode. On Windows, you can do this by right-clicking the program shortcut and choosing "Run as Administrator". On Linux and Mac, it means you must launch the program such that it is run by an administrative user.

Plugins and modules are installed and uninstalled using the Workbench Plugin Manager. To open the Plugin Manager, click on the **Plugins** () button in the top Toolbar, or go to the menu option:

Utilities | Manage Plugins... ()

The Plugin Manager has two tabs at the top:

- **Manage Plugins** An overview of your installed plugins and modules is provided under this tab.
- **Download Plugins** Plugins and modules available to download and install are listed in this tab.

To install a plugin, click on the **Download Plugins** tab (figure 6.1). Select a plugin. Information about it will be shown in the right hand panel. Click on the **Download and Install** button to install the plugin.

Accepting the license agreement

The End User License Agreement (EULA) must be read and accepted as part of the installation process. Please read the EULA text carefully, and if you agree to it, check the box next to the text **I accept these terms**. If further information is requested from you, please fill this in before clicking on the **Finish** button.

Installing a cpa file

If you have a .cpa installer file for Multilocus Sequence Typing Module, you can install it by clicking on the **Install from File** button at the bottom of the Plugin Manager.

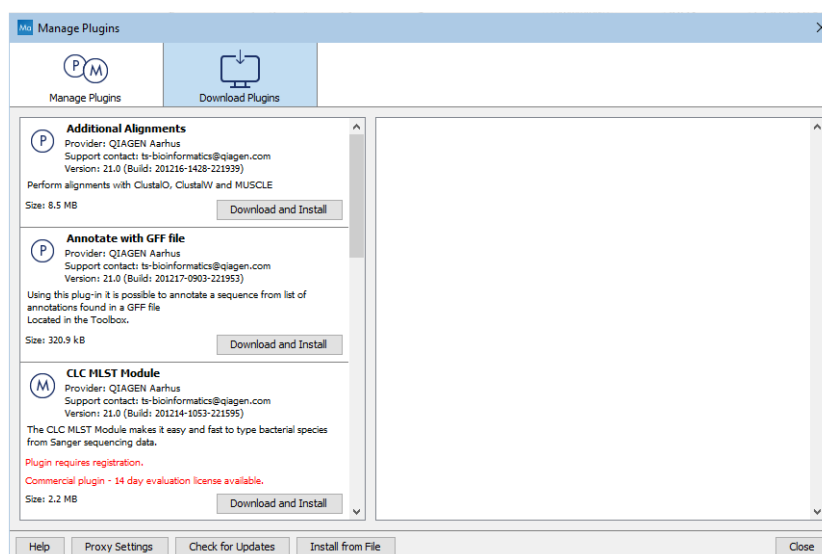


Figure 6.1: Plugins and modules available for installation are listed in the Plugin Manager under the Download Plugins tab.

If you are working on a system not connected to the internet, plugin and module .cpa files can be downloaded from <https://digitalinsights.qiagen.com/products-overview/plugins/> using a networked machine, and then transferred to the non-networked machine for installation.

Restart to complete the installation

Newly installed plugins and modules will be available for use after restarting the software. When you close the Plugin Manager, a dialog appears offering the opportunity to restart the *CLC Workbench*. When you have installed the Multilocus Sequence Typing Module and start a tool from that module for the first time, the License Assistant will open (figure 6.2).

The License Assistant can also be launched by opening the Workbench Plugin Manager, selecting the installed module from under the Manage Plugins tab, and clicking on the button labeled *Import License*.

To install a license, the *CLC Workbench* must be run in administrator mode. On Windows, you can do this by right-clicking the program shortcut and choosing "Run as Administrator". On Linux and Mac, it means you must launch the program such that it is run by an administrative user.

The following options are available:

- **Request an evaluation license.** Request a fully functional, time-limited license.
- **Download a license.** Use the license order ID received when you purchased the software to download and install a license file.
- **Import a license from a file.** Import an existing license file, for example a file downloaded from the web-based licensing system.
- **Configure License Server connection.** If your organization has a *CLC Network License Manager* (or CLC License Server), select this option to configure the connection to it.

