

Семинар 5. Программа MEGA. Выравнивание последовательностей на примере коровых белков (ATP, DNA-polymerase, 16S-RNA).

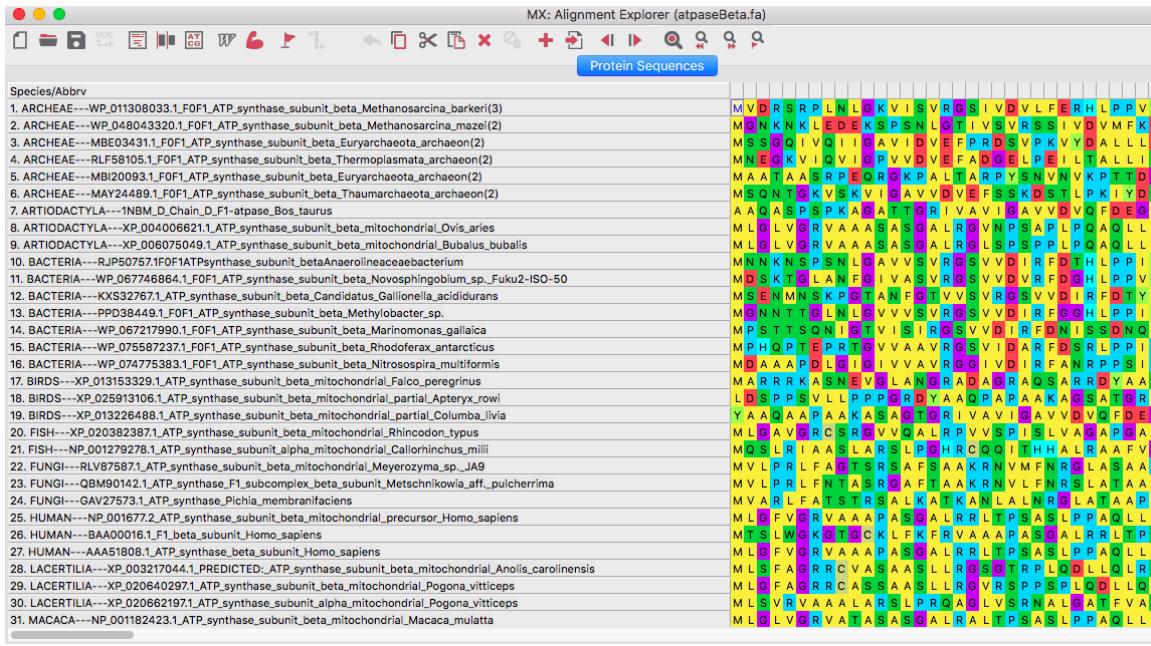
Устанавливаем пакет MEGA и выравниваем последовательности из файлов материалов семинара (субъединицы ДНК- и РНК-полимераз из разных видов по всему дереву жизни).

