Вердикт:	WrongAnswer
Контест:	72730: Финал агротех 9
Участник:	Klaimiks
Задача:	1. Задание 1. "Фермер и плесень" (10 баллов)
Баллы:	0
Компилятор:	No compiler
IP пользователя:	85.143.3.38
GlobalID:	10000067-5ecc-ef0a-e4cb-a7a504e4eee9

Тесты Логи Исходный код

¶ 0xFF

```
0
```

Вердикт:	OK
Контест:	72730: Финал агротех 9
Участник:	Klaimiks
Задача:	2. Задание 2. "Нанотехнологии покоряют мир" (5 баллов)
Баллы:	5
Компилятор:	No compiler
IP пользователя:	85.143.3.38
GlobalID:	10000067-5ecc-40a5-2e3a-bad3dee80312

Тесты Логи Исходный код

¶ 0xFF

```
12
```

Вердикт:	OK
Контест:	72730: Финал агротех 9
Участник:	Klaimiks
Задача:	4. Задание 4. "Юный натуралист" (10 баллов)
Баллы:	10
Компилятор:	No compiler
IP пользователя:	85.143.3.38
GlobalID:	10000067-5ebc-fe6b-77d1-2a1bfbf1ec7a

Тесты Логи Исходный код

¶ 0xFF

```
6
```

Вердикт: WrongAnswer

Контест: 72730: Финал агротех 9

Участник: Klaimiks

Задача: 5. Задание 5. "Амбар" (15 баллов)

Баллы: 0

Компилятор: No compiler

IP пользователя: 85.143.3.38

GlobalID: 10000067-5eca-b754-a371-73c47ca24826

Тесты Логи Исходный код

0xFF

```
4
```

Участник: Klaimiks

Задача: 6. Задание 6. "Петя" (10 баллов)

Баллы: 8

Компилятор: Python 3.12.3

IP пользователя: 85.143.3.38

GlobalID: 10000067-5ed8-d021-14bd-483186fb6776

Тесты Логи Исходный код

0xFF

```
count = int(input())
seasons = list(map(int, input().split()))
final = 0
sames = []
for i in range(2, len(seasons)+1):
    temporary = seasons[0:i]
    if (min(temporary) == temporary[i-1] or max(temporary) == temporary[i-1]) and temporary[i-1] not in sames:
        final +=1
    if seasons.count(temporary[i-1]) > 1:
        sames.append(temporary[i-1])
print(final)
```

Вердикт: Crash

Контест: 72730: Финал агротех 9

Участник: Klaimiks

Задача: 7. Задание 7. "Сломалась нанопора" (10 баллов)

Баллы: 0

Компилятор: Python 3.12.3

IP пользователя: 85.143.3.38

GlobalID: 10000067-5edc-0efb-8d18-ba104f5e100a

Тесты Логи Исходный код

0xFF

```
inp = input()
count = []
for i in range(len(inp)):
    if inp[i] == 'A':
        count.append(i)

print(inp[:count[0]] + inp[count[0]:(count[len(count)-1])+1][::-1] + inp[count[len(count)-1]+1:])
```

Вердикт: RuntimeError

Контест: 72730: Финал агротех 9

Участник: Klaimiks

Задача: 8. Задание 8. "Юный предприниматель" (10 баллов)

Баллы: 0

Компилятор: Python 3.12.3

IP пользователя: 85.143.3.38

GlobalID: 10000067-5ee2-37fc-7f81-eba915b26bf0

Тесты Логи Исходный код

0xFF

```
from math import floor
x, p, y = map(int, input().split())
years = 0
while x < y:
    x += floor(x * (1/p))
    years += 1
print(years)
```

Участник: Klaimiks

Задача: 9. Задание 9. "Грядка" (15 баллов)

Баллы: 0

Компилятор: Python 3.12.3

IP пользователя: 85.143.3.38

GlobalID: 10000067-5ee1-1ed6-153f-2847e89b54f2

Тесты Логи Исходный код

0xFF

```
n = int(input())
k = int(input())
f = 0
final = []
while len(final) != n:
    if f + k <= n and f+k not in final:
        f += k
        final.append(f)
        print(f)
    else:
        f = ((k + f) % n) - k
```

Контест: 72730: Финал агротех 9

Участник: Klaimiks

Задача: 10. Задание 10. "Рис или кукуруза????" (5 баллов)

