



# H3ABioNet

Pan African Bioinformatics Network for H3Africa

**Introduction to Bioinformatics online course: IBT**

## **Multiple Sequence Alignment**

### ***Lec2: Loading your own sequence***

By

**Ahmed Mansour Alzohairy**

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

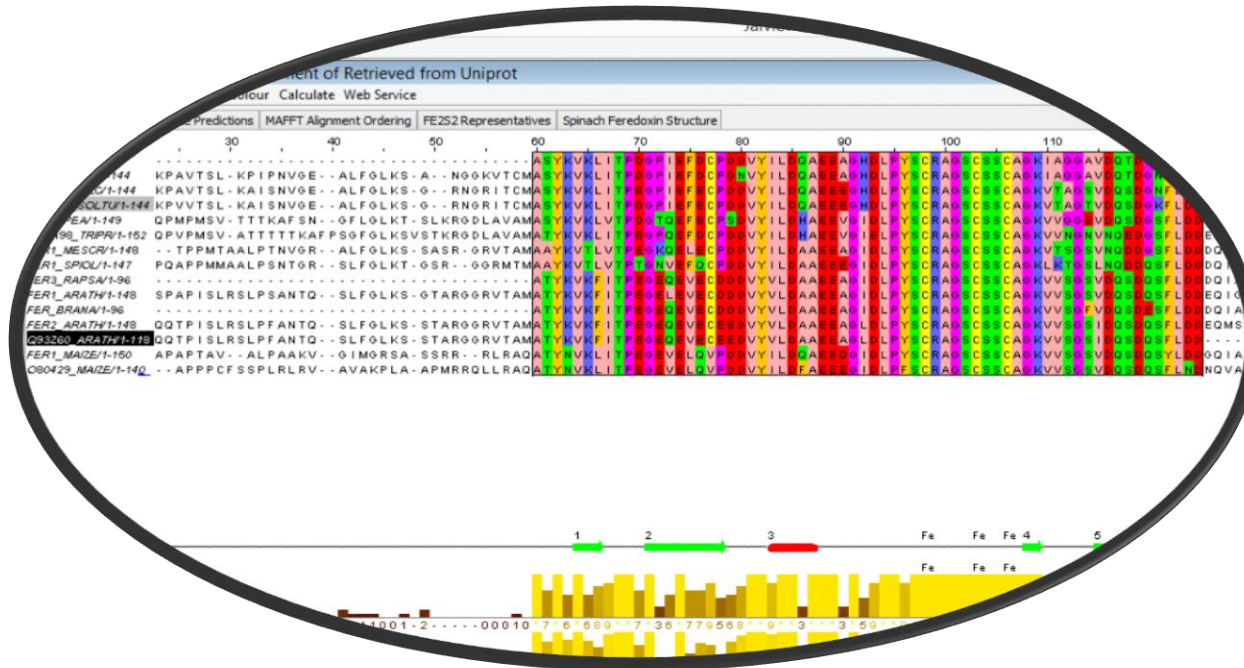




Source: Jaview Online Training Video

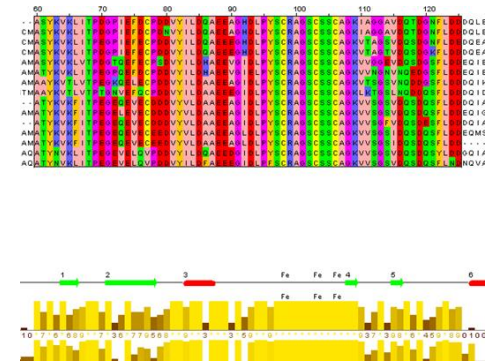
<https://www.jalview.org/>

## *Loading your own sequence*



# Jalview provides many ways to load - your own sequences.

- Drag and Drop
- From a File
- From a URL
- Cut and Paste
- From a public database



# Reading Sequence Files

## Typical sequence alignment formats:

- Fasta (.mfa, .fastq,.fasta,.fa)
- Clustal (.aln)
- PFAM (.pfam)
- MSF (.msf)
- PIR (.pir)
- BLA (.blc)
- AMSA (.amsa)
- Jalview (.jvp)
- Stockholm (.stk, .sto)

