

Compendium - CLC Main Workbench Workshop

CLC bio
Finlandsgade 10-12 · 8200 Aarhus N · Denmark
Telephone: +45 70 22 55 09 · Fax: +45 70 22 55 19
www.clcbio.com · support@clcbio.com

CLC Academy

Contents

1 Tutorial: Open a DNA sequence **4**

2 Tutorial: Adjust and save view settings **6**

2.1 Saving the settings in the Side Panel 7

2.2 Applying saved settings 8

3 Tutorial: Primer design **10**

3.1 Specifying a region for the forward primer 10

3.2 Examining the primer suggestions 11

3.3 Calculating a primer pair 12

4 Tutorial: In silico cloning cloning work flow **15**

4.1 Locating the data to use 15

4.2 Add restriction sites to primers 15

4.3 Simulate PCR to create the fragment 18

4.4 Specify restriction sites and perform cloning 18

5 Tutorial: Find restriction sites **21**

5.1 The Side Panel way of finding restriction sites 21

5.2 The Toolbox way of finding restriction sites 21

6 Tutorial: Working with annotations **24**

6.1 Browsing and viewing annotations in sequence views 24

6.2 Adding and editing annotations 24

6.3 Copying annotations 27

7 Tutorial: Align protein sequences **28**

7.1 The alignment dialog 28

8 Tutorial: BLAST search **31**

8.1 Performing the BLAST search 31

8.2 Inspecting the results 32

8.3 Using the BLAST table view 33

9 Gateway Cloning **35**

9.1 Importing Gateway Cloning vectors 35

9.2 Adding attB sites 36

9.3 Creation of an entry vector 39

9.4 Creating an expression vector (LR) 41

9.5 Short suggestion for a MultiSite Gateway Workflow using 2 fragments 42

CLC Academy

Chapter 1

Tutorial: Open a DNA sequence

This brief tutorial will take you through some different ways to display a sequence in the program. The tutorial introduces zooming on a sequence, dragging tabs, and opening selection in new view.

We will be working with the sequence called *pcDNA3-atp8a1* located in the 'Cloning' folder in the Example data. Double-click the sequence in the **Navigation Area** to open it. The sequence is displayed with annotations above it. (See figure 1.1).



Figure 1.1: *Sequence pcDNA3-atp8a1 opened in a view.*

As default, *CLC Main Workbench* displays a sequence with annotations (colored arrows on the sequence like the green promoter region annotation in figure 1.1) and zoomed to see the residues.

In this tutorial we want to have an overview of the whole sequence. Hence;

> **click Zoom Out ( ) in the Toolbar | click the sequence until you can see the whole sequence**

This sequence is circular, which is indicated by << and >> at the beginning and the end of the sequence.

In the following we will show how the same sequence can be displayed in two different views - one linear view and one circular view. First, zoom in to see the residues again by using the **Zoom In** ( ) or the **100%** ( ). Then we make a split view by:

> **press and hold the Ctrl-button on the keyboard (⌘ on Mac) | click Show as Circular ( ) at the bottom of the view**

This opens an additional view of the vector with a circular display, as can be seen in figure 1.2.

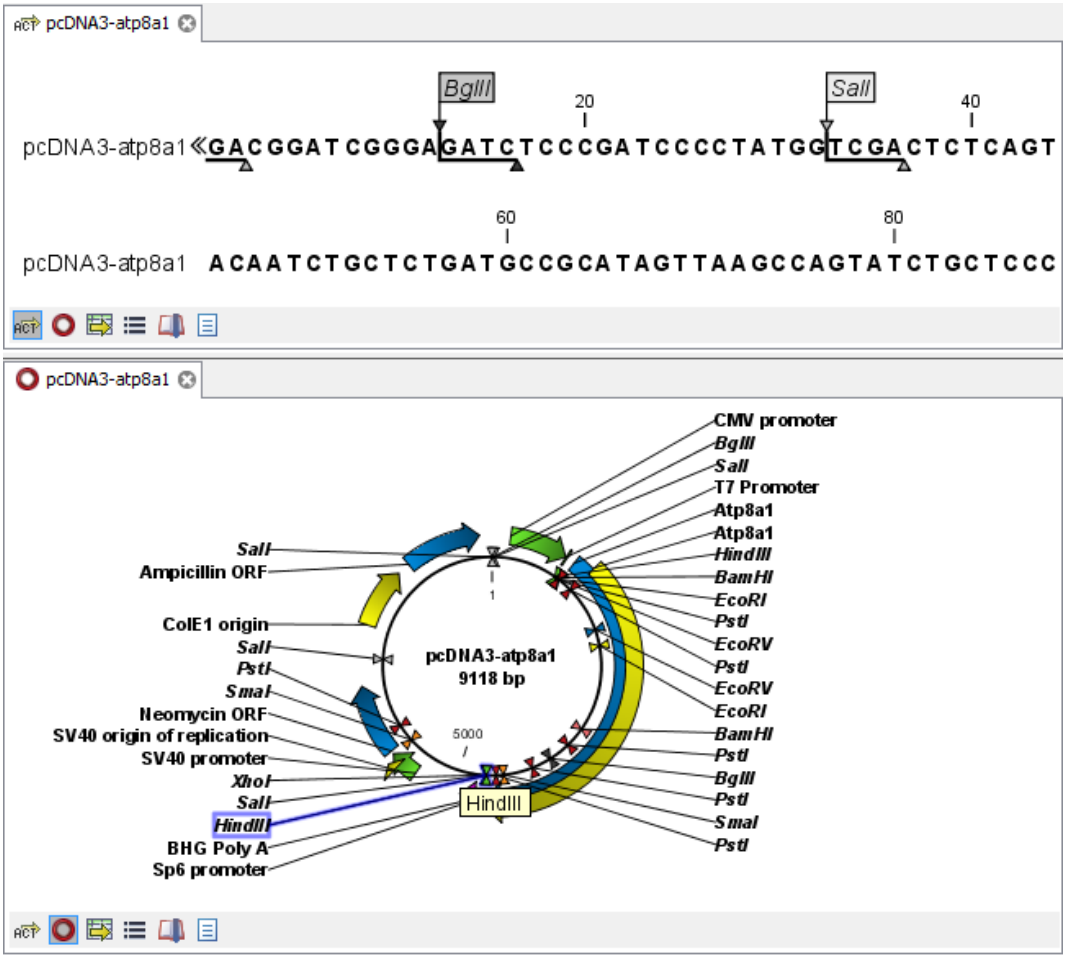


Figure 1.2: *The resulting two views which are split horizontally.*

Make a selection on the circular sequence (remember to switch to the **Selection** ( ) tool in the tool bar) and note that this selection is also reflected in the linear view above.

Chapter 2

Tutorial: Adjust and save view settings

This brief tutorial will show you how to use the **Side Panel** to change the way your sequences, alignments and other data are shown. You will also see how to save the changes that you made in the **Side Panel** .

Open the protein alignment located under *Protein orthologs* in the **Example data**. The initial view of the alignment has colored the residues according to the Rasmol color scheme, and the alignment is automatically wrapped to fit the width of the view (shown in figure 2.1).

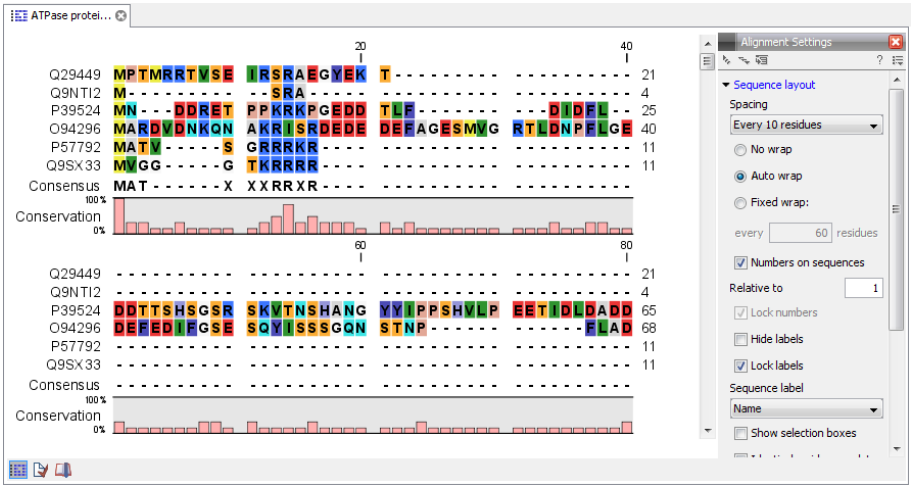


Figure 2.1: *The protein alignment as it looks when you open it with background color according to the Rasmol color scheme and automatically wrapped.*

Now, we are going to modify how this alignment is displayed. For this, we use the settings in the **Side Panel** to the right. All the settings are organized into groups, which can be expanded / collapsed by clicking the name of the group. The first group is **Sequence Layout** which is expanded by default.

First, select **No wrap** in the **Sequence Layout**. This means that each sequence in the alignment is kept on the same line. To see more of the alignment, you now have to scroll horizontally.

Next, expand the **Annotation Layout** group and select **Show Annotations**. Set the **Offset** to "More offset" and set the **Label** to "Stacked".

Expand the **Annotation Types** group. Here you will see a list of the types annotation that are

carried by the sequences in the alignment (see figure 2.2).

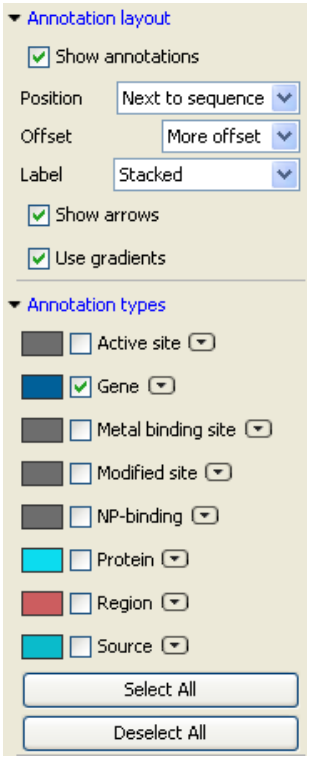


Figure 2.2: *The Annotation Layout and the Annotation Types in the Side Panel.*

Check the "Region" annotation type, and you will see the regions as red annotations on the sequences.

Next, we will change the way the residues are colored. Click the **Alignment Info** group and under **Conservation**, check "Background color". This will use a gradient as background color for the residues. You can adjust the coloring by dragging the small arrows above the color box.

2.1 Saving the settings in the Side Panel

Now the alignment should look similar to figure 2.3.

At this point, if you just close the view, the changes made to the **Side Panel** will not be saved. This means that you would have to perform the changes again next time you open the alignment. To save the changes to the **Side Panel**, click the **Save/Restore Settings** button ( ) at the top of the **Side Panel** and click **Save Settings** (see figure 2.4).

This will open the dialog shown in figure 2.5.

In this way you can save the current state of the settings in the **Side Panel** so that you can apply them to alignments later on. If you check **Always apply these settings**, these settings will be applied every time you open a view of the alignment.

Type "My settings" in the dialog and click **Save**.

