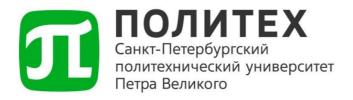
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО» ВШ программной инженерии



Отчет по дисциплине «Программирование биоинформатических приложений на суперкомпьютере»

Выполнил

студент гр. № 3540202/00201

Е.В. Ковальчук

Руководитель

А.П. Маслаков

Санкт-Петербург 2021 г.

Оглавление

Постановка задачи	3
Выбор сборки генома коронавируса	
Получение белков генома и расчет молекулярной массы	∠
Построение модели белка	5
Заключение	7
Список использованных источников	5

Постановка задачи

В рамках работы были поставлены следующие задачи:

- 1. Выбрать любую сборку генома коронавируса;
- 2. Получить все возможные белки, содержащиеся в геноме;
- 3. Подсчитать для каждого белка молекулярную массу;
- 4. Получить трехмерную модель одного из найденных белков.

Для решения основных задач необходимо ознакомиться и решить задачи в Rosalind [1]:

- 1. Расшифровка DNA в RNA;
- 2. Перевод RNA в белок;
- 3. ORF и другие.

Выбор сборки генома коронавируса

На сайте [2] представлена таблица, была выбрана первая запись данной таблицы. После выбора сборки генома (рис. 1) необходимо перейти во вкладку fasta для получения самого генома.

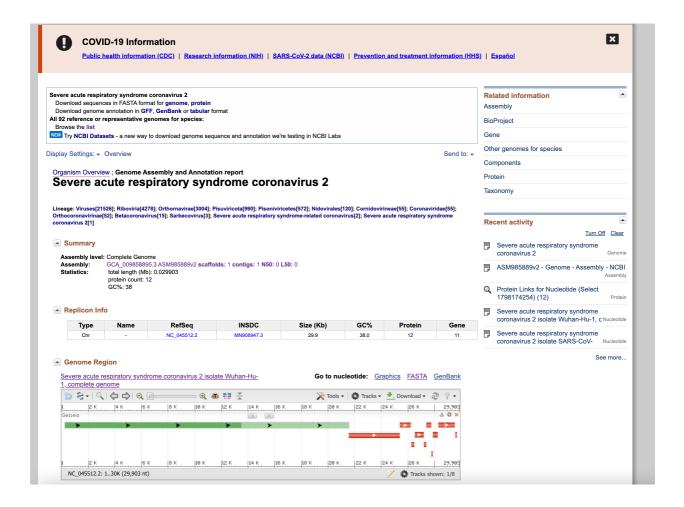


Рисунок 1.

Получение белков генома и расчет молекулярной массы

На данном этапе необходимо найти белки, которые содержатся в геноме, и их молекулярную массу.

Каждому символу белка соответствует некоторое значение. Таким образом для получения молекулярной массы белка необходимо сложить значения символов, входящих в этот белок.

Большинство геномов, в том числе геном человека и геномы всех остальных клеточных форм жизни, построены из ДНК.

Был разработан алгоритм, который находит все возможные белки и для каждого из них находит молекулярную массу. Часть результата алгоритма представлена на рисунке 2, где первая строчка - это найденный белок, вторая - молекулярная масса.

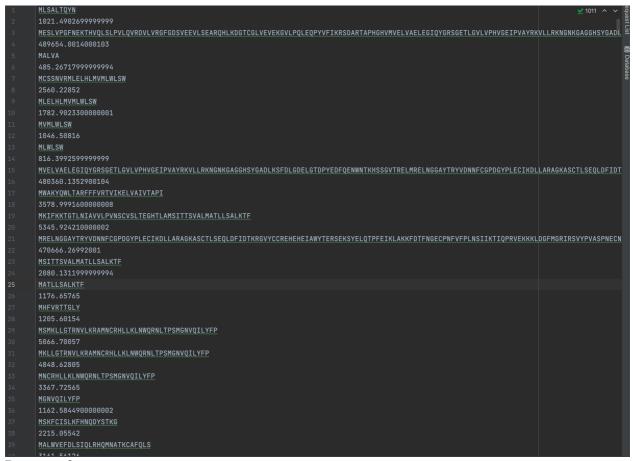


Рисунок 2.

Построение модели белка

Модель белка можно построить с помощью сервиса [3].

Для того, чтобы построить модель, необходимо выбрать белок и вставить последовательность белка в окно ввода, затем построить модель. Модель будет доступна в правой части окна.

Для построения трехмерной модели был выбран следующий белок: MTTFLKYSKKRKSTSILLVTLNLMKRSPLFWHLFLLPQVLLWKL

Модель выбранного белка представлена на рисунке 3. С полученной моделью можно взаимодействовать.

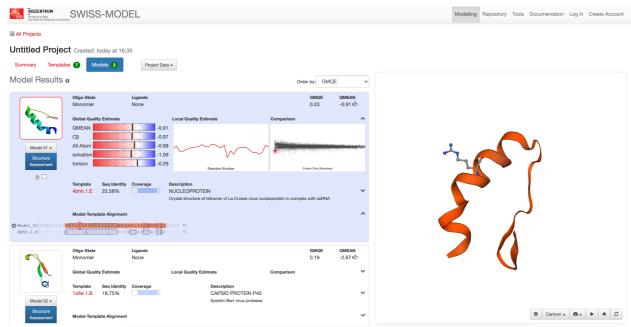


Рисунок 3.

Заключение

В ходе данной работы из генома коронавируса были получены все возможные белки и рассчитаны их молекулярные массы. Для одного из найденных белков была построена трехмерная модель.

Также были решены соответствующие задачи в Rosalind.

Список использованных источников

- 1. Rosalind [Электронный ресурс]. URL: http://rosalind.info/problems/list-view/ (Дата обращения: 15.05.2021).
- 2. Severe acute respiratory syndrome coronavirus 2 [Электронный ресурс]. URL: https://www.ncbi.nlm.nih.gov/genome/browse/#!/viruses/86693/ (Дата обращения: 17.05.2021).
- 3. Swiss-model [Электронный ресурс]. URL: https://swissmodel.expasy.org// (Дата обращения: 19.05.2021).