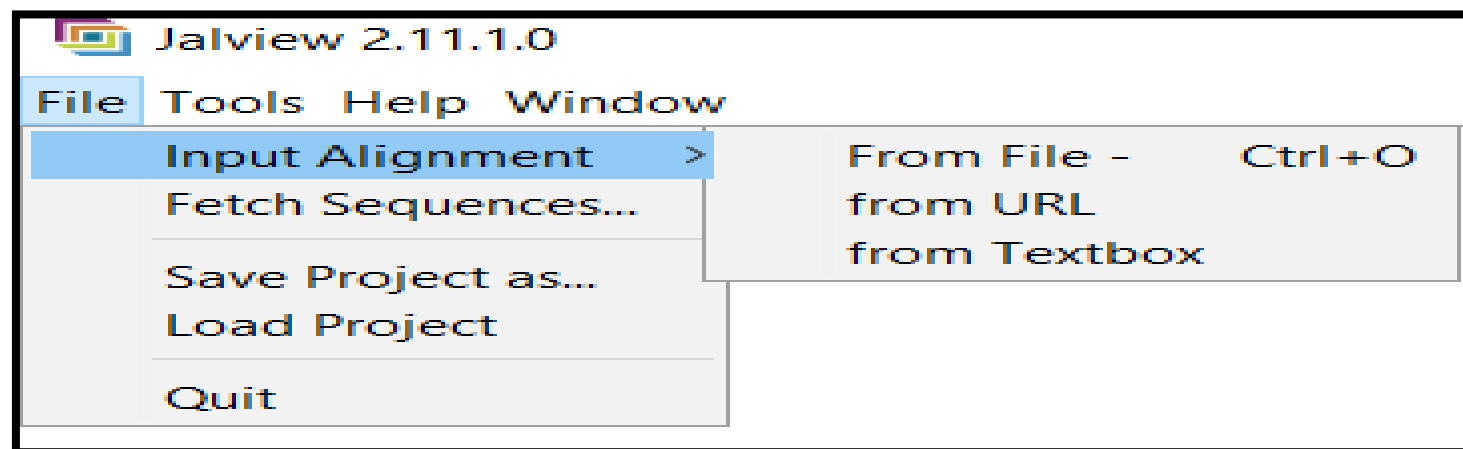
Source: Jaview Online Training Video

<https://www.jalview.org/>

# From a File

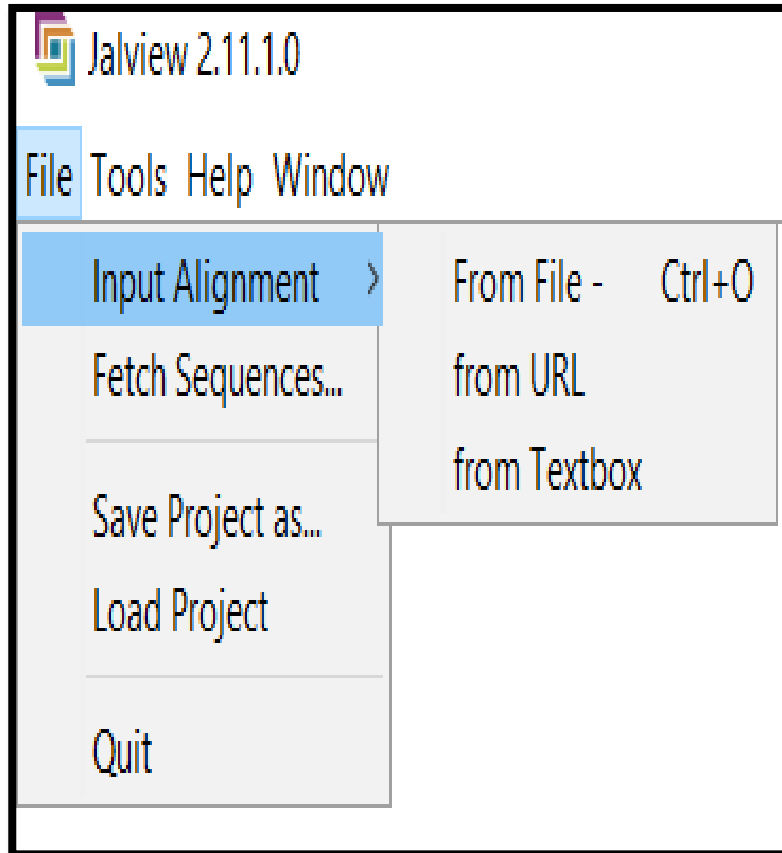
Jalview can read sequence alignments from a sequence alignment file. This is a **text file**, **NOT** a **word processor** document.

Select **File**  $\Rightarrow$  **Input Alignment**  $\Rightarrow$  **From File** from the main menu. Remember to select the **appropriate file type**. Jalview can automatically identify some sequence file formats.



