H3ABioNet

Bioinformatics Training Unit

H3ABioNet

Pan African Bioinformatics Network for H3Africa

Introduction to Bioinformatics online course: IBT

Multiple Sequence Alignment

Lec6:Selecting and editing sequences

By

Ahmed Mansour Alzohairy

Professor (Full) at Department of Genetics,
Zagazig University, Zagazig, Egypt

Selecting and editing sequences

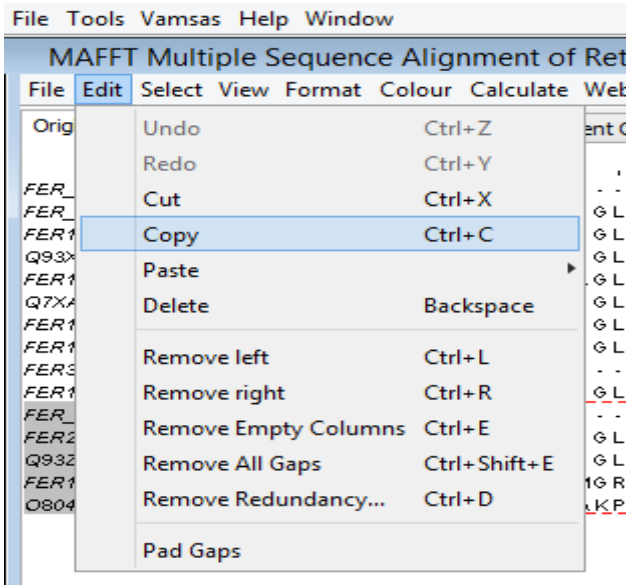
Selecting parts of an alignment

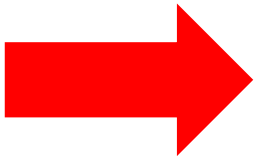
Selections can be of **arbitrary regions** in an alignment, one or more **complete columns**, or one or more complete sequences.

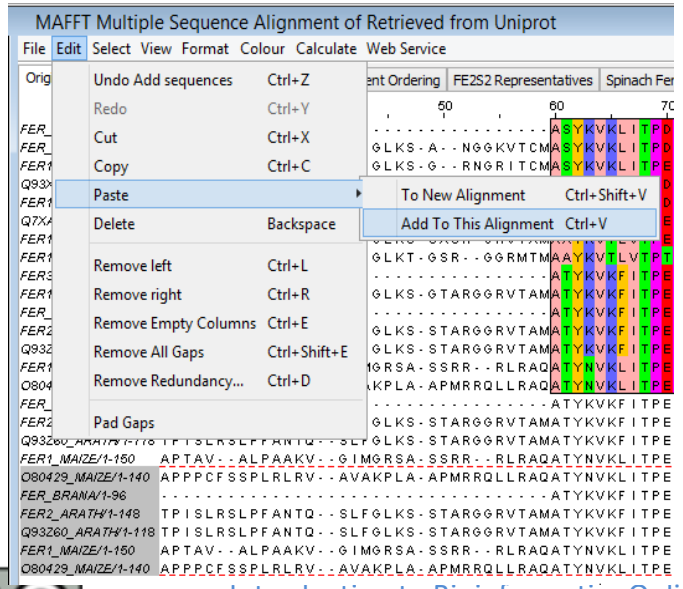
A **selected region** can be copied and pasted as a new alignment using the:

Edit ⇒ Copy* and *Edit ⇒Paste ⇒ To New Alignment

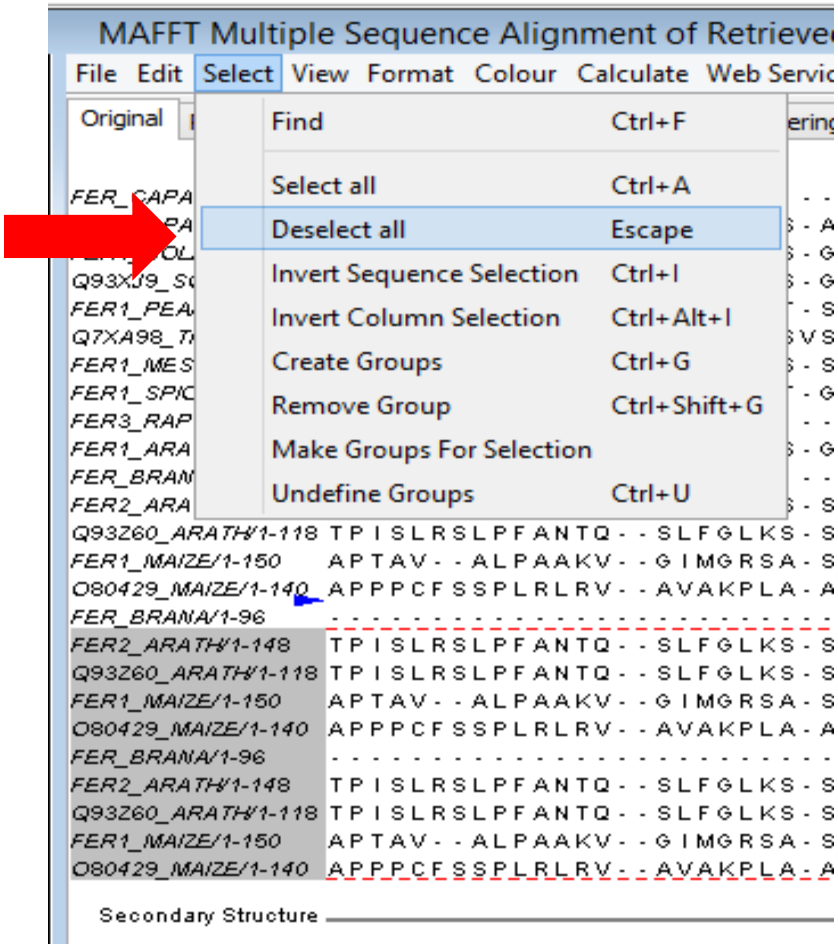
in the alignment window menu options..