1. MEGA

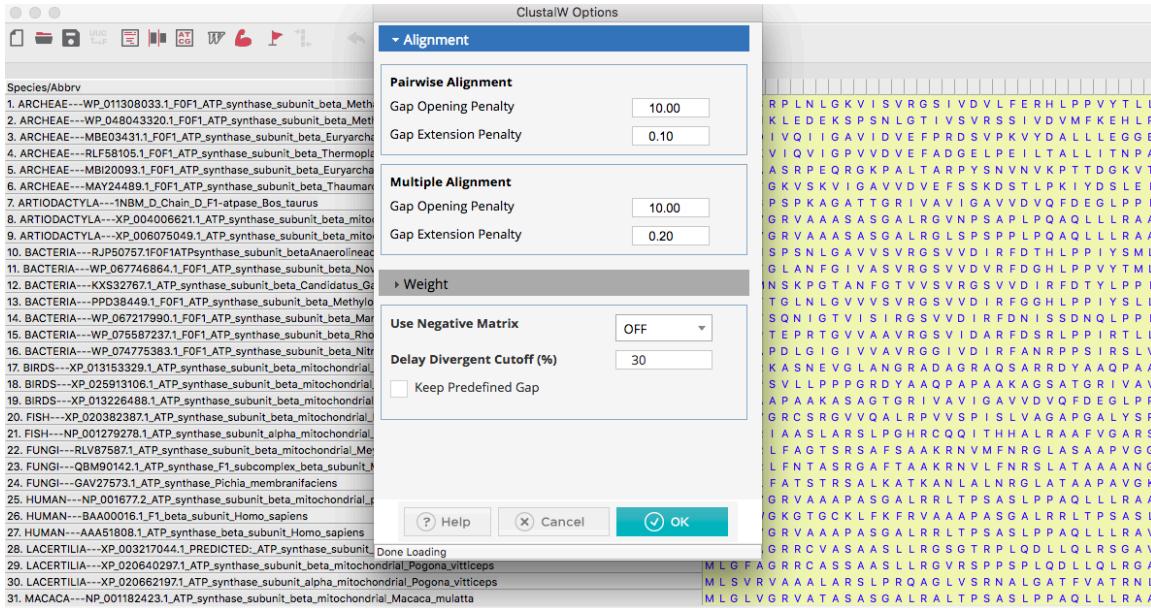
<https://www.megasoftware.net/>

The screenshot shows the official website for MEGA (Molecular Evolutionary Genetics Analysis) at <https://www.megasoftware.net/>. The page features a dark header with the MEGA logo and navigation links for 'home', 'features', 'publications', 'manual', and 'feedback'. Below the header is a decorative banner with various biological icons representing different organisms. A prominent red circular icon with a white bell contains a red 'X'. To its right, text reads: 'Volunteers are needed for testing the new macOS version of MEGA X. If you are a macOS user and would like to help, please [sign up here](#). Note: this version is only for 64-bit macOS systems and is only intended for testing.' Below this message is a bold statement: 'Sophisticated and user-friendly software suite for analyzing DNA and protein sequence data from species and populations.' At the bottom of the main content area are four dropdown menus: 'macOS', 'Graphical (GUI)', 'MEGA 7 (32-bit)', and a green 'DOWNLOAD' button with a download icon. Below these buttons are three sections: 'Sequence Analyses' (Phylogeny Inference, Model Selection, Dating and Clocks), 'Statistical Methods' (Maximum Likelihood, Distance Methods, Ordinary Least Squares), and 'Powerful Visual Tools' (Alignment/Trace Editor, Tree Explorer, Data Explorers).

Открываем fasta-file из файлов семинара (например, с субъединицей бета ATP-beta)