You need a license...

In order to load the plugin "CLC Genome Finishing Module" you need a valid license.
Please choose how you would like to obtain a license for this plugin.


- ☒ **Request an evaluation license**
Choose this option if you would like to try out the plugin for 14 days.
Please note that only a single evaluation license will be allowed for each computer.
- ☐ **Download a license**
Use a license order ID to download a static license.
- ☐ **Import a license from a file**
Import a static license from an existing license file.
- ☐ **Configure License Server connection**
Configure the necessary connection for the software to connect to a CLC License Server that hosts network license(s) for this product. This option also allows you to alter or disable an existing configuration.

Figure 6.2: The License Assistant provides options for licensing modules installed on the Workbench.

These options are described in detail in sections under http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Workbench_Licenses.html.

To download licenses, including evaluation licenses, your machine must have access to the external network. To install licenses on non-networked machines, please see http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Download_static_license_on_non_networked_machine.html.

6.2 Uninstalling modules

Plugins and modules are uninstalled using the Workbench Plugin Manager. To open the Plugin Manager, click on the **Plugins** () button in the top Toolbar, or go to the menu option:

Utilities | Manage Plugins... ()

This will open the Plugin Manager (figure 6.3). Installed plugins and modules are shown under the Manage Plugins tab of the Plugins Manager.

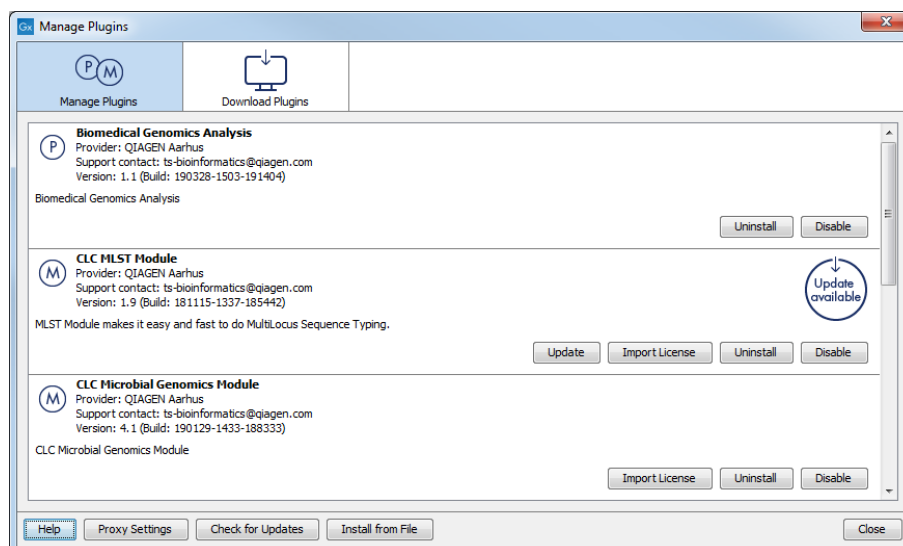


Figure 6.3: Installed plugins and modules are listed in the Plugins Manager under the Manage Plugins tab.

To uninstall a plugin or module, click on its entry in the list, and click on the **Uninstall** button.

Plugins and modules are not uninstalled until the Workbench is restarted. When you close the Plugin Manager, a dialog appears offering the opportunity to restart the *CLC Workbench*.

Disabling a plugin without uninstalling it

If you do not want a plugin to be loaded the next time you start the Workbench, select it in the list under the Manage Plugins tab and click on the **Disable** button.

Chapter 7

Licensing requirements for the Multilocus Sequence Typing Module

To use tools delivered by the Multilocus Sequence Typing Module, you will need a license. Multilocus Sequence Typing Module can be installed in Viewing Mode on a *CLC Workbench* to access data created using the module's tools without requiring a license.