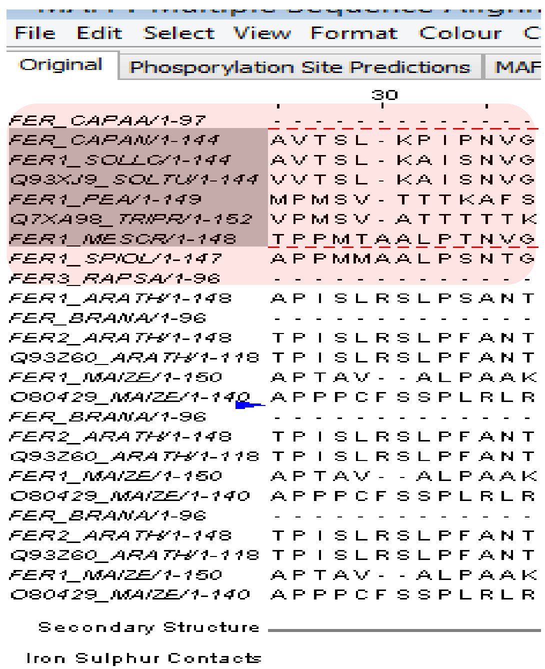


To clear (Deselect) the selection press the [ESC] (Escape) key



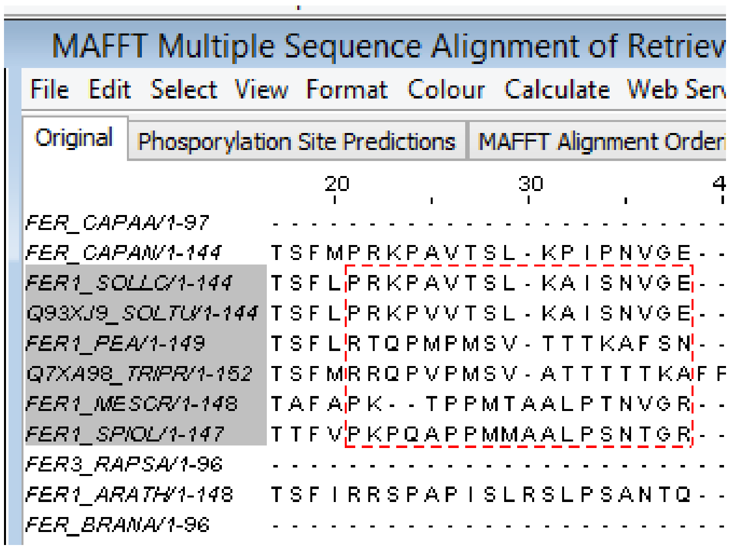
Selecting arbitrary regions

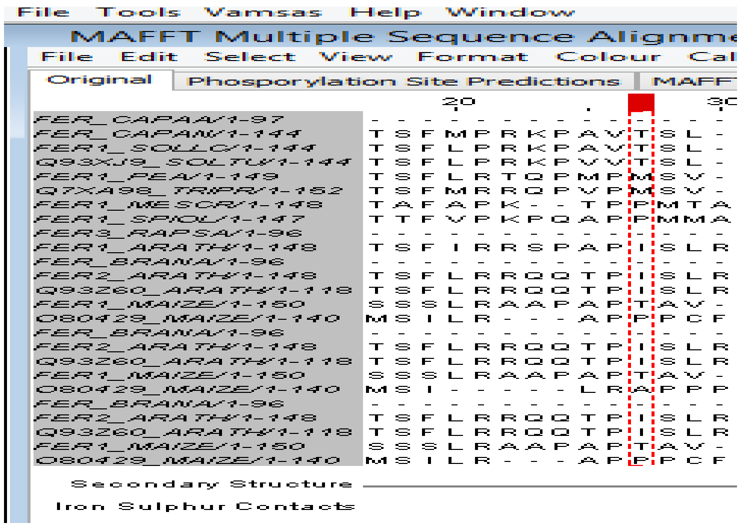
To select part of an alignment, place the mouse at the **top left corner** of the region you wish to select. **Press and hold** the mouse button and **drag** the mouse to the **bottom right corner** of the chosen region then release the mouse button. A dashed red box appears around the selected region .**Selecting a region in an alignment**



