H3ABioNet

H3ABioNet
Pan African Bioinformatics Network for H3Africa

Introduction to Bioinformatics Online Course: IBT

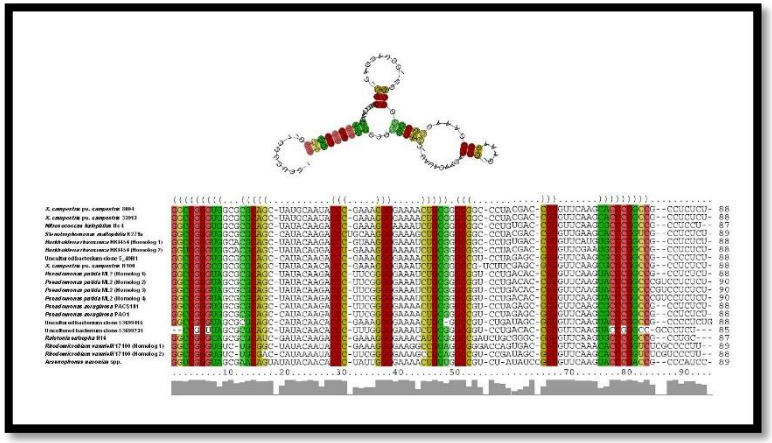
Multiple Sequence Alignment
Building Multiple Sequence Alignment
Lec4: Coloring MSA for Publication

CC BY NC SA

Preparing Your Multiple Alignment for Publication

- You want to show this alignment to your colleagues, you want to **include** it in **publications**

- In short, you need a **high-impact picture** to **convince people** that your research is going well.



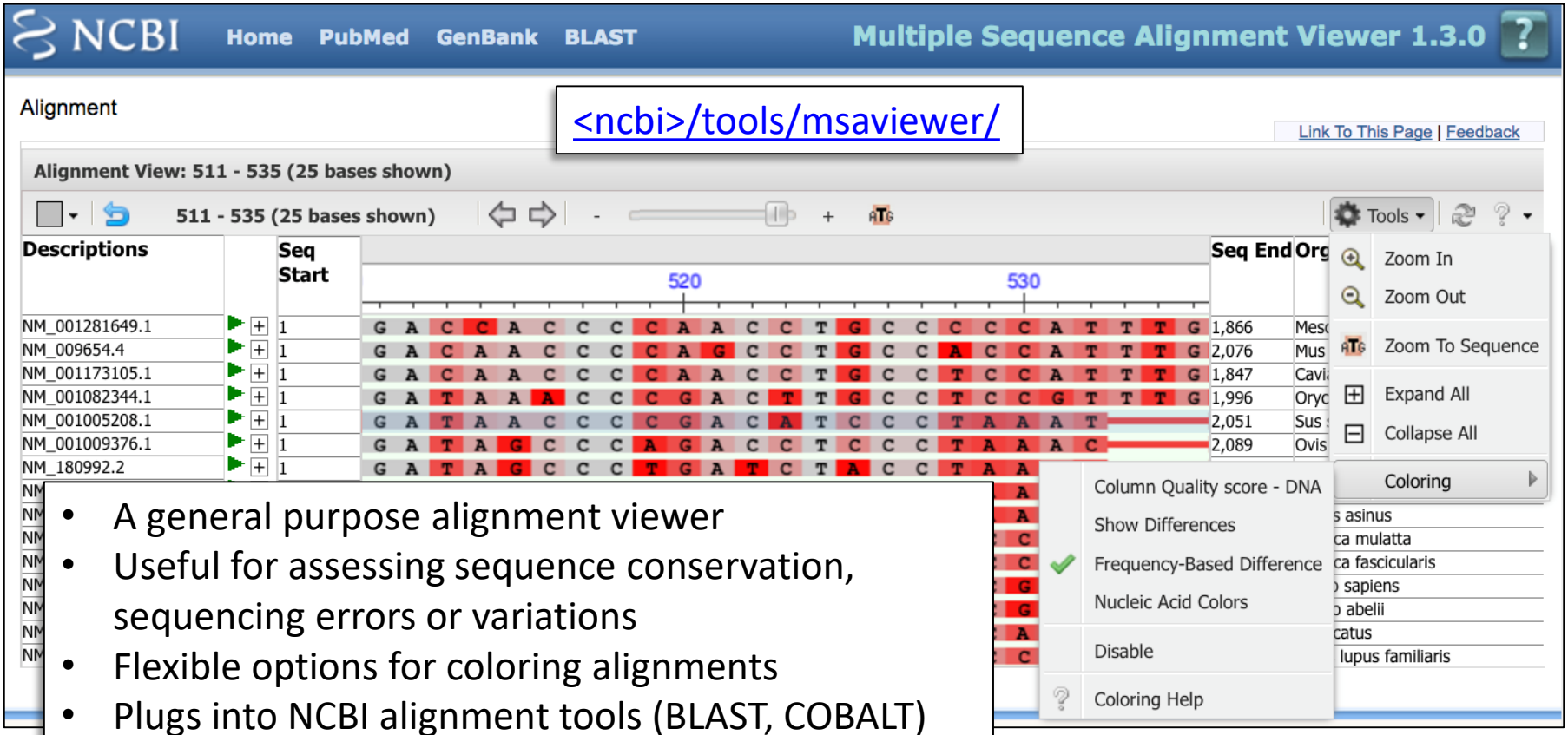
Claverie J, Notredame C (2007). Bioinformatics for Dummies (2nd Edn). *Wiley publishing, Inc*. 436 pp.

CC BY NC SA

Introducing the Multiple Sequence Alignment Viewer (MSAV)

CC BY NC SA

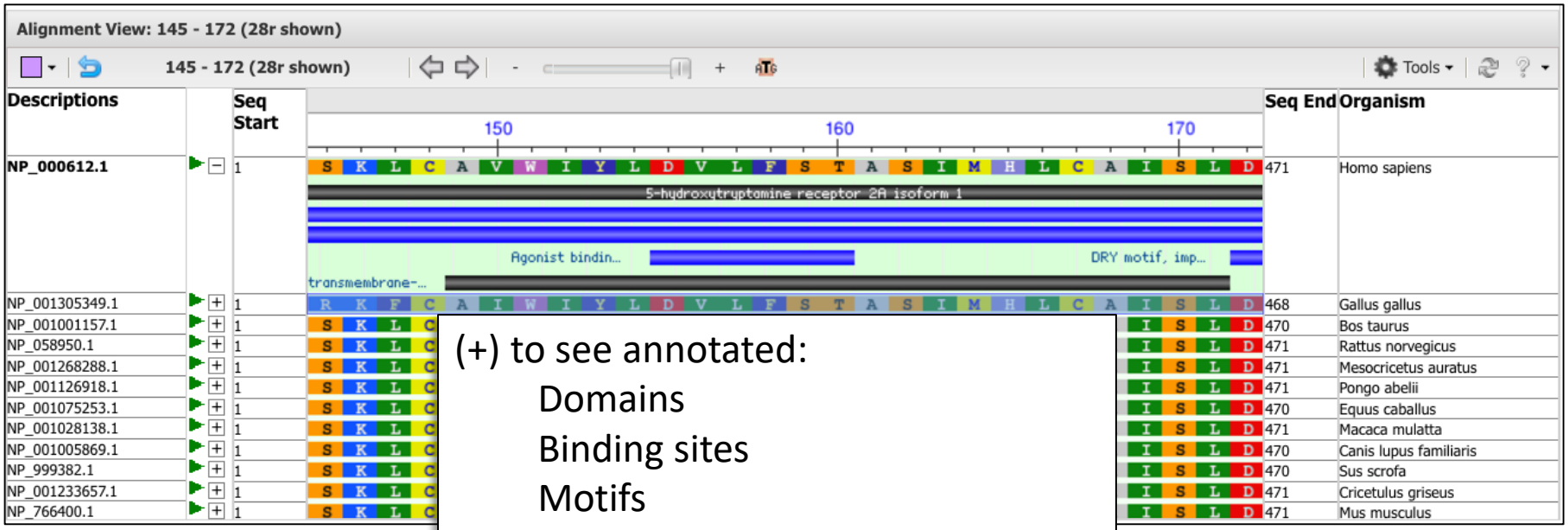
The MSA Viewer



- A general purpose alignment viewer
- Useful for assessing sequence conservation, sequencing errors or variations
- Flexible options for coloring alignments
- Plugs into NCBI alignment tools (BLAST, COBALT)
- Also accepts custom alignment data
- Easily embedded in your own web pages

CC BY NC SA

Display Feature: Annotations



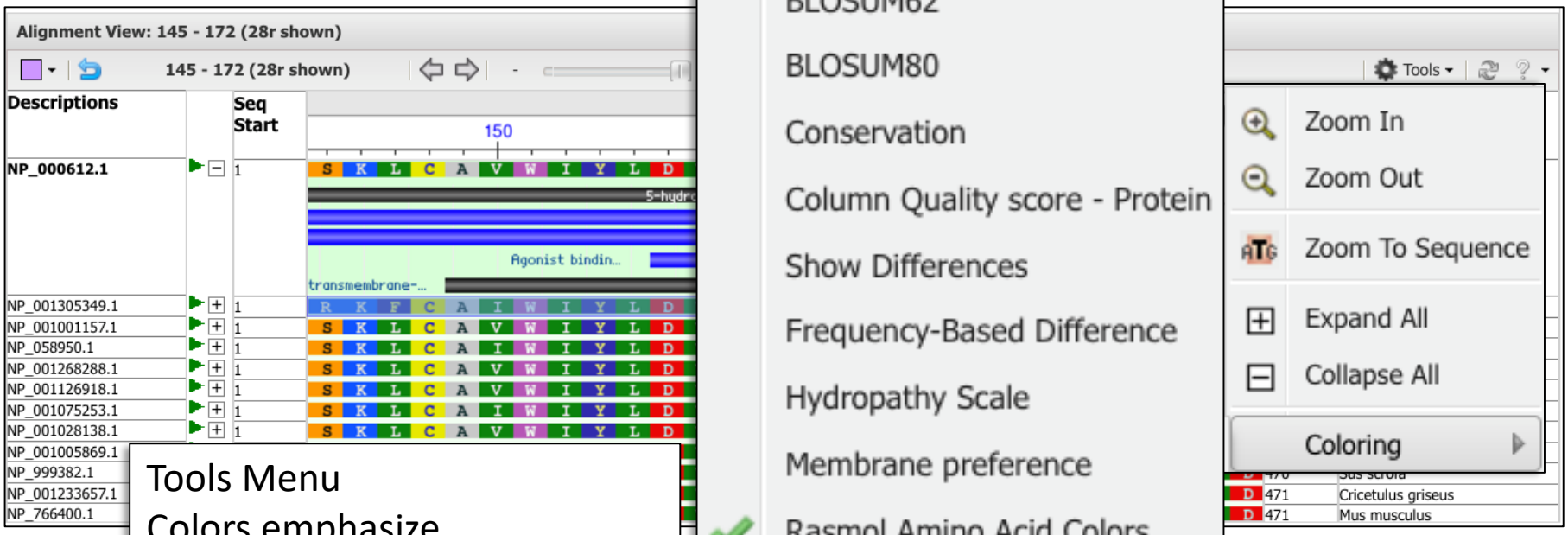
(+) to see annotated:

