**Домашнее задание 6**

**Выбор геномов**

Выполнил выбор геномов Rattus norvegicus GCF\_036323735.1 и Caenorhabditis elegans GCF\_000002985.6.

Эти эукариотические организмы достаточно далеки друг от друга, поэтому решил использовать домен, связанный с репликацией ДНК (например, PF00476 для домена DNA polymerase).

Провёл поиск похожих доменов в геномах и получил списки белков, выбрав по 2 лучших совпадения.

Для **Rattus norvegicus**:

* DNA polymerase subunit gamma-1 NP\_445980.2
* DNA polymerase theta NP\_001099348.1

Хотя наилучшие совпадения были с HAUS augmin-like complex, решил его не использовать, так как этот белок участвует в работе микротрубочек, а выравнивание с полимеразами не дало бы интересных результатов.

Для **Caenorhabditis elegans**:

* DNA-directed DNA polymerase NP\_496592.1
* DNA polymerase theta NP\_001367994.1

Далее провёл выравнивание четырёх последовательностей с помощью Clustal Omega.

Команда hmmemit PF00476.hmm выдала ожидаемую консенсусную последовательность, которую добавил в выравнивание для поиска совпадений.

>DNA\_pol\_A-sample1

SSLSKELGATAQRRLQAKAAMTYMVRVPQFSLLEPKQYGQGLFYQLRIEAPESTDIAFST

DDSALSLLKKEFIISMILLQHYKSADYDAKVASFYKQRLTKECKTEISFAGAATNRLSSP

NINLKKITIDTETNARYRVVLIPKNVVETAIDYGQVELRDLAAQGGAPRMIRMYLQREDL

HRFLARRIVQTNQVAALVEGAKTVSYGLVYGGTAMGIGKEDHTMQEHADNGLGDFVETYD

MIKKVFMFDEFDVKHDQGYQETTFGRGQYNKFAQRITAVVGSVRIYGAINNEDLQGTVAS

IIRVGLIEIAAFLENKDYQRAVYLSQVHDNGVHSMQGRIVIAYTKLVRYHMEEEHDAYIT

VILRWEDDLGGCYDE