Баллы: 2

Компилятор: Python 3.12.3

IP пользователя: 85.143.3.38

GlobalID: 10000067-5ee5-ad61-f9b3-3ecbd0b95a37

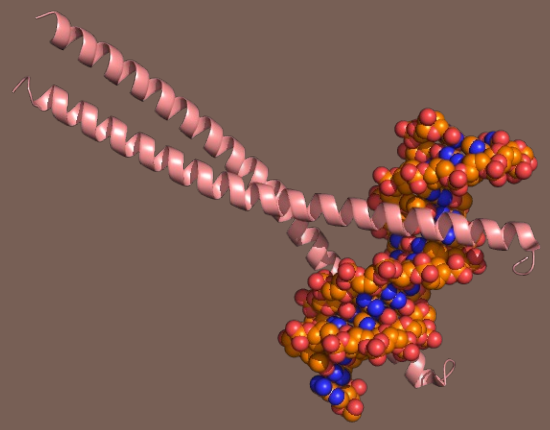
Тесты Логи Исходный код

0xFF

```
n = int(input())
l = 0
for i in range(2, 7):
    if n % i == 0 and i not in [2, 3, 5, 7]:
        l += 1
if l == 0:
    print('rice')
else:
    print('corn')
```

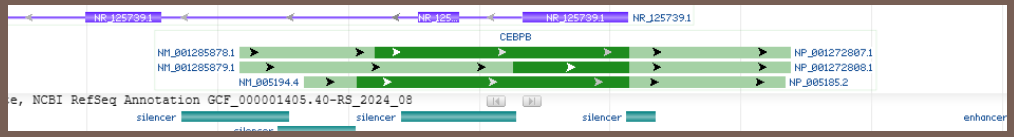
ГЕН “CEBPB” (ПОЛНОЕ НАЗВАНИЕ - CCAAT ENHANCER BINDING PROTEIN BETA)

Фактор транскрипции, участвующий в кодировании белка. Расположен на 20 хромосоме.
Его для регуляции генов, вовлеченных в иммунные и воспалительные реакции, среди прочих процессов. Кроме того, он напрямую влияет на старение, замедляя его. CEBPB способен увеличить экспрессию нескольких целевых генов. Среди них некоторые играют определенную роль в нервной системе



Взаимодействует с 1477 генами: DDIT3, CEBPA, HOMER3 и другими.

Ген имеет один экзон, но интронов не содержит вовсе, что видно на фотографии ниже.



- Ниже представлены последовательности данного гена

Нуклеотидовая последовательность (часть)

ACATTAAAAAATCAGCTGGGTGTGGTGGCATGCCCCTGT
GGTCCCAGCTACTTGGGAGGTTGAGGTGGGA
GGATCACCTGAGCCCAAGAAGCAGAGGCTGCAGTGAG
CAGAGATCGTGCCACTGCACTCCAGCAGTGAGA
CCCTGTCTCAAAAAACAAACAAACAAAAAACCCCAGA
AAACAAAAAATCAAATCTCCCAGAACTCCTGCT
CTGCTCCCCTGCTGTATTTTCCTCCATAGTGCTGACCAGAA
TCTAATACAATCTATACTTTAATCATTGA
TCTTGTTTAGTGCTTGTCTTCCCCCAGGAGAATTTCAGCTC
CACGAGAACAGGAATTGCTTGTCTTGTCC
ATGAATGTGTCTTTAGGGCCAGAATTACAGCATGTGCCA
AGGAACATTCGATAAATGTGAAGGTTGAGCA
CTGTTTTCTGTTCGTTGGACACTCTGGTTCCTCTCGGCGGC
CTGCTGATGCCCTGTGTGCATCTATCACA
TCCATGGCACCGGCCTTTCCTGATCCCAGAGCAAGGGCG
CCCAGCGGATCTTCACGCATGTGAATGGCTC
CTGCAGTCCCAGAACCTGCCACCTGGGGGCAGTATATAC
TCAGGCAAGTTTCACTACCGGTCCTGAGCAC
CTCCCGTCTCCAAAATCCTGGATCTCAGAACCTCGGAGC
CTGGCAGCCCTTGGGGATCTATCTCTGCCAG

Аминокислотная последовательность

AYLGYQAVPSGSSGSLSTSSSSSPPGTPSP
ADAKAPPTACYAGAAPAPSQVKSKAKK
TVDKHSDEYKIRR
ERNNIAVRKSRDKAKMRNLETQHKVLEL
TAENERLQKKVEQLSRELSTLRNLFKQLPE
PLLASSGHC

Для поиска информации о гене, в том числе и для поиска последовательностей, использовался крупнейший в мире источник генетических данных NCBI. В числе используемых баз данных : RefSeq, PubMED и Protein, а в числе поисковых инструментов – BLAST.