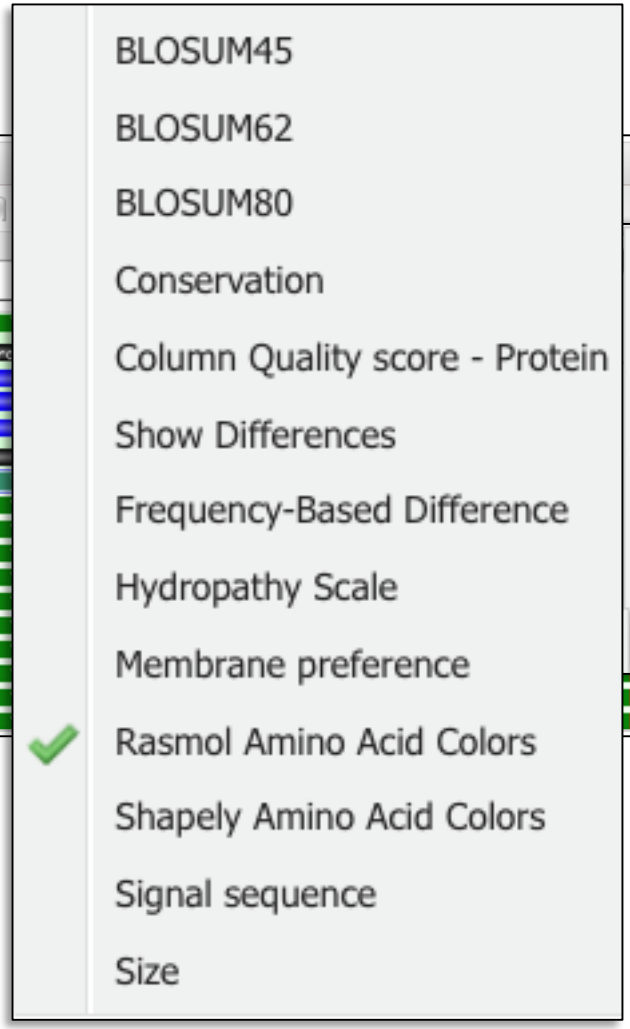
- Domains
- Binding sites
- Motifs
- Other features

Requires NCBI accessioned sequences

CC BY NC SA

Display Feature: Colors

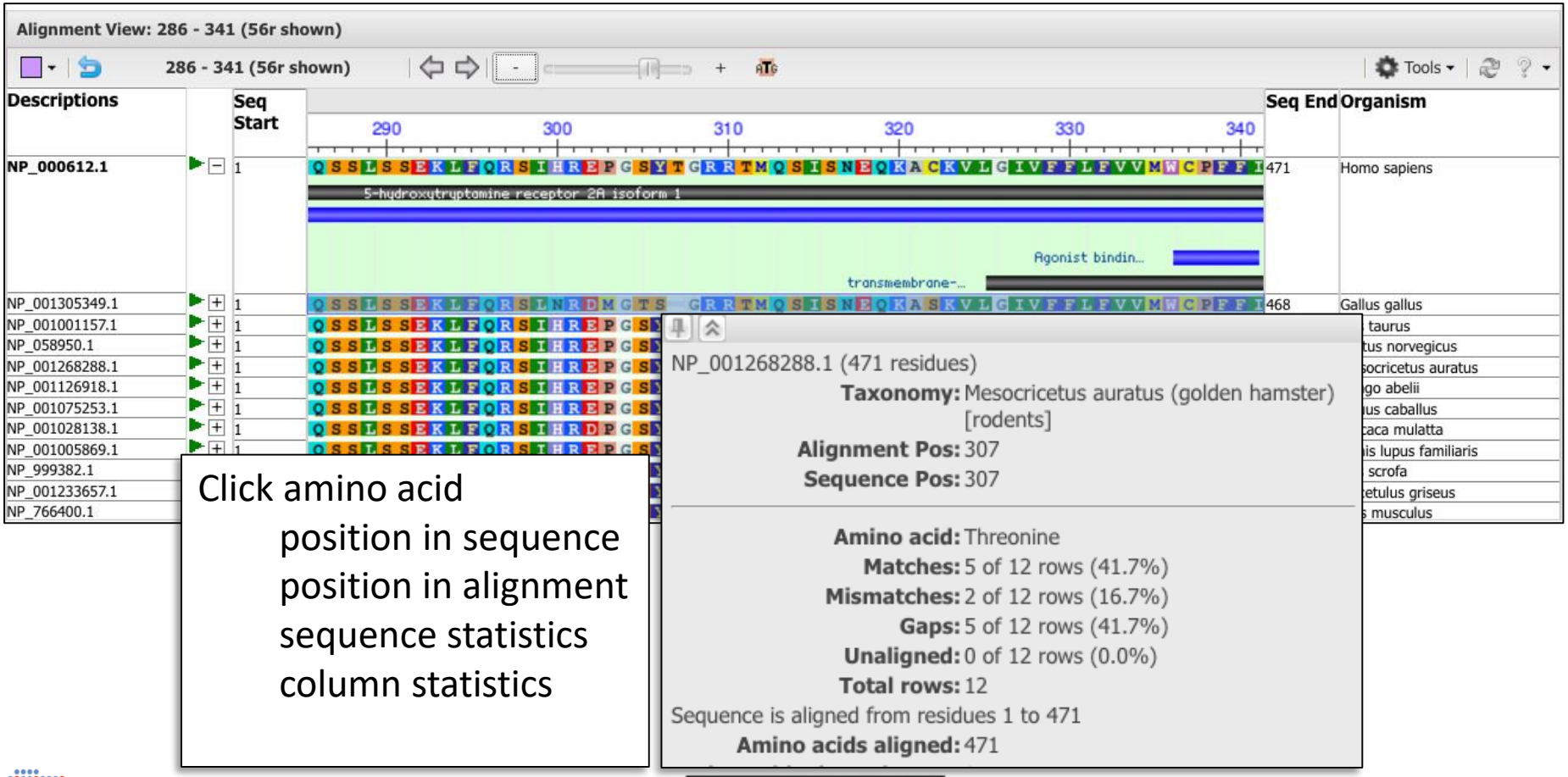




Tools Menu
Colors emphasize
- amino acid properties
- conservation
  - column average
  - compared to Master