Selecting columns

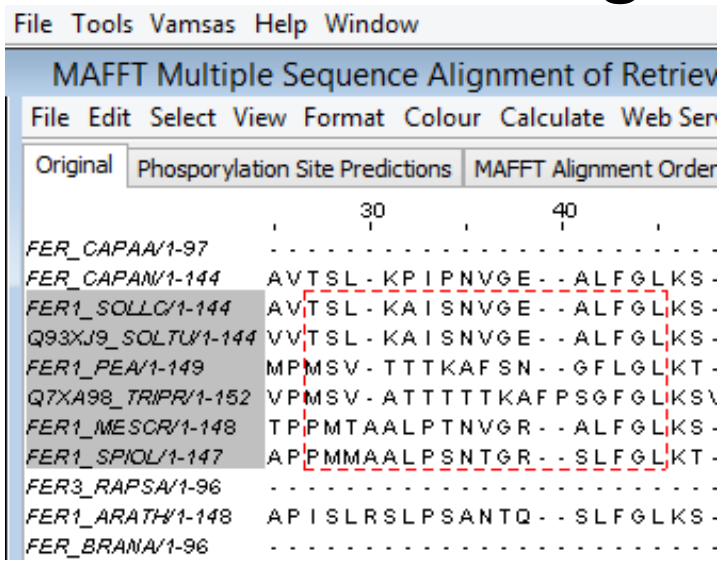
- To select the **same residues** in **all sequences**, **click and drag along** the alignment **ruler**. This selects the entire height of the alignment.
- **Ranges of positions** can also be selected by **clicking on the first** position then **holding down** the **[SHIFT]** key whilst **clicking the other end** of the selection.
- **Discontinuous regions** can be selected by **holding down** [**CTRL]** and **clicking on positions to add** to the column selection. Note that each **[CTRL]** Selected columns are indicated by red highlighting in the ruler bar





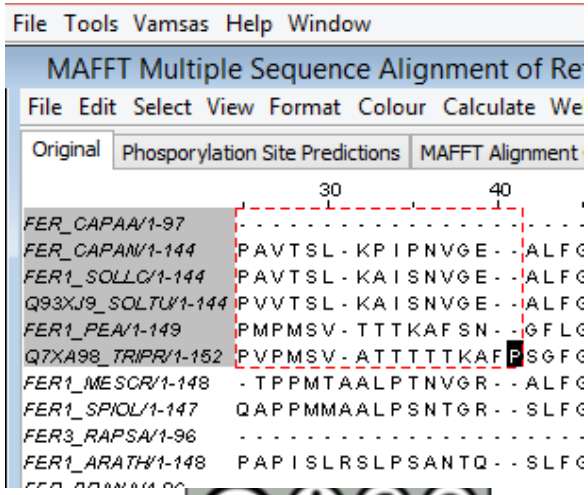
Selecting sequences

To select **multiple complete sequences,** click and **drag** the mouse **down** the sequence **ID panel**. The same technique as used for columns above can be used with **[SHIFT]-**Click for **continuous** and **[CTRL]-**Click to select **discontinuous** ranges of sequences.



Making selections in Cursor mode

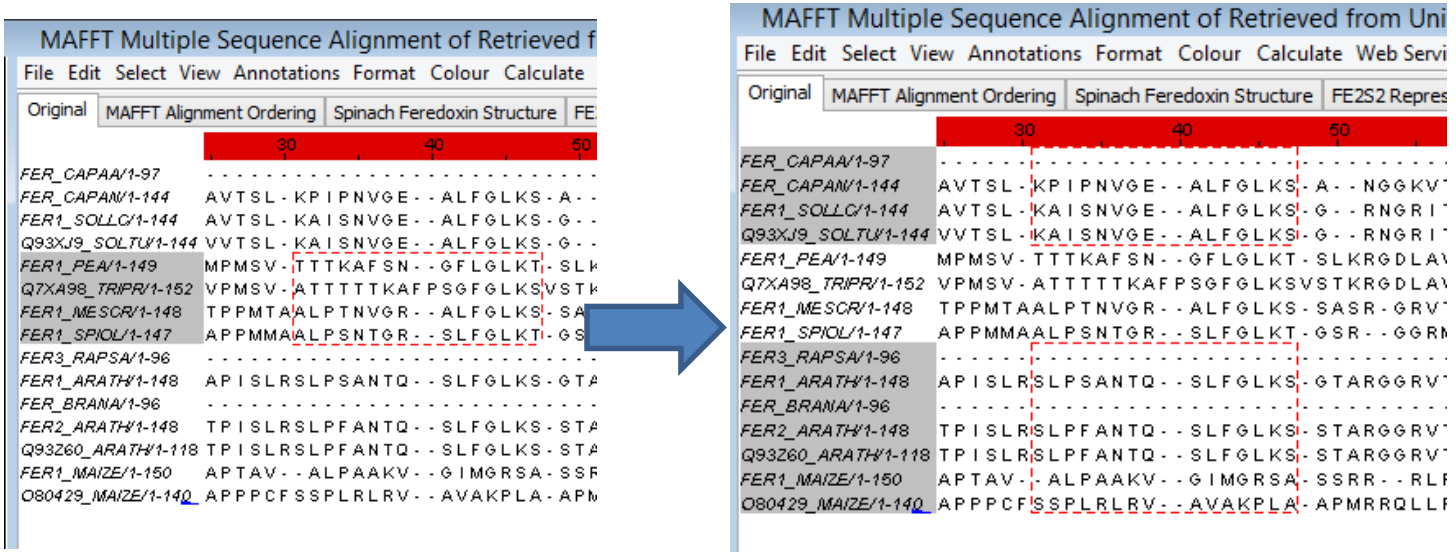
To define a selection in cursor mode (which is enabled by pressing **[F2]** when the alignment window is selected), navigate to the top left corner of the proposed selection (using the mouse, the arrow keys, or the keystroke commands described). Pressing the **[Q]** key **marks** this as the **corner**. A red outline appears around the cursor . Navigate to the **bottom** right corner of the proposed selection and press the **[M]** key. This marks the bottom right corner of the selection. The selection can then be treated in the same way as if it had been created in normal mode.