7.1 Licensing modules on a Workbench

The License Assistant can be launched by opening the Workbench Plugin Manager, selecting the installed module from under the Manage Plugins tab, and clicking on the button labeled *Import License*.

To install a license, the *CLC Workbench* must be run in administrator mode. On Windows, you can do this by right-clicking the program shortcut and choosing "Run as Administrator". On Linux and Mac, it means you must launch the program such that it is run by an administrative user.

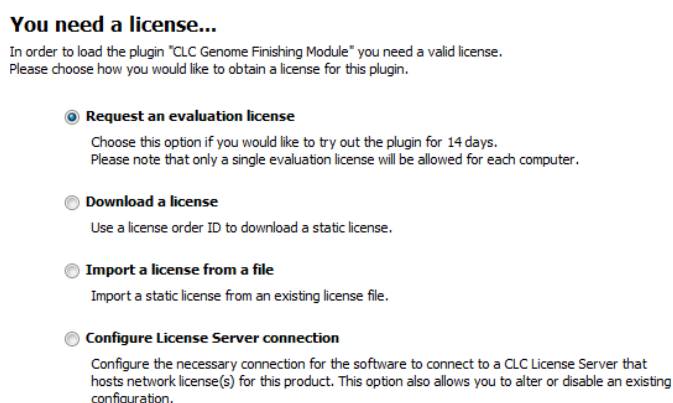


Figure 7.1: The License Assistant provides options for licensing modules installed on the Workbench.

7.1.1 Request an evaluation license

We offer a fully functional version of the Multilocus Sequence Typing Module free of charge for a 14 day period for evaluation purposes. The 14 day period commences when the evaluation license is downloaded. If you have questions about Multilocus Sequence Typing Module features

or product licensing options, please send an email to bioinformaticssales@qiagen.com.

When you choose the option **Request an evaluation license**, the dialog shown in figure 7.2 opens.

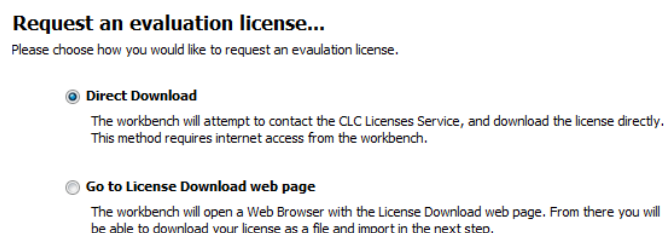


Figure 7.2: Choosing between direct download or going to the license download web page.

In this dialog, there are two options:

- **Direct Download.** Download the license directly. This method requires that the Workbench has access to the external network.
- **Go to CLC License Download web page.** The online license download form will be opened in a web browser. This option is suitable for when downloading a license for use on another machine that does not have access to the external network, and thus cannot access the QIAGEN Aarhus servers.

After selecting your method of choice, click on the button labeled **Next**.

Direct download

After choosing the **Direct Download** option and clicking on the button labeled **Next**, a dialog similar to that shown in figure 7.3 will appear if the license is successfully downloaded and installed.

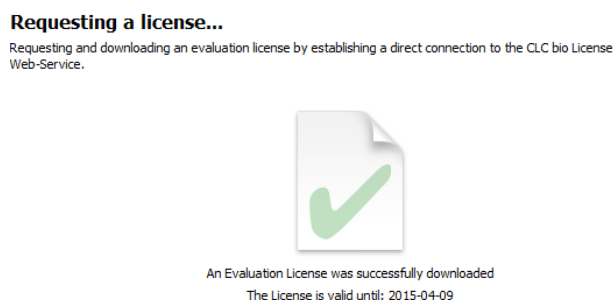


Figure 7.3: A license has been successfully downloaded and installed for use.

When the license has been downloaded and installed, the **Next** button will be enabled.

If there is a problem, a dialog will appear indicating this.

Go to license download web page

After choosing the **Go to CLC License Download web page** option and clicking on the button labeled **Next**, the license download form will be opened in a web browser, as shown in figure 7.4.