CC BY NC SA

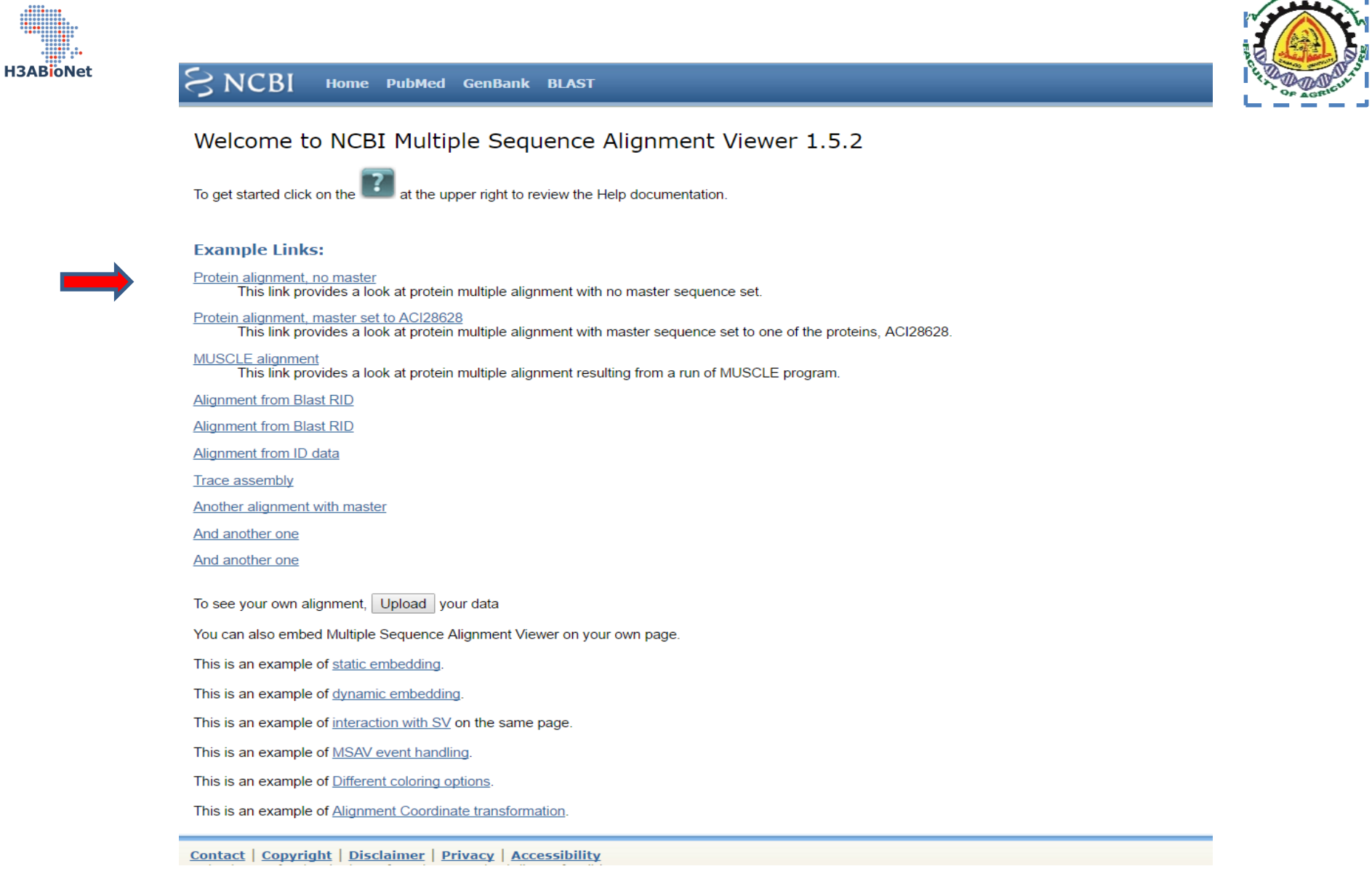
Display Feature: Residue Context

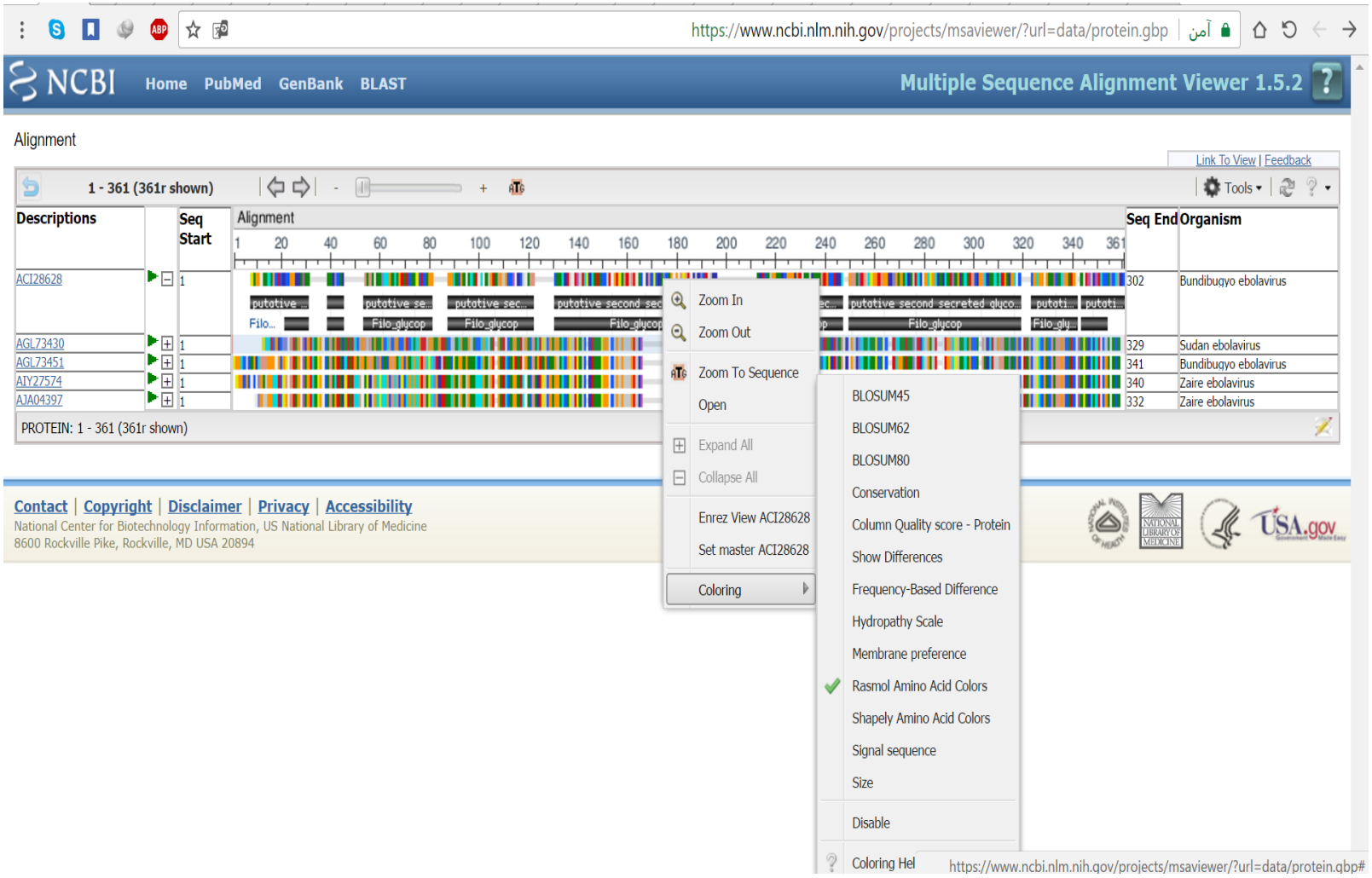


Click amino acid

- position in sequence
- position in alignment
- sequence statistics
- column statistics

CC BY NC SA


NCBI Home PubMed GenBank BLAST
Welcome to NCBI Multiple Sequence Alignment Viewer 1.5.2
To get started click on the ? at the upper right to review the Help documentation.
Example Links:
Protein alignment, no master
This link provides a look at protein multiple alignment with no master sequence set.
Protein alignment, master set to ACI28628
This link provides a look at protein multiple alignment with master sequence set to one of the proteins, ACI28628.
MUSCLE alignment
This link provides a look at protein multiple alignment resulting from a run of MUSCLE program.
Alignment from Blast RID
Alignment from Blast RID
Alignment from ID data
Trace assembly
Another alignment with master
And another one
And another one
To see your own alignment, Upload your data
You can also embed Multiple Sequence Alignment Viewer on your own page.
This is an example of static embedding.
This is an example of dynamic embedding.
This is an example of interaction with SV on the same page.
This is an example of MSAV event handling.
This is an example of Different coloring options.
This is an example of Alignment Coordinate transformation.
Contact | Copyright | Disclaimer | Privacy | Accessibility


https://www.ncbi.nlm.nih.gov/projects/msaviewer/?url=data/protein.gbp
NCBI
Home
PubMed
GenBank
BLAST
Multiple Sequence Alignment Viewer 1.5.2
Alignment
Link To View | Feedback
1 - 361 (361r shown)
Tools
Descriptions
Seq Start
Alignment
Seq End
Organism
ACI28628
AGL73430
AGL73451
AIY27574
AJA04397
302 Bundibugyo ebolavirus
329 Sudan ebolavirus
341 Bundibugyo ebolavirus
340 Zaire ebolavirus
332 Zaire ebolavirus
PROTEIN: 1 - 361 (361r shown)
Zoom In
Zoom Out
Zoom To Sequence
Open
Expand All
Collapse All
Enrez View ACI28628
Set master ACI28628
Coloring
BLOSUM45
BLOSUM62
BLOSUM80
Conservation
Column Quality score - Protein
Show Differences
Frequency-Based Difference
Hydropathy Scale
Membrane preference
Rasmol Amino Acid Colors
Shapely Amino Acid Colors
Signal sequence
Size
Disable
Coloring Hel
Contact | Copyright | Disclaimer | Privacy | Accessibility
National Center for Biotechnology Information, US National Library of Medicine
8600 Rockville Pike, Rockville, MD USA 20894
USA.gov

NCBI Home PubMed GenBank BLAST

Welcome to NCBI Multiple Sequence Alignment Viewer 1.5.2

To get started click on the ? at the upper right to review the Help documentation.

Example Links:

Protein alignment, no master
This link provides a look at protein multiple alignment with no master sequence set.

Protein alignment, master set to ACI28628
This link provides a look at protein multiple alignment with master sequence set to one of the proteins, ACI28628.

MUSCLE alignment
This link provides a look at protein multiple alignment resulting from a run of MUSCLE program.

Alignment from Blast RID

Alignment from Blast RID

Alignment from ID data

Trace assembly

Another alignment with master

And another one

And another one

To see your own alignment, Upload your data

You can also embed Multiple Sequence Alignment Viewer on your own page.

This is an example of static embedding.

This is an example of dynamic embedding.

This is an example of interaction with SV on the same page.