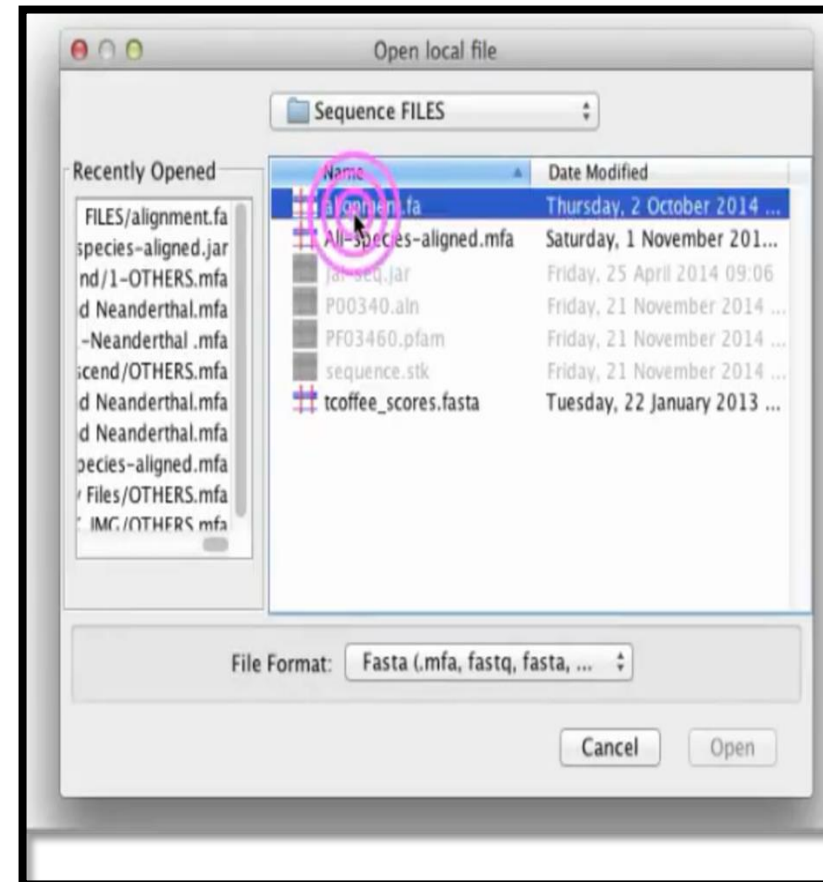


**File >> Input Alignment >> From File**



***From a File***

***Choose the right File format***



**Source: Jaview Online Training Video**

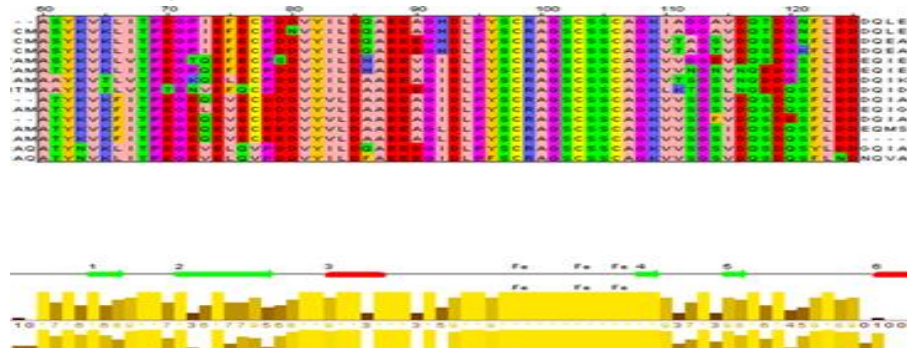
<https://www.jalview.org/>

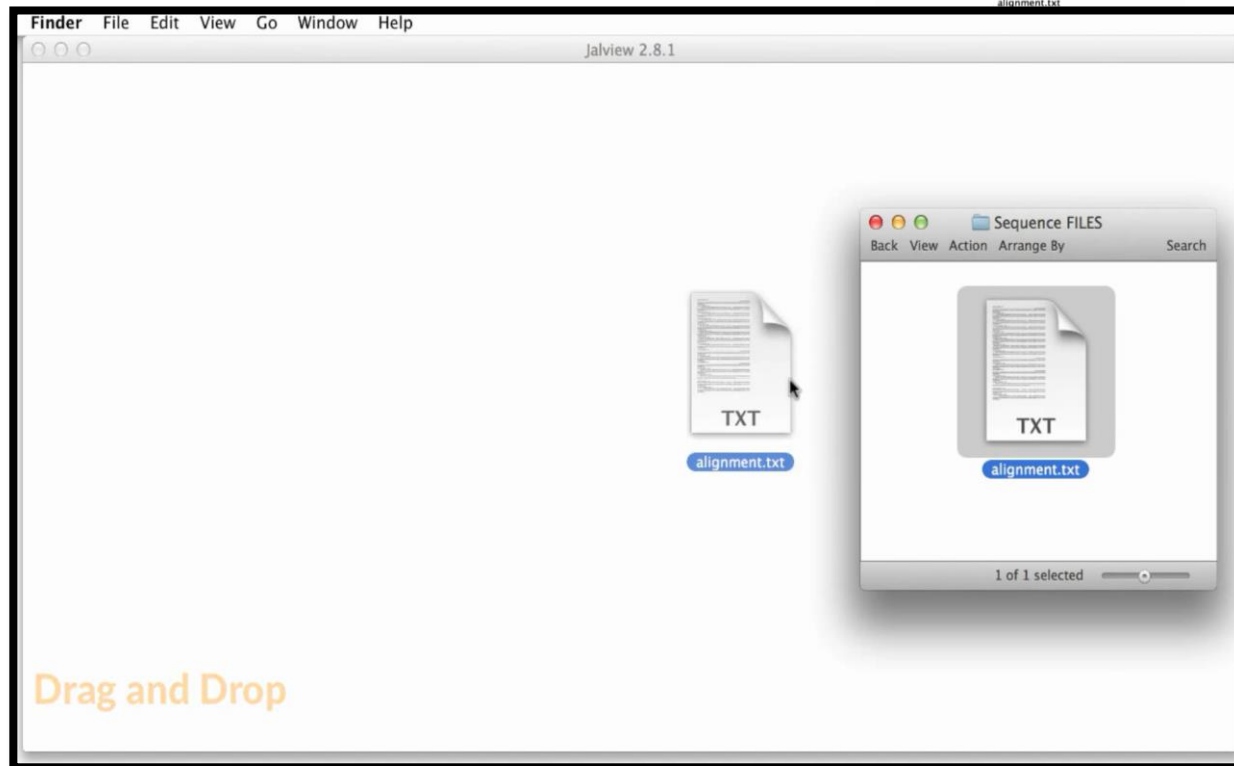
# Drag and Drop



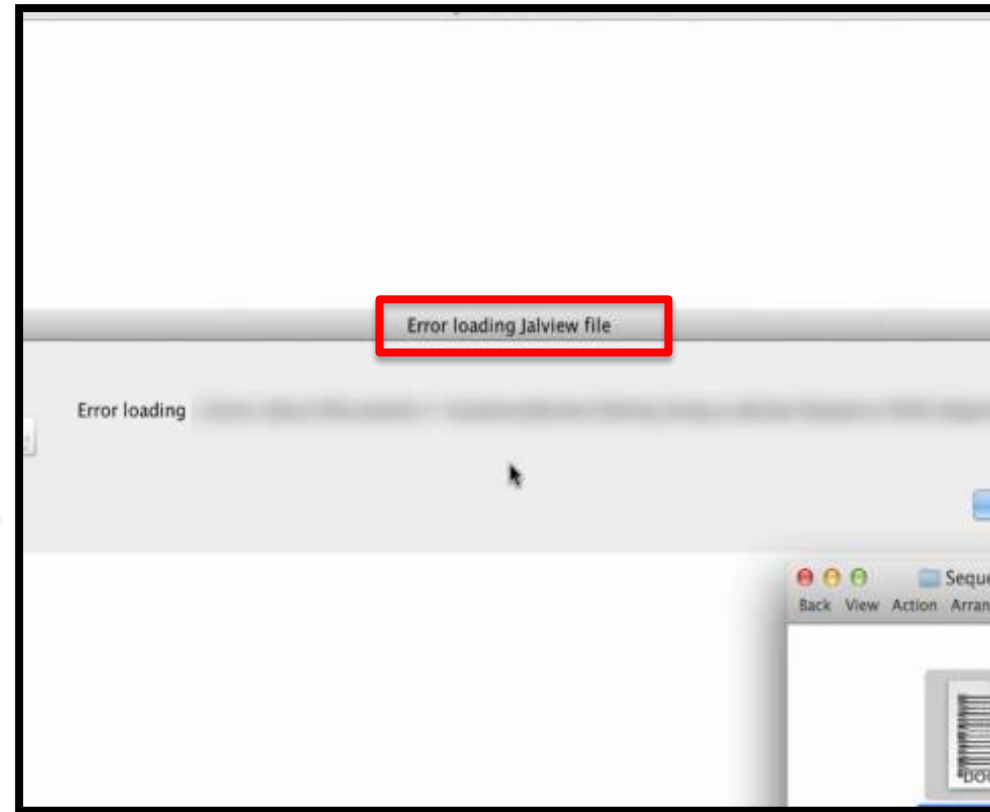
# Drag and Drop

- In most operating systems you can just **drag** a **file** icon *from a file browser window* and **drop it** on an open Jalview application window.
- Drag and drop also **works** when loading data *from a URL - simply drag the link or URL* from the address panel of your browser on to an alignment or the Jalview desktop background and Jalview will load data from the URL directly.





Source: Jaview Online Training Video  
<https://www.jalview.org/>



alignment.docx

Source: Jaview Online Training Video

<https://www.jalview.org/>

# Load using via url

Source: Jaview Online Training Video

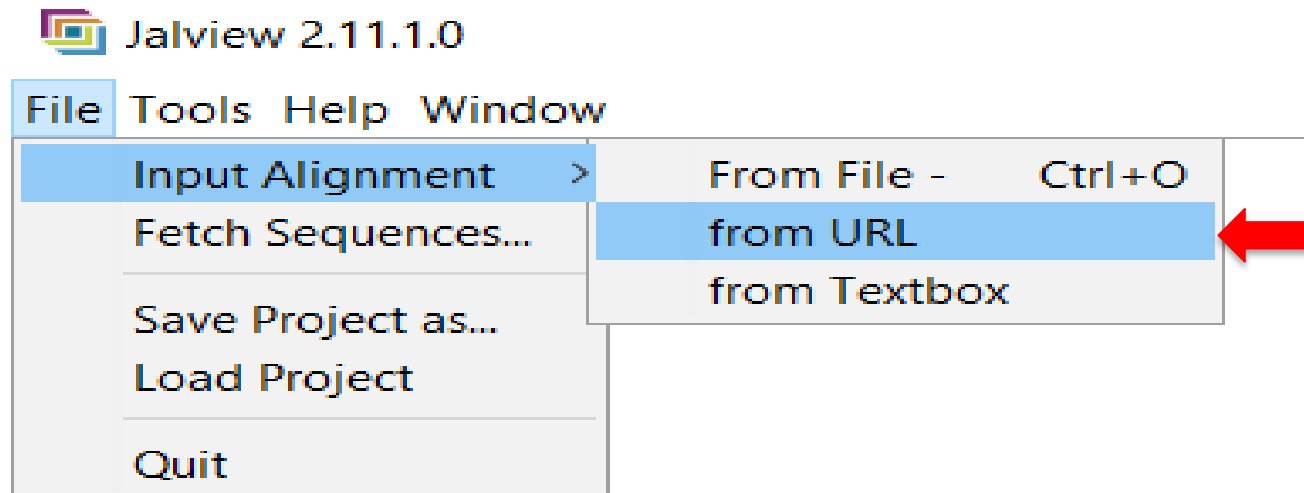
<https://www.jalview.org/>

## From a URL

Jalview can read sequence alignments directly **from a URL**. Please note that the files must be in a sequence alignment format - **an HTML alignment or graphics file cannot** be read by Jalview.