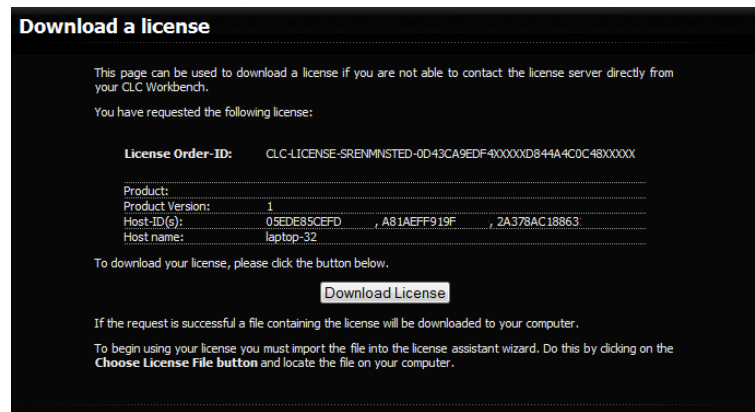


Figure 7.4: The license download form opened in a web browser.

Click on the **Download License** button and then save the license file.

Back in the Workbench window, you will now see the dialog shown in 7.5.

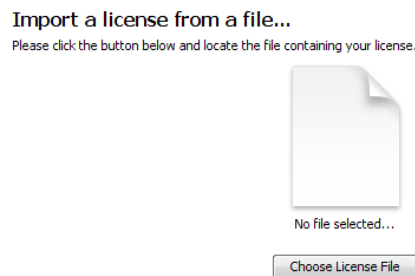


Figure 7.5: Importing the license file downloaded from the web page.

Click on the **Choose License File** button, find the saved license file and select it. Then click on the **Next** button.

7.1.2 Download a license using a license order ID

Using a license order ID, you can download a license file via the Workbench or using an online form. When you have chosen this option and clicked **Next** button, you will see the dialog shown in 7.6. Enter your license order ID into the text field under the title License Order-ID. (The ID can be pasted into the box after copying it and then using menus or key combinations like Ctrl+V on some system or ⌘ + V on Mac).

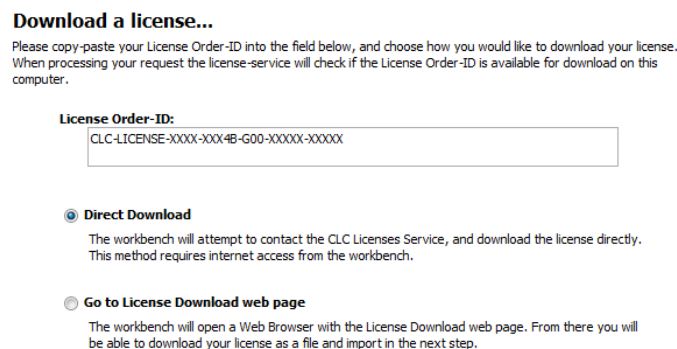


Figure 7.6: Enter a license order ID for the software.

In this dialog, there are two options:

- **Direct Download.** Download the license directly. This method requires that the Workbench has access to the external network.
- **Go to CLC License Download web page.** The online license download form will be opened in a web browser. This option is suitable for when downloading a license for use on another machine that does not have access to the external network, and thus cannot access the QIAGEN Aarhus servers.

After selecting your method of choice, click on the button labeled **Next**.

Direct download

After choosing the **Direct Download** option and clicking on the button labeled **Next**, a dialog similar to that shown in figure 7.7 will appear if the license is successfully downloaded and installed.

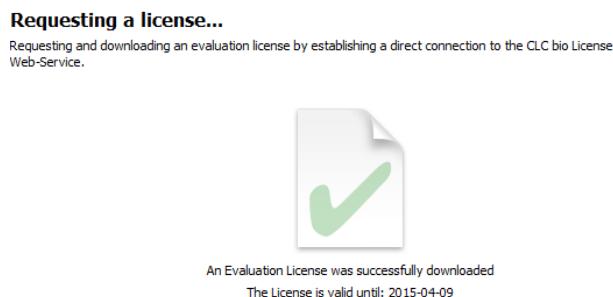


Figure 7.7: A license has been successfully downloaded and installed for use.

