Introduction to Bioinformatics
4th practice
Multiple alignment and visualisation

Necessary files:

- Input.fasta (input file for multiple alignments)
- result.clustalw_aln (result of the 3D T-Coffee Expresso)
- result_pdb1.template_list (result of the 3D T-Coffee Expresso)

In order to avoid overloading of the online servers, the students will be divided into 4 groups designed with letters A-D. The groups will use different servers for the multiple alignment tasks.

1. Multiple Alignment

This practice focuses on the multiple alignment of sequences. We will get familiar with some multiple alignment programs using 4 protein sequences of proteins associated with quorum sensing in bacteria.

1.1. ClustalW

This program is one of the widely used tools for aligning multiple sequences and constructing a phylogenetic tree based on the alignment.

I. Multiple alignment with ClustalW

Group A

- Follow the link:

http://www.genome.jp/tools/clustalw/

and upload Input.fasta to the server with the “Browse” buttons, then click “Execute Multiple alignment”.

- You will see an output of the results. At the first section, you can see a general logging of the output. The second section is the alignment. Right click the filename “clustalw.aln” to save the alignment to your computer. Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). What do you see in the multiple alignment? Are there many gaps?

Group B

- Follow the link:

http://embnet.vital-it.ch/software/ClustalW.html

Copy the content of Input.fasta into the input field.
Click “Run ClustalW”.

- After some time you will see a table with two columns: the first is the alignments in various formats, the second is the tree. From the first column, right click the first file ending with “aln” to save it to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). What do you see in the multiple alignment? Are there many gaps?

Group C

- Follow the link:

http://clustalw.ddbj.nig.ac.jp/

- In the “Sequences” section, choose “Protein” from the list (the default is DNA).
- Upload Input.fasta with the “Browse” button. At the “Submit” section click the “Send to ClustalW” button.
- After some time you will see the output. The first section is a general report and log. The second section is the alignment. Click the button “Download Alignment file” to save it to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). What do you see in the multiple alignment? Are there many gaps?

Group D

- Follow the link:

https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html
Copy the content of Input.fasta into the input field.

- At Pairwise alignment type choose “Fast” from the list (the default is slow). Click the “Submit” button. Here the program allows you to make predictions and viewing options, but in this practical we are not using these.
- At the middle of the page you can see the colorful multiple alignment. At the top of the “Garnier parameters” section, there is a text “Result files (text) : ” righ tclick the “ClustalW” link below that to save the alignment to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). What do you see in the multiple alignment? Are there many gaps?

1.2. T-Coffee

T-Coffee is also a very frequently used program to make multiple alignments of sequences. We will use the same sequences as in the case of ClustalW.
Note that generally this calculation will take more time than the one with ClustalW.

II. Multiple alignment with T-Coffee

Group A

- Follow the link:

http://tcoffee.vital-it.ch/apps/tcoffee/do:regular
Upload “Input.fasta” by clicking the link “Click here to upload a file” using a button or dragging and dropping the file. Click “Select”.
You can give an e-mail address to the system so it can send you an e-mail if the alignment is ready. It is more useful if you have more and longer sequences, because in these cases the alignment can take much time.

- The results are shown as a colorful alignment. At the “Result files” section right click the “clustalw_aln_file” in the “Multiple Alignment” line to save the alignment to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). Is it different from the one obtained with ClustalW?

Group B

- Follow the link:

http://tcoffee.crg.cat/apps/tcoffee/do:regular
Upload “Input.fasta” by clicking the link “Click here to upload a file” using a button or dragging and dropping the file. Click “Select”.

You can give an e-mail address to the system so it can send you an e-mail if the alignment is ready. It is more useful if you have more and longer sequences, because in these cases the alignment can take much time.

- The results are shown as a colorful alignment. At the “Result files” section right click the “clustalw_aln_file” in the “Multiple Alignment” line to save the alignment to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). Is it different from the one obtained with ClustalW?

Group C

- Follow the link:

https://toolkit.tuebingen.mpg.de/t_coffee
Upload the Input.fasta with the “Browse” button.
Click the “Submit job” button.

- The result alignment will appear on your screen. Click the “Save” button at the top of the page to save the clustalw formatted alignment to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). Is it different from the one obtained with ClustalW?

