Брянцев Всеволод Александрович

ПММ, 4 курс, 61 группа

Отчет по лабораторной работе №2

**Вариант 4**

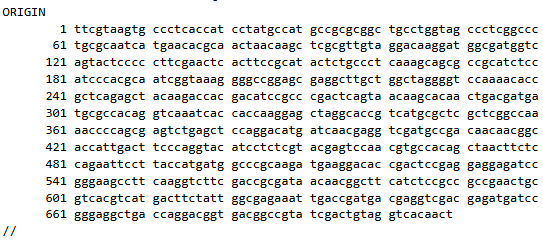
* ***Alternaria sp. strain YZU 191419 calmodulin (calmodulin) gene, partial cds (гриб альтернария)***
* ***Stemphylium vesicarium isolate SVA1 calmodulin (calmodulin) gene, partial cds; mitochondrial (гриб стемфилиум)***

**Задание 1.**

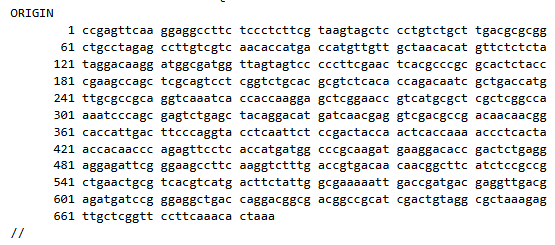
В базе данных GenBank найти нуклеотидные последовательности в соответствии с вариантом. Привести каждую последовательность и определить число пар оснований каждой последовательности.

***Решение:***

Искомая мРНК гриба альтернария содержит 709 пар оснований:



Искомая мРНК гриба стемфилиум содержит 685 пар оснований:



**Задание 2.**

Описать какие последовательности вы нашли, за что они отвечают и каким организ мам принадлежат.

***Решение:***

1. Это частичная последовательность гена, кодирующего кальмодулин, из штамма Alternaria sp. YZU 191419. Кальмодулин является важным кальций-связывающим белком, участвующим в регуляции различных клеточных процессов, включая сигнальные пути и контроль клеточного цикла. Организм относится к роду Alternaria, который включает в себя виды грибов, часто встречающихся в окружающей среде и известных как фитопатогены.
2. Это частичная последовательность гена кальмодулина (calmodulin) митохондриального происхождения, выделенная из изолята SVA1 гриба Stemphylium vesicarium. Кальмодулин является кальций-связывающим белком, играющим ключевую роль в регуляции различных клеточных процессов, включая сигнальные пути и контроль клеточного цикла. Stemphylium vesicarium — фитопатогенный гриб, известный как возбудитель болезней растений, таких как груша и спаржа.

**Задание 3.**

Построить точечную матрицу сходства между двумя последовательностями в формате FASTA. Сделать вывод о сходстве или несходстве последовательностей.

***Решение:***

Для визуализации точечной матрицы был использован скрипт на языке программирования Python с использованием библиотеки matplotlib:

import matplotlib.pyplot as plt

# считываем файлы с последовательностями

with open('seq1.txt', 'r') as f1, open('seq2.txt', 'r') as f2:

    seq1 = f1.read()

    seq2 = f2.read()

# формируем матрицу с размерами последовательностей

matrix = [[1 if seq1[i] == seq2[j] else 0 for j in range(len(seq2))] for i in range(len(seq1))]

# строим диаграмму

plt.imshow(matrix, cmap='binary', origin='lower')

plt.xlabel('Stemphylium vesicarium isolate SVA1 calmodulin (calmodulin) gene, partial cds; mitochondrial')

plt.ylabel('Alternaria sp. strain YZU 191419 calmodulin (calmodulin) gene, partial cds')

plt.show()

# вычисляем процент идентичности последовательностей

matches = sum(1 for a, b in zip(seq1, seq2) if a == b)

identity = (matches / min(len(seq1), len(seq2))) \* 100

print(f"Identity: {identity:.2f}%")

Рассматриваемые последовательности записаны в текстовые файлы seq1 и seq