Деревья паралогов и ортологов для гена CEBPB

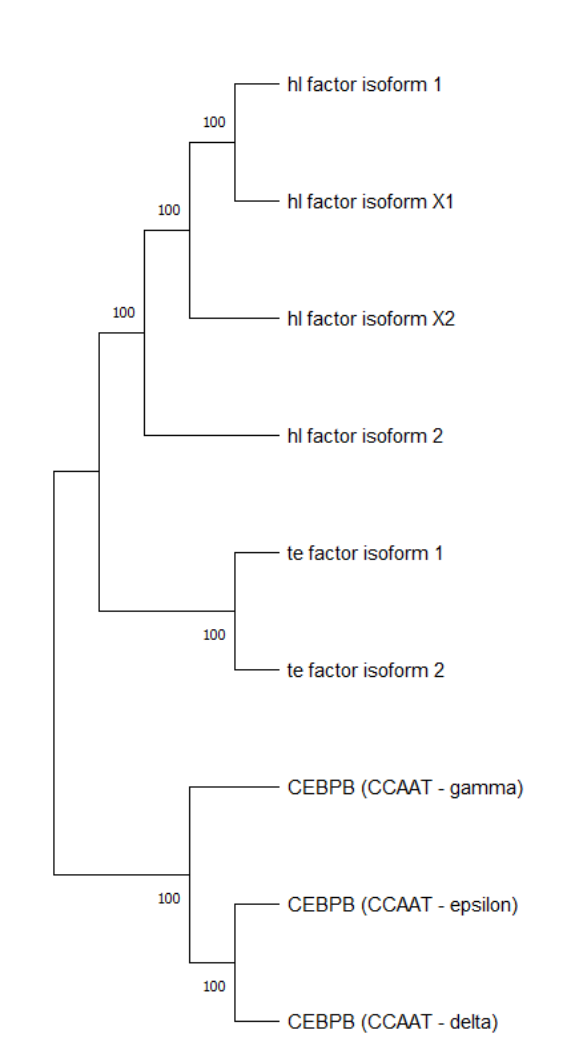
Поиск паралогов с помощью нуклеотидовой последовательности осуществлен в матрице BLOSUM62. Так мне удалось обнаружить всего три. При этом они не целиком являются паралогами, а лишь благодаря определенным доменам, поэтому их связь очень отдалена. Последним минусом оказалось отсутствие возможности воспользоваться методом Бутстрапа, используемым для определения устойчивости родства между генами в ходе эволюции. По этим причинам я решила использовать аминокислотную последовательность для создания дерева.

Для начала, FASTA файлы всех аминокислотных последовательностей преобразованы в единый текстовый файл. После изменения формата на FASTA, произведено выравнивание, затем уже само создание дерева. Оба процесса происходили в программе MEGA.

Метод объединения дерева выбран мною для обоих деревьев, т.к по моему мнению он нагляднее других показывает связи между генами.