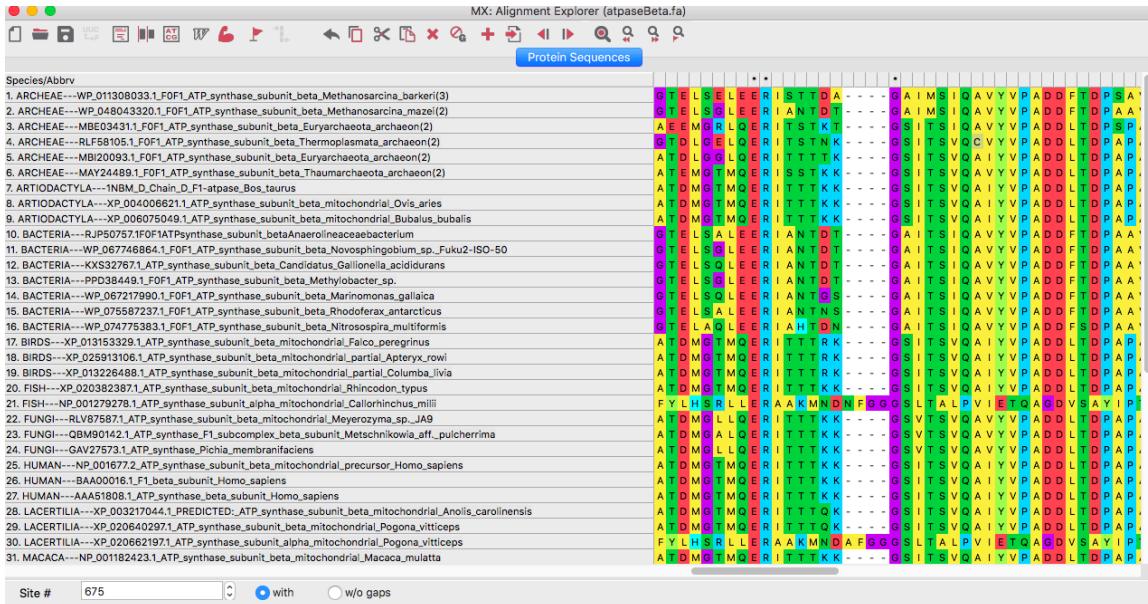


В меню Alignment -> align with ClustalW (или with Muscle – это два разных алгоритма выравнивания)



Не меняем параметры – используем по умолчанию

Получаем итоговое выравнивание



1. Файл CytBProt содержит аминокислотные последовательности белков цитохрома В митохондриального генома 16 видов позвоночных. Последовательности помечены названием вида. Взгляните на эти последовательности.

- Какой формат используется для представления последовательностей?
- Потребуются ли очень длинные гэпы для выравнивания этих последовательностей?
- Почему?

2. Используйте Clustal в MEGA или на EBI, чтобы выполнить множественное выравнивание этих белковых последовательностей. Сначала выполните выравнивание с помощью матрицы BLOSUM, а затем с помощью матрицы идентичности.

- Каким образом использование разных матриц оценок влияет на ваше выравнивание?
- Можете ли вы идентифицировать консервативные области, длина которых превышает 10 аминокислот?

3. Файл CytBDNA содержит нуклеотидные последовательности белков цитохрома В того же вида. Постройте выравнивание последовательностей ДНК (выравнивание ДНК займет больше времени).

Не забудьте установить тип последовательности - ДНК .

- Получили вы то, что ожидали?
- Что увидели неожиданное (если есть)?
- Например, есть ли в последовательностях промежутки, и если да, то насколько они велики? Почему?
- Какие параметры можно изменить, чтобы исправить возможную ошибку?

4. В чем разница между выравниванием ДНК и белков? Как это объяснить?
5. Давайте теперь вернемся к выравниванию цитохрома В и посмотрим на последовательность белка цитохрома В из другого царства, например, из *Arabidopsis thaliana*. Найдите эту последовательность в NCBI (последовательность белка CAA47966.1). Постройте выравнивание последовательностей позвоночных с этой последовательностью растений.
 - Сохраняются ли регионы, которые вы ранее определили как консервативные?
 - Изучите запись Entrez для CAA47966.1. Можете ли вы сделать какие-либо выводы о функциональных свойствах консервативных областей?
6. Гексокиназы - это ферменты, фосфорилирующие гексозу (в основном глюкозу). После фосфорилирования сахар может вступить в некоторые внутриклеточные метаболические процессы.

Файл Hexokinases.fasta содержит аминокислотные последовательности гексокиназ человека и собаки. Выполните выравнивание. Посмотрите на таблицу результатов и изучите попарные расстояния между последовательностями (например, гексокиназа_1 человека и гексокиназа_1 собаки, гексокиназа_2 человека и гексокиназа_2 собаки и т. д.).

- Какие пары последовательностей больше всего похожи друг на друга?
7. Основываясь на этих ограниченных наборах данных, является ли эволюция гексокиназ или цитохрома В быстрее?

Использованы материалы

<https://tbb.bio.uu.nl/BPA/COO/COO1.html>

<https://teaching.healthtech.dtu.dk/>