Select **File  $\Rightarrow$  Input Alignment  $\Rightarrow$  From URL**

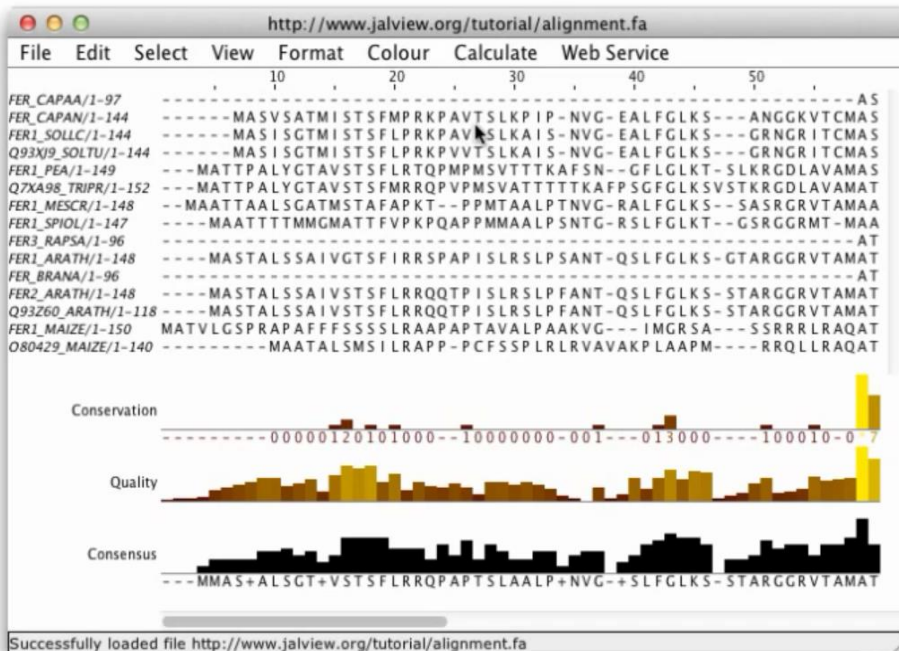
Jalview will attempt to automatically discover the file format





# *From a URL*





```

>FER_CAPAA/1-97
.....ASYKVKLITPDGPI
EFDCPPDDVYILDAAEEAGHDLFYSCRAGSCSSCAGKIAGGAVDQTDGNFLDDQLEEGVWLTCVAYPQSDVT
IETHKEEELVG.
>FER_CAPA/1-144
.....MASVSATMISTFSMPRKPAPVTSLKPIF.NVG.EALFGLKS...ANGGKVTCMASYKVKLITPDGPI
EFDCPPDDVYILDAAEEAGHDLFYSCRAGSCSSCAGKIAGGAVDQTDGNFLDDQLEEGVWLTCVAYPQSDVT
IETHKEEELVG.
>FER1_SOLLC/1-144
.....MASISGTMISTFSFLRPKPAPVTSLKAFS.NVG.EALFGLKS...GRNGRITCMASYKVKLITPEGPI
EFDCPPDDVYILDAAEEEGHDLFYSCRAGSCSSCAGKVTAGSVQSDGNFLDDEQAEAGVLTTCVAYPKGOVT
IETHKEEELTA.
>Q93XJ9 SOLTU/1-144
.....MASISGTMISTFSFLRPKPAPVTSLKAFS.NVG.EALFGLKS...GRNGRITCMASYKVKLITPDGPI
EFDCPPDDVYILDAAEEEGHDLFYSCRAGSCSSCAGKVTAGTVQSDGKFLDDQAEAGVLTTCVAYPKGOVT
IETHKEEELTA.
>FER1_PEA/1-149
.....MATTPALYGTAVSTFSFLRTPQPMMSVTTTKAFSN..GFLGLKT.SLKRGLDLAVAMASYKVKLITPDGTQ
EFDCPSDDVYILDAAEEVGLDLYSCRAGSCSSCAGKVGSEVQSDGSFLDDEQIEAGVLTTCVAYPTSDDV
IETHKEEELTA.
>Q7XA98 TRIPR/1-152
.....MATTPALYGTAVSTFSFMRRPQVPMMSVATTTTTKAPPSGFLKS SVSTKRGLDLAVAMASYKVKLITPEGPQ
EFDCPPDDVYILDAAEEVGLIELFYSCRAGSCSSCAGKVGNGVQSDGSFLDDEQIEGGVLTTCVAFPTSDDV
IETHKEEELTA.
>FER1_MESCR/1-148
.....MAATTAALSGATMSTFAFAPK..PPMTAALPTNVG.RALFGLKS...SASRGRTVAMAYKVLITVTEPGKQ
ELECPDDVYILDAAEEAGIDLYSCRAGSCSSCAGKVTSGSVQSDGSFLDDQIKEGVLTTCVAYPTGOVT
IETHKEEELTA.
>FER1_SPIOL/1-147
.....MAATTTMGMATVTFVFKQAPPMMAALPSNTG.RSLFGLKT..GSRGRMT.MAAYKVLITVTPGTNV
EFQCPDDVYILDAAEEEGIDLYSCRAGSCSSCAGKLKTGSLNQDDQSFLLDDQIDEGVLTTCAYPVSDVT
IETHKEEELTA.
>FER3_RAPSA/1-96
.....ATYKVKFITPEGEQ
EVECDDDVYILDAAEEAGIDLYSCRAGSCSSCAGKVSGSVQSDQSFLLDDQIAEGVLTTCAYPTSDDV
IETHREEDMV..
>FER1_ARATH/1-148
.....MASTALSAIVTSQFIRRSFAPISLRLSPANT.QSLFGLKS.GTAGRGRTVAMATYKVKFITPEGEL
EVECDDDVYILDAAEEAGIDLYSCRAGSCSSCAGKVSGSVQSDQSFLLDDEQIEGVLTCAYPTSDDV
IETHKEEELV..
>FER3_BRANA/1-96
.....ATYKVKFITPEGEQ
EVECDDDVYILDAAEEAGIDLYSCRAGSCSSCAGKVSGFVQSDQSFLLDDQIAEGVLTTCAYPTSDDV
IETHKEEELV..
>FER2_ARATH/1-148
.....MASTALSAIVTSFSLRRQOTPISLRLSPANT.QSLFGLKS.STARGGRVTAMATYKVKFITPEGEQ
EVECEEDVYILDAAEEAGIDLYSCRAGSCSSCAGKVSGSIDQSDQSFLLDEQMSGVLTTCVAYPTSDDV
IETHKEEELAM..
>Q93Z60 ARATH/1-118
.....MASTALSAIVTSFSLRRQOTPISLRLSPANT.QSLFGLKS.STARGGRVTAMATYKVKFITPEGEQ
EVECEEDVYILDAAEEAGIDLYSCRAGSCSSCAGKVSGSIDQSDQSFLLD.....
>FER1_MAIZE/1-150
MATVLSGRAPPAFFSSSLRAAPAPTALPAKVG...IMGRSA...SSRRRLRAQATYNVKLITPEGEV
ELQVPDDVYILDAAEEEDIDLYSCRAGSCSSCAGKVSGSVQSDQSYLDDQIADGVLTTCAYPTSDDV
IETHKEEELTGA