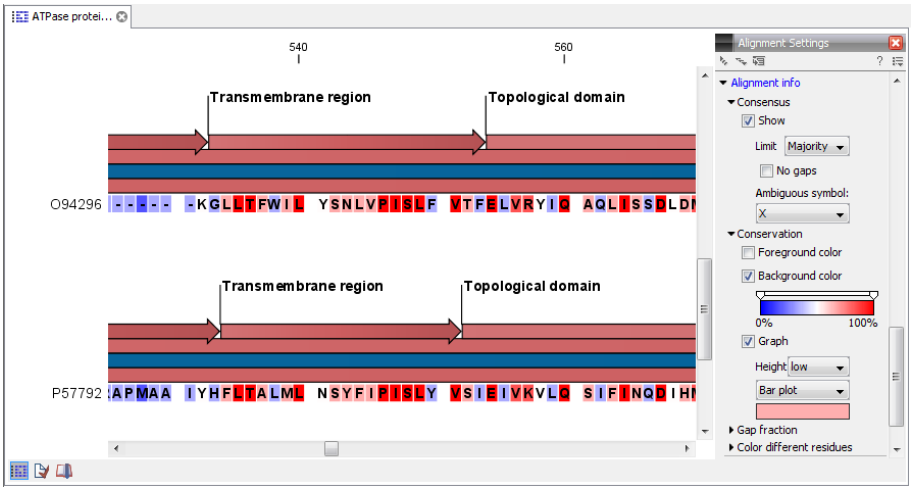


Figure 2.3: *The alignment when all the above settings have been changed.*

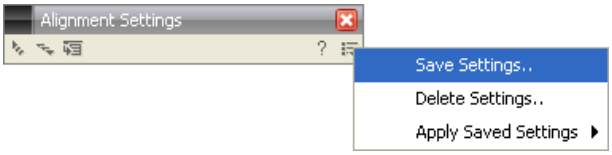


Figure 2.4: *Saving the settings of the Side Panel.*

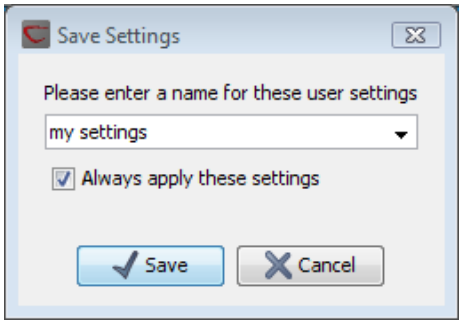


Figure 2.5: *Dialog for saving the settings of the Side Panel.*

2.2 Applying saved settings

When you click the **Save/Restore Settings** button ( ) again and select **Apply Saved Settings**, you will see "My settings" in the menu together with some pre-defined settings that the *CLC Main Workbench* has created for you (see figure 2.6).

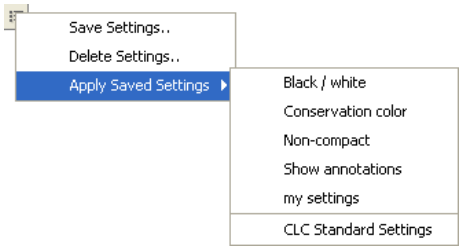


Figure 2.6: *Menu for applying saved settings.*

Whenever you open an alignment, you will be able to apply these settings. Each kind of view has its own list of settings that can be applied.

At the bottom of the list you will see the "CLC Standard Settings" which are the default settings

for the view.

CLC Academy

Chapter 3

Tutorial: Primer design

In this tutorial, you will see how to use the *CLC Main Workbench* to find primers for PCR amplification of a specific region.

We use the pcDNA3-atp8a1 sequence from the 'Primers' folder in the Example data. This sequence is the pcDNA3 vector with the atp8a1 gene inserted. In this tutorial, we wish design primers that would allow us to generate a PCR product covering the insertion point of the gene. This would let us use PCR to check that the gene is inserted where we think it is.

First, open the sequence in the Primer Designer:

Select the pcDNA3-atp8a1 sequence | Show ( ) | Primer Designer ( )

Now the sequence is opened and we are ready to begin designing primers.

3.1 Specifying a region for the forward primer

First zoom out to get an overview of the sequence by clicking **Fit Width** ( ). You can now see the blue gene annotation labeled Atp8a1, and just before that there is the green CMV promoter. This may be hidden behind restriction site annotations. Remember that you can always choose not to Show these by altering the settings in the right hand pane.

In this tutorial, we want the forward primer to be in a region between positions 600 and 900 - just before the gene (you may have to zoom in ( ) to make the selection). Select this region, right-click and choose "Forward primer region here" ( ) (see figure 3.1).

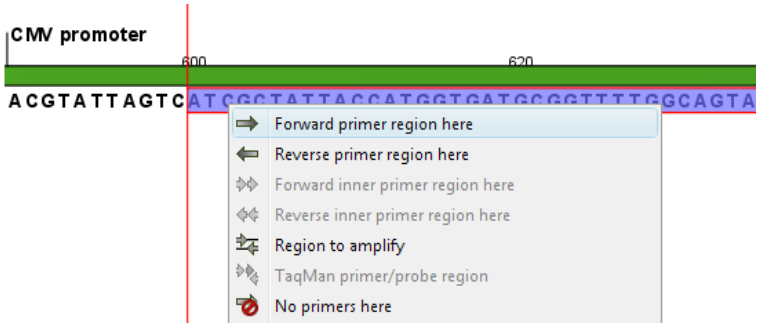


Figure 3.1: *Right-clicking a selection and choosing "Forward primer region here".*

This will add an annotation to this region, and five rows of red and green dots are seen below as shown in figure 3.2:



Figure 3.2: *Five lines of dots representing primer suggestions. There is a line for each primer length - 18bp through to 22 bp.*

3.2 Examining the primer suggestions

Each line consists of a number of dots, each representing the *starting point* of a possible primer. E.g. the first dot on the first line (primers of length 18) represents a primer starting at the dot's position and with a length of 18 nucleotides (shown as the white area in figure 3.3):

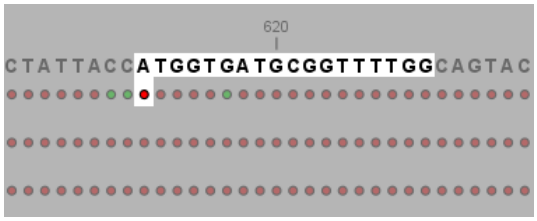


Figure 3.3: *The first dot on line one represents the starting point of a primer that will anneal to the highlighted region.*

Position the mouse cursor over a dot. A box will appear, providing data about this primer. Clicking the dot will select the region where that primer would anneal. (See figure 3.4):

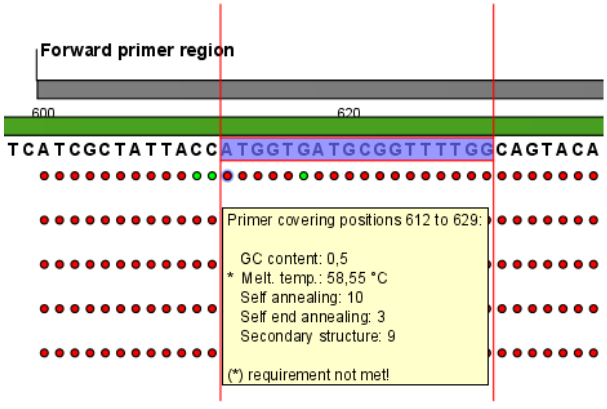


Figure 3.4: *Clicking the dot will select the corresponding primer region. Hovering the cursor over the dot will bring up an information box containing details about that primer.*

Note that some of the dots are colored red. This indicates that the primer represented by this dot does not meet the requirements set in the **Primer parameters** (see figure 3.5):

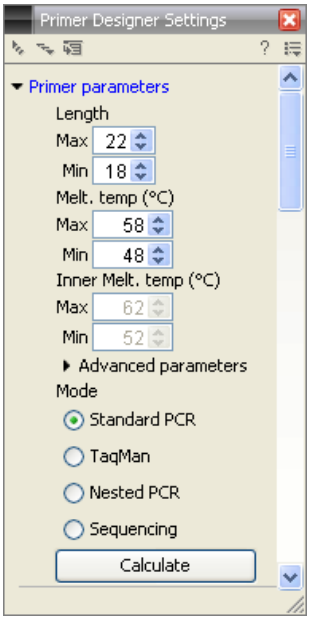


Figure 3.5: *The Primer parameters.*

The default maximum melting temperature is 58. This is the reason why the primer in figure 3.4 with a melting temperature of 58.55 does not meet the requirements and is colored red. If you raise the maximum melting temperature to 59, the primer will meet the requirements and the dot becomes green.

In figure 3.4 there is an asterisk (*) before the melting temperature. This indicates that this primer does not meet the requirements regarding melting temperature. In this way, you can easily see why a specific primer (represented by a dot) fails to meet the requirements.

By adjusting the **Primer parameters** you can define primersto meet your specific needs. Since the dots are dynamically updated, you can immediately see how a change in the primer parameters affects the number of red and green dots.

3.3 Calculating a primer pair

Until now, we have been looking at the forward primer. To mark a region for the reverse primer, make a selection from position 1200 to 1400 and:

Right-click the selection | Reverse primer region here (⇐)

The two regions should now be located as shown in figure 3.6:

Now, you can let *CLC Main Workbench* calculate all the possible primer pairs based on the **Primer parameters** that you have defined:

Click the Calculate button (right hand pane) | Modify parameters regarding the combination of the primers (for now, just leave them unchanged)| Calculate

This will open a table showing the possible combinations of primers. To the right, you can specify the information you want to display, e.g. showing **Fragment length** (see figure 3.7):

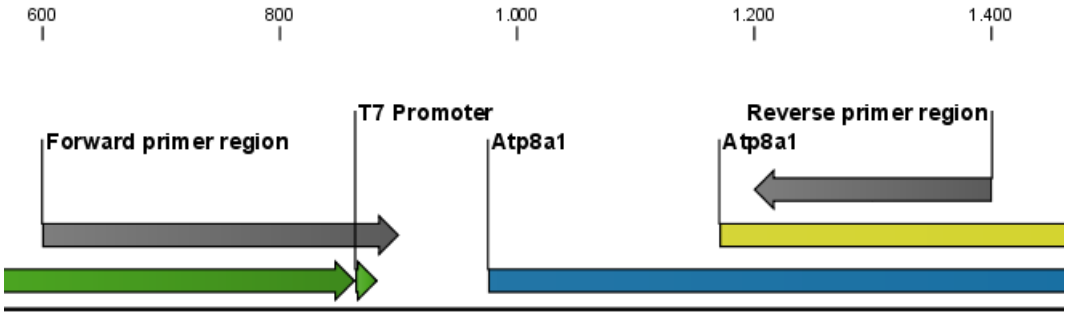
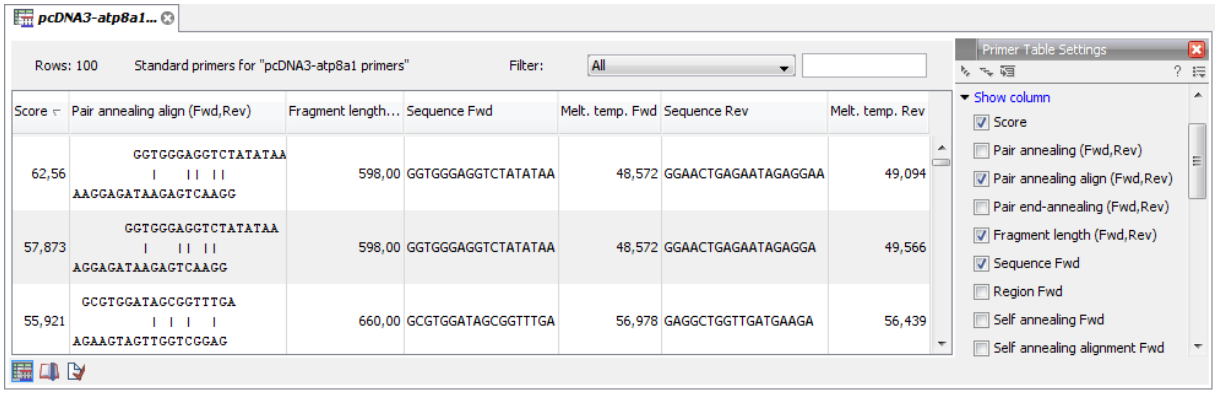


Figure 3.6: *A forward and a reverse primer region.*



Score	Pair annealing align (Fwd,Rev)	Fragment length...	Sequence Fwd	Melt. temp. Fwd	Sequence Rev	Melt. temp. Rev
62,56	GGTGGGAGGTCTATATAA AAGGAGATAAGAGTCAAGG	598,00	GGTGGGAGGTCTATATAA	48,572	GGAACTGAGAATAGAGGAA	49,094
57,873	GGTGGGAGGTCTATATAA AGGAGATAAGAGTCAAGG	598,00	GGTGGGAGGTCTATATAA	48,572	GGAACTGAGAATAGAGGA	49,566
55,921	GCGTGGATAGCGGTTTGA AGAAGTAGTTGGTCGGAG	660,00	GCGTGGATAGCGGTTTGA	56,978	GAGGCTGGTTGATGAAGA	56,439

Figure 3.7: *A list of primers. To the right are the Side Panel showing the available choices of information to display.*

Clicking a primer pair in the table will make a corresponding selection on the sequence in the view above. At this point, you can either settle on a specific primer pair or save the table for later. If you want to use e.g. the first primer pair for your experiment, right-click this primer pair in the table and save the primers.

