

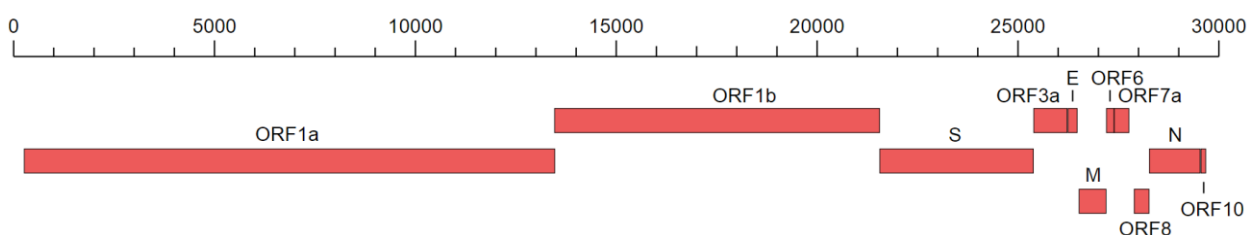
**Задания практического тура заключительного этапа XXXVIII  
Всероссийской олимпиады школьников по биологии.**

**2021-22 уч. год. 11 класс**

**БИОИНФОРМАТИКА. Максимум 50 баллов**

За последние два с небольшим года коронавирус SARS-CoV-2 привлек к себе значительное внимание исследователей, разместивших миллионы его последовательностей в биоинформатических базах данных. (NB: в базах данных последовательности РНК приводятся с тиминном вместо урацила для единообразия с последовательностями ДНК). Задание практического тура посвящено применению биоинформатических методов анализа SARS-CoV-2 и вакцин против него.

1. На рисунке ниже показана схема организации генома SARS-CoV-2 (ORF – открытая рамка считывания, соответствует кодирующей последовательности):



Откройте файл Wuhan.doc и скопируйте из него последовательность исходного Уханьского вируса целиком с заголовком (это называется «последовательность в FASTA-формате») во вкладку браузера с инструментом ORF Finder. Чтобы выделить последовательность целиком, можно использовать сочетание клавиш Ctrl-A, чтобы скопировать – Ctrl-C, чтобы вставить – Ctrl-V.

Очистите окно ввода используя кнопку Clear. Затем укажите в настройках (см. рис. ниже):  
стартовый кодон ATG (выбрав «atg» в строке «ORF can begin with:»),  
поиск по всем рамкам считывания (выбрав «1, 2, and 3» в строке «Search for ORFs in reading frame»)  
минимальную длину рамки 100 кодонов (наберите «100» в строке «Only return ORFs that are at least»).

Submit Clear Reset

- ORFs can begin with: .
- Search for ORFs in reading frame  on the  strand.
- Only return ORFs that are at least  codons long.
- Use the  genetic code.

Вставьте последовательность генома исходного коронавируса (Wuhan.doc) и нажмите кнопку «Submit», в новой вкладке откроется результат поиска. Запишите найденные координаты генов *ORF1a*, *ORF1b*, *S*, *M*, и *N* на Листе ответов.

**При распознавании генов ориентируйтесь на приведенную выше схему организации генома, инструмент выдает рамки считывания не по порядку!**

На самом деле у коронавируса SARS-CoV-2 при трансляции происходит запрограммированный сдвиг рибосомы (13468 нуклеотид участвует в двух кодонах),

поэтому рибосома не доходит до стоп-кодона *ORF1a* и начинает трансляцию раньше рамки *ORF1b*. Таким образом на основе двух рамок считывания образуется единый белок ORF1, который затем разрезается протеазами на несколько частей.

Рассчитайте длину белка ORF1 (в аминокислотах) на основе найденных Вами координат *ORF1a* и *ORF1b*. Укажите на Листе ответов размер белков:

S (белок «шипов», взаимодействующий с рецепторами клетки-мишени),

M (главный мембранный белок сложного капсида) и

N (белок нуклеокапсида).

Скопируйте аминокислотные последовательности белков S, M и N в пустой файл Proteins.doc (для аминокислот используются однобуквенные обозначения) из вкладки результатов ORF Finder, **файл Вам понадобится в дальнейшем.**

2. Рекомбинантные вакцины от коронавируса как правило используют для иммунизации антигены S-белка. Откройте файл Gam\_COVID\_Vac.doc, содержащий последовательность аденовирусного вектора (первого компонента вакцины Спутник V). Найдите последовательность S-белка с помощью ORF Finder (**важно: настройки ORF Finder нужно изменить по сравнению с первым поиском** – выбрать минимальную длину рамки считывания 1000 вместо 100 и обратную нить вместо прямой (выбрать «reverse» вместо «direct» в строке «Strand»)).