```

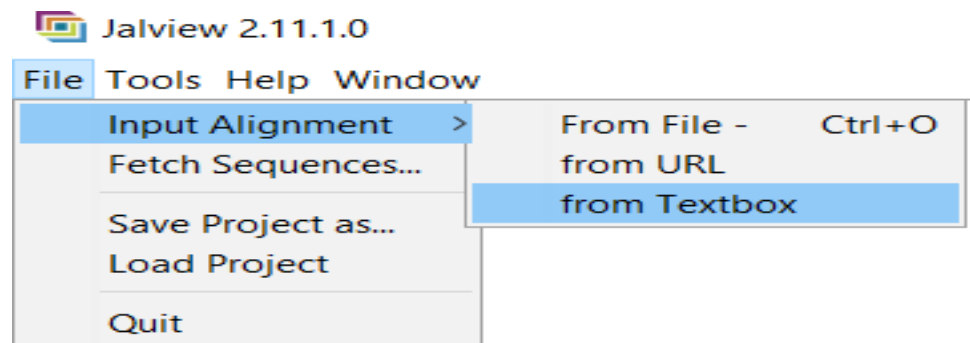


# *From Textbox*

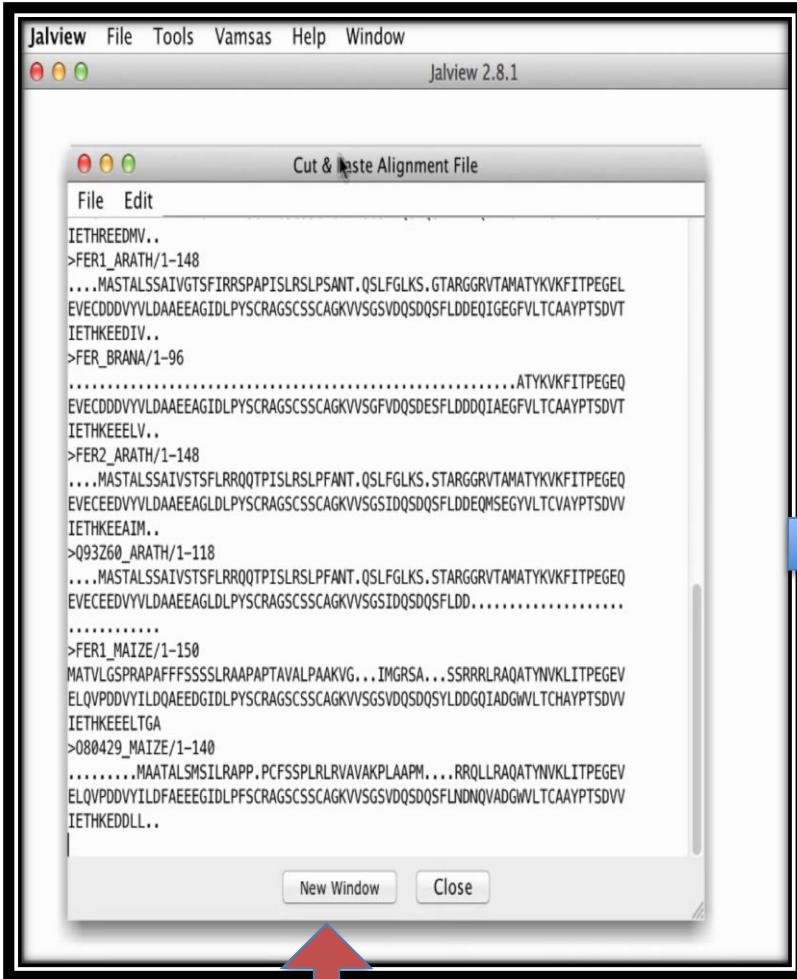
## (Copy and Paste)

# Copy and Paste

- Documents such as those produced by Microsoft Word cannot be readily understood by Jalview. The way to read sequences from these documents is to select the data from the document and copy it to the clipboard. There are two ways to do this:
- 1- One is to **right-click on the desktop background**, and select the **‘Paste to new alignment’** option in the menu that appears.
- 2- The other is to select ***File ⇒ Input Alignment ⇒ From Textbox*** from the main menu, and paste the sequences into the textbox window that will appear . In both cases, sequences should be in the right format







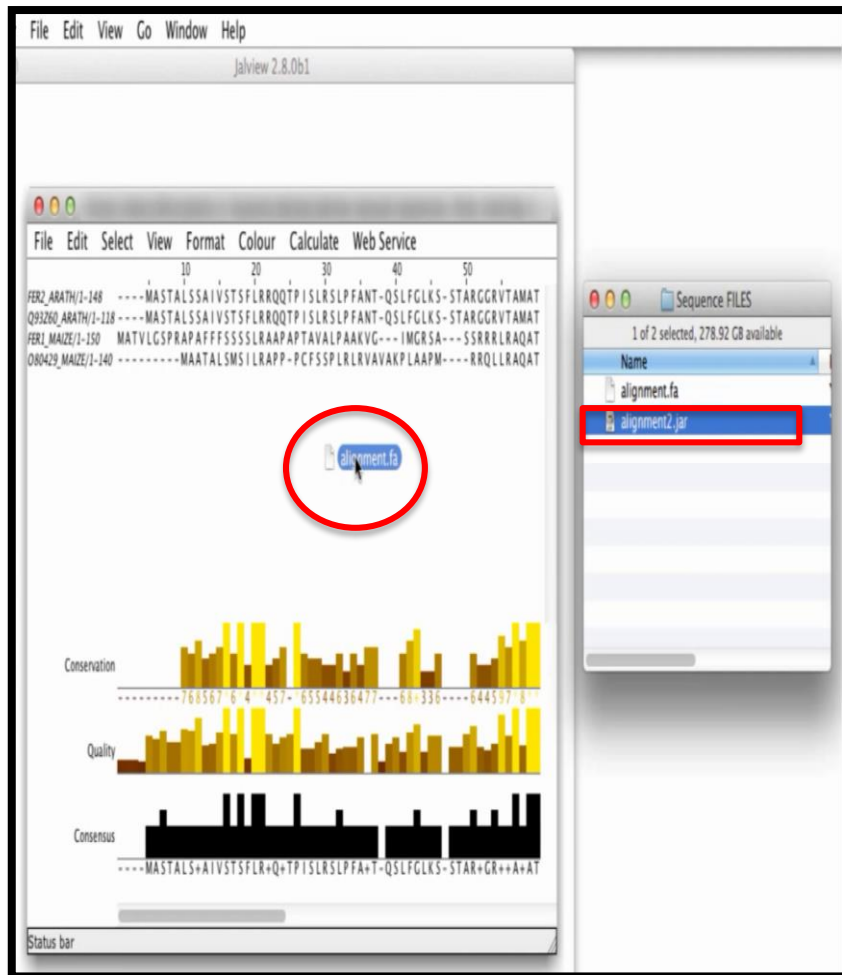
Paste sequence  
(⌘+V or CTRL+V)



Source: Jalview Online Training Video

<https://www.jalview.org/>

# Add more sequences into your alignment



Jalview 2.8.0b1

File Edit View Go Window Help

File Edit Select View Format Colour Calculate Web Service

10 20 30 40 50

ER2\_ARATH/1-148 ---MASTALSSAIVSTSFLLRQQTPISLRSLPFANT-QSLFGLKS-STARGGRVTAMAT  
Q93260\_ARATH/1-118 ---MASTALSSAIVSTSFLLRQQTPISLRSLPFANT-QSLFGLKS-STARGGRVTAMAT  
FER1\_MAIZE/1-150 MATVLGSPRAPAFFSSSLRAAPATAVALPAKVG---IMGRSA---SSRRRLRAQAT  
O00429\_MAIZE/1-140 -----MAATALSMSILRAPP-PCFSSPLRLRVAVAKPLAAPM---RRQLLRAQAT

