

Отчет по практикуму 3 четвертого семестра

Дорофеева Вероника, группа 201

1. Реконструкция дерева по нуклеотидным последовательностям

В рамках практической работы была проведена реконструкция филогенетического дерева для ранее выбранных животных на основе последовательностей митохондриального гена 12S rRNA. Сначала для каждого организма были найдены записи в базе ENA, содержащие информацию о полном митохондриальном геноме. Поиск выполнялся либо через ссылки со страниц Uniprot, посвящённых цитохрому B соответствующего организма, либо с использованием Advanced Search на сайте ENA по базе Nucleotide sequences с указанием нужного таксона и митохондриальной кольцевой ДНК.

Из найденных записей были определены координаты гена 12S rRNA, после чего последовательности извлекались при помощи программы seqret. Например, для *Canis lupus* использовалась следующая команда:

```
seqret embl:MH746950[70:1023] CANLU_12S_rRNA.fasta
```

Полученные последовательности были объединены в один FASTA-файл, переименованы и выровнены с помощью программы muscle со стандартными параметрами:

```
muscle -align all_12S_rRNA_mnemo.fasta -output 12S_alignment.fasta
```

Далее выравнивание было преобразовано в формат phylip-relaxed при помощи специального скрипта. На заключительном этапе с использованием программы IQ-TREE было реконструировано филогенетическое дерево по последовательностям 12S rRNA. Программа автоматически определяла тип данных как нуклеотидное выравнивание:

```
iqtree -s 12S.phy
```

Дерево визуализировано в iTOL.

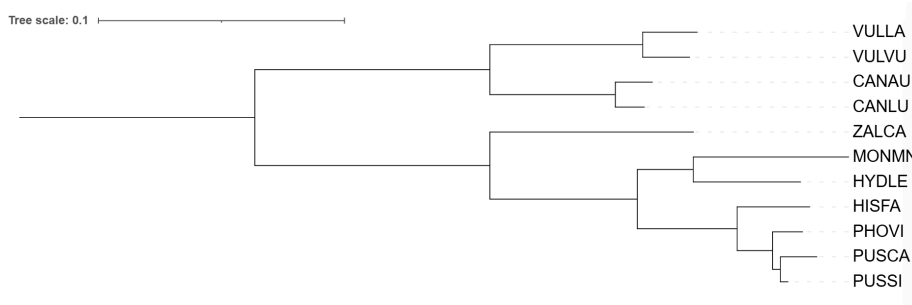


Рис 1. Реконструкция дерева по последовательностям 12S рРНК (малая РНК митохондриальных рибосом) программой IQtree, переукоренённое относительно ветви, разделяющей *Canidae* и *Pinnipedia*.

Программа автоматически выбрала модель эволюции TIM2+F+R2:

- TIM2 (Transition Model 2) - модель, предполагающая разные частоты транзиций (A-G и C-T) и транспозиций
- +F - означает, что программа не предполагает равного распределения (25% A, 25% C, 25% G, 25% T), а берет эмпирические частоты оснований напрямую из анализируемого участка ДНК (A: 0.366, C: 0.229, G: 0.176, T: 0.228)
- +R2 - Это модель гетерогенности скоростей по сайтам с двумя категориями. Вместо обычного гамма-распределения, эта модель позволяет каждой нуклеотидной позиции эволюционировать со своей собственной скоростью, разбивая все сайты выравнивания на определенные группы (в данном случае на 2 дискретных класса).

Программа использует алгоритм максимальной экономии для построения начального дерева, для оптимизации NNI (Nearest Neighbor Interchange) и ML (Maximum Likelihood).

Отличный результат, сравнивая с [прак2](#). У нас тут дерево практически полностью повторяет полученное в [практикуме1](#). За исключением разве что клады HISFA и PHOVI, которую ошибочно не объединили. Эволюционные расстояния тут смотрятся куда симпатичнее полученных на цитохроме B. Интересно, что так сложилось, хотя выравнивание получилось очень качественным.