Найденная Вами последовательность S-белка окажется на несколько аминокислот короче, чем рассчитанная для Wuhan.doc, поскольку реальная трансляция S-белка начинается не с первого метионинового кодона последовательности, в остальном последовательности идентичны. Найдите настоящую длину S-белка с помощью инструмента Pairwise Align Protein. Для этого в верхнее окно вставьте последовательность S-белка исходного коронавируса, в нижнее окно – последовательность, найденную в вакцине. **Перед вставкой оставьте у последовательностей, найденных ORF Finder, заголовок, но удалите последний символ – звездочку, означающую стоп-кодон.** Запишите результаты на листе ответов.

3. Иммунный ответ после аденовирусной вакцины проходит через несколько определенных стадий. Пациент был вакцинирован аденовирусной вакциной, затем заразился и выздоровел. Расположите события А - К по порядку от наиболее ранних к наиболее поздним. **Обратите внимание, некоторые стадии лишние.**

А – нейтрализация SARS-CoV-2 антителами, убийство зараженных вирусом клеток цитотоксическими лимфоцитами

Б – синтез лимфоцитами огромного количества провоспалительных цитокинов

В – появление в организме наивных лимфоцитов, специфичных к антигенам коронавируса

Г – гибель клеток-мишеней, зараженных рекомбинантным аденовирусом

Д – попадание в организм коронавируса SARS-CoV-2

Е – внутримышечное введение рекомбинантного аденовируса (вакцины)

Ж – размножение лимфоцитов, специфичных к антигенам S-белка, появление антиген-специфических клеток памяти

З – перенос антигенов из мышечной ткани в лимфоузел

И – синтез S-белка в зараженных рекомбинантным аденовирусом клетках

К – репликация рекомбинантного аденовируса в организме

4. В ходе эволюции вирусы постепенно накапливают мутации, на которые может действовать естественный отбор. Откройте файл Delta.doc, содержащий геномную последовательность штамма «дельта» SARS-CoV-2. Найдите с помощью ORF Finder (**не забудьте выставить верные настройки**) в его геноме последовательность, кодирующую S-белок (обратите внимание, геном более короткий и поэтому координаты генов в нём немного отличаются). Откройте в браузере инструмент Ident and Sim, вставьте в окно две последовательности в FASTA-формате S-белка штамма дельта и Уханьского вируса (**последовательности белков исходного вируса Вы сохраняли во временном файле, считайте что они уже выравнены**). Сравните эти последовательности, нажав кнопку «Submit». Не меняйте никаких настроек инструмента. На сколько аминокислот различаются S-белки этих вирусов? Повторите сравнение для белков N и M. Запишите результаты (% идентичности белков и число различающихся аминокислот) на Листе ответов. Сохраните последовательности S, N и M белков в файл Proteins.doc.

Откройте файл Omicron.doc, содержащий геномную последовательность штамма «омикрон» SARS-CoV-2. Сравните его S, N и M белки с белками исходного штамма. Для S-белка сначала используйте инструмент Pairwise Align Protein, поскольку кроме ряда замен в этом белке есть две «выпадающие» аминокислоты, поэтому Ident and Sim сразу не сработает. **Не меняйте никаких настроек инструмента.** Не закрывайте вкладку результатов ORF Finder, они понадобятся в задании 5.

Запишите % идентичности и число различий аминокислот каждого белка на Листе ответов. Какой из рассмотренных Вами белков наименее консервативен? Чем это можно объяснить?

5. Приобретенный иммунитет распознает отдельные высокоиммуногенные участки вирусных белков – эпитопы. Какие именно эпитопы окажутся важны для искоренения вируса, зависит от генотипа локуса МНС, отвечающего за представление фрагментов вирусных белков Т-лимфоцитам. Ниже приведены последовательности шести распространенных Т-клеточных эпитопов SARS-CoV-2, часто вызывающих иммунный ответ.

Найдите с помощью сочетания клавиш Ctrl-F последовательности эпитопов 1-6 в белках S, M и N в штамме омикрон. Запишите на Листе ответов, какому белку соответствует каждый из эпитопов. Напишите, есть ли эти эпитопы в изученных Вами вирусах, используя символ «С» для вакцины «Спутник V», символ «Δ» для варианта «дельта» и символ «W» для исходного Уханьского варианта.

Последовательности эпитопов (**можно копировать из pdf-файла заданий**):

- 1) LLNKHIDAY
- 2) CVADYSVLY
- 3) FAYANRNRF
- 4) LPQGFSAL
- 5) SPRWYFYLY
- 6) LPQGFSAL

**Не забудьте закрыть лишние окна и вкладки браузера, удалите временные файлы!**

**Желаем Вам удачи!**