Conservation

Quality

Consensus

---MASTALS+AIVSTSFLLR+Q+TPISLRS LPFANT-QSLFGLKS-STARGGR+AAAT

Status bar

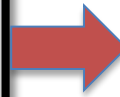
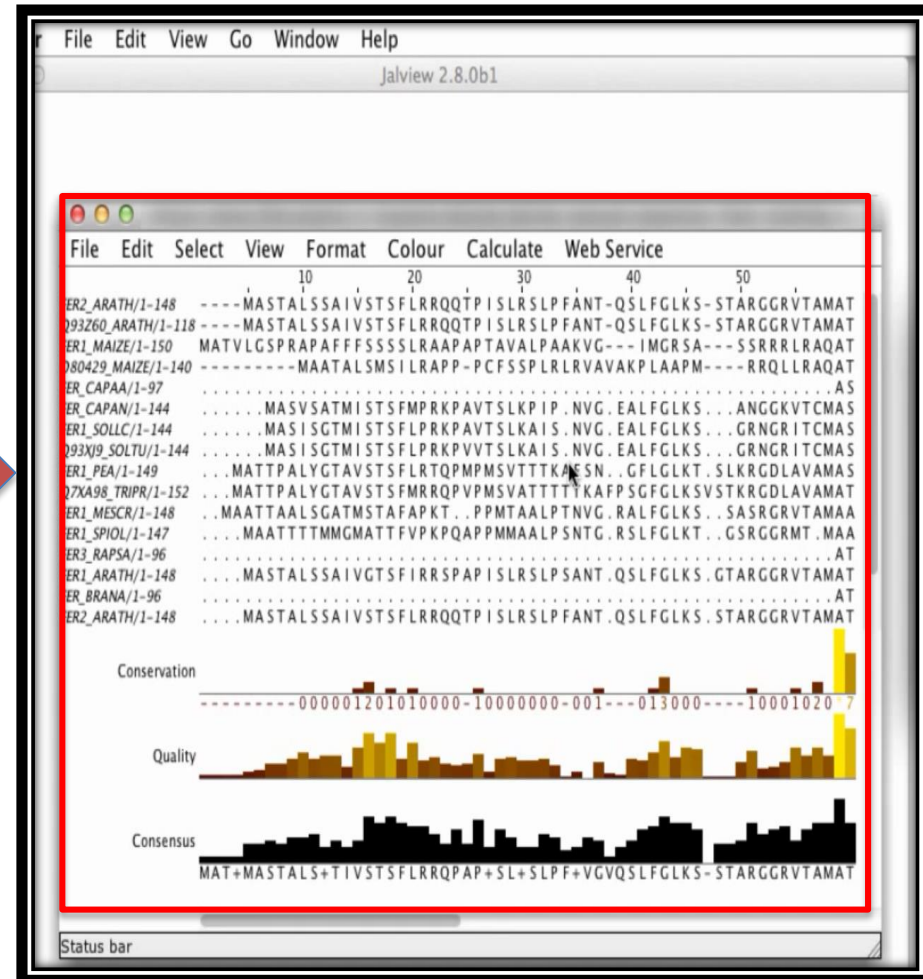
Sequence FILES

1 of 2 selected, 278.92 GB available

Name

alignment1

alignment2.jar

Jalview 2.8.0b1

File Edit View Go Window Help

File Edit Select View Format Colour Calculate Web Service

10 20 30 40 50

ER2\_ARATH/1-148 ---MASTALSSAIVSTSFLLRQQTPISLRSLPFANT-QSLFGLKS-STARGGRVTAMAT  
Q93260\_ARATH/1-118 ---MASTALSSAIVSTSFLLRQQTPISLRSLPFANT-QSLFGLKS-STARGGRVTAMAT  
ER1\_MAIZE/1-150 MATVLGSPRAPAFFSSSLRAAPATAVALPAKVG---IMGRSA---SSRRRLRAQAT  
O00429\_MAIZE/1-140 -----MAATALSMSILRAPP-PCFSSPLRLRVAVAKPLAAPM---RRQLLRAQAT  
ER\_CAPAA/1-97 .....AS  
ER\_CAPAN/1-144 .....MASVSATMISTSFMPKPAVTS LKPIP.NVG.EALFGLKS...ANGGKVTCMAS  
ER1\_SOLLIC/1-144 .....MASISGMTISTSF LPRKPAVTS LKAI S.NVG.EALFGLKS...GRNGRITCMAS  
Q93X19\_SOLTU/1-144 .....MASISGMTISTSF LPRKPVVTS LKAI S.NVG.EALFGLKS...GRNGRITCMAS  
ER1\_PEA/1-149 .....MATTPALYCTAVSTSF LRTQPMPSVTTTKASN.GFLGLKT.SLKRGLD LAVAMAS  
Q7XA98\_TRIPR/1-152 .....MATTPALYCTAVSTSFMRQPVPMSVATTTTKAFPSGFLKSVSTKRGLD LAVAMAT  
ER1\_MESCR/1-148 .....MAATTAALSGATMSTAFAPKT.PPMTAALPTNVG.RALFGLKS.SASRGRVTAMAA  
ER1\_SPIOL/1-147 .....MAATTTTMMGMATTFVPKQPAPPMMAALPSNTG.RSLFGLKT.GSRGGRMT.MAA  
ER3\_RAPSA/1-96 .....AT  
ER1\_ARATH/1-148 .....MASTALSSAIVGTSFIRRSAPISLRS LPSANT.QSLFGLKS.GTARGGRVTAMAT  
ER\_BRANA/1-96 .....AT  
ER2\_ARATH/1-148 .....MASTALSSAIVSTSFLLRQQTPISLRSLPFANT.QSLFGLKS.STARGGRVTAMAT

Conservation

Quality

Consensus

MAT+MASTALS+TIVSTSFLLRQPAP+SL+SLPF+VGVQSLFGLKS-STARGGRVTAMAT

Status bar



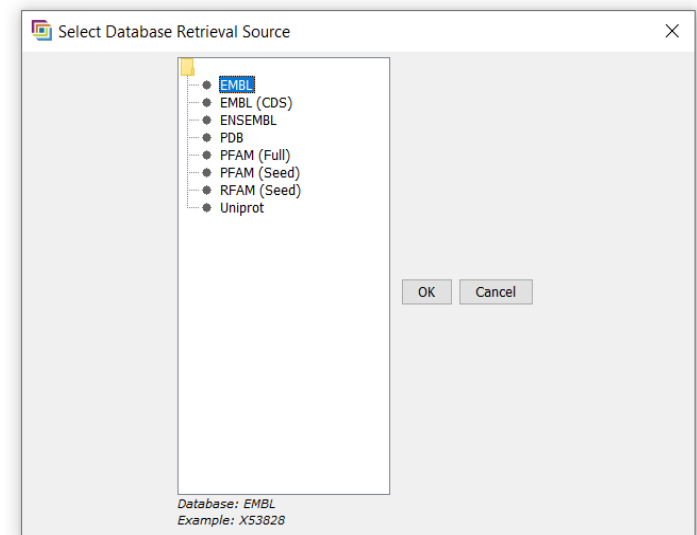
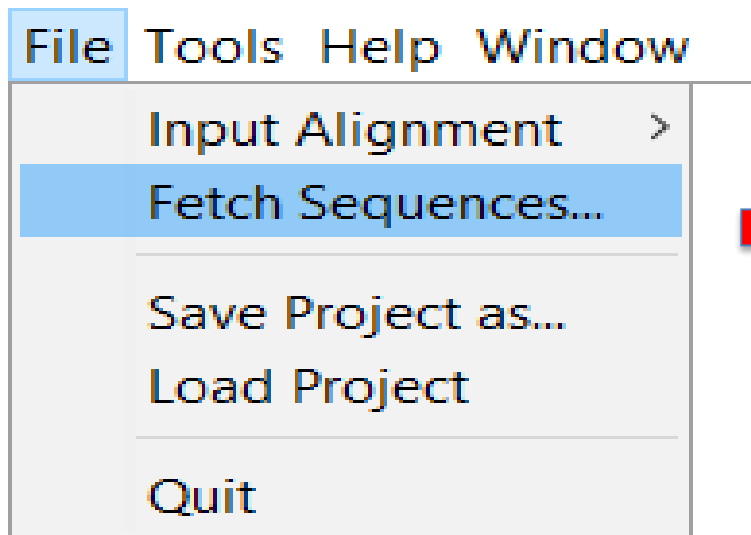
# Accessing public databases