When the license has been downloaded and installed, the **Next** button will be enabled.

If there is a problem, a dialog will appear indicating this.

Go to license download web page

After choosing the **Go to CLC License Download web page** option and clicking on the button labeled **Next**, the license download form will be opened in a web browser, as shown in figure 7.8.

Click on the **Download License** button and then save the license file.

Back in the Workbench window, you will now see the dialog shown in 7.9.

Click on the **Choose License File** button, find the saved license file and select it. Then click on the **Next** button.

7.1.3 Import a license from a file

If you already have a license file associated with the host ID of your machine, it can be imported using this option.

When you have clicked on the **Next** button, you will see the dialog shown in 7.10.

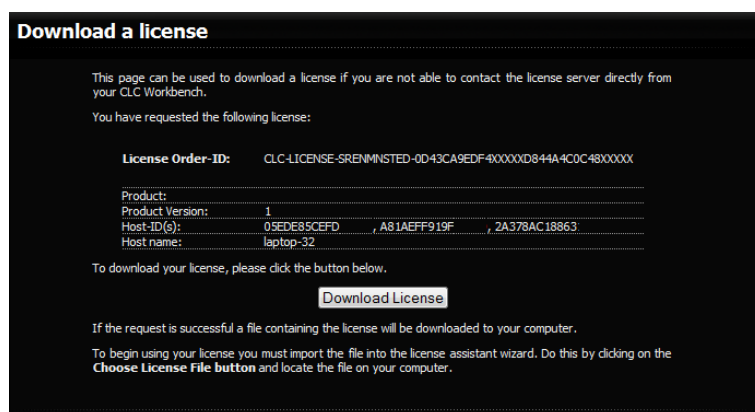


Figure 7.8: The license download form opened in a web browser.

Import a license from a file...

Please click the button below and locate the file containing your license.

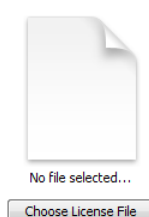


Figure 7.9: Importing the license file downloaded from the web page.

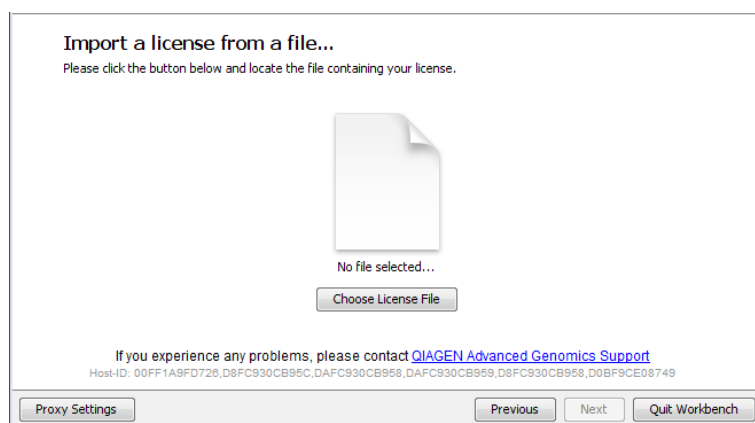


Figure 7.10: Selecting a license file.

Click the **Choose License File** button and browse to find the license file. When you have selected the file, click on the **Next** button.

7.1.4 Configure License Server connection

If your organization is running a *CLC Network License Manager* or *CLC License Server*, you can configure your Workbench to connect to it to get a license for the module.

To configure the Workbench to connect to a *CLC Network License Manager* or *CLC License Server*, select the **Configure License Server connection** option and click on the **Next** button. A dialog appears, as shown in figure 7.11.

Configure license manager connection...

Please choose how you would like to connect to your CLC Network License Manager.