You can also mark the position of the primers on the sequence by selecting **Mark primer annotation on sequence** in the right-click menu (see figure 3.8):

This tutorial has shown some of the many options of the primer design functionalities of *CLC Main Workbench*. You can read much more using the program's **Help** function (?) or in the *CLC Main Workbench* user manual, linked to on this webpage: http://www.clcbio.com/download.

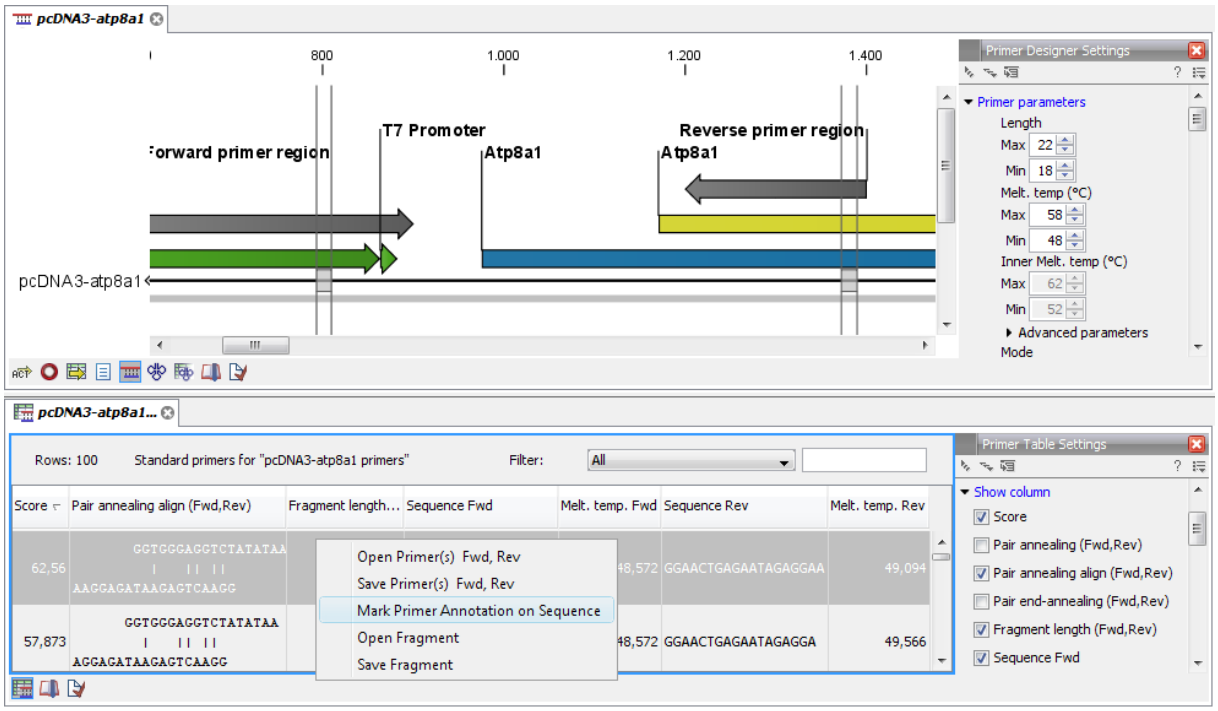


Figure 3.8: *The options available in the right-click menu. Here, "Mark primer annotation on sequence" has been chosen, resulting in two annotations on the sequence above (labeled "Oligo").*

Chapter 4

Tutorial: In silico cloning cloning work flow

In this tutorial, the goal is to virtually PCR-amplify a gene using primers with restriction sites at the 5' ends, and insert the gene into a multiple cloning site of an expression vector. We start off with a set of primers, a DNA template sequence and an expression vector loaded into the Workbench.

This tutorial will guide you through the following steps:

1. Adding restriction sites to the primers
2. Simulating the effect of PCR by creating the fragment to use for cloning.
3. Specifying restriction sites to use for cloning, and inserting the fragment into the vector

4.1 Locating the data to use

Open the `Example data` folder in the **Navigation Area**. Open the `Cloning` folder, and inside this folder, open the `Primer` folder.

If you do not have the example data, please go to the **Help** menu to import it.

The data to use in these folders is shown in figure 4.1.

Double-click the `ATP8a1 mRNA` sequence and zoom to **Fit Width** ( ) and you will see the yellow annotation which is the coding part of the gene. This is the part that we want to insert into the `pcDNA4_TO` vector. The primers have already been designed using the primer design tool in *CLC Main Workbench* (to learn more about this, please refer to the Primer design tutorial).

4.2 Add restriction sites to primers

First, we add restriction sites to the primers. In order to see which restriction enzymes can be used, we create a split view of the vector and the fragment to insert. In this way we can easily make a visual check to find enzymes from the multiple cloning site in the vector that do not cut in the gene of interest. To create the split view:

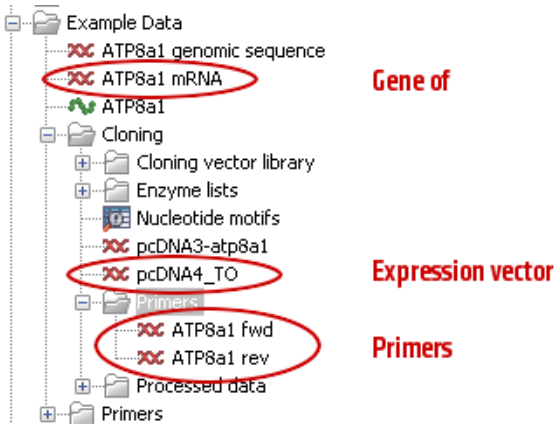


Figure 4.1: *The data to use in this tutorial.*

double-click the pcDNA4_TO sequence | View | Split Horizontally ( )

Note that this can also be achieved by simply dragging the `pcDNA4_TO` sequence into the lower part of the open view.

Switch to the **Circular** ( ) view at the bottom of the view.

Zoom in ( ) on the multiple cloning site downstream of the green `CMV` promoter annotation. You should now have a view similar to the one shown in figure 4.2.

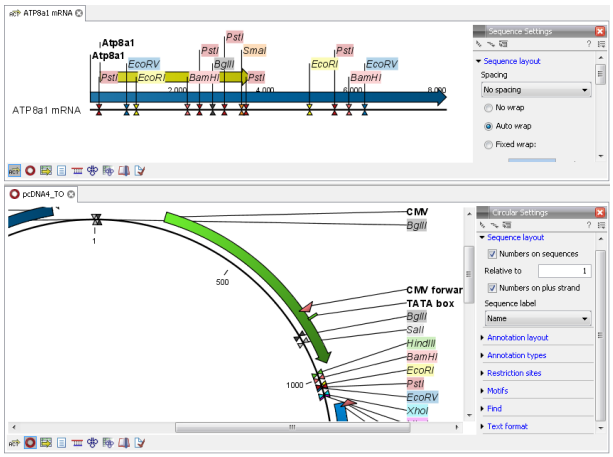


Figure 4.2: *Check cut sites.*

By looking at the enzymes we can see that both *HindIII* and *XhoI* cut in the multi-cloning site of the vector and not in the Atp8a1 gene. Note that you can add more enzymes to the list in the **Side Panel** by clicking **Manage Enzymes** under the **Restriction Sites** group.

Close both views and open the `ATP8a1 fwd` primer sequence. When it opens, double-click the name of the sequence to make a selection of the full sequence. If you do not see the whole sequence turn purple, please make sure you have the Selection Tool chosen, and not one of the other tools available from the top right side of the Workbench (e.g. Pan, Zooming tools, etc.)

Once the sequence is selected, right-click and choose to **Insert Restriction Site Before Selection** as shown in figure 4.3.

In the **Filter** box enter *HindIII* and click on it. At the bottom of the dialog, add a few extra bases

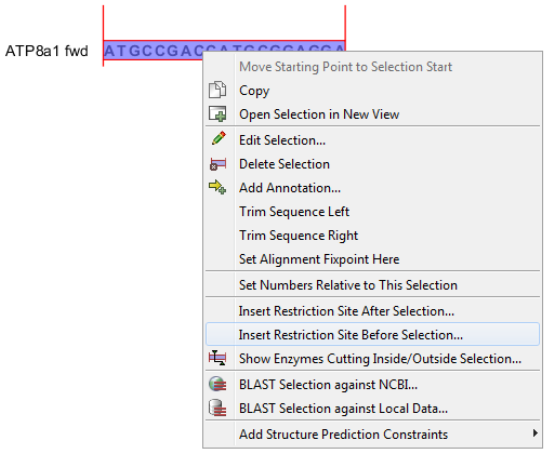


Figure 4.3: *Adding restriction sites to a primer.*

5' of the cut site (this is done to increase the efficiency of the enzyme) as shown figure 4.4.

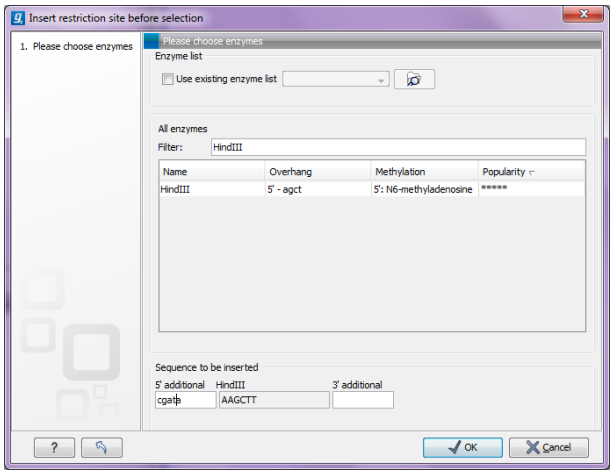


Figure 4.4: *Adding restriction sites to a primer.*

Click **OK** and the sequence will be inserted at the 5' end of the primer as shown in figure figure 4.5.



Figure 4.5: *Adding restriction sites to a primer.*

Perform the same process for the `ATP8a1 rev` primer, this time using *XhoI* instead. This time, you should also add a few bases at the 5' end as was done in figure4.4 when inserting the *HindIII* site.

Note! The `ATP8a1 rev` primer is designed to match the negative strand, so the restriction site should be added at the 5' end of this sequence as well (**Insert Restriction Site before Selection**).

Save () the two primers and close the views and you are ready for next step.

4.3 Simulate PCR to create the fragment

Now, we want to extract the PCR product from the template `ATP8a1 mRNA` sequence using the two primers with restriction sites:

Toolbox | **Primers and Probes** ( ) | **Find Binding Sites and Create Fragments** ( )

Select the `ATP8a1 mRNA` sequence and click **Next**. In this dialog, use the **Browse** ( ) button to select the two primer sequences. Click **Next** and adjust the output options as shown in figure 4.6.

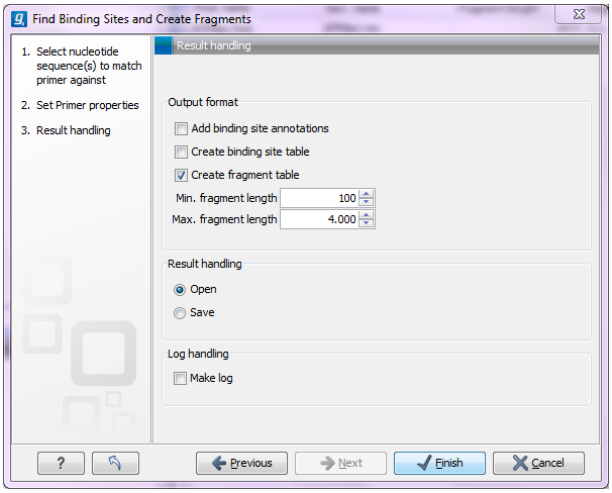


Figure 4.6: *Creating the fragment table including fragments up to 4000 bp.*

Click **Finish** and you will now see the fragment table displaying the PCR product.

In the **Side Panel** you can choose to show information about melting temperature for the primers.

Right-click the fragment and select **Open Fragment** as shown in figure 4.7.

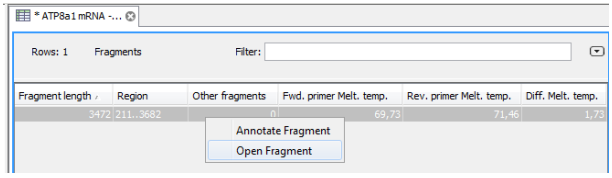


Figure 4.7: *Opening the fragment as a sequence.*

This will create a new sequence representing the PCR product. **Save** ( ) the sequence in the `Cloning` folder and close the views. You do not need to save the fragment table.