Source: Jaview Online Training Video

<https://www.jalview.org/>

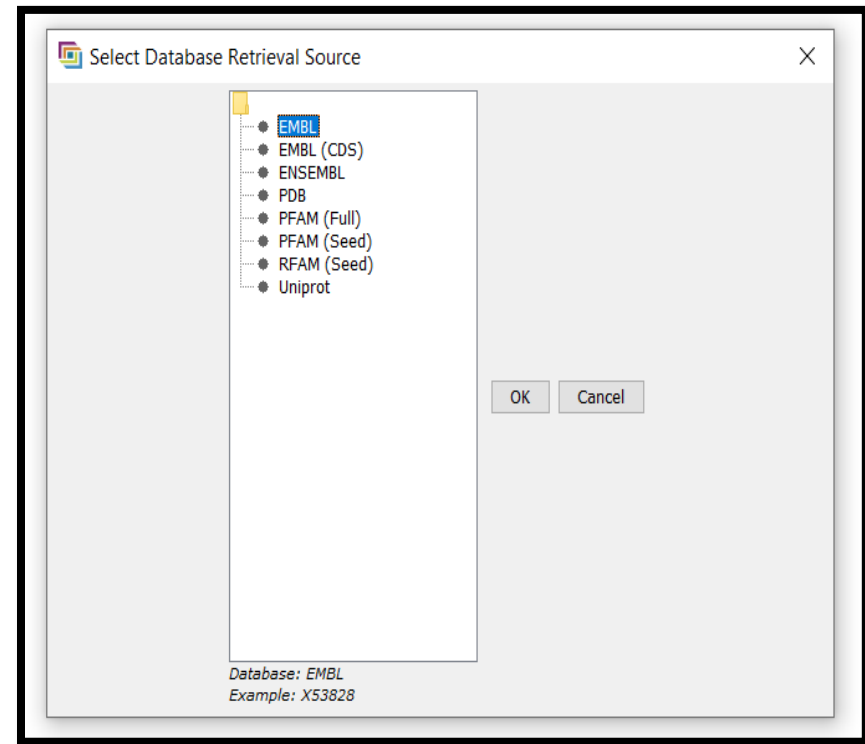
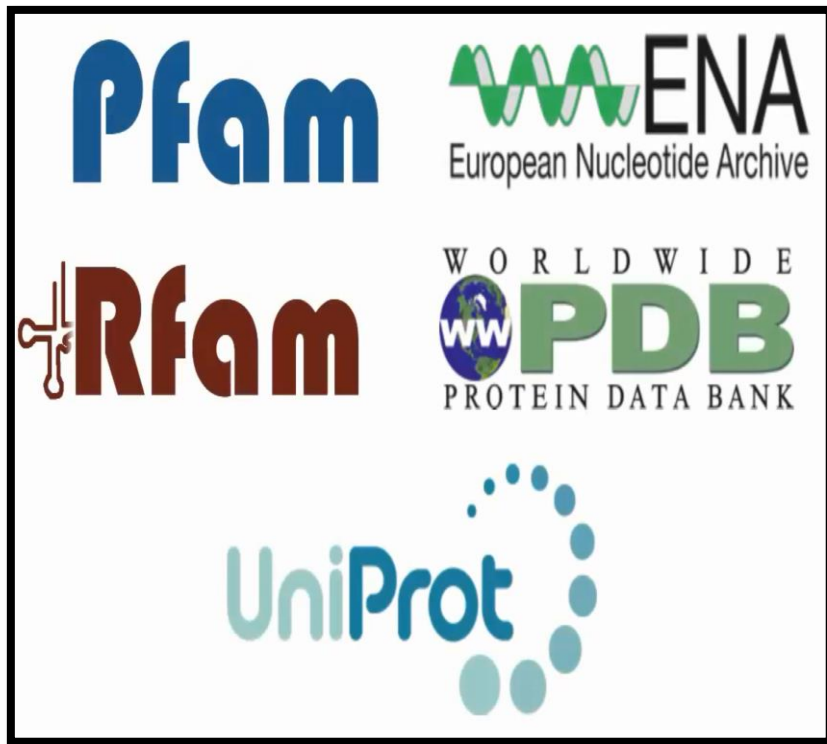
# From a public database

- Jalview's sequence fetching capabilities allow you to avoid having to manually locate and save sequences from a web page before loading them into Jalview. It also allows Jalview to gather additional metadata provided by the source, such as annotation and database cross references. **Select File ⇒ Fetch Sequence(s)**





- **Jalview** can **retrieve sequences** and sequence alignments from the public databases housed at the European Bioinformatics Institute EMBL, including **Uniprot, Pfam, Rfam** and the **PDB**, as well as any DAS sequence server registered at the configured DAS registry.



## Select Database

Jalview 2.11.1.0

File Tools Help Window

Input Alignment >

Fetch Sequences...

Save Project as...