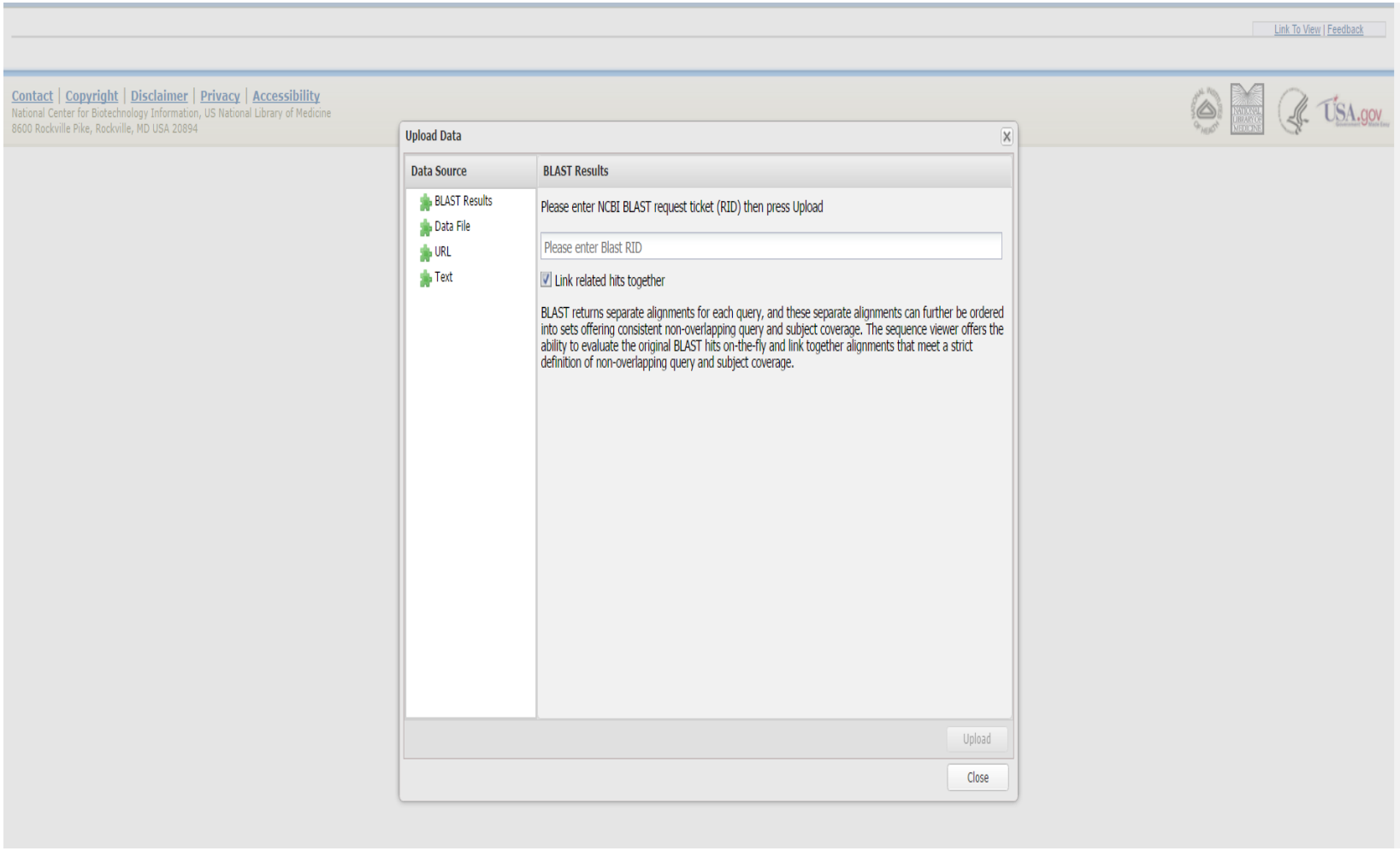
This is an example of MSAV event handling.

This is an example of Different coloring options.

This is an example of Alignment Coordinate transformation.

Contact | Copyright | Disclaimer | Privacy | Accessibility

CC BY NC SA


Link To View | Feedback
Contact | Copyright | Disclaimer | Privacy | Accessibility
National Center for Biotechnology Information, US National Library of Medicine
8600 Rockville Pike, Rockville, MD USA 20894
USA.gov
Upload Data
Data Source
BLAST Results
BLAST Results
Data File
URL
Text
Please enter NCBI BLAST request ticket (RID) then press Upload
Please enter Blast RID
Link related hits together
BLAST returns separate alignments for each query, and these separate alignments can further be ordered into sets offering consistent non-overlapping query and subject coverage. The sequence viewer offers the ability to evaluate the original BLAST hits on-the-fly and link together alignments that meet a strict definition of non-overlapping query and subject coverage.
Upload
Close

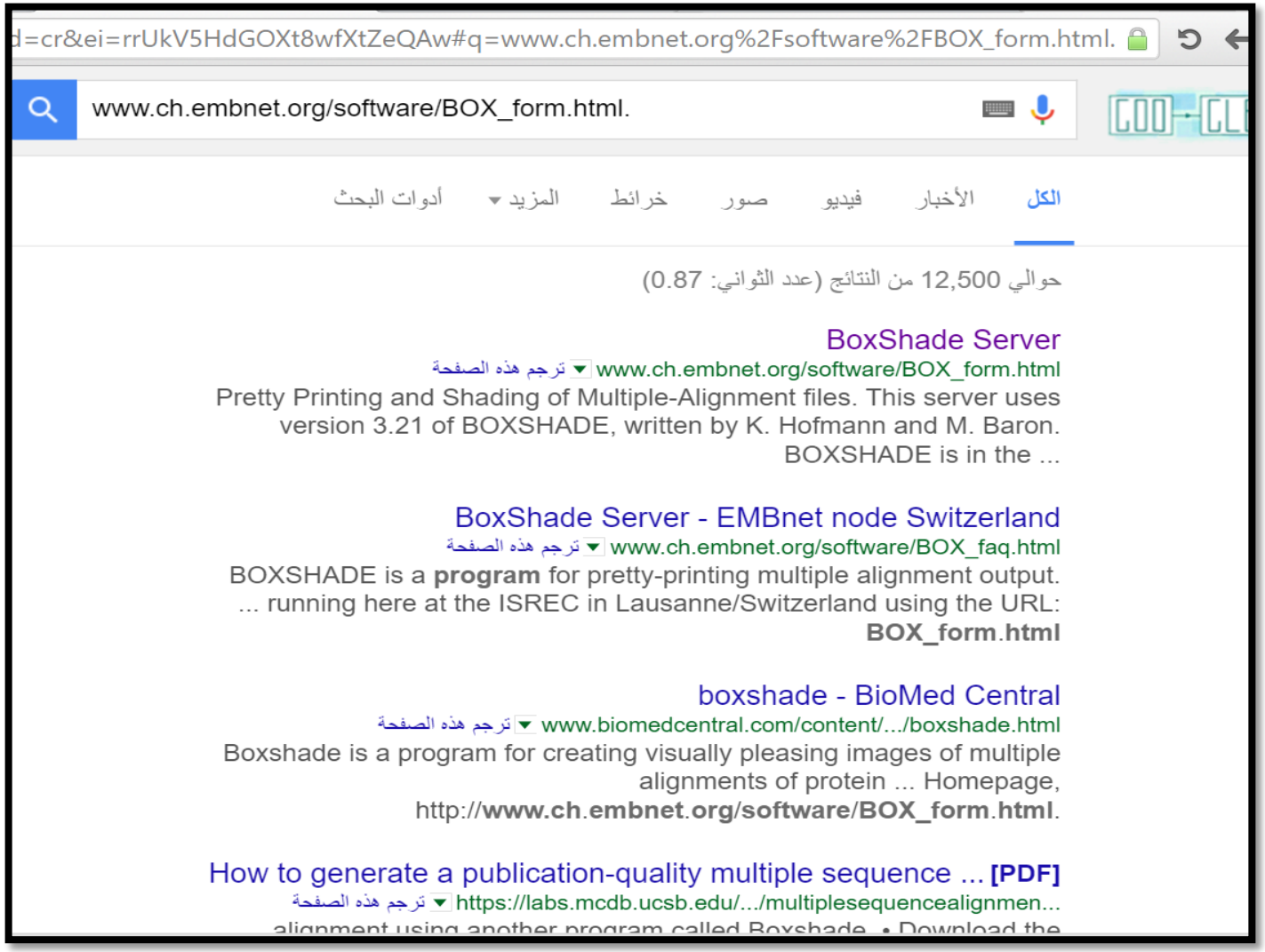
cc BY NC SA

Using Boxshade

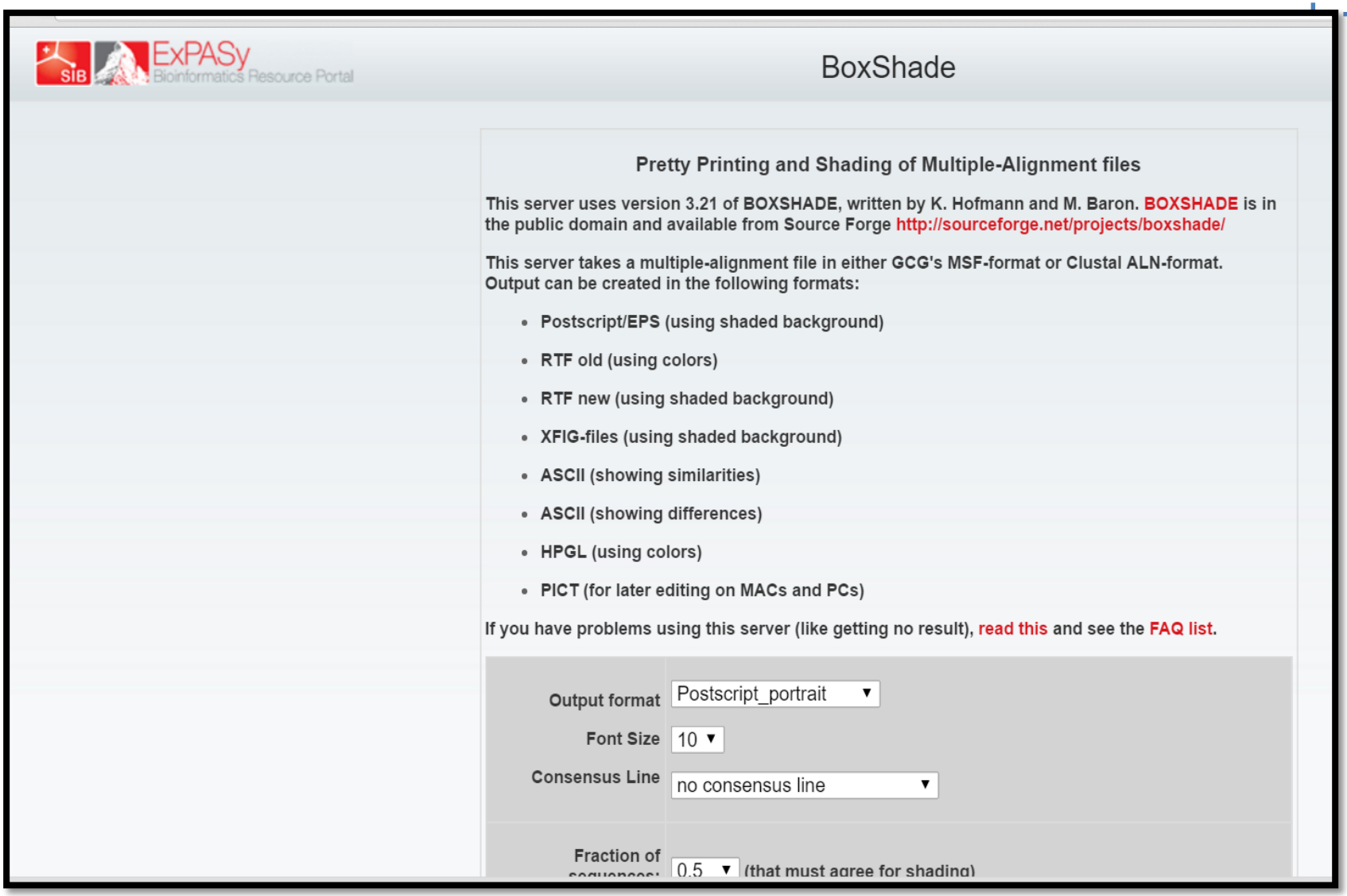
Boxshade is a utility that allows you to put some life into your alignment. It **shades columns** according to their **level of conservation** and produces files that you can easily manipulate for **inclusion** in **reports or articles**.