4.4 Specify restriction sites and perform cloning

The final step in this tutorial is to insert the fragment into the cloning vector:

Toolbox in the Menu Bar | **Cloning and Restriction Sites** ( ) | **Cloning** ( )

Select the `Fragment (ATP8a1 mRNA (ATP8a1 fwd - ATP8a1 rev))` sequence you just saved and click **Next**. In this dialog, use the **Browse** ( ) button to select `pcDNA4_TO` cloning vector also located in the `Cloning` folder. Click **Finish**.

You will now see the cloning editor where you will see the `pcDNA4_TO` vector in a circular view. Press and hold the Ctrl (⌘ on Mac) key while you click first the *HindIII* site and next the *XhoI* site (see figure 4.8).

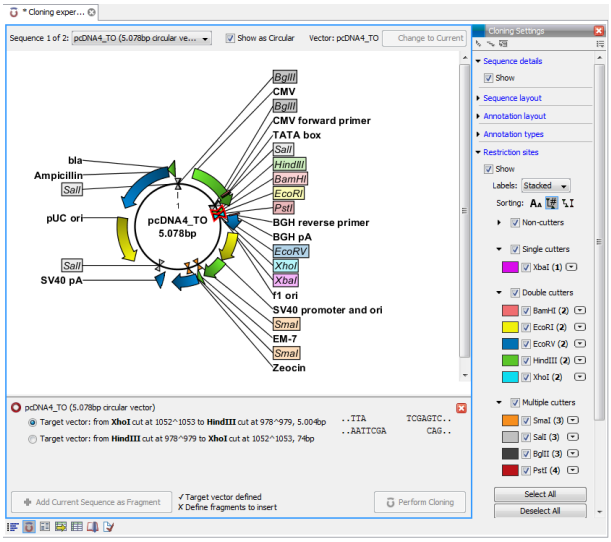


Figure 4.8: *Press and hold the Ctrl key while you click first the* HindIII *site and next the* XhoI *site.*

At the bottom of the view you can now see information about how the vector will be cut open. Since the vector has now been split into two fragments, you can decide which one to use as the target vector. If you selected first the *HindIII* site and next the *XhoI* site, the *CLC Main Workbench* has already selected the right fragment as the target vector. If you click one of the vector fragments, the corresponding part of the sequence will be high-lighted.

Next step is to cut the fragment. At the top of the view you can switch between the sequences used for cloning (at this point it says `pcDNA4_TO 5.078bp circular vector`). Switch to the fragment sequence and perform the same selection of cut sites as before while pressing the Ctrl (⌘ on Mac) key. You should now see a view identical to the one shown in figure 4.9.

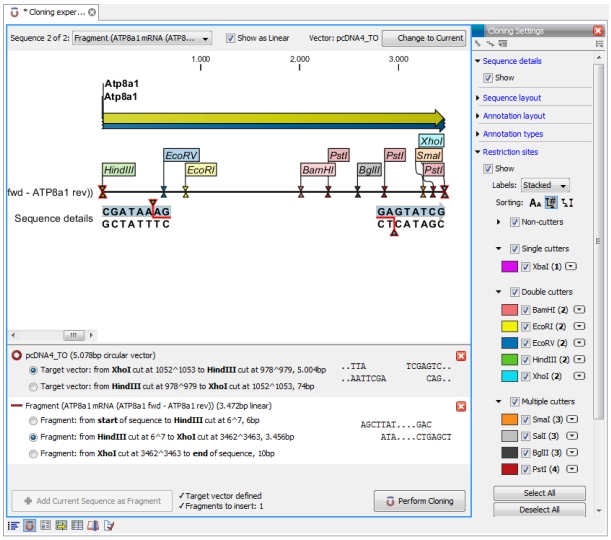


Figure 4.9: *Press and hold the Ctrl key while you click first the* HindIII *site and next the* XhoI *site.*

When this is done, the **Perform Cloning** ( ) button at the lower right corner of the view is active

because there is now a valid selection of both fragment and target vector. Click the **Perform Cloning** ( ) button and you will see the dialog shown in figure 4.10.

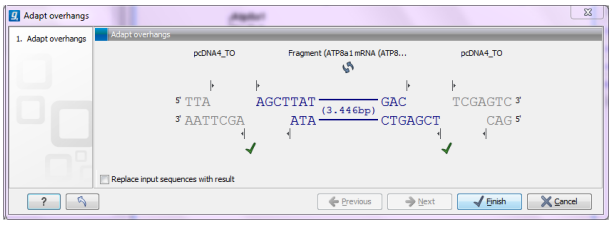


Figure 4.10: *Showing the insertion point of the vector*

This dialog lets you inspect the overhangs of the cut site, showing the vector sequence on each side and the fragment in the middle. The fragment can be reverse complemented by clicking the **Reverse complement fragment** ( ) but this is not necessary in this case. Click **Finish** and your new construct will be opened.

When saving your work, there are two options:

- Save the Cloning Experiment. This is saved as a sequence list, including the specified cut sites. This is useful if you need to perform the same process again or double-check details.
- Save the construct shown in the circular view. This will only save the information on the particular sequence including details about how it was created (this can be shown in the **History** view).

You can, of course, save both. In that case, the history of the construct will point to the sequence list in its own history.

The construct is shown in figure 4.11.

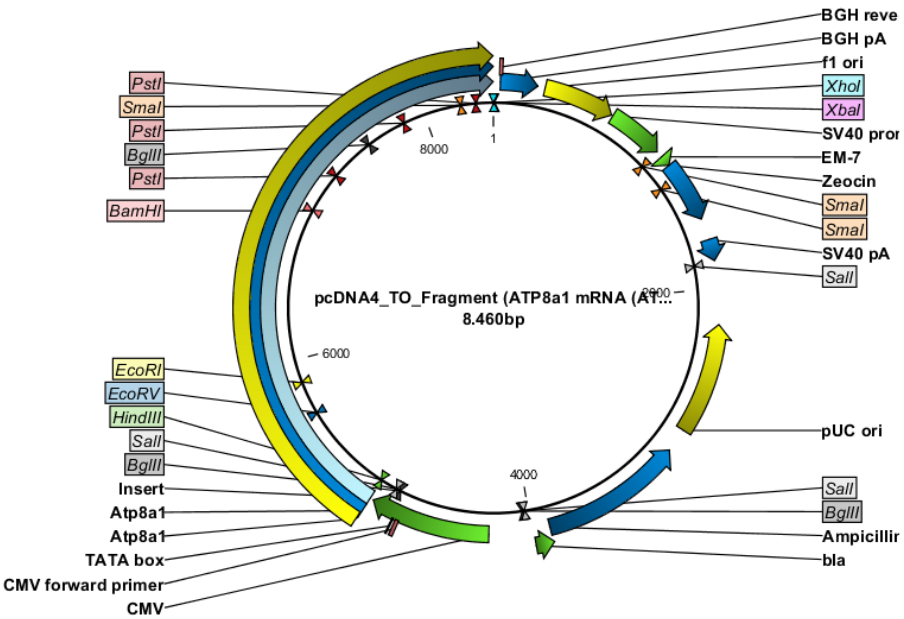


Figure 4.11: *The Atp8a1 gene inserted after the CMV promoter*

Chapter 5

Tutorial: Find restriction sites

This tutorial will show you how to find restriction sites and annotate them on a sequence.

There are two ways of finding and showing restriction sites. In many cases, the dynamic restriction sites found in the **Side Panel** of sequence views will be useful, since it is a quick and easy way of showing restriction sites. In the **Toolbox** you will find the other way of doing restriction site analyses. This way provides more control of the analysis and gives you more output options, e.g. a table of restriction sites and a list of restriction enzymes that can be saved for later use. In this tutorial, the first section describes how to use the Side Panel to show restriction sites, whereas the second section describes the restriction map analysis performed from the **Toolbox**.

5.1 The Side Panel way of finding restriction sites

When you open a sequence, there is a **Restriction sites** setting in the **Side Panel**. By default, 10 of the most popular restriction enzymes are shown (see figure 5.1).

The restriction sites are shown on the sequence with an indication of cut site and recognition sequence. In the list of enzymes in the **Side Panel**, the number of cut sites is shown in parentheses for each enzyme (e.g. *SalI* cuts three times). If you wish to see the recognition sequence of the enzyme, place your mouse cursor on the enzyme in the list for a short moment, and a tool tip will appear.

You can add or remove enzymes from the list by clicking the **Manage enzymes** button.

5.2 The Toolbox way of finding restriction sites

Suppose you are working with sequence 'ATP8a1 mRNA' from the example data, and you wish to know which restriction enzymes will cut this sequence exactly once and create a 3' overhang. Do the following:

> **select the ATP8a1 mRNA sequence | Toolbox in the Menu Bar | Cloning and Restriction Sites ( ) | Restriction Site Analysis ( )**

Click **Next** to set parameters for the restriction map analysis.

In this step first select **Use existing enzyme list** and click the **Browse for enzyme list** button



Figure 5.1: *Showing restriction sites of ten restriction enzymes.*

(). Select the 'Popular enzymes' in the Cloning folder under Enzyme lists.

Then write 3' into the filter below to the left. Select all the enzymes and click the **Add** button (). The result should be like in figure 5.2.

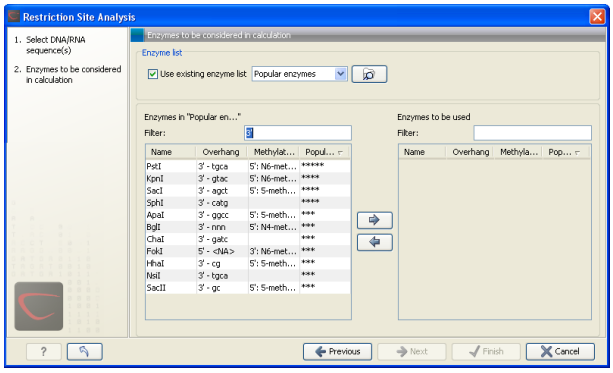


Figure 5.2: *Selecting enzymes.*

Click **Next**. In this step you specify that you want to show enzymes that cut the sequence only once. This means that you should de-select the **Two restriction sites** checkbox.

Click **Next** and select that you want to **Add restriction sites as annotations on sequence** and **Create restriction map**. (See figure 5.3).

Click **Finish** to start the restriction map analysis.

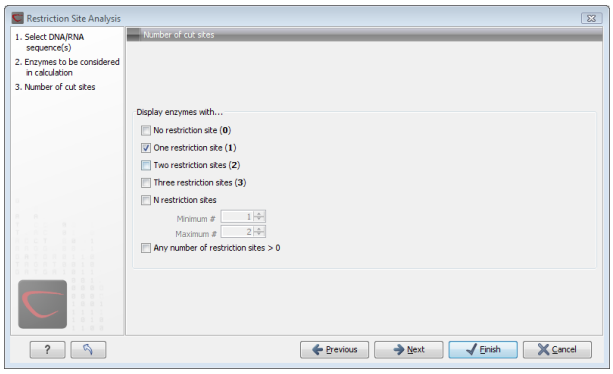


Figure 5.3: *Selecting output for restriction map analysis.*

5.2.1 View restriction site

The restriction sites are shown in two views: one view is in a tabular format and the other view displays the sites as annotations on the sequence.