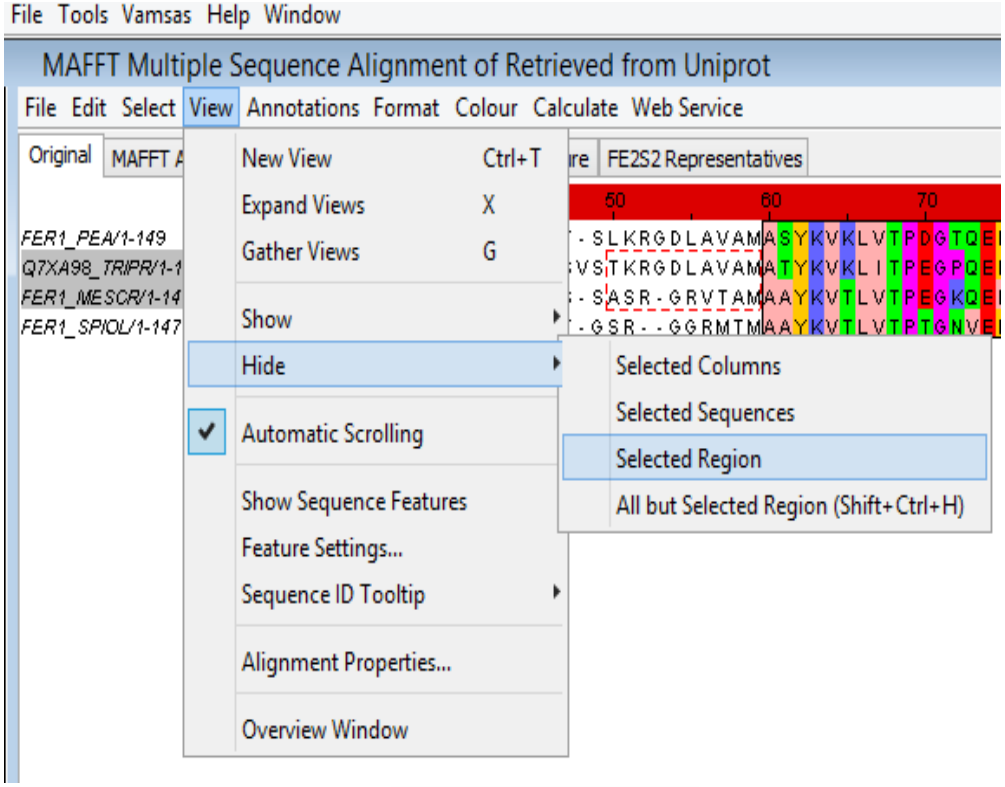
Inverting the current selection

The current sequence or column selection can be inverted, using ***Select ⇒ Invert Sequence/Column Selection*** in the alignment window.

Inverting the selection is useful when selecting large regions in an alignment, simply select the region that is to be kept unselected, and then invert the selection.



This may also be useful when hiding large regions in an alignment. Instead of selecting the columns and rows that are to be hidden, simply select the region that is to be kept visible, invert the selection, then select ***View ⇒ Hide ⇒ Selected Region .***



Creating groups

Selections are lost as soon as a **different region is selected**. Groups can be created which are **labeled regions** of the alignment.