☒ Enable license manager connection

☐ Automatically detect license manager.

☒ Manually specify license manager:

Hostname/IP-address:

Port:

☐ Use custom username when requesting a license

Username:

☐ Disable license borrowing

If you choose this option, users of this computer will not be able to borrow licenses from the license manager.

Figure 7.11: Connecting to a CLC Network License Manager or CLC License Server.

The options in that dialog are:

- **Enable license manager connection.** This box must be checked for the Workbench is to contact the *CLC Network License Manager* or CLC License Server to get a license for the CLC Workbench.
- **Automatically detect license manager.** By checking this option the Workbench will look for a *CLC Network License Manager* or CLC License Server accessible from the Workbench. Automatic server discovery sends UDP broadcasts from the Workbench on port 6200. Available license servers respond to the broadcast. The Workbench then uses TCP communication for to get a license, if one is available. Automatic server discovery works only on local networks and will not work on WAN or VPN connections. Automatic server discovery is not guaranteed to work on all networks. If you are working on an enterprise network on where local firewalls or routers cut off UDP broadcast traffic, then you may need to configure the details of the *CLC Network License Manager* or CLC License Server using the **Manually specify license manager** option instead.
- **Manually specify license manager.** Select this option to enter the details of the machine the *CLC Network License Manager* or CLC License Server software is running on, specifically:
 - **Host name.** The address of the machine the *CLC Network License Manager* or CLC License Server software is running on.
 - **Port.** The port used by the *CLC Network License Manager* or CLC License Server to receive requests.
- **Use custom username when requesting a license.** Optional. If this is checked, a username can be entered that will be used when requesting a network license instead of the username of the account being used to run the Workbench.
- **Disable license borrowing on this computer.** Check this box if you do not want users of the computer to borrow a license. See section 7.1.4 for further details.

Special note on modules needing a license

A valid module license is needed to start a module tool, or a workflow including a module tool. Network licenses for modules are valid for four hours after starting the tool or the workflow. A process started (whether a module tool or a workflow including a module tool) will always be completed, even if its completion exceeds the four hours period where the license is valid.

If the tool or the workflow completes before the four hour validity period, it is possible to start a new tool or a workflow, and this will always refresh the validity of the license to a full four hours period. However, if the tool or the workflow completes after the four hour validity period, a new license will need to be requested after that to start the next tool or workflow.

These measures ensure that more licenses are available to active users, rather than blocked on an inactive computer, i.e., where the workbench would be open but not in use.

Borrowing a license

A CLC Workbench using a network license normally needs to maintain a connection to the *CLC Network License Manager* or CLC License Server. However, if allowed by the network license administrator, network licenses can be *borrowed* for offline use. During the period a license has been borrowed, there will be one less network license available for other users.

If administrator has chosen not to allow the borrowing of network licenses, then the information in this section is not relevant.

The Workbench must be connected to the *CLC Network License Manager* or CLC License Server at the point when the license is borrowed. The procedure for borrowing a license is:

1. Go to the Workbench menu option:

Help | License Manager

2. Click on the "Borrow License" tab to display the dialog shown in figure 7.12.
3. Select the license(s) that you wish to borrow by clicking in the checkboxes in the Borrow column in the License overview panel.
4. Choose the length of time you wish to borrow the license(s) for using the drop down list in the Borrow License tab. By default the maximum is 7 days, but network license administrators can specify a lower limit than this.
5. Click **Borrow Selected Licenses**.
6. Close the License Manager when you are done.

You can now go offline and continue working with the CLC Workbench. When the time period you borrowed the license for has elapsed, the network license will be again made available for other users. To continue using CLC Workbench with a license, you will need to connect to the network again so the Workbench can request another license.

You can return borrowed licenses early if you wish by started up the **License Manager**, opening the "Borrow License" tab, and clicking on the **Return Borrowed Licenses** button.