The result is shown in figure 5.4. The restriction map at the bottom can also be shown as a

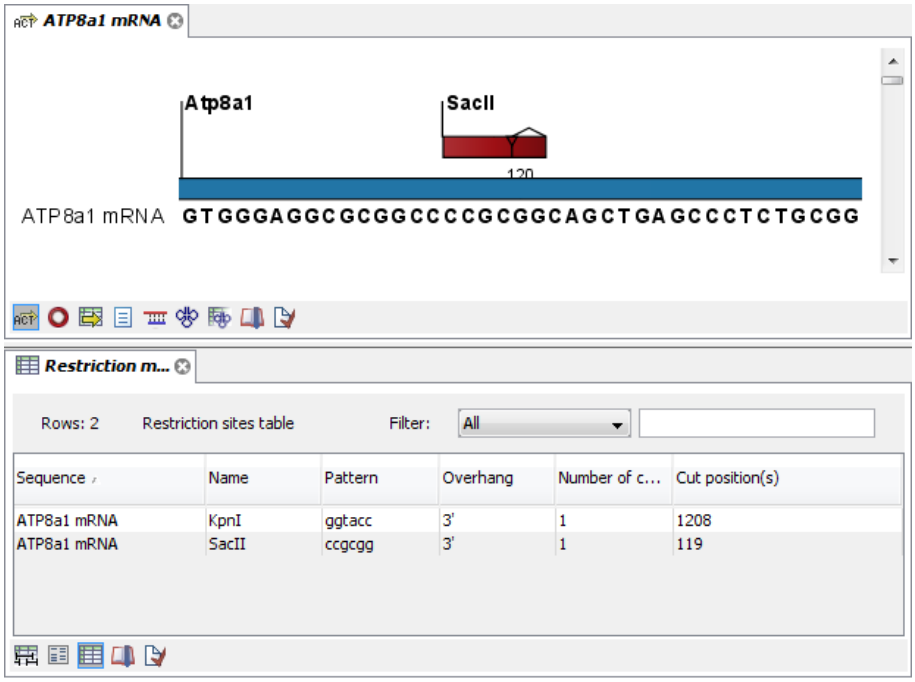


Figure 5.4: *The result of the restriction map analysis is displayed in a table at the bottom and as annotations on the sequence in the view at the top.*

table of fragments produced by cutting the sequence with the enzymes:

Click the Fragments button () at the bottom of the view

In a similar way the fragments can be shown on a virtual gel:

Click the Gel button () at the bottom of the view

Chapter 6

Tutorial: Working with annotations

Annotations are the basis of many of the analyses in the Workbench. Once an annotation has been added to a sequence, it stays there even when the sequence is transformed to be part of an alignment, a BLAST result or a sequence list.

Because annotations are so fundamental to many of the tasks performed in the Workbench, there are a number of ways to browse, search for, view and edit annotations. This tutorial takes you through a lot of different areas of the Workbench to show some of the places where you can work with annotations.

6.1 Browsing and viewing annotations in sequence views

Open sequence *ATP8a1 mRNA* from the Example data by double-clicking the sequence. Click **Annotation types** in the **Side Panel** and you will see a list of the types of annotations on this sequence (shown in figure 6.1).

Click the *Exon* annotation type. This will show exon annotations on the sequence. As this sequence is very long, you cannot see the whole sequence, and the exons may therefore not be visible. In this situation, you can either Zoom out to find the annotations, or you can use the browse button () next to the annotation type to jump directly to one of the annotations (see figure 6.2).

Click *Exon 2*, and the beginning of the annotation will be visible in the view (see figure 6.3).

Annotations are displayed as arrows, going from left to right on the positive strand, and going from right to left on the negative strand. Placing your mouse cursor on the arrow will show additional information about the annotation. Try to place the mouse cursor on the blue gene annotation, and you can see more information about it (see figure 6.4).

This way of displaying and accessing annotations is similar for both circular views, alignments, and all other views displaying sequence residues.

6.2 Adding and editing annotations

You can add your own annotations to a sequence. A new annotation is most easily added if you first select a region on the sequence. In the upper view, select a region from residue 10 to 26

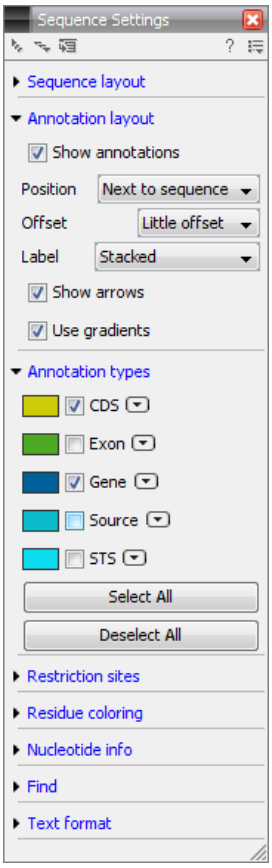


Figure 6.1: *Annotation types of the ATP8a1 mRNA sequence.*

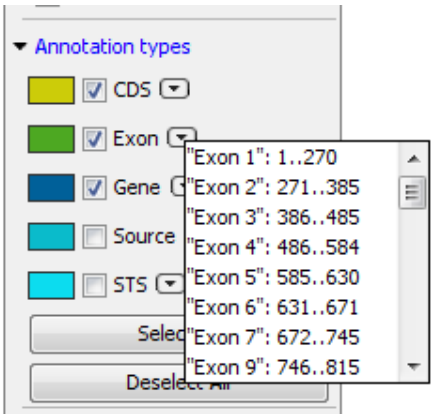


Figure 6.2: *Click in the list to go directly to this annotation.*

(see the status bar in the lower right corner of the Workbench). Then:

right-click the selection | Add Annotation ()

This will display a dialog where you can enter more information about the annotation. Enter *Test* as the name of the annotation (see figure 6.5).

Click **OK** and the annotation is added to the sequence (see figure 6.6).

Now click the **Show Annotation Table** () icon at the bottom of the view. This will display a list of all the annotations on the sequence.

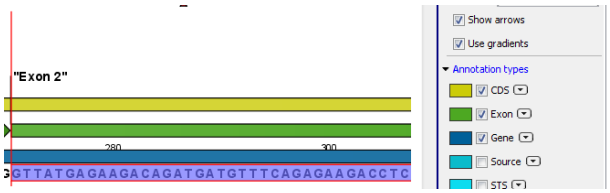


Figure 6.3: *The view jumps to the beginning of exon 2.*

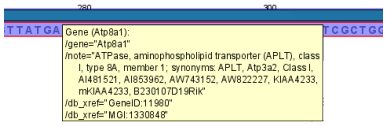


Figure 6.4: *Information about the annotation.*

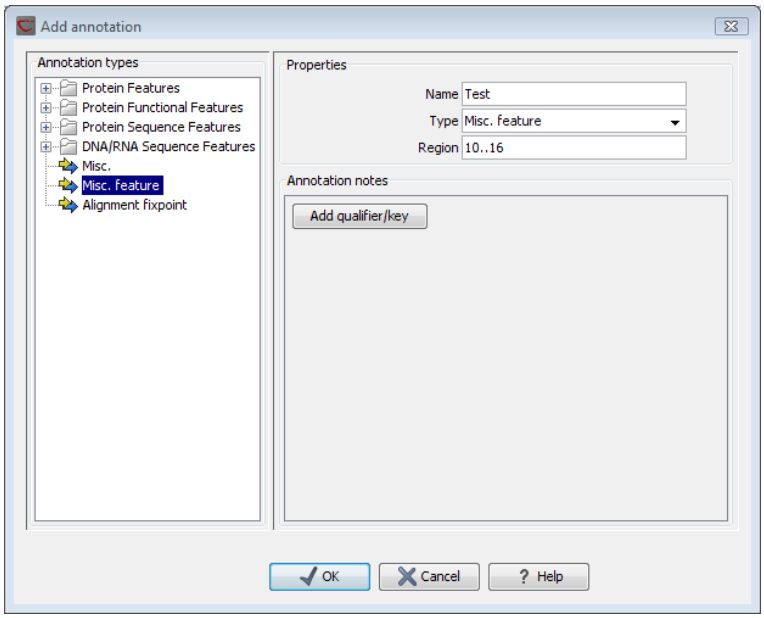


Figure 6.5: *The Add Annotation dialog.*

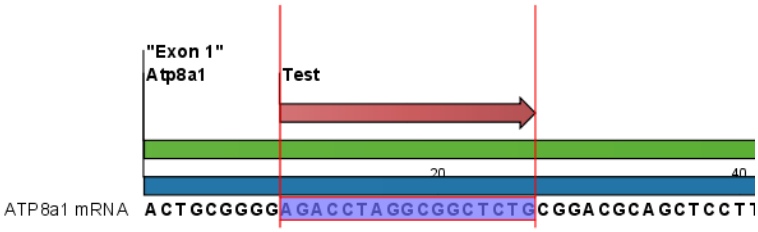


Figure 6.6: *The annotation has been added.*

Enter *Test* into the **Filter** in the annotation table, and your newly added annotation will appear.

You can now double-click any of the **Name**, **Type** or **Region** fields to edit the annotation. Double-click the **Type** which is now "Misc. feature" and enter *My type* and press **Enter**.

The table should now look as shown in figure 6.7.

In this way, you can quickly edit annotations. If you double-click the **Qualifiers** cell, a dialog similar to the one in figure 6.5 will be shown.

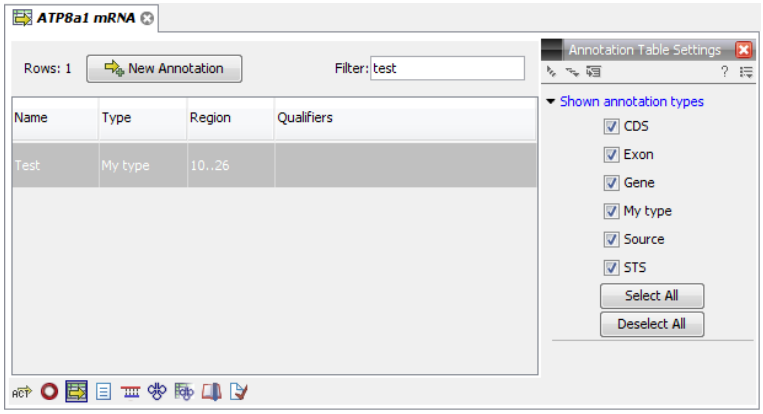


Figure 6.7: *The type of the annotation is now My type.*

6.3 Copying annotations

One of the more advanced features of the Workbench is the ability to copy and paste annotations. Annotations can be copied from one sequence to another or you can create a duplicate of the annotation by pasting it back on the same sequence. Because it is very easy to edit the region of the annotation, this can be a quick way of adding a number of similar annotations.

Select the *Test* **annotation in the table** | **Copy (****)** | **Paste (****)**

Double-click the region of the new annotation and type *30..46* and press **Enter**. You now have two annotations as shown in figure 6.8.

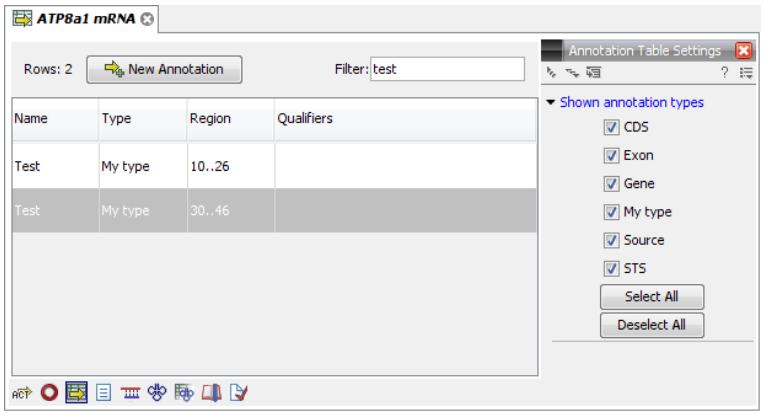


Figure 6.8: *Two copies of the Test annotation.*

Chapter 7

Tutorial: Align protein sequences

This tutorial outlines some of the alignment functionality of the *CLC Main Workbench*. In addition to creating alignments of nucleotide or peptide sequences, the software offers several ways to view alignments. The alignments can then be used for building phylogenetic trees.

Sequences must be available via the **Navigation Area** to be included in an alignment. If you have sequences open in a View that you have not saved, then you just need to select the view tab and press Ctrl + S (or ⌘ + S on Mac) to save them.

In this tutorial six protein sequences from the Example data folder will be aligned. (See figure 7.1).

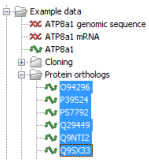


Figure 7.1: *Six protein sequences in 'Sequences' from the 'Protein orthologs' folder of the Example data.*

To align the sequences:

> **select the sequences from the 'Protein' folder under 'Sequences'** | **Toolbox** | **Alignments and Trees ( ) | Create Alignment ( )**

7.1 The alignment dialog

This opens the dialog shown in figure 7.2.

It is possible to add and remove sequences from **Selected Elements** list. Since we had already selected the eight proteins, just click **Next** to adjust parameters for the alignment.