To create a group,

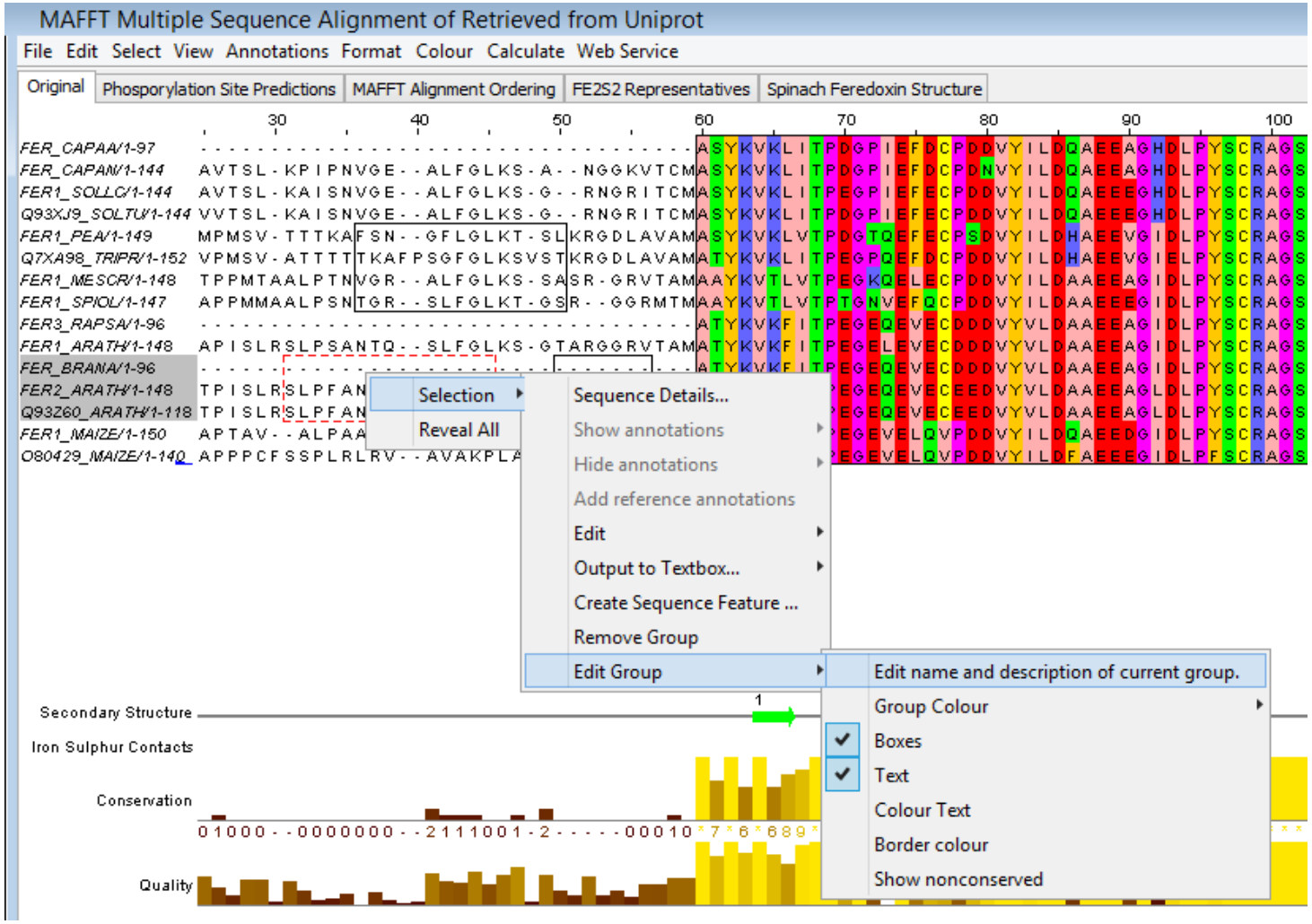
- **first select the region** which is to comprise the group.
- **Then click the right mouse** button on the selection to bring up a context menu.

Select ***Selection ⇒ Group ⇒ Edit***

- *name and description of current group*15
- then enter a name for the group in the dialogue box which appears.

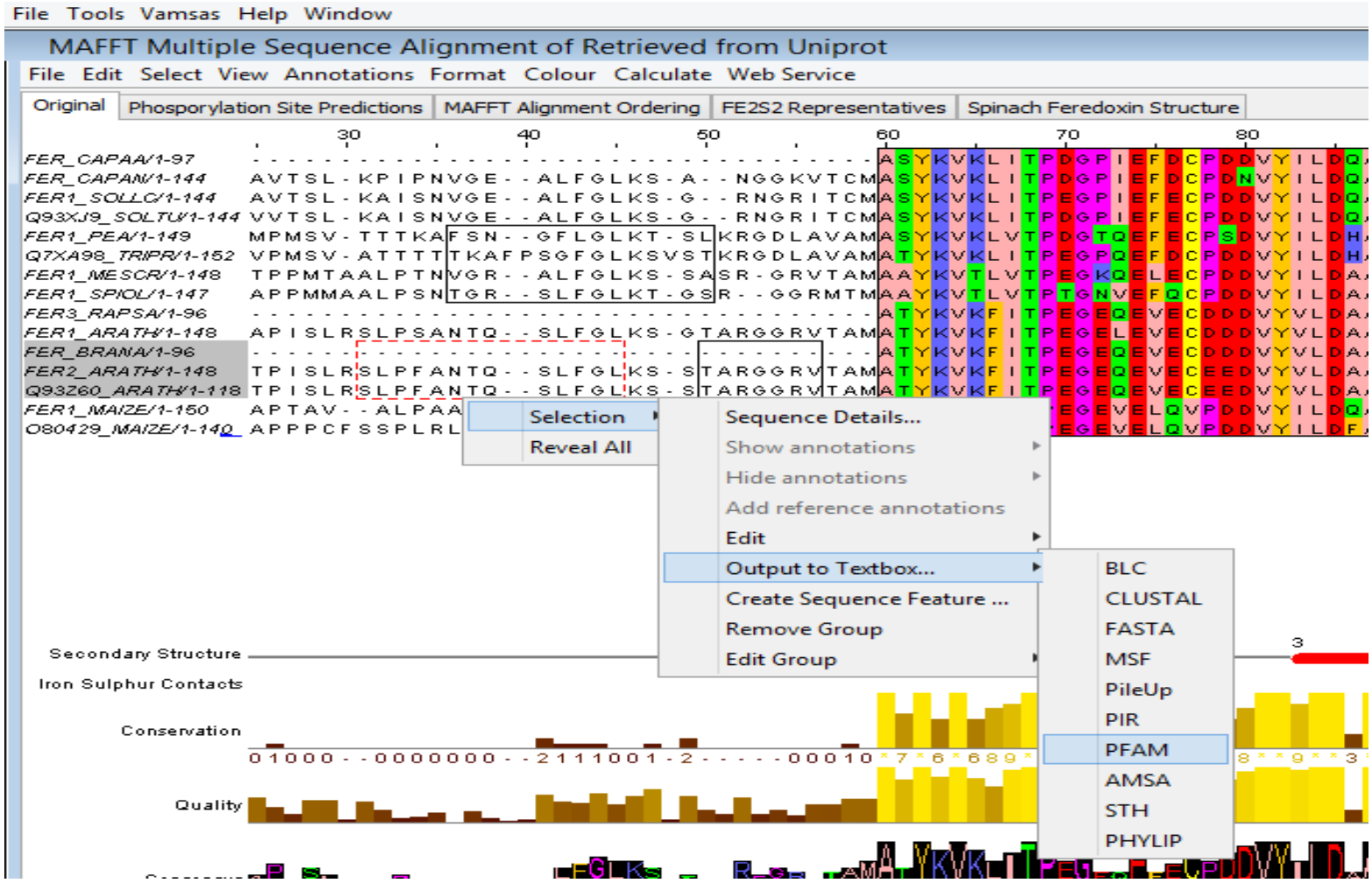
By default the new group will have a box drawn around it. The appearance of the group can be changed .