***Hl- hepatic leukemia**

***Te - thyrotroph embryonic**



Деревья паралогов и ортологов для гена CEBPB

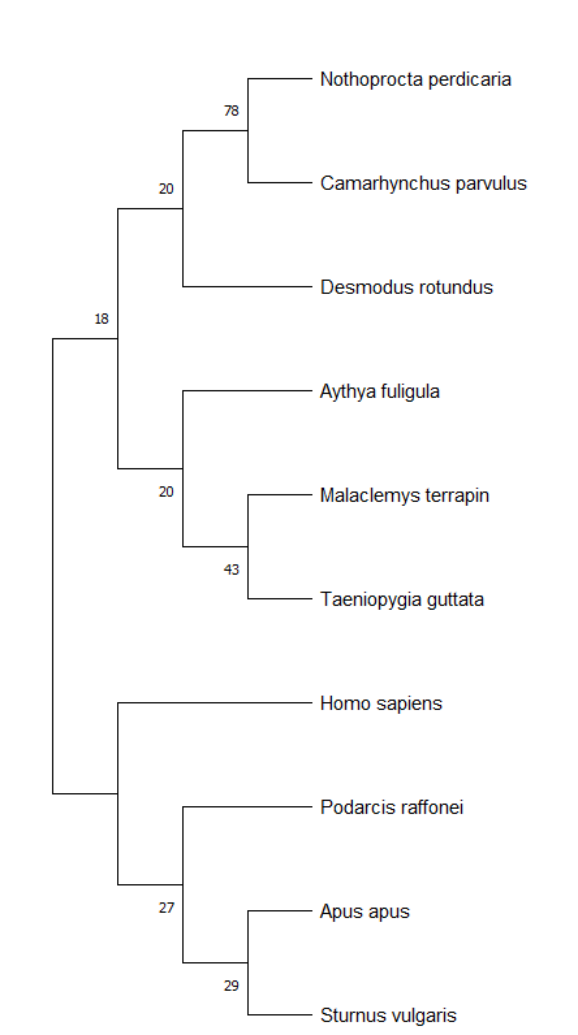
Для создания дерева ортологов использовались следующие виды:

Desmodus rotundus
Podarcis raffonei
Pantherophis guttatus
Apus apus
Aythya fuligula
Sturnus vulgaris
Malaclemys terrapin
Taeniopygia guttata
Nothoprocta perdicaria
Camarhynchus parvulus

Принцип образования дерева тот же, что и у прошлого, но в этот раз я решила использовать нуклеотидовую последовательность.

Можно подметить, что чаще всего этот ген встречался у птиц

Данный ген не был найден ни у кого, кроме животных, а это значит, что он не является древним.



ВЫВОД ПО ПРАКТИЧЕСКОЙ РАБОТЕ.

Результаты данной работы помогли мне выстроить полное представление о функциях и строении генома, а также выстроить представление об его сравнительно недолгом эволюционном пути:

- Его активность важна для усиления работы имунной системы
- Ген не обладает интронами
- Несмотря на большое количество данных, его изучение продолжается до сих пор.

олимпиада иннагрика

2024 Финал

профиль АгроТех

Кодеры

Команда Ц9-040 9-й класс Кейс №Ц9-2

Анализ нуклеотидных и аминокислотных последовательностей

Кодеры

Андрей Кустышев

Первый ~~питошка~~ программист, командир

Максим Касьянов

Визуализатор, разработчик презентации

Даша Макарова

Разработчик презентации, генератор идей

Лайла Бокоева

Второй ~~питошка~~ программист, биолог

Оля Надточиева

Биолог

Наставник - Долгих Владислав Андреевич

Проблематика

Сложность анализа данных

Использование библиотек

Error , defauldict

Многомерность данных

Специфичны по преобразованию и большие по объему

Сложность алгоритмов

Использование алгоритмов для анализа биологических последовательностей

Последствия решения

Автоматизация анализа геномных данных для дальнейшей работы с ними

Задание 1

Вычисление GC состава последовательности ДНК

- Вычисление длины последовательности ДНК
- Вычисление кол-во гуанина и цитозина
- Расчет GC состава по следующей формуле:

$$GC\% = \frac{G + C}{L} \times 100$$

Задание 1

Вводные данные

TTCCACTCGGTTGGGCCGGCTAGGCCTCCCTGCCCGGAGTTTCGGCGGACTGCTGCCGACACCCGGGCATTGTCTTAGGGGGGTTATTCGAGGGCACTCGCA
GCCAACTTGTCGGGACCAGCCGGGCTGGTCATCGGGCTTATACAGCGAAATGCCGAGGACCCGGCCCCGCGCTATGGAACGTCTTTAGCTCCGGCAGGCAAT
TAAGGACAACGCAAGCATGGCGGATATAAACAGAGAAACGGGCGAATACACCTGTTCGTGTCGTATCGGTAAATAGCCTCGCGGAGGCATGTGCCATGCTGG
CCTGCGGAGCACTCTGGTTATGCATATGGTCCACAGGACACTCGTCGCTTCCGGGTTTGCCCTCTATGTGACGGTCTTTAGGCGCACTTATGCTCAGCACCGTT
TAAACCAGACCGACACCAGATCTGTAAGGTCCGCCACGCAGACGAGAGCGCACGGAGACCACCGAGCGATCTACCTGATCGGCGACCATTTGTGTGGTACTG
GGGCCGAGAGGTAACTACGGTGCCGCTAACAGCCCCTCGGTCGTCGCTGACGTCTGTAGTCTAGCCTCATTATGATCGTACGCTATTCAGGGATTGACTGATAC
CGGAAGACATCTCAGTTGAAGTGGTGTATGCGACAGAGACCGTGCACCTACCAAACCTCCTTAGTCTAAGTTCAGACCAATTGGTAGTTTGTCCAGAACTCAG
ATTTTATCACCAGAGGACGCACGCCCTACCTCCATGATCCACTGACGTCCCTGAGGCTGCAATACATGCAACCAGGCAGTCTCCGCGGTAAGTCCTAGTGCAAT
GGGGCTTTTTTTCCCTGGTCCTCGAGAAGAGGGGACGCCGGTCCAGACATCTCTAATGTGGTAATTGGGAGGACTCTTGGCCCTCCGCCCTTAGGCAGTGCAT
ACTCTTCCATAAACGGGCTGTTAGTTATGGGGTCCGAGGATTGAAAAAGGTGAGCGAACTCGGCCGAACCGGAG

```
Введите долю GC:
70
Задайте длину участка
6
```