Clicking **Next** opens the dialog shown in figure 7.3.

Leave the parameters at their default settings. An explanation of the parameters can be found by clicking the help button (?). Alternatively, a tooltip is displayed by holding the mouse cursor on the parameters.

Click **Finish** to start the alignment process which is shown in the **Toolbox** under the **Processes**

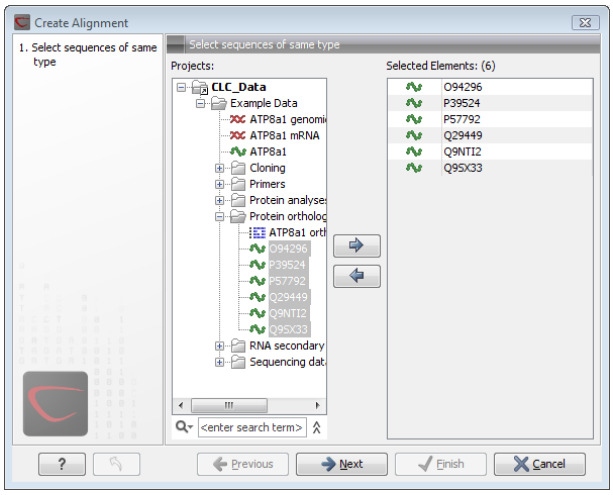


Figure 7.2: *The alignment dialog displaying the six protein sequences.*

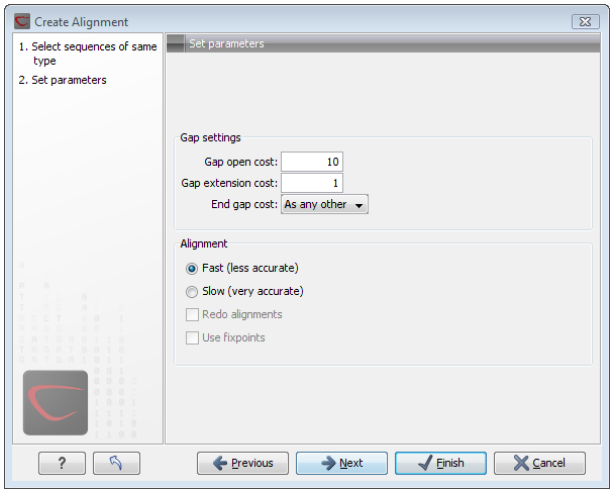


Figure 7.3: *The alignment dialog displaying the available parameters which can be adjusted.*

tab. When the program is finished calculating it displays the alignment (see fig. 7.4):

Note! The new alignment is not saved automatically.

To save the alignment, drag the tab of the alignment view into the **Navigation Area**.

Installing the Additional Alignments plugin gives you access to other alignment algorithms: ClustalW (Windows/Mac/Linux), Muscle (Windows/Mac/Linux), T-Coffee (Mac/Linux), MAFFT (Mac/Linux), and Kalign (Mac/Linux). The Additional Alignments Module can be downloaded from `http://www.clcbio.com/plugins`. Note that you will need administrative privileges on your system to install it.

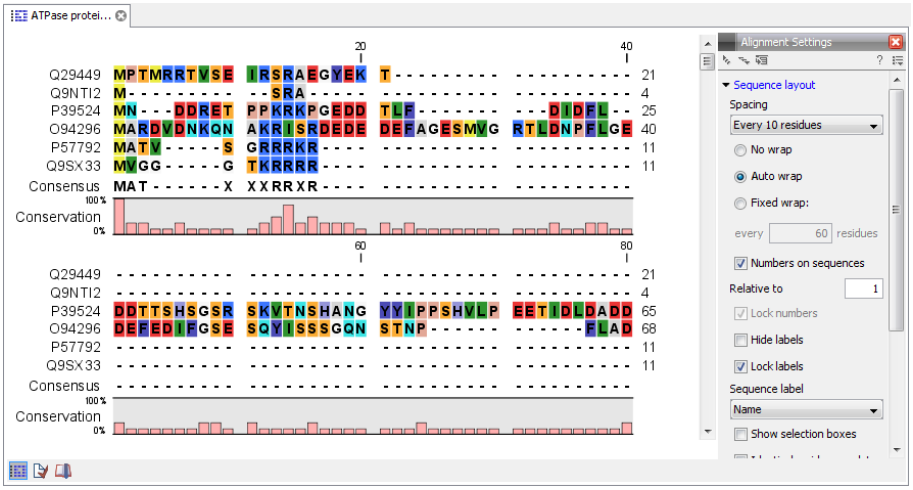


Figure 7.4: *The resulting alignment.*

Chapter 8

Tutorial: BLAST search

BLAST is an invaluable tool in bioinformatics. It has become central to identification of homologues and similar sequences, and can also be used for many other different purposes. This tutorial takes you through the steps of running a blast search in CLC Workbenches. If you plan to use blast for your research, we highly recommend that you read further about it. Understanding how blast works is key to setting up meaningful and efficient searches.

Suppose you are working with the ATP8a1 protein sequence which is a phospholipid-transporting ATPase expressed in the adult house mouse, *Mus musculus*. To obtain more information about this molecule you wish to query the peptides held in the Swiss-Prot* database to find homologous proteins in humans *Homo sapiens*, using the **Basic Local Alignment Search Tool** (BLAST) algorithm.

This tutorial involves running BLAST remotely using databases housed at the NCBI. Your computer must be connected to the internet to complete this tutorial.

8.1 Performing the BLAST search

Start out by:

select protein ATP8a1 | **Toolbox** | **BLAST Search (****)** | **NCBI BLAST (****)**

In **Step 1** you can choose which sequence to use as query sequence. Since you have already chosen the sequence it is displayed in the **Selected Elements** list.

Click **Next**.

In **Step 2** (figure 8.1), choose the default BLAST program: **blastp: Protein sequence and database** and select the **Swiss-Prot** database in the **Database** drop down menu.

Click **Next**.

In the **Limit by Entrez query** in **Step 3**, choose **Homo sapiens[ORGN]** from the drop down menu to arrive at the search configuration seen in figure 8.2. Including this term limits the query to proteins of human origin.

Choose to **Open** your results.

Click **Finish** to accept the parameter settings and begin the BLAST search.

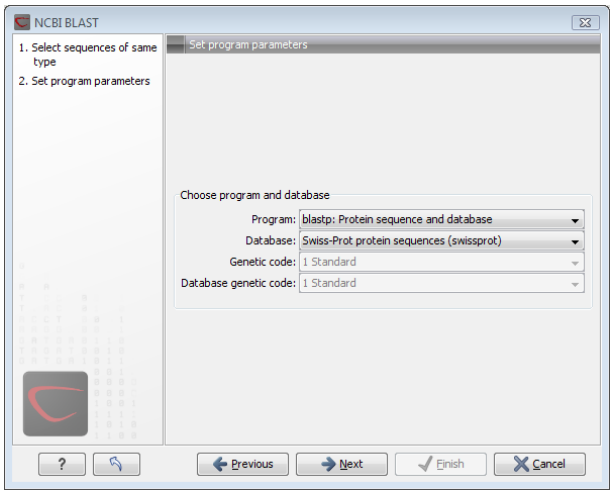


Figure 8.1: *Choosing BLAST program and database.*

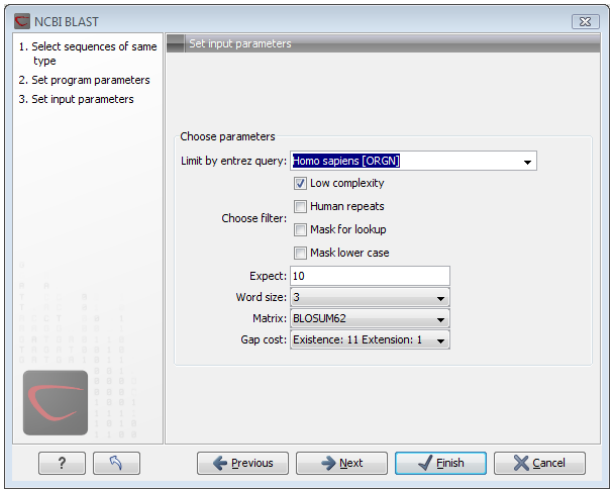


Figure 8.2: *The BLAST search is limited to homo sapiens[ORGN]. The remaining parameters are left as default.*

The computer now contacts NCBI and places your query in the BLAST search queue. After a short while the result should be received and opened in a new view.

8.2 Inspecting the results

The output is shown in figure 8.3 and consists of a list of potential homologs that are sorted by their BLAST match-score and shown in descending order below the query sequence.

Try placing your mouse cursor over a potential homologous sequence. You will see that a context box appears containing information about the sequence and the match-scores obtained from the BLAST algorithm.

The lines in the BLAST view are the actual sequences which are downloaded. This means that you can zoom in and see the actual alignment:

> **Zoom in in the Tool Bar ( ) | Click in the BLAST view a number of times until you see the residues**

Now we will focus our attention on sequence Q9Y2Q0 - the BLAST hit that is at the top of the

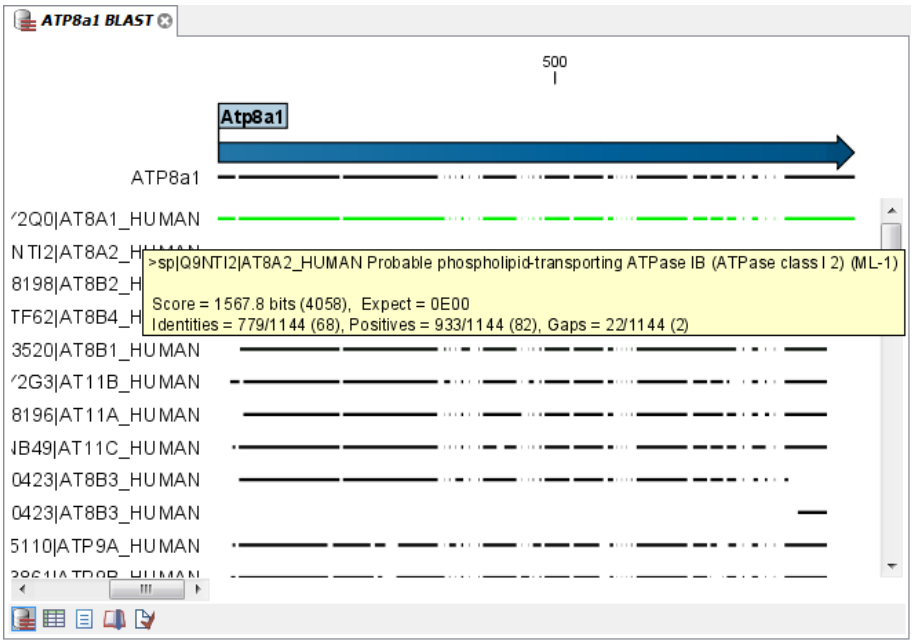


Figure 8.3: *Output of a BLAST search. By holding the mouse pointer over the lines you can get information about the sequence.*

list. To download the full sequence:

right-click the line representing sequence Q9Y2Q0 | Download Full Hit Sequence from NCBI

This opens the sequence. However, the sequence is not saved yet. Drag and drop the sequence into the **Navigation Area** to save it. This homologous sequence is now stored in the *CLC Main Workbench* and you can use it to gain information about the query sequence by using the various tools of the workbench, e.g. by studying its textual information, by studying its annotation or by aligning it to the query sequence.

8.3 Using the BLAST table view

As an alternative to the graphic BLAST view, you can click the Table View ( ) at the bottom. This will display a tabular view of the BLASt hits as shown in figure 8.4.

This view provides more statistics about the hits, and you can use the filter to search for e.g. a specific type of protein etc. If you wish to download several of the hit sequences, this is easily done in this view. Simply select the relevant sequences and drag them into a folder in the **Navigation Area**.

