## 1. Цели

- ullet Сравнить результаты работы алгоритмов выравнивания ClustalW и Muscle на примере Ribosomal protein L7
- Сравнить топологии деревьев, полученные в результате различных методов построения
- Сравнить бутстрэп значения
- Сравнить деревья с принятыми деревьями видов

## 2. Скачивание данных

#### О данных

Данные используемые в данной работе взяты с сайта Национального центра биотехнической информации США NCBI.

Данные об исследуемом белке могут быть найдены через поисковит белков NCBI Также доступна документация.

#### Скачивание

Скачаем последовательности в формате FASTA с NCBI:

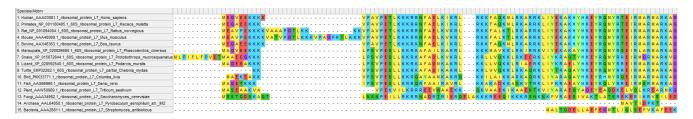
- (1) Человек Homo sapiens
- (2) Обезьяна Macaca mulatta
- (3) Крыса Rattus norvegicus
- (4) Мышь Mus musculus
- (5) Копытное Bos taurus
- (6) Сумчатое Phascolarctos cinereus
- (7) Змея Protobothrops mucrosquamatus
- (8) Ящерица Podarcis muralis
- (9) Черепаха Chelonia mydas
- (10) Птица Columba livia
- (11) Рыба Danio rerio
- (12) Растение Triticum aestivum
- (13) Пекарские дрожжи Saccharomyces cerevisiae
- (14) Apxeя Pyrobaculum aerophilum
- (15) Бактерия Streptomyces antibioticus

# 3. Обработка данных

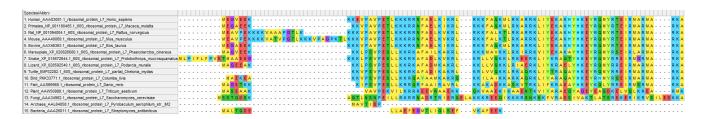
#### Выравнивание

Выравнием последовательности методом ClustalW и Muscle, получим выровненные последовательности  $\frac{\text{ClustalW}}{\text{UNSCle}}$  и  $\frac{\text{Muscle}}{\text{ClustalW}}$ 

#### Заметим, что они различны:



#### . ClustalW



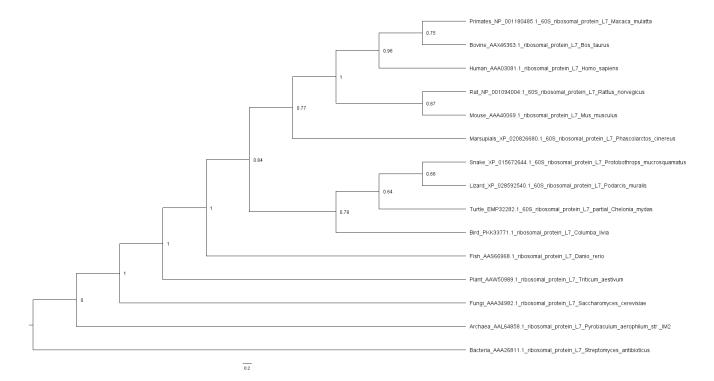
. Muscle

### Деревья

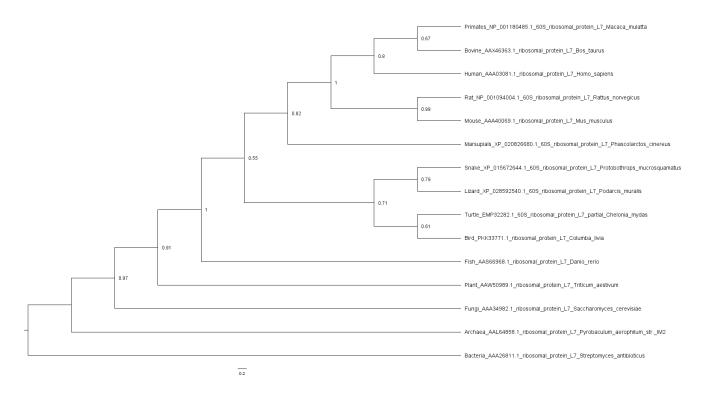
Построим деревья, используя MEGA X

Выставим параметры Test of Phylogeny – Bootstrap method; No. of Bootstrap Replications – 100. Затем экспортируем полученные деревья в Newick и отформатируем их в Fig tree. Для каждого из деревьев приведена ссылка на исходный Newick.