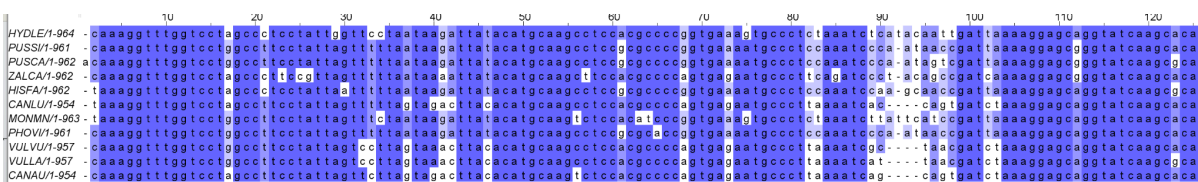


Рис 2. Выравнивание последовательностей программой muscle, открытое в программе Jalview и окрашенное по проценту идентичности.

2. Укоренение во внешнюю группу

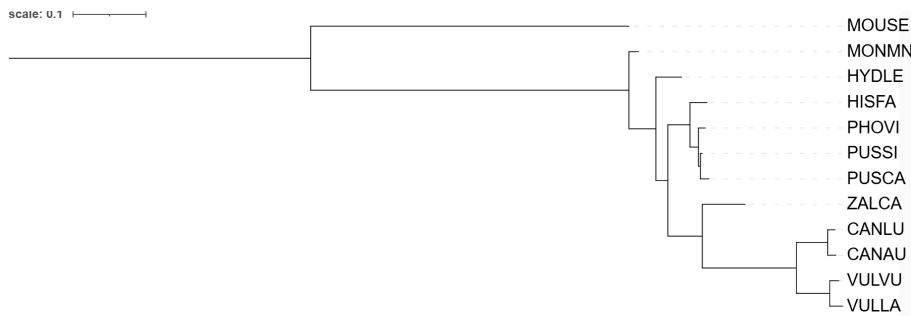


Рис 3. Реконструкция филогенетического дерева программой IQtree по последовательностям малой РНК митохондриальных хромосом (12S rRNA). Укоренение по внешнюю группу.

Теперь построим дерево (рис3), добавив к нашим родственникам внешнюю группу(мышь). Таксономия для нее представлена ниже в таблице. Из хорошего - дерево правильно укоренилось. В остальном беды, особенно с эволюционными расстояниями. Почему-то тюлени вдруг стали древнее волков и лис, хотя в практикуме 2 было показано обратное. Топология тоже странная, правильная только ветвь с волками и лисами.

Таблица 1. Таксономия для *mus musculus*

Вид	Таксономия	Мнемоника
<i>Mus musculus</i>	Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus	MOUSE

3. Бутстреп

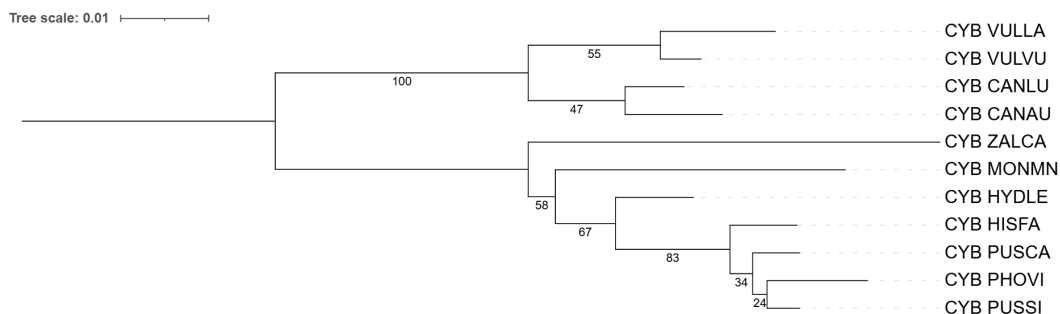


Рис. 4 Реконструкция филогенетического дерева *Caniformia* программой fastME с оценкой эволюционных расстояний с помощью модели MtREV по аминокислотным последовательностям цитохрома b. Цифрами под ветвями обозначена их бутстреп-поддержка.

Дерево повторяет полученное в [прак2](#)(рис3). Как видим, почти все числа отражают достоверность полученных ветвлений и для ошибочно построенных действительно

числа меньше. Как например, для листа, отделяющего *monachus monachus* или для *pusa sibirica*, *phoca vitulina* и *pusa caspica*. Однако, например, для неправильного ветвления получили значение 58. Это может быть связано с тем, что мы рассматриваем очень близкородственные организмы.