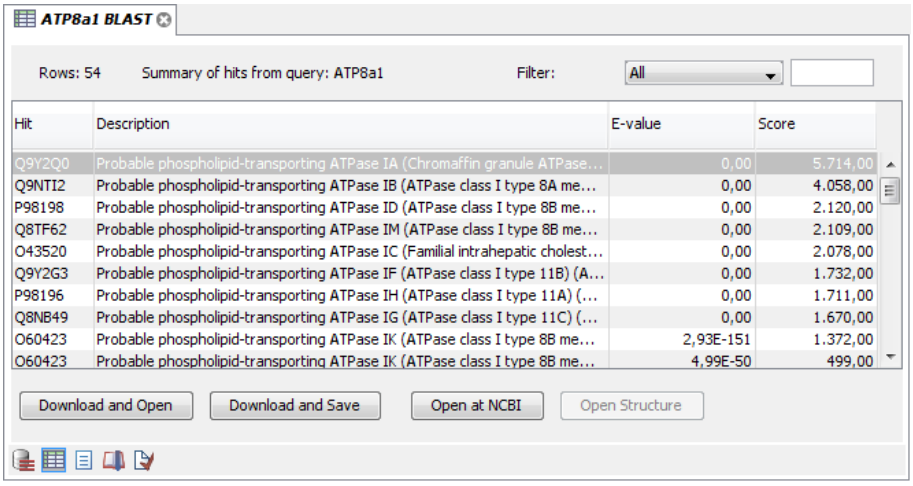


Figure 8.4: *Output of a BLAST search shown in a table.*

Chapter 9

Gateway Cloning

This tutorial will show you how to use the Workbench tool for in silico Gateway®Cloning. With this you can perform both standard (1 fragment) and mult-site Gateway Cloning. Here we focus on the 1-fragment cloning approach, but we provide tips to help you move onto multisite Gateway Cloning easily.

For information on the Gateway Cloning system please visit:

http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cloning/Gateway-Cloning.html.

After you have the Gateway Cloning vectors imported into your Workbench (section 9.1), you work through the three Gateway Cloning tools, which imitate the steps of the Gateway Cloning lab procedure. You will:

- Add attB sites to the fragment of interest, as would be done in the lab using PCR. (Section 9.2)
- Construct an entry vector by an attB/attP recombination event (BP reaction) in which the attB-flanked fragment is exchanged for the attP-flanked sequence of the chosen donor vector. (Section 9.3)
- Construct the final expression vector by an attL/attR recombination event (LR reaction) in which the target fragment of the entry vector is exchanged for the attR-flanked sequence of a chosen destination vector. (Section 9.4)

9.1 Importing Gateway Cloning vectors

1. To perform the in silico Gateway Cloning we first need to import donor and destination vectors. A file of these can be downloaded from the Invitrogen web site:

   http://tools.invitrogen.com/downloads/Gatewayvectors.ma4
2. Once downloaded, create a folder in your Workbench to import this file into.
3. Select this new folder by clicking on it.
4. Now, import the vector sequences into the Workbench using the **standard import tool. (****)**

The vector sequences should now be imported into your new folder.

For future reference, information on the vectors for MultiSite Gateway Cloning can be found on the Invitrogens web site: http://www.invitrogen.com/search/global/searchAction.action?query=MultiSite+Gateway%C2%AE+Vectors&resultPage=1&resultsPerPage=15&navigator=seproductcategorynavigator&modifier=DNA+Vectors%2FMultiSite+Gateway%26reg%3B+Vectors&seSeeallNavigator=true.

If you choose which type of MultiSite Cloning you wish to perform (number of fragments), then the relevant vectors can be listed. Vector information can then be downloaded from the linked areas on the site.

9.2 Adding attB sites

This step mimics the lab PCR procedure of adding attB sites to the target fragment making it eligible for recombination into the donor vector. The output can also include a set of primers, although we do not do that in this tutorial.

Our target fragment is the coding sequence (CDS) of ATP8a1, the mRNA of which is found in the standard example data you can download into the Workbench. If you do not already have the example data go to the **Help** menu in the Workbench, and choose to **Import Example Data**.

Now:

1. Extract the CDS sequence from the ATP8a1 mRNA file. To do this,
   - Double click on the sequence object called **ATP8a1 mRNA**. This opens it in the viewing area of the Workbench.
   - Right-click the CDS annotation, which by default will be coloured in yellow, and chose **Open annotation in New View**
   - Save the new data object that has just opened in the new view. There are different ways to do this. Two options are:
     (a) Right click on the tab at the top of the new view, and choose Save As... from the menu that appears, or
     (b) Move the cursor over the tab at the top of the new view, depress the left mouse button. Keeping it depressed, drag the tab over into a folder in the Navigation area of the Workbench. The data will then be saved into that folder.
2. Add the attB sites. To do this,
   - Start the **Gateway Cloning** tool:
     Toolbox | **Cloning and Restriction Sites** () | **Gateway Cloning** | **Add attB Sites** ()
   - In the wizard, select the ATP8a1 CDS fragment you just saved. When the wizard starts up, that fragment should be in the list of **Selected Elements**. If it is not, please select it in the wizard, and click on the right hand arrow to add it to the **Selected Elements** list.
   - Click on the button marked **Next**.

- It is in this step of the wizard you choose the attB sites to be added to your fragment. This is shown in figure 9.1. The sites to choose depend on your choice of donor vector. We will be going with the default option of attB1 and attB2. This ensures compatibility with our downstream choice of donor vector, which will be **pDONR221**.
- Click on the button marked **Next**.

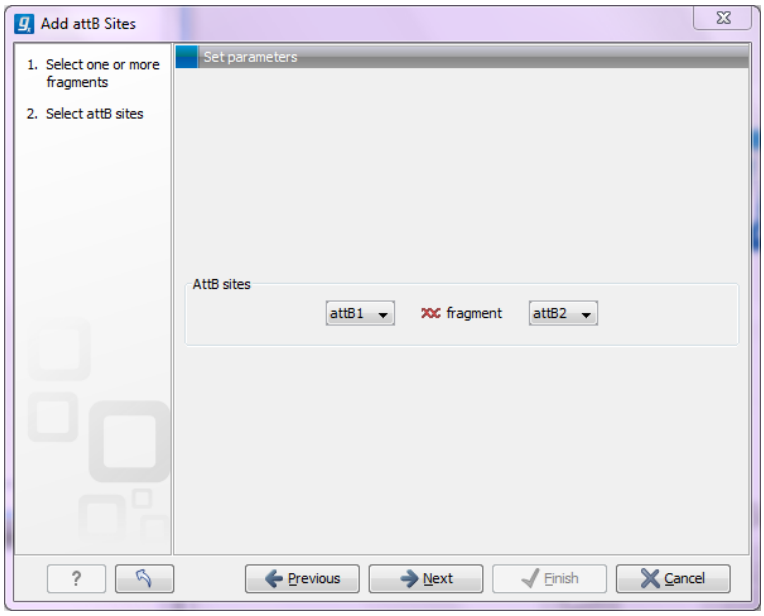


Figure 9.1: *Adding attB sites.*

If you have selected more than one fragment these will all be given the same combination of attB sites. If you wish to add different combination of sites you will have to run the tool once for each combination.

3. Next you will be given the option to insert additional elements downstream and upstream of the first and second attB site, respectively. That is, to insert elements immediately 5 prime and 3 prime of the fragment.

   We will add a Shine-Dalgarno site 5 prime of ATP8a1. Click in the field **Forward primer additions** and press on the keys Shift+F1 to bring up a list of elements you can choose from. Use either the mouse or the up and down arrow keys and enter to select the Shine-Dalgarno site. See figure 9.2.

4. Insert 8 random nucleotides after the Shine-Dalgarno sequence. This is done by simply typing in the bases. See figure 9.3. Doing this ensures that the spacing between the Shine-Dalgarno sequence and the initiation codon is correct.

   You will find a preview of the final PCR product below. Any preset additional elements appear in green whereas the bases typed in manually are displayed in blue.

5. Click on the button marked **Next**.

6. In this step, you can choose the length of the template-specific part of your primers. See figure 9.4.

   This template-specific fragment is joined with the attB site and any additional elements you choose to make up the primer. Here, we will just keep the defaults.

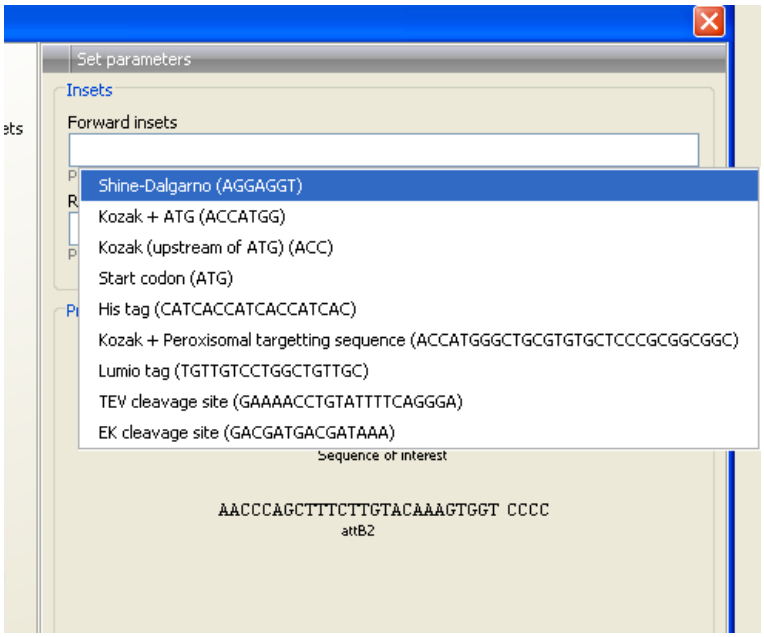


Figure 9.2: *Choose Shine-Dalgarno.*

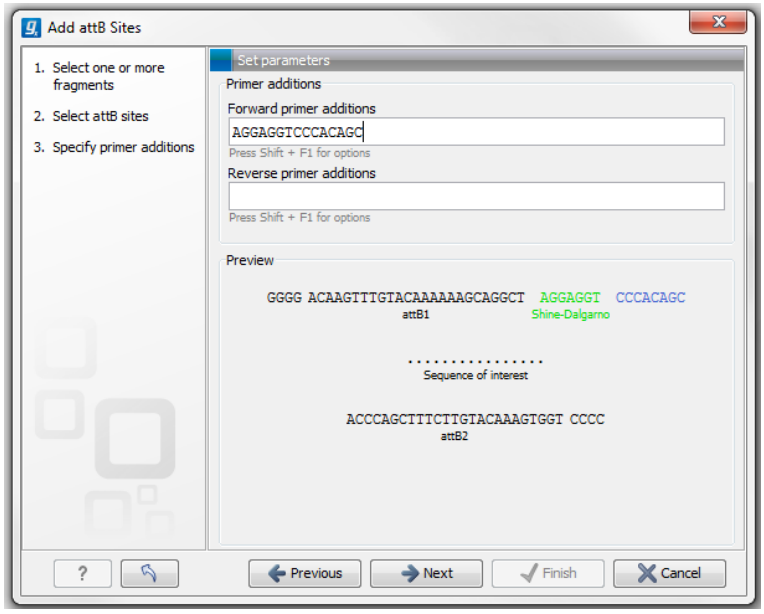


Figure 9.3: *Add bases to ensure correct spacing between the Shine-Dalgarno sequence and the initiation codon.*

7. Click on the button marked **Next**.
8. Choose to **Save** the results. Uncheck the **create primer list** option as we will not be needing this in this tutorial. See figure 9.5.
9. Click on the button marked **Next** and choose where to **Save** the output.
10. Open up the result **ATP8a1 CDS (attB1 attB2)** to have a look. See figure 9.6.
11. In the side panel called **Sequenc Settings**, find the section called **Annotation types**. Click on it to expand it.