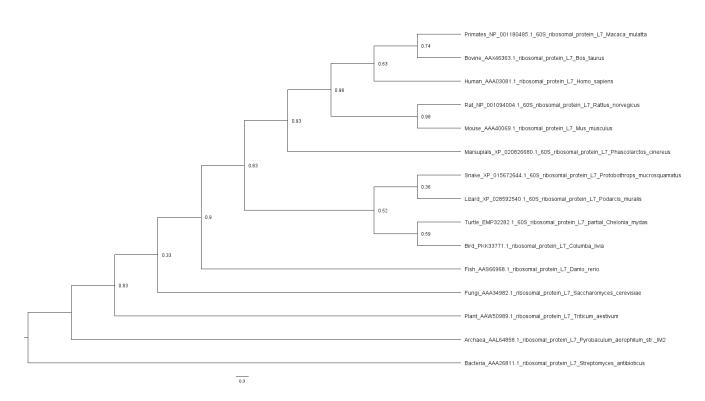
#### ClustalW



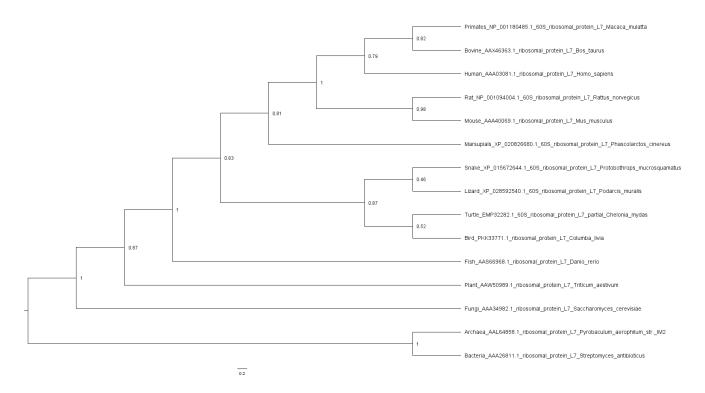
. UPGMA



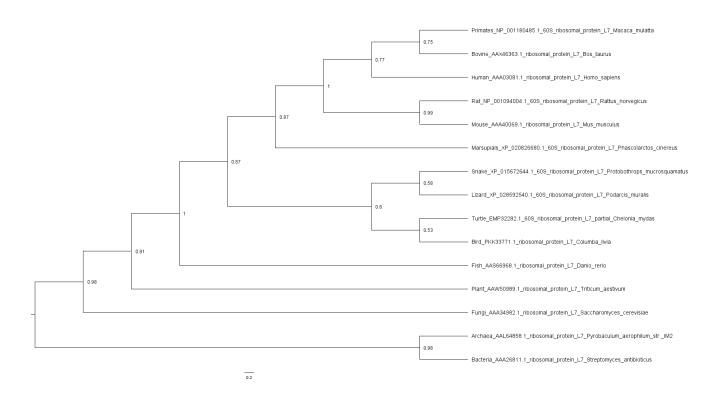
## . Neighbor Joining



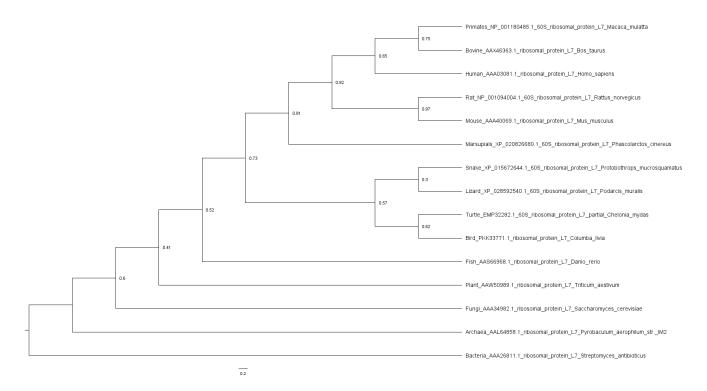
. Maximum Likelihood



#### UPGMA



Neighbor Joining



. Maximum Likelihood

# 4. Выводы

- Muscle, так как он помимо глобального выравнивания, выравнивает локально Но при построении эти различия выявились только в методе UPGMA. Возможно, при большем количестве видов и большем значении bootstrap различия будут более заметными.
- Нет, топология деревьев различается. Дерево, полученное в результате сортировки ClustalW и метода построения UPGMA отличается от всех остальных расположением птицы(Columba livia) и черепахи(Chelonya mydas). Все остальные деревья изоморфны
- Нет, значения различны.
- Нет, не совпадают, например во всех деревьях копытное к примату ближе, чем человек.