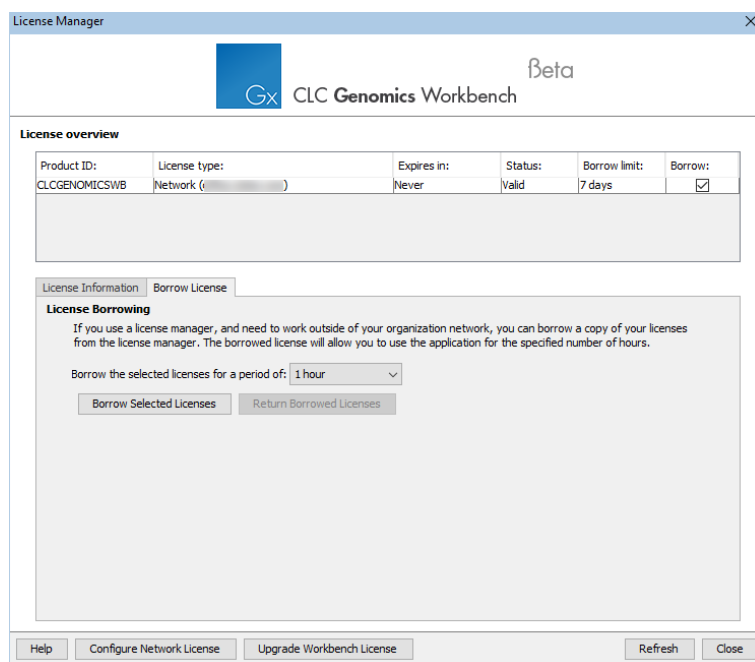


Figure 7.12: Borrow a license.

Common issues when using a network license

- No license available at the moment If all licenses are in use, you will see a dialog like that shown in figure 7.13 when you start up the Workbench.

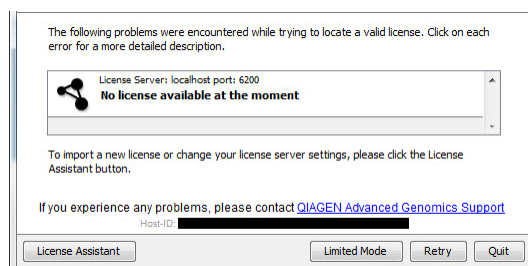


Figure 7.13: This window appears when there are no available network licenses for the software you are running.

You will need to wait for at least one license to be returned before you can continue to work with a fully functional copy of the software. If running out of licenses is a frequent issue, you may wish to discuss this with your administrator.

Clicking on the **Viewing Mode** button in the dialog allows you to run the CLC Workbench for viewing data, and for basic analyses, import and export.

- Lost connection to the CLC Network License Manager or CLC License Server If the Workbench connection to the *CLC Network License Manager* or CLC License Server is lost, you will see a dialog like that shown in figure 7.14.

If you have chosen the option to **Automatically detect license manager** and you have not succeeded in connecting to the *CLC Network License Manager* or CLC License Server before, please check with your local IT support that automatic detection will be possible to do at your site. If it is not, you will need to specify the settings, as described earlier in this section.

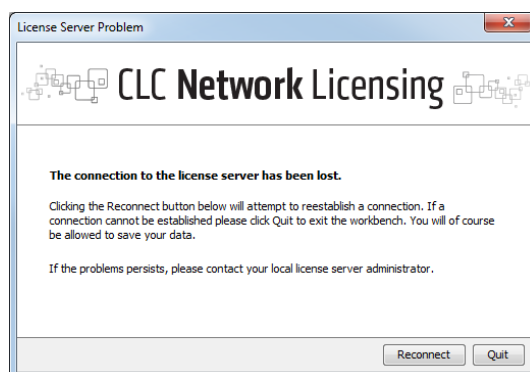


Figure 7.14: Here, the Workbench is unable to establish a connection to obtain a network license.

If you have successfully contacted the *CLC Network License Manager* or *CLC License Server* from your Workbench previously, please consider discussing this issue with your administrator, for example, making sure that the *CLC Network License Manager* or *CLC License Server* is running and that your Workbench is able to connect to it.