Вывод

```
В богатой последовательности 14 богатых участков и 22 бедных, длиною 6
```

Время выполнения:
0,21 мс

Задание 2

Алгоритм преобразования нуклеотидной последовательности в аминокислотную

Начинаем преобразование с кодона ATG и заканчиваем одним из стоп-кодонов

Определение молекулярной массы в аминокислотных последовательностях

Сортировка полипептидных цепочек длиной 5 аминокислот

Рассчитали молекулярную массу самых легких и самых тяжелых изоформ белка титина

Задание 2

Вводные данные

ATGGCGGGTCTTCGCAAGTCCAGGCCGGGTTAG

ATG	GCG	GGT	CTT	CGC	AAG	TCC	AGG	CCG	GGT	TAG
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
M(149)	A(89)	G(75)	L(131)	R(174)	K(146)	S (105)	R (174)	P (115)	G (75)	STOP

Вывод

['AAGTC', 'AGGCC', 'AGTCC', 'ATGGC', 'CAAGT', 'CAGGC', 'CCAGG', 'CCGGG', 'CGCAA', 'CGGGT', 'CTTCG', 'GCAAG', 'GCCGG', 'GCGGG', 'GGCCG', 'GGCGG', 'GGGTC', 'GGGTT', 'GGTCT', 'GGTTA', 'GTCCA', 'GTCTT', 'GTTAG', 'TCCAG', 'TCGCA', 'TCTTC', 'TGGCG', 'TTCGC']

Молекулярная масса цепочки: 2938

Аминокислотная последовательность: MAGLRKSRPG

Время выполнения:
0,5 мс

Задание 3

Обратная трансляция

- это процесс декодирования последовательности аминокислот в соответствующие кодоны

Восстановление цепочки ДНК и определение GC состава для нее

Определение GC и AT состава для тугоплавких и легкоплавких последовательностей соответственно

Вводные данные: MVGRRALFNA

легкоплавкость

тугоплавкость

Вывод

Последовательность ДНК: ATGGTTGGTAGAAGAGCTTTATTTAATGCT

GC-состав: 33.33%

Последовательность ДНК: ATGGTCGGCCGCCGCGCCCTCTTCAACGCC

GC-состав: 73.33%

Время выполнения:
0,35 мс

Задание 4

Оценка степени сходства нуклеотидных последовательностей

- Мы взяли 5 нуклеотидных последовательностей, написали программу, которая ищет p-расстояние нуклеотидной последовательности
- Далее, происходит сортировка внутри каждой строки матрицы
- В самом конце, происходит парсинг данных и вывод самой матрицы

Задание 4

Вводные данные

5 последовательностей одинаковой длины таксонов Reptilia, Mammalia и Aves

Вывод

Матрица p-расстояний:

	1	2	3	4	5
1	0.0000	0.2238	0.3147	0.3287	0.2517
2	0.2238	0.0000	0.2354	0.3333	0.2681
3	0.3147	0.2354	0.0000	0.2657	0.3450
4	0.3287	0.3333	0.2657	0.0000	0.1189
5	0.2517	0.2681	0.3450	0.1189	0.0000

Время выполнения:
0,28 мс

Задание 5

Консенсусная последовательность ДНК

Консенсусная последовательность - определенная нуклеотидная последовательность, регулярно встречающаяся (с небольшими вариациями по отдельным нуклеотидам) в данном генетическом элементе.

