# Дедлайн 18:10 01.03.22

# ДЗ №3 BLAST

**1**. Программой **BLAST** найдите в банке **Swiss-Prot** два белка с последовательностями, наиболее сходными с вашей. Если BLAST найдёт не одно, а больше выравниваний с одним и тем же белком, отметьте это и приведите характеристики обоих выравниваний.

Для каждой из двух лучших находок приведите:

* ID и AC в Swiss-Prot,
* организм, из которого взят белок,
* проценты идентичных и сходных (Positives) остатков,
* длину участков в обоих белках
* число гэпов
* вес (Score): обычный (в скобках) и в битах
* E
* процент покрытия вашего белка выравниванием

Кроме этого

* сколько всего находок;
* сколько находок с E-value менее одной тысячной;
* каково максимальное E-value находки, выданное программой;
* чем оказался лимитирован список находок: порогом на E-value или максимальным числом находок? Если список лимитирован максимальным числом находок, повторите поиск, увеличив максимальное число; постарайтесь добиться того, чтобы находок было меньше заданного максимума; в отчёте опишите, что делали и что получили (в частности, каково теперь максимальное E-value)

## **2.** Измените длину слова с **6** на **2** и повторите поиск.

## Опишите отличия результата (если они есть), по тем же пунктам:

* общее число находок
* число находок с E-value<0.001
* максимальное E-value.

Для следующего задания: возьмите одну из находок данного поиска с достаточно хорошим (<0.001), **но ненулевым** E-value. Зайдите в соответствующую запись Swiss-Prot и спишите таксономию организма, из которого взят белок. Выберите один из достаточно "высоких" таксонов (но не самый высокий; лучше всего второй сверху).

Запишите E-value и Score (обычный и битовый) выравнивания с этой находкой.

## **3.** Повторите предыдущий поиск, ограничив его выбранным таксоном (**банк тот же, длину слова оставьте равной 2**). Найдите выбранную находку и сравните её E-value и Score при этом поиске с E-value и Score, полученными при предыдущем поиске. Сделайте вывод о количестве последовательностей белков из данного таксона в банке.

Может ли из-за изменения объема поиска измениться Score (обычный или битовый)?