Group D

- Follow the link:

http://www.ebi.ac.uk/Tools/msa/tcoffee/
Upload the Input.fasta with the “Browse” button.
Click the “Submit job” button.

- The result alignment will appear on your screen. Click the “Download alignment file” button to save the alignment in ClustalW format to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). Is it different from the one obtained with ClustalW?

1.3. Muscle

Muscle is another widely used multiple alignment program.
Here we also use the same input.

III. Multiple alignment with Muscle

Group A and B

- Follow the link:

http://www.ebi.ac.uk/Tools/msa/muscle/
Upload the Input.fasta with the “Browse” button.
Click the “Submit” button.

- The result alignment will appear on your screen. Click the “Download alignment file” button to save the alignment in ClustalW format to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). Is it different from the ones obtained with T-Coffe or ClustalW?

Group C and D

- Follow the link:

https://toolkit.tuebingen.mpg.de/muscle
Upload the Input.fasta with the “Browse” button.

In the “Select output format” choose “ClustalW”.
Click the “Submit job” button.

- The result alignment will appear on your screen. Click the “Save” button at the top of the page to save the clustalw formatted alignment to your computer.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). Is it different from the one obtained with ClustalW or T Coffee?

1.4. Expresso

Expresso is a 3D extension of T Coffee, which uses structural information in the alignment. It searches a template structure for each sequence and aligns the structures themselves, not the sequences as it is in the other programs.

IV. Multiple Alignment with Expresso

In order not to overload the server, you are provided with the output of the multiple alignment:

- the result.clustalw_aln file containing the resulting alignment
- and the result_pdb1.template_list file containing the resulting template structures for each sequence.
- Visualise the alignment file with MViewer (see section 2.1) and JalView (see section 2.2). What is the difference between this alignment and the other alignments?

The following webserver was used for the calculation:
http://tcoffee.crg.cat/apps/tcoffee/do:expresso

2. Visualisation of the alignments

Some of the multiple alignment programs also makes visualisation of alignments such as coloring, although some of them does not have this feature. There are programs dedicated to this tasks. In this section of the practical we will get familiar with them.

2.1. MSA Viewer

This is an online web applet for visualising multiple sequence alignments.

- Follow the link:

http://msa.biojs.net/app/

You can load alignment files with the “Import” button.

- The program displays your alignment with colors, each amino acid having a different color. At the top of the alignment there is a profile of the sequence with the size of the letters corresponding to the consensus. The alignment might be too long to display on your screen. In this case you can scroll the alignment with the gray bar at the top of it. You can export the resulting alignment with the “Export” button in various formats, such as a png image.

2.2. Jalview

Installing:

Linux

On linux you have to install the icedtea-8-plugin package.
You can do this with the Aptitude package manager with the following commans:
sudo apt-get install icedtea-8-plugin
than typing your sudo password.

After that, follow the link:
http://www.jalview.org/
Click the “Launch Jalview Desctop” button.
Click “Open with” and select “Iced tea Java Web Start (default)” if anything else is selected.
Click “Ok”.
You can untick “Always trust content of this publisher” if you want, but it is not necessary.
Click “Run”.
Click “Proceed”.
You can click “No” to gathering usage statistics, but “Yes” also suffices.

Windows
You have to have Java 1.6 or newer to run this application both JRE and JDK.
Follow the link:
http://www.jalview.org/
Click the “Launch Jalview Desctop” button.
Save the file and open it.
Click “Run”.
You can click “No” to gathering usage statistics, but “Yes” also suffices.

Using the program:

- You can close the “News from http://www.jalview.org” window inside the main window.
- Import sequences with the uppermost File menu -> Input alignment -> From file.
- In the window showing the alignment click the “color” menu and select “Taylor”.
- You can export the colored alignment as a PNG image with the File menu of the alignment window -> Export image -> PNG option.
- You can associate structures by right clicking the sequence name and clicking 3D Structure data. In this case it automatically fetches a corresponding 3D structure if it exists. By selecting one of the structures, click “View”. In the JMol viewer you can see this structure. In the colour menu of the JMol you can choose the same color schemes as in the sequence window.