Для начала, мы взяли последовательности разных таксонов из 4 задания

Для начала, циклом мы перебираем последовательность циклом, попарно добавляя каждый следующий элемент в общий список

Далее, мы заменяем элементы на соответствующие символы из списка

Затем мы написали программу которая строит консенсус

Задание 5

Вводные данные

```
Введите вашу последовательность для консенсуса. После ввода последней, введите "конец"
TGGTGCGGAGGCCCTGGAGAGGATGTTCCTGTCCTTCCCCACCACCAAGACCTACTTCCCGCACTTCGACCTGAGCCACGGCTCTGCCCAGGTTAAGTC
Введите вашу последовательность для консенсуса. После ввода последней, введите "конец"
TGGAGGAGAAACCCTGGGAAGGCTCCTGGTTGTCTACCCATGGACCCAGAGGTTCTTTGACAGCTTTGGCAACCTGTCCTCTGCCTCTGCCATCATGAA
Введите вашу последовательность для консенсуса. После ввода последней, введите "конец"
TGGTGGTGAGGCCCTGGGCAGATTACTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCTCTCCTGATGCTGTTATGAT
Введите вашу последовательность для консенсуса. После ввода последней, введите "конец"
конец
```

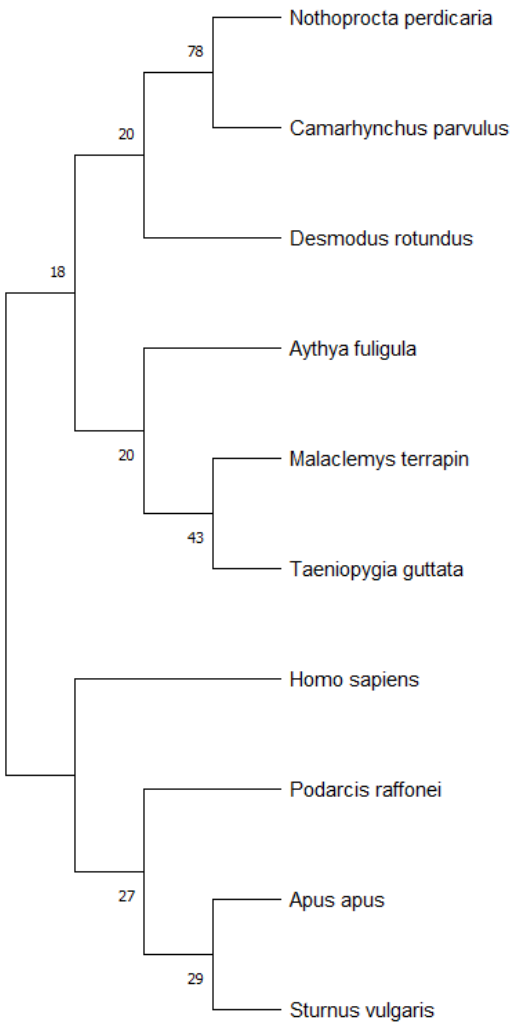
Вывод

```
Консенсус данных последовательностей:
 TGGWGSDGARRCCCTGGRVAGRHTVYTSSTKKYCTWCCCHWSSACCMAGASSTWCTTYSMSHVCTTYGRSVWBMKSYMCKSYBCYKMYSMBRTYAWGWH
```

Время выполнения:
0,1 мс

Практическая работа

- **Тема практической работы – изучение выбранного гена.**
- Мы проводили исследование гена с помощью NCBI, NCBI BLAST и построили филогенетическое дерево в MEGA.
- Работая над практическим заданием мы углубили свои навыки в работе с биологическими последовательностями. Эта работа помогла лучше понять их специфику и помогла в работе над кейсом



Филогенетическое дерево

Вывод

Выводы

- Мы освоили основные операции с биологическими данными, что позволило нам эффективно работать с нуклеотидными и аминокислотными последовательностями, включая их анализ и обработку.
- Использовали биоинформационную базу данных NCBI для поиска данных.
- Создали комплекс специализированных программ, которые автоматизируют и упрощают многие процессы работы с геномной информацией. Это поможет быстрее изучать свойства растений и животных, важных для сельского хозяйства.

Спасибо за внимание