This group will stay defined even when the selection is removed.


MAFFT Multiple Sequence Alignment of Retrieved from Uniprot
File Edit Select View Annotations Format Colour Calculate Web Service
Original
Phosporylation Site Predictions
MAFFT Alignment Ordering
FE2S2 Representatives
Spinach Feredoxin Structure
30
40
50
60
70
80
90
100
FER_CAPAA/1-97
FER_CAPAN/1-144
FER1_SOLLC/1-144
Q93XJ9_SOLTU/1-144
FER1_PEA/1-149
Q7XA98_TRIPR/1-152
FER1_MESCR/1-148
FER1_SPIOL/1-147
FER3_RAPSA/1-96
FER1_ARATH/1-148
FER_BRANA/1-96
FER2_ARATH/1-148
Q93Z60_ARATH/1-118
FER1_MAIZE/1-150
O80429_MAIZE/1-140
Selection
Reveal All
Sequence Details...
Show annotations
Hide annotations
Add reference annotations
Edit
Output to Textbox...
Create Sequence Feature ...
Remove Group
Edit Group
Edit name and description of current group.
Group Colour
Boxes
Text
Colour Text
Border colour
Show nonconserved
Secondary Structure
Iron Sulphur Contacts
Conservation
Quality

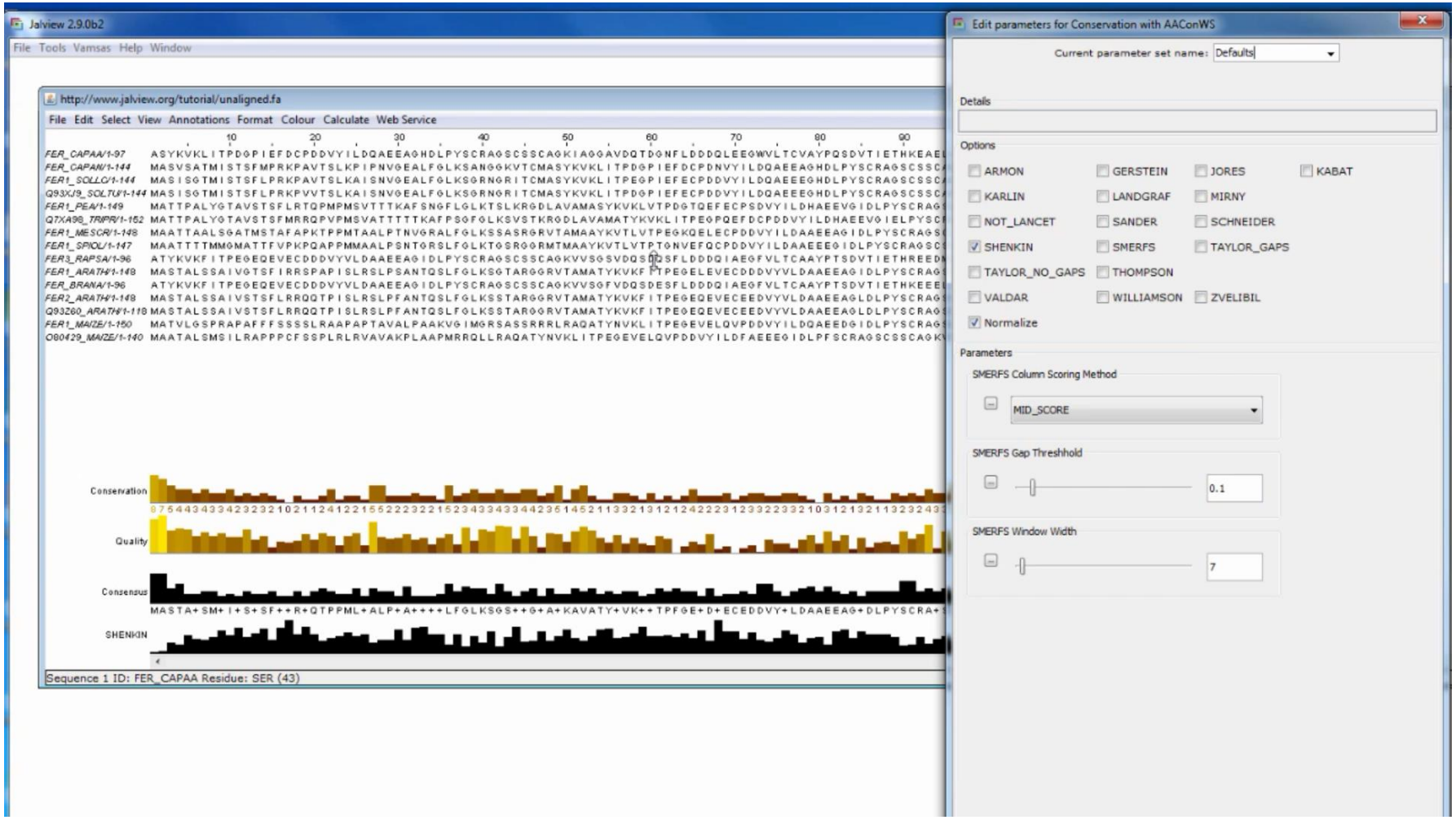
Exporting the current selection

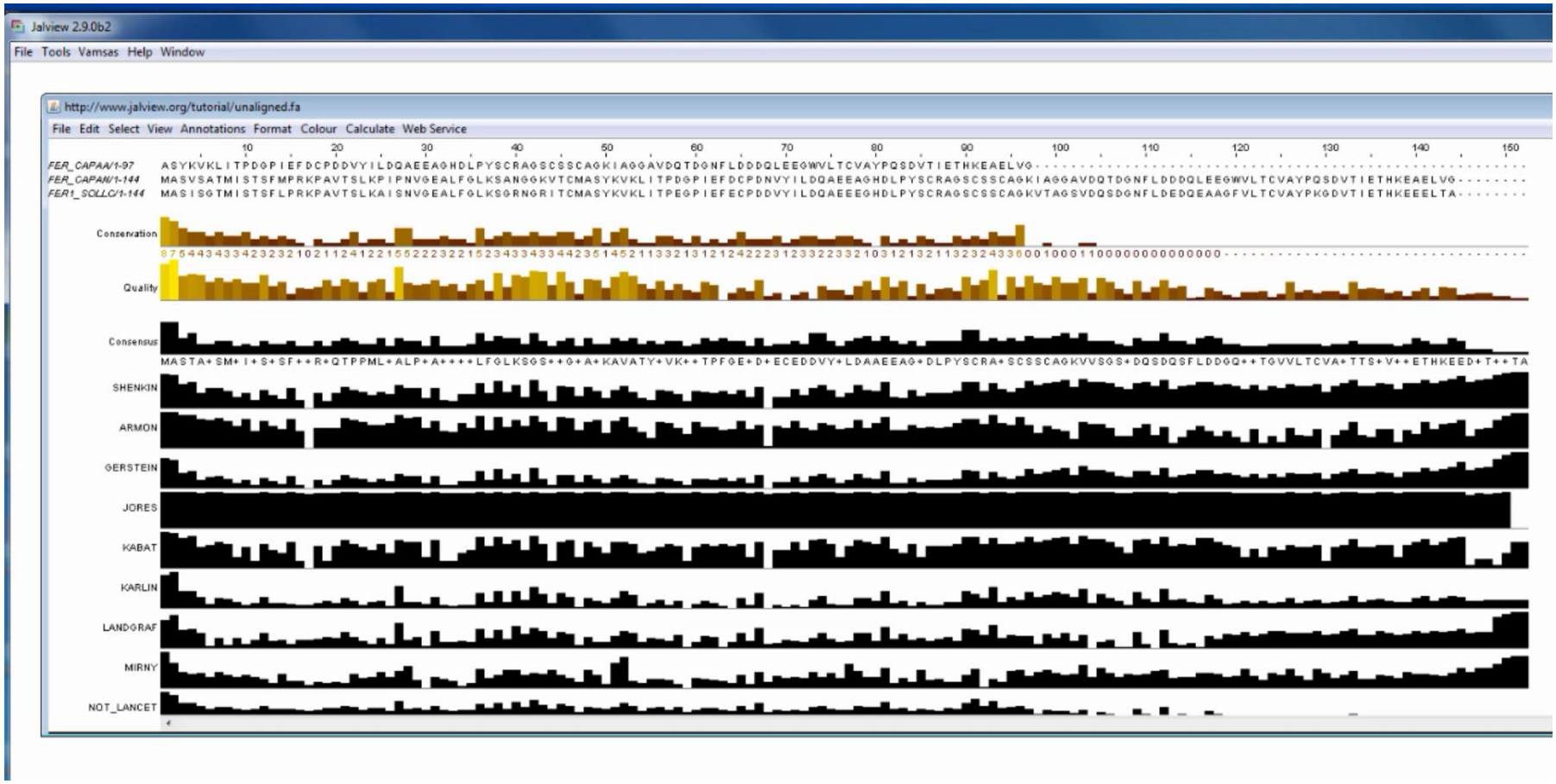
- The **current selection** can be copied to the **clipboard** (in PFAM format).
- It can also be **output to a textbox** using the output functions in the pop-up menu obtained by right clicking the current selection.
- The **textbox enables** quick manual editing of the alignment prior to importing it into a new window (using the [New Window] button) or saving to a file with the *File ⇒ Save As* pulldown menu option from the text box.