Claverie J, Notredame C (2007). Bioinformatics for Dummies (2nd Edn). *Wiley publishing, Inc*. 436 pp.

CC BY NC SA


d=cr&ei=rrUkV5HdGOXt8wfXtZeQAw#q=www.ch.embnet.org%2Fsoftware%2FBOX_form.html.
www.ch.embnet.org/software/BOX_form.html.
الكل
الأخبار
فيديو
صور
خرائط
المزيد
أدوات البحث
حوالي 12,500 من النتائج (عدد الثواني: 0.87)
BoxShade Server
www.ch.embnet.org/software/BOX_form.html ترجم هذه الصفحة
Pretty Printing and Shading of Multiple-Alignment files. This server uses version 3.21 of BOXSHADE, written by K. Hofmann and M. Baron. BOXSHADE is in the ...
BoxShade Server - EMBnet node Switzerland
www.ch.embnet.org/software/BOX_faq.html ترجم هذه الصفحة
BOXSHADE is a program for pretty-printing multiple alignment output. ... running here at the ISREC in Lausanne/Switzerland using the URL: BOX_form.html
boxshade - BioMed Central
www.biomedcentral.com/content/.../boxshade.html ترجم هذه الصفحة
Boxshade is a program for creating visually pleasing images of multiple alignments of protein ... Homepage, http://www.ch.embnet.org/software/BOX_form.html.
How to generate a publication-quality multiple sequence ... [PDF]
https://labs.mcdb.ucsb.edu/.../multiplesequencealignmen... ترجم هذه الصفحة

CC BY NC SA


ExPASy
Bioinformatics Resource Portal
BoxShade
Pretty Printing and Shading of Multiple-Alignment files
This server uses version 3.21 of BOXSHADE, written by K. Hofmann and M. Baron. BOXSHADE is in the public domain and available from Source Forge http://sourceforge.net/projects/boxshade/
This server takes a multiple-alignment file in either GCG's MSF-format or Clustal ALN-format. Output can be created in the following formats:
Postscript/EPS (using shaded background)
RTF old (using colors)
RTF new (using shaded background)
XFIG-files (using shaded background)
ASCII (showing similarities)
ASCII (showing differences)
HPGL (using colors)
PICT (for later editing on MACs and PCs)
If you have problems using this server (like getting no result), read this and see the FAQ list.
Output format
Postscript_portrait
Font Size
10
Consensus Line
no consensus line
Fraction of
0.5
(that must agree for shading)

CC BY NC SA

Align Example - Notepad

File Edit Format View Help

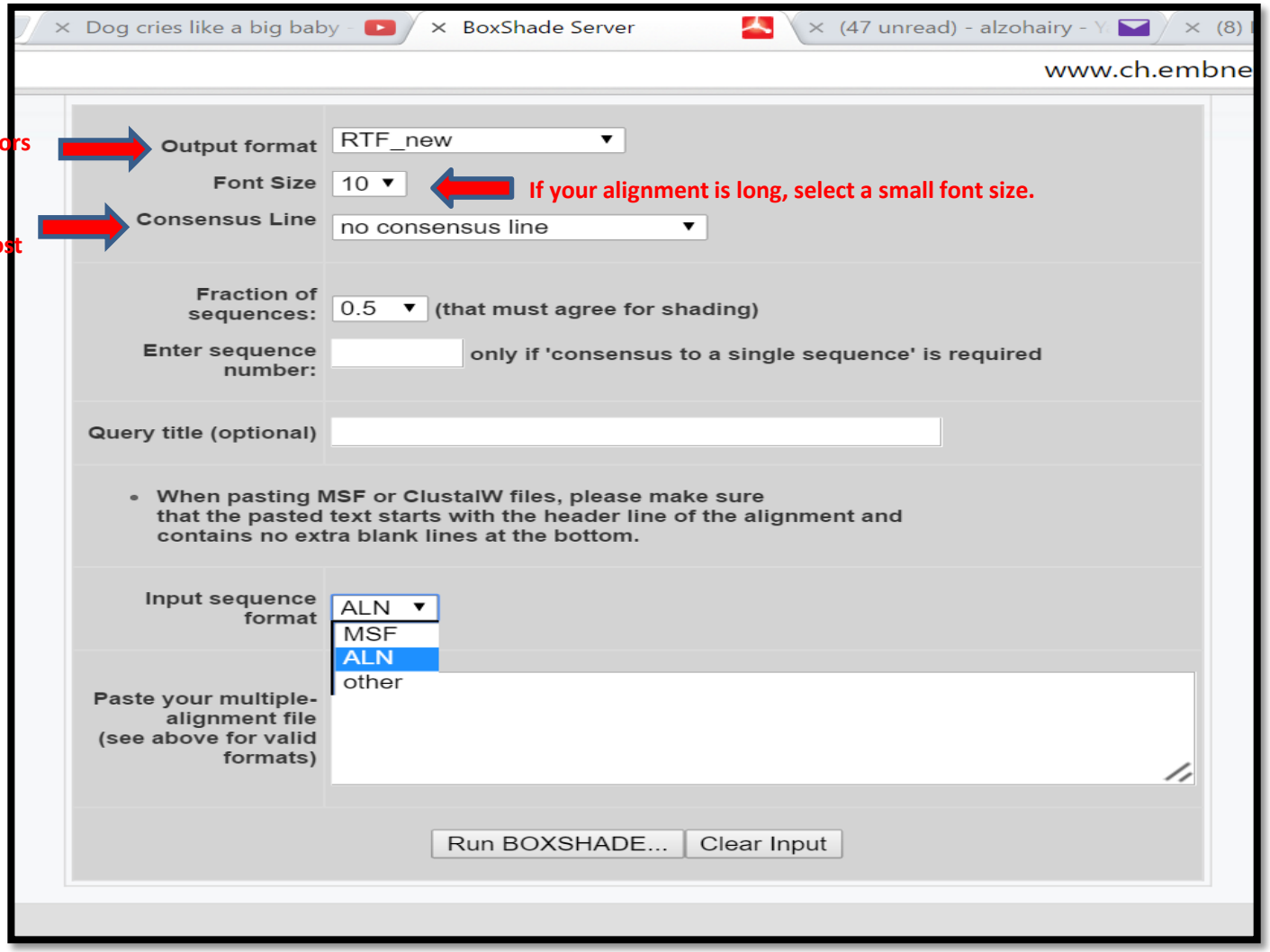
```
CLUSTAL FORMAT for T-COFFEE Version_4.26 [http://www.tcoffee.org], CPU=3.97 sec, SCORE=69, Nseq=9, Len=158

sp|P00028|CYC_LAMTR     ----------------------GDVEKGKKVFVQKCSQCHTVEKAG------KHKTGPNL
sp|P62898|CYC_RAT       ----------------------GDVEKGKKIFVQKCAQCHTVEKGG------KHKTGPNL
sp|P99999|CYC_HUMAN     ----------------------GDVEKGKKIFIMKCSQCHTVEKGG------KHKTGPNL
sp|P00036|CYC_LUCCU     ------------------GVPAGDVEKGKKIFVQRCAQCHTVEAGG------KHKVGPNL
sp|P00044|CYC1_YEAST    -----------------TEFKAGSAKKGATLFKTRCLQCHTVEKGG------PHKVGPNL
sp|P00048|CYC_NEUCR     ------------------GFSAGDSKKGANLFKTRCAQCHTLEEGG------GNKIGPAL
sp|Q96VP3|CYC_CURLU     -----------------MGFEQGDAKKGANLFKTRCAQCHTLKAGE------GNKIGPEL
sp|P00076|CYC_EUGGR     ----------------------GDAERGKKLFESRAAQCHSAQKGV-------NSTGPSL
sp|Q00499|CY550_PARVE   MKISIYATLAALSLALPAVAQEGDAAKGEKEF-NKCKACHMVQAPDGTDIVKGGKTGPNL
                                            *.  :* . *  :.  **  :           . ** *

sp|P00028|CYC_LAMTR     SGLFGRKTGQAPGFSYTD-----ANKSKGIVWNQETLFVYLENPKKYI--------PGTK
sp|P62898|CYC_RAT       HGLFGRKTGQAAGFSYTD-----ANKNKGITWGEDTLMEYLENPKKYI--------PGTK
sp|P99999|CYC_HUMAN     HGLFGRKTGQAPGYSYTA-----ANKNKGIIWGEDTLMEYLENPKKYI--------PGTK
sp|P00036|CYC_LUCCU     HGLFGRKTGQAPGFAYTN-----ANKAKGITWQDDTLFEYLENPKKYI--------PGTK
sp|P00044|CYC1_YEAST    HGIFGRHSGQAEGYSYTD-----ANIKKNVLWDENNMSEYLTNPKKYI--------PGTK
sp|P00048|CYC_NEUCR     HGLFGRKTGSVDGYAYTD-----ANKQKGITWDENTLFEYLENPKKYI--------PGTK
sp|Q96VP3|CYC_CURLU     HGLFGRKTGSVAGYSYTD-----ANKQKGIEWNHDTLFEYLENPKKYI--------PGTK
sp|P00076|CYC_EUGGR     WGVYGRTSGSVPGYAYSN-----ANKNAAIVWEEETLHKFLENPKKYV--------PGTK
sp|Q00499|CY550_PARVE   YGVVGRKIASVEGFKYGDGILEVAEKNPDMVWSEADLIEYVTDPKPWLVEKTGDSAAKTK
                         *: **  ... *: *       *:    : * .  :  :: :** ::        . **

sp|P00028|CYC_LAMTR     MIFAGIKKEGERKDLIAYLKKSTSE-------------
sp|P62898|CYC_RAT       MIFAGIKKKGERADLIAYLKKATNE-------------
sp|P99999|CYC_HUMAN     MIFVGIKKKEERADLIAYLKKATNE-------------
sp|P00036|CYC_LUCCU     MIFAGLKKPNERGDLIAYLKSATK--------------
sp|P00044|CYC1_YEAST    MAFGGLKKEKDRNDLITYLKKACE--------------
sp|P00048|CYC_NEUCR     MAFGGLKKDKDRNDIITFMKEATA--------------
sp|Q96VP3|CYC_CURLU     MAFGGLKKPKDRNDLITFLEQETK--------------
sp|P00076|CYC_EUGGR     MAFAGIKAKKDRQDIIAYMKTLKD--------------
sp|Q00499|CY550_PARVE   MTF---KLGKNQADVVAFLAQHSPDAGAEAAPAEGAAN
                         * *   *   :: *:::::
```