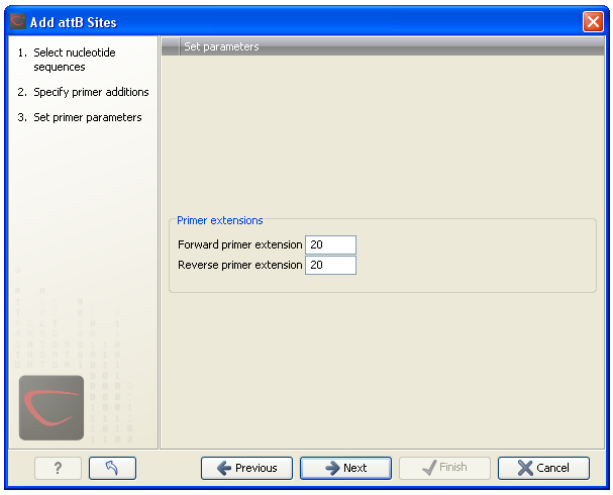


Figure 9.4: *Set the appropriate primer lengths.*

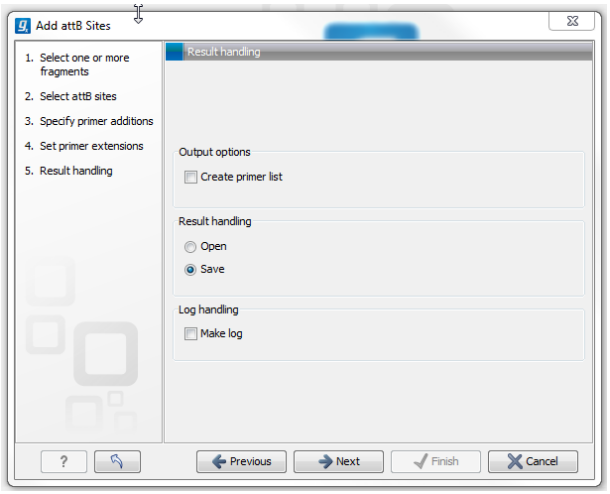


Figure 9.5: *Choose to save your results.*

12. Click on the box beside **RBS** to mark it, so that the Shine-Dalgarno ribosomal binding site annotation is displayed. Note also that you can see annotations for primer regions and the attB sites.

9.3 Creation of an entry vector

This step recombines the attB-flanked ATP8a1 into the donor vector, which is **pDONR221**, resembling the BP-reaction carried out in the lab.

1. Choose **Create Entry Clone** from under the Gateway Cloning menu:

   Toolbox | **Cloning and Restriction Sites** ( ) | **Gateway Cloning** | **Create Entry Clone** ( )

2. In the dialog select the result that you generated earlier: **ATP8a1 CDS (attB1 attB2)**. Then click on the button marked **Next**.

3. Click the **Browse button** ( ) and find the Gateway folder of vectors, which was previously imported.

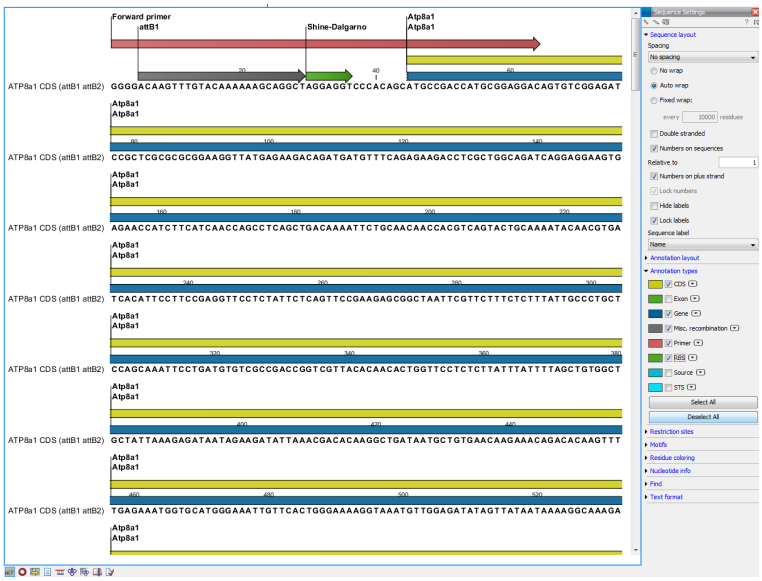


Figure 9.6: *A view of the results: ATP8a1 CDS (attB1 attB2).*

4. Select donor vector **pDONR221** as shown in figure 9.7 and click on the button labelled **OK**.

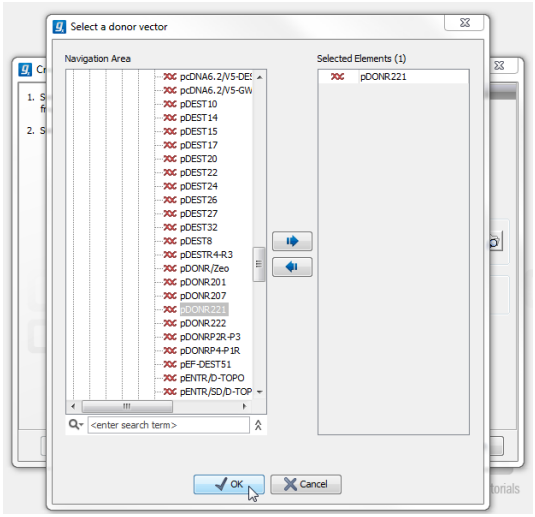


Figure 9.7: *Select the donor vector, which in this case is pDONR221.*

5. Click on the button labelled **Next**.

   Note: The Workbench only checks the chosen entry clone for valid attP sites. It does **not** check that these are compatible with the attB sites present in the fragment. You need to make sure that a valid entry vector is used for the attB-fragment of interest. If the right combination of attB and attP sites is not found, no entry clones will be produced.

6. Choose to **Save** the result and click on the button labelled **Next** to be able to specify where to save the results to.

7. Click on the button labelled **Finish**.

8. Open the output **pDONR221 (ATP8a1 CDS (attB1 attB2))** to have a look at it. To see a view like that shown in figure 9.8:

- Go to circular view ( ).
- Check annotation type **RBS** in the side panel to get a better view. The attB/attP recombination event has resulted in the fragment of interest being inserted into the vector. The fragment is flanked by attL sites.

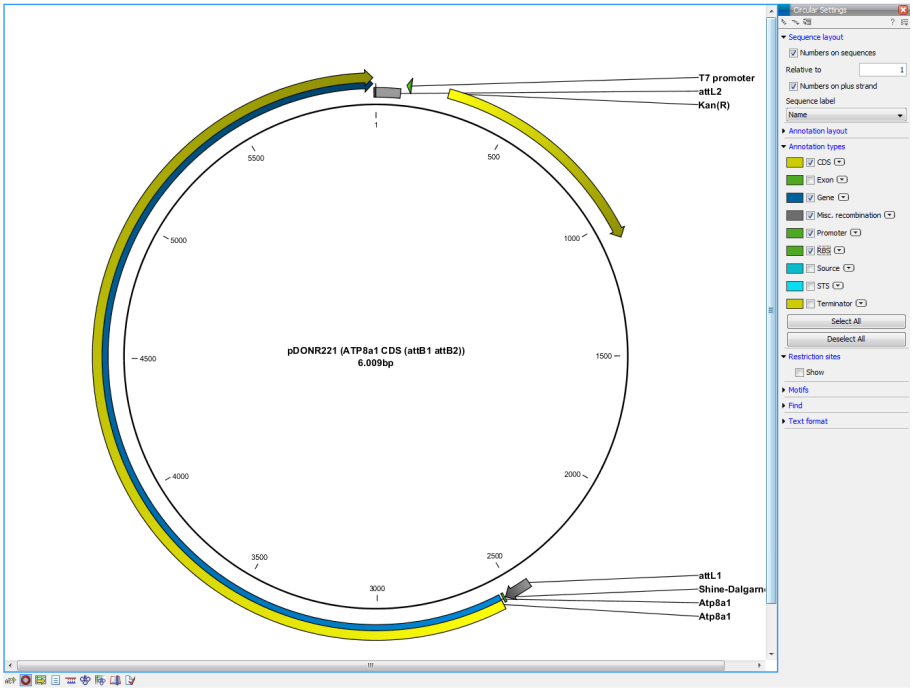


Figure 9.8: *View the donor vector, pDONR221, after the entry clone is createdXXX.*

If several fragments were selected for input (using the batch option), the output would be one entry clone per fragment, all based on the same donor vector.

9.4 Creating an expression vector (LR)

This last step in a Gateway Cloning workflow mimics the lab LR recombination reaction. That is, the fragment of interest is recombined from the entry vector into a destination vector (the LR reaction) to create the final expression vector.

In this tutorial we are carrying out a standard, single fragment Gateway Cloning task. Thus we will choose only one entry clone as input during this step. For MultiSite Gateway Cloning, we would choose all the entry clones to be joined in the final expression vector. Note though that if you do this, you need to ensure that the entry clones contain a valid combination of attL sites. For more information on this please have a look at the MultiSite Gateway explanation on the Invitrogen webpage:

http://tools.invitrogen.com/downloads/gateway-multisite-seminar.html

The steps to create an expression clone are:

1. Choose **Create Expression Clone** from under the Gateway Cloning menu:

   Toolbox | **Cloning and Restriction Sites** ( ) | **Gateway Cloning** | **Create Expression Clone** ( )

2. In the dialog box select the entry clone you created in the previous step and click on the button marked **Next**.

3. In the next window, click the **Browse** button and navigate to the Gateway vector folder. Select the destination vector: **pDEST14** and click on the button labelled **OK**.

4. Click on the button labelled **Next**.

   The Workbench only checks for valid attR sites in the destination vectors. You, yourself, need to make sure that the attL and attR sites are compatible. If not, no expression clone will be created.

5. In the next window choose to open the result. If you would like to also have a look at the byproduct, check the corresponding box. Click on the button labelled **Finish**.

The output is the result of the recombination between the attL and attR sites. The fragment of interest is now inserted in the expression vector flanked by attB sites, see figure 9.9

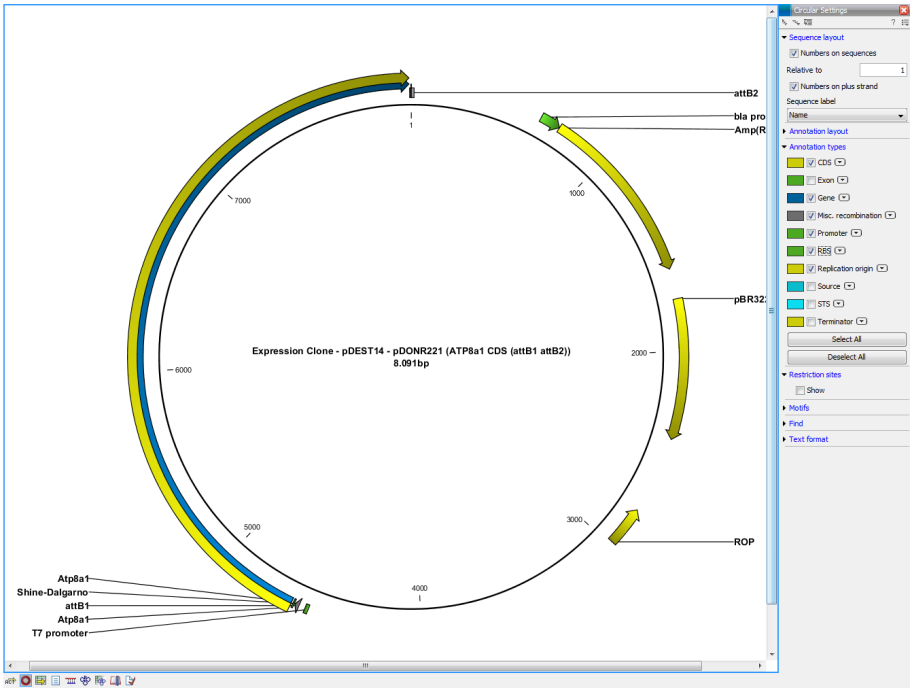


Figure 9.9: *View the finished expression vector.*

This concludes the Gateway Cloning tutorial. Below is a section outlining the general steps you could take in the case of MultiSite Gateway Cloning.

9.5 Short suggestion for a MultiSite Gateway Workflow using 2 fragments

The vectors mentioned in this section can be downloaded from Invitrogens web page:

http://tools.invitrogen.com/downloads/Gatewayvectors.ma4

1. Choose two fragments, which you would like to have in the final vector.

2. Create the following fragments:
    - Add attB1 and attB5r sites to sequence of interest 1 –> sequence of interest 1 (attB1 attB5r)
    - Add attB5 and attB2 sites to sequence of interest 2 –> sequence of interest 2 (attB5 attB2)
3. Create two entry vectors:
    - sequence of interest 1 (attB1attB5r) + pDONR221-P1P5r
    - sequence of interest 2 (attB5 attB2) + pDONR221-P5P2
4. Create the final expression vector by selecting:
    - the first entry clone: pDONR221-P1P5r (fragment of interest 1(attB1 attB5r)), and
    - the second entry clone: pDONR221-P1P5r (fragment of interest 1(attB1 attB5r))
    - the destination vector: pDEST14

CLC Academy