File Tools Vamsas Help Window
MAFFT Multiple Sequence Alignment of Retrieved from Uniprot
File Edit Select View Annotations Format Colour Calculate Web Service
Original
Phosporylation Site Predictions
MAFFT Alignment Ordering
FE2S2 Representatives
Spinach Feredoxin Structure
FER_CAPAA/1-97
FER_CAPAN/1-144
FER1_SOLLC/1-144
Q93XJ9_SOLTU/1-144
FER1_PEA/1-149
Q7XA98_TRIPR/1-152
FER1_MESCR/1-148
FER1_SPIOL/1-147
FER3_RAPSA/1-96
FER1_ARATH/1-148
FER_BRANA/1-96
FER2_ARATH/1-148
Q93Z60_ARATH/1-118
FER1_MAIZE/1-150
O80429_MAIZE/1-140
Selection
Reveal All
Sequence Details...
Show annotations
Hide annotations
Add reference annotations
Edit
Output to Textbox...
Create Sequence Feature ...
Remove Group
Edit Group
BLC
CLUSTAL
FASTA
MSF
PileUp
PIR
PFAM
AMSA
STH
PHYLIP
Secondary Structure
Iron Sulphur Contacts
Conservation
Quality

AACon Protein Alignment Conservation Services

Web Services >> Conservation >> Change AACon Settings


Jalview 2.9.0b2
File Tools Vamsas Help Window
http://www.jalview.org/tutorial/unaligned.fa
File Edit Select View Annotations Format Colour Calculate Web Service
Conservation
Quality
Consensus
SHENKIN
Sequence 1 ID: FER_CAPAA Residue: SER (43)
Edit parameters for Conservation with AAConWS
Current parameter set name: Defaults
Details
Options
ARMON
GERSTEIN
JORES
KABAT
KARLIN
LANDGRAF
MIRNY
NOT_LANCET
SANDER
SCHNEIDER
SHENKIN
SMERFS
TAYLOR_GAPS
TAYLOR_NO_GAPS
THOMPSON
VALDAR
WILLIAMSON
ZVELIBIL
Normalize
Parameters
SMERFS Column Scoring Method
MID_SCORE
SMERFS Gap Threshhold
0.1
SMERFS Window Width
7


Jalview 2.9.0b2
File Tools Vamsas Help Window
http://www.jalview.org/tutorial/unaligned.fa
File Edit Select View Annotations Format Colour Calculate Web Service
FER_CAPAA/1-97
FER_CAPAN/1-144
FER1_SOLLC/1-144
Conservation
Quality
Consensus
SHENKIN
ARMON
GERSTEIN
JORES
KABAT
KARLIN
LANDGRAF
MIRNY
NOT_LANCET

Practical

"To try on your own time"

Exercise : Making selections and groups

1- **Close all windows in Jalview and load the ferredoxin alignment (PFAM ID PF03460).**

2- Choose a residue and place the mouse cursor on it. Click and drag the mouse cursor to create a selection. As you drag, a red box will 'rubber band' out to show the extent of the selection. Release the mouse button and a red box should border the selected region.

3 - Now press [ESC] to clear the selection.

- **Select one sequence by clicking on the ID panel. Note that the sequence ID takes on a highlighted background and a red box appears around the selected sequence.**
- **Now hold down [SHIFT] and click another sequence ID a few positions above or below. Note how the selection expands to include all the sequences between the two positions on which you clicked.**
- **Now hold down [CTRL] and click on several sequences ID's both selected and unselected. Note how unselected IDs are individually added to the selection and previously selected IDs are individually deselected.**

4- **Repeat the step above but selecting columns by clicking on the ruler bar instead of selecting rows by clicking on the sequence ID.**

5- Press [F2] to enter Cursor mode. Navigate to column 59, row 1 by pressing *5 9 , 1 [RETURN]*.

- Press *Q* to mark this position. Now navigate to column 65, row 8 by pressing *6 5 , 8 [RETURN]*. Press *M* to complete the selection.
- Open the popup menu by right-clicking the selected region with the mouse. Open the *Selection ⇒ Group ⇒ Group Colour* menu and select 'Percentage Identity'.
- This will turn the selected region into a group and color it accordingly.

- Hold down [CTRL] and use the mouse to select and deselect sequences by clicking on their Sequence ID label. Note how the group expands to include newly selected sequences, and the 'Percentage Identity' coloring changes.
- Use the mouse to click and drag the right-hand edge of the selected group. Note again how the group resizes.

Right click on the text area to open the selection popup-menu. Follow the menus and pick an output format from the *Selection ⇒ Output to Textbox . . .* submenu. Try manually editing the alignment and then press the [New Window] button to import the file into a new alignment window.


BTU
BIOINFORMATICS
TRAINNING UNIT
GENOMICS
1st Module What Bioinformatics Can Do for You
2nd Module Manipulation of Biological Sequences
3rd Module Working with single DNA sequence
4th Module How to Build a Multiple Sequence Alignment?
5th Module Inferring Phylogenetics analysis using Jalview
6th Module Advanced Molecular Concepts
7th Module Inferring Protein Sequence (structure & Function)
8th Module RNAanalysis and Function
9th Module Editing and Publishing Alignments in your Manuscript
10th Module Building and Publishing Phylogenetic Trees
11th Module Working with Protein 3-D Structures
12th Module Advanced Bioinformatics Using R
Bioinformatics Workshops Series
EST by, Dr. Ahmed Mansour Alzohairy
Call us: 01026060324 / 01000727270
Facebook: https://www.facebook.com/Bioinformaticsunit

By

Ahmed Mansour Alzohairy

Department of Genetics, Zagazig University, Zagazig, Egypt