When pasting MSF or ClustalW files, please make sure that the pasted text starts with the header line of the alignment and contains no extra blank lines at the bottom.

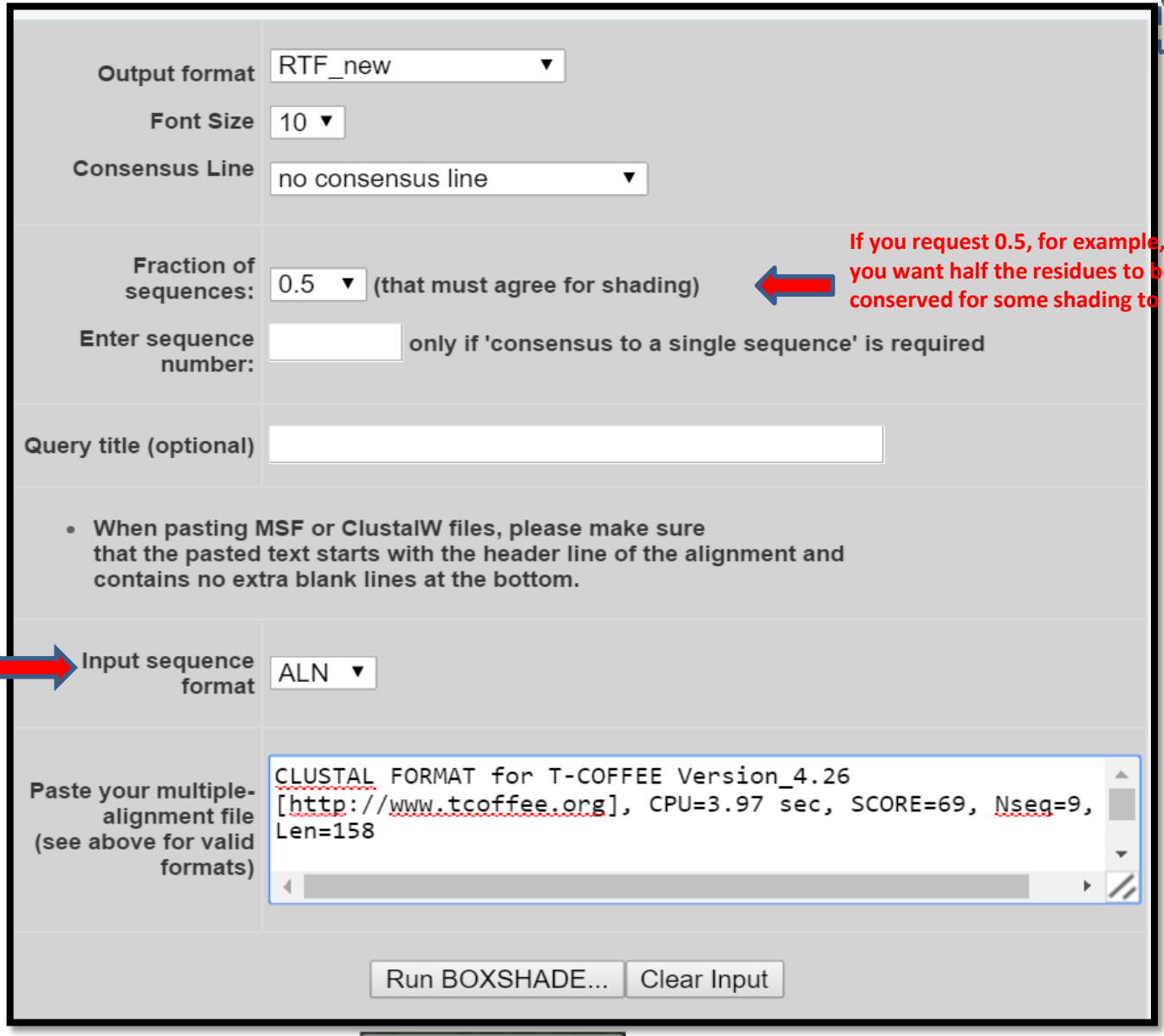
CC BY NC SA

Generates a file that most word processors (such as Microsoft Word) can read

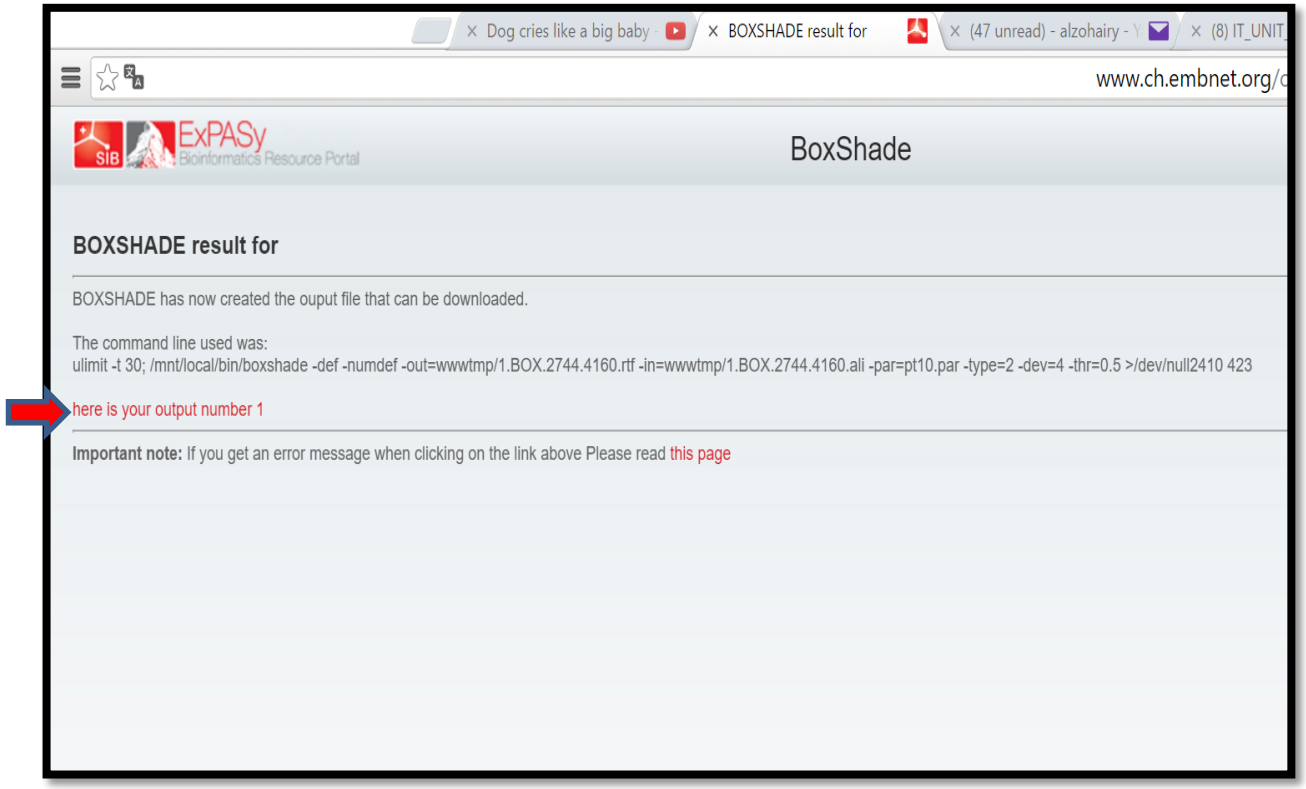
consensus sequence that contains the most common amino acid for each column.