Load Project

Quit

**New Sequence Fetcher**

EMBL

**Example query: X53828**  
Enter one or more accession IDs separated by a semi-colon ";"

☐ Replace commas with semi-colons

X53828

Example Clear OK Cancel



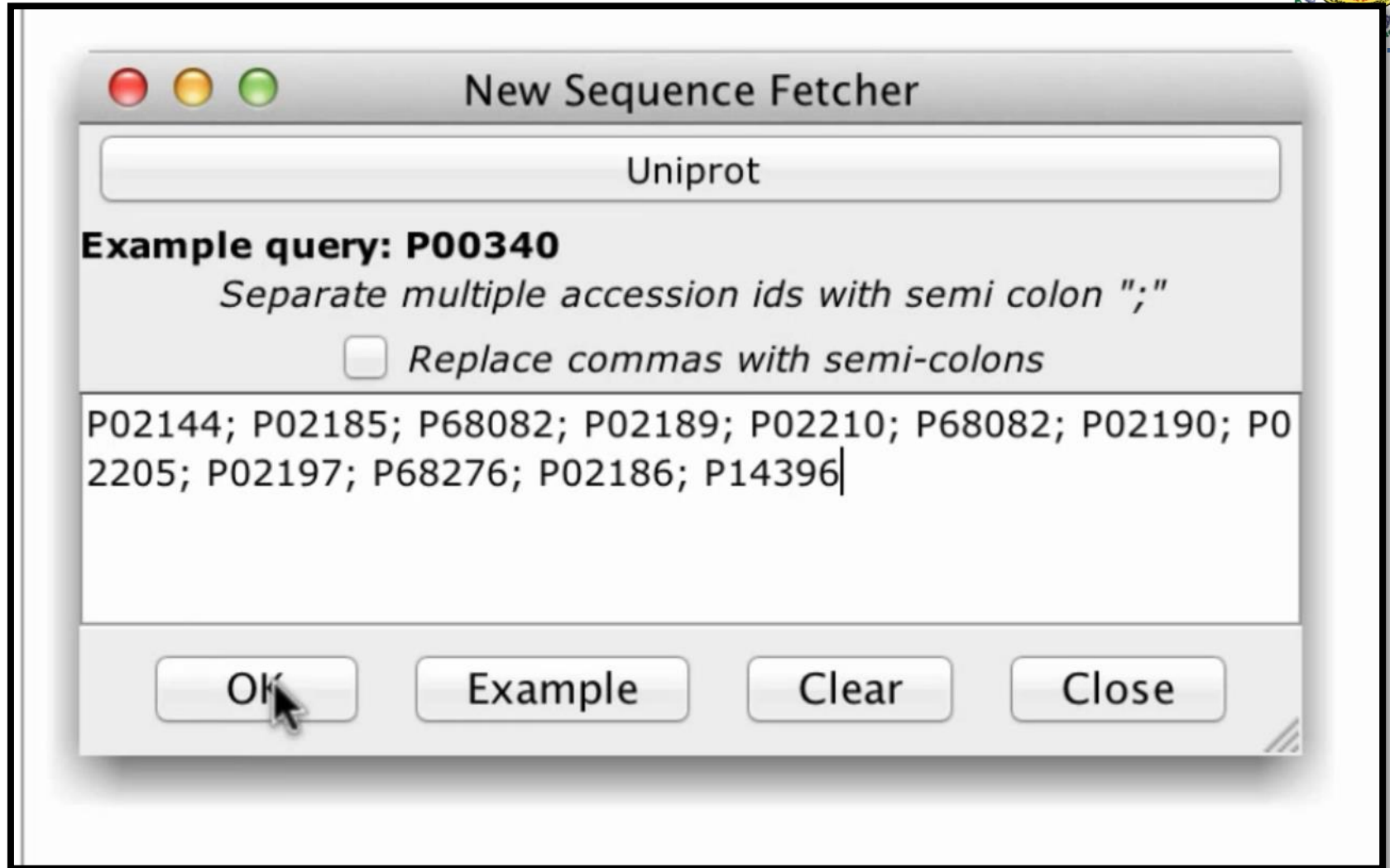
**Select Database Retrieval Source**

- EMBL
- EMBL (CDS)
- ENSEMBL
- PDB
- PFAM (Full)
- PFAM (Seed)
- RFAM (Seed)
- Uniprot

OK Cancel

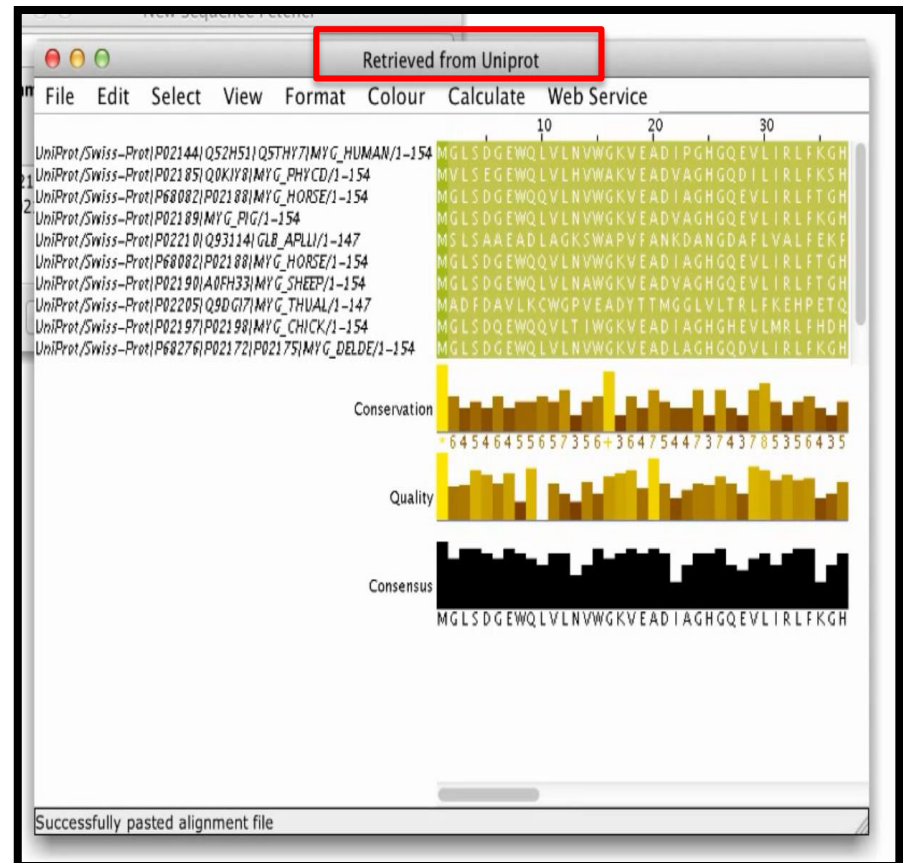
Database: EMBL  
Example: X53828

- Pressing the **database selection** button in the dialog box opens a new window showing all the database sources Jalview can access (*grouped by the type of database*).
- Once you've selected the **appropriate database**, hit (**OK**) close the database selection window, and then enter one or **several database IDs or accession numbers** separated by a **semicolon** and press (**OK**).
- (Jalview) will then attempt to **retrieve** them from the chosen database.



Source: Jaview Online Training Video

<https://www.jalview.org/>



## Exercise :

### ***“To try on your own time”***

#### ***Loading sequences.***

- Start Jalview then close all windows (if necessary) by selecting **Window ⇒ Close All** from the Desktop window.
- Select **File ⇒ Input Alignment ⇒ From URL** from the Desktop and enter **<http://www.jalview.org/tutorial/alignment.fa>** in the box. Click **OK** and the alignment should load
- Close all windows using the **Window ⇒ Close All** menu option from the Desktop. Point your web browser at the same URL (**<http://www.jalview.org/tutorial/alignment.fa>**) and save the file to your desktop.
- Open this file in Jalview by selecting **File ⇒ Input Alignment ⇒ From File** from the main menu and selecting the file from your **desktop**. Click OK and load the alignment.
- Select **Desktop ⇒ Window ⇒ Close All** and drag the **alignment.fa** file from the desktop onto the Jalview window. The alignment should open. Try dragging onto an empty Jalview and onto an existing alignment and observe the results. You can also try dragging the URL directly onto Jalview



## Exercise : Continue

### To try on your own time

- **Select File ⇒ Fetch Sequence(s)**. from the Desktop. Select the **PFAM** seed database and enter the accession number **PF03460**. Click **OK**. An alignment of about **107 sequences** should load.
- **Open <http://www.jalview.org/tutorial/alignment.fa>** in a web browser. Note: If the URL is downloaded instead of opened in the browser, then locate the file in your download directory and open it in a text editor.
- Select and copy the entire text to the clipboard (usually via the browser's Edit ⇒ Copy menu option).
- Ensure Jalview is running and **select File ⇒ Input Alignment ⇒ From Textbox**.
- Paste the clipboard into the large window using the **Edit ⇒ Paste** text box menu option. Click New Window and the alignment will be loaded.



**BTU BIOINFORMATICS TRAINING UNIT**

**Bioinformatics Workshops Series**

EST by, Dr. Ahmed Mansour Alzohairy  
Call us: 01026060324 / 01000727270  
Facebook: <https://www.facebook.com/Bioinformaticsunit>  
design by: Mohamed Tarek, 01026060324



By

**Ahmed Mansour Alzohairy**

**Department of Genetics, Zagazig University,  
Zagazig, Egypt**