There may be situations where you wish to use a different license or view information about the license(s) the Workbench is currently using. To do this, open the License Manager using the menu option:

Help | License Manager (📄)

The license manager is shown in figure 7.15.

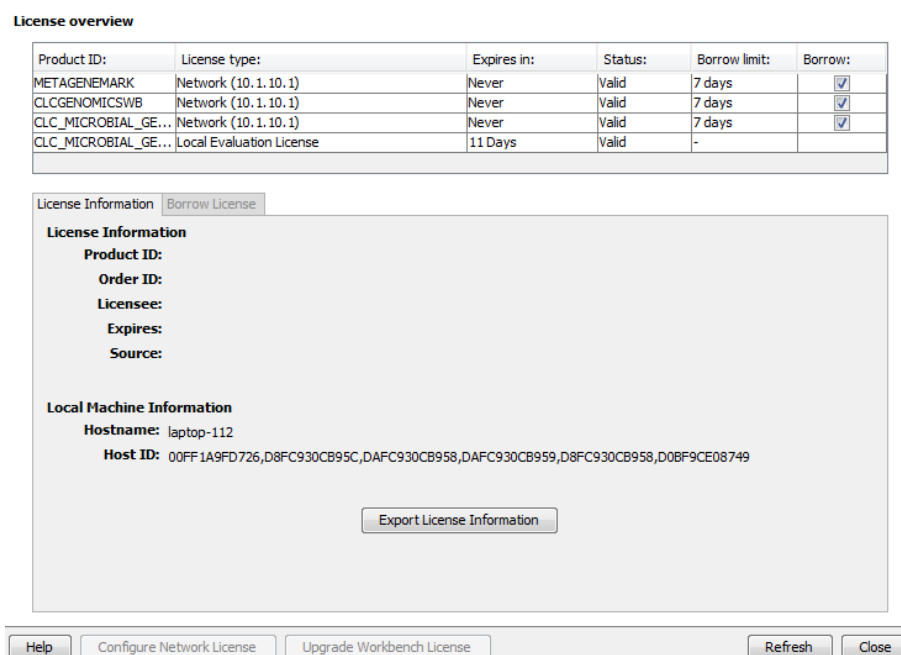


Figure 7.15: The license manager.

This dialog can be used to:

- See information about the license (e.g. what kind of license, when it expires)

- Configure how to connect to a license server (**Configure License Server** the button at the lower left corner). Clicking this button will display a dialog similar to figure 7.11.
- Upgrade from an evaluation license by clicking the **Upgrade license** button. This will display the dialog shown in figure 6.2.
- Export license information to a text file.
- Borrow a license

If you wish to switch away from using a network license, click on the button to **Configure License Server** and uncheck the box beside the text **Enable license server connection** in the dialog. When you restart the Workbench, you can set up the new license as described in section ??.

7.1.5 Download a static license on a non-networked computer

To download a static license for a machine that does not have direct access to the external network, you can follow the steps below:

- Install the Multilocus Sequence Typing Module on the machine you wish to run the software on.
- Start up the software as an administrative user and find the host ID of the machine that you will run the CLC Workbench on. You can see the host ID the machine reported at the bottom of the License Manager window in grey text.
- Make a copy of this host ID such that you can use it on a machine that has internet access.
- Go to a computer with internet access, open a browser window and go to the network license download web page:
<https://secure.clcbio.com/LmxWSv3/GetLicenseFile>
- Paste in your license order ID and the host ID that you noted down in the relevant boxes on the webpage.
- Click 'download license' and save the resulting .lic file.
- Open the Workbench on your non-networked machine. In the Workbench license manager choose 'Import a license from a file'. In the resulting dialog click 'choose license file' to browse the location of the .lic file you have just downloaded.

If the License Manager does not start up by default, you can start it up by going to the Help menu and choosing License Manager.

- Click on the **Next** button and go through the remaining steps of the license manager wizard.

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.