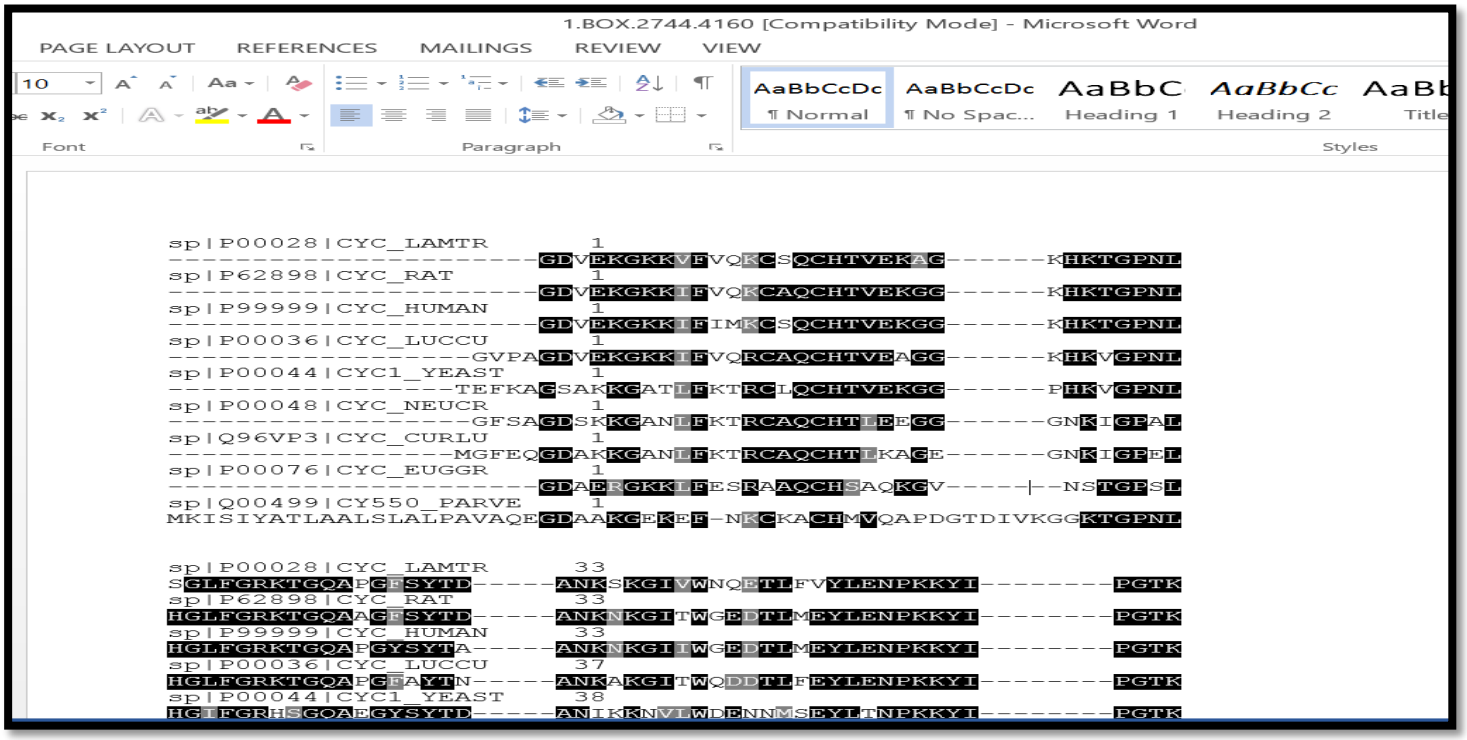
CC BY NC SA


Output format
RTF_new
Font Size
10
Consensus Line
no consensus line
Fraction of sequences:
0.5
(that must agree for shading)
If you request 0.5, for example, it means you want half the residues to be conserved for some shading to occur.
Enter sequence number:
only if 'consensus to a single sequence' is required
Query title (optional)
When pasting MSF or ClustalW files, please make sure that the pasted text starts with the header line of the alignment and contains no extra blank lines at the bottom.
Select the format of the multiple sequence alignment
Input sequence format
ALN
Paste your multiple-alignment file (see above for valid formats)
CLUSTAL FORMAT for T-COFFEE Version_4.26
[http://www.tcoffee.org], CPU=3.97 sec, SCORE=69, Nseq=9, Len=158
Run BOXSHADE...
Clear Input

CC BY NC SA


Dog cries like a big baby
BOXSHADE result for
(47 unread) - alzohairy -
(8) IT_UNIT
www.ch.embnet.org/
ExPASy
Bioinformatics Resource Portal
BoxShade
BOXSHADE result for
BOXSHADE has now created the ouput file that can be downloaded.
The command line used was:
ulimit -t 30; /mnt/local/bin/boxshade -def -numdef -out=wwwtmp/1.BOX.2744.4160.rtf -in=wwwtmp/1.BOX.2744.4160.ali -par=pt10.par -type=2 -dev=4 -thr=0.5 >/dev/null2410 423
here is your output number 1
Important note: If you get an error message when clicking on the link above Please read this page

CC BY NC SA



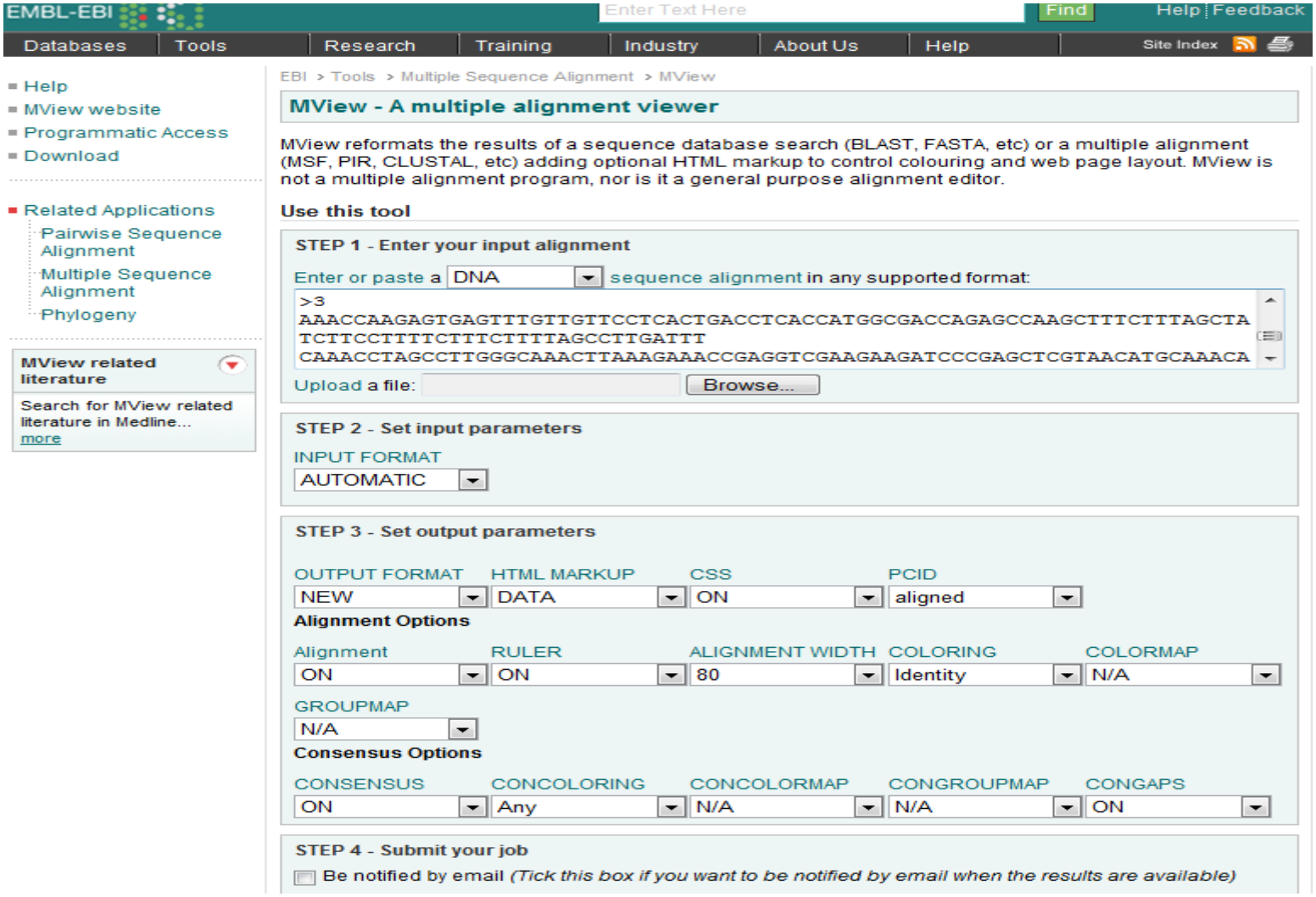
Conservation **doesn't necessarily mean** ***identity*** **in Boxshade. Similar residues, such as isoleucine and valine, also account for conservation. Two types of shading exist:**

- ***Black:*** **Identical amino acids or nucleotides**
- ***Gray:*** **Similar amino acids**

CC BY NC SA

MView for colorful MSA

CC BY NC SA



EMBL-EBI | Enter Text Here | Find | Help | Feedback

Databases | Tools | Research | Training | Industry | About Us | Help | Site Index

- Help
- MView website
- Programmatic Access
- Download

- Related Applications
  - Pairwise Sequence Alignment
  - Multiple Sequence Alignment
  - Phylogeny

MView related literature

Search for MView related literature in Medline... more

EBI > Tools > Multiple Sequence Alignment > MView

MView - A multiple alignment viewer

MView reformats the results of a sequence database search (BLAST, FASTA, etc) or a multiple alignment (MSF, PIR, CLUSTAL, etc) adding optional HTML markup to control colouring and web page layout. MView is not a multiple alignment program, nor is it a general purpose alignment editor.

Use this tool

STEP 1 - Enter your input alignment

Enter or paste a DNA sequence alignment in any supported format:

```
>3
AAACCAAGAGTGAGTTTGTTGTTCCTCACTGACCTCACCATGGCGACCAGAGCCAAGCTTTCTTTAGCTA
TCTTCCTTTTCTTTCTTTTAGCCTTGATTT
CAAACCTAGCCTTGGGCAAACTTAAAGAAACCGAGGTCGAAGAAGATCCCGAGCTCGTAACATGCAAACA
```

Upload a file: Browse...

STEP 2 - Set input parameters

INPUT FORMAT: AUTOMATIC

STEP 3 - Set output parameters

OUTPUT FORMAT	HTML MARKUP	CSS	PCID
NEW	DATA	ON	aligned

Alignment Options

Alignment	RULER	ALIGNMENT WIDTH	COLORING	COLORMAP
ON	ON	80	Identity	N/A

GROUPMAP: N/A

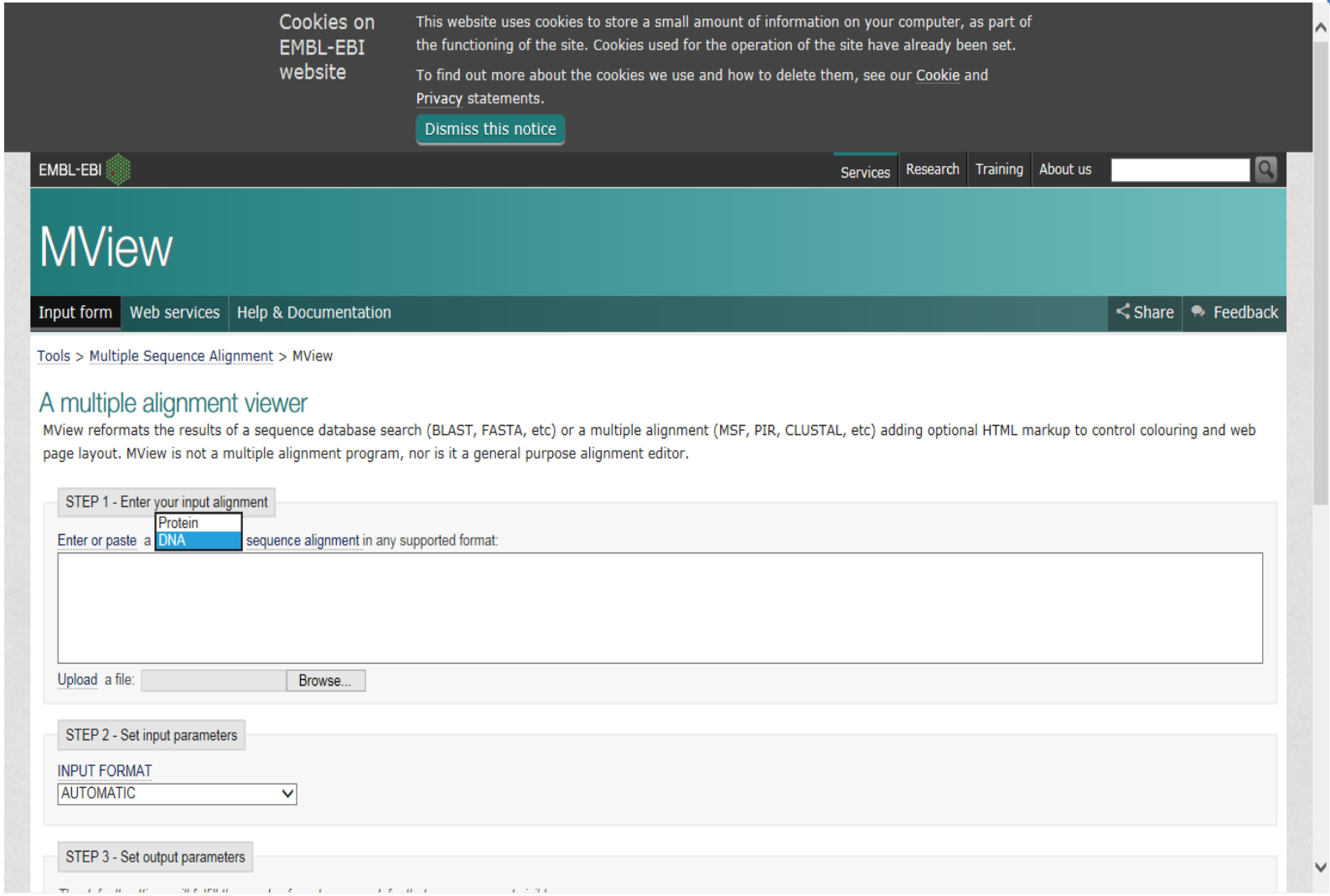
Consensus Options

CONSENSUS	CONCOLORING	CONCOLORMAP	CONGROUPMAP	CONGAPS
ON	Any	N/A	N/A	ON

STEP 4 - Submit your job

Be notified by email *(Tick this box if you want to be notified by email when the results are available)*

CC BY NC SA


Cookies on EMBL-EBI website
This website uses cookies to store a small amount of information on your computer, as part of the functioning of the site. Cookies used for the operation of the site have already been set.
To find out more about the cookies we use and how to delete them, see our Cookie and Privacy statements.
Dismiss this notice
EMBL-EBI
Services
Research
Training
About us
MView
Input form
Web services
Help & Documentation
Share
Feedback
Tools > Multiple Sequence Alignment > MView
A multiple alignment viewer
MView reformats the results of a sequence database search (BLAST, FASTA, etc) or a multiple alignment (MSF, PIR, CLUSTAL, etc) adding optional HTML markup to control colouring and web page layout. MView is not a multiple alignment program, nor is it a general purpose alignment editor.
STEP 1 - Enter your input alignment
Enter or paste a Protein DNA sequence alignment in any supported format:
Upload a file:
Browse...
STEP 2 - Set input parameters
INPUT FORMAT
AUTOMATIC
STEP 3 - Set output parameters

CC BY NC SA

- Related Applications
  - Multiple Sequence Alignment
  - Phylogeny

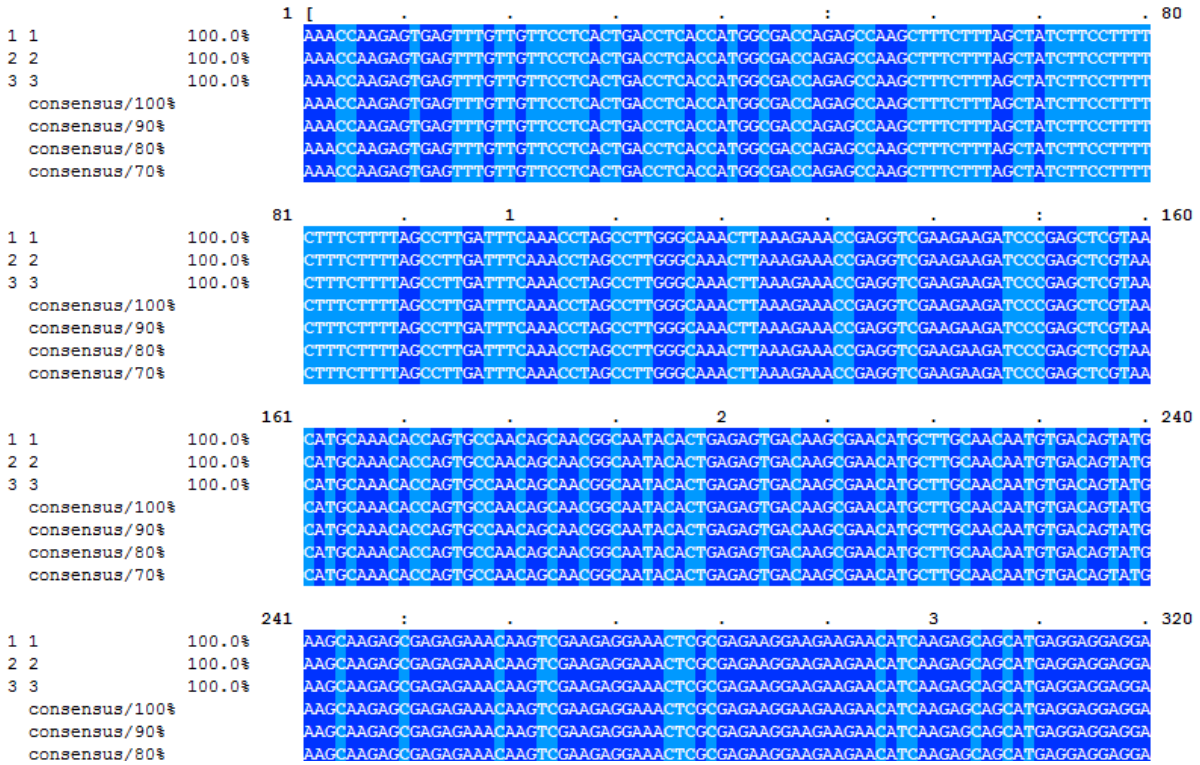
MView Results

Alignments | Submission Details | Submit Another Job

Alignment

Download Alignment File

Reference sequence (1): 1
Identities normalised by aligned length.
Colored by: identity + property



CC BY NC SA

Practical

(to try in your own time)

1. **Point your browser to** www.ch.embnet.org/software/BOX_form.html.
2. **Choose RTF_new from the Output Format drop-down menu.**
3. **Select the font size you want.**
4. **Choose Add a Consensus Line with Letters from the Consensus Line drop-down menu.**
5. **Select the fraction of sequences you would like shaded.**
6. **Select the format of the multiple sequence alignment you want to use.**
7. **Paste your multiple sequence alignment into the Sequence window.**
8. **Click the Run Boxshade button.**
9. **Click the** here is your output **link, save to a local file, then open the local file with Word.**

CC BY NC SA

BTU BIOINFORMATICS
TRAINNING UNIT
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

CC BY NC SA

By

Ahmed Mansour Alzohairy

Department of Genetics, Zagazig University, Zagazig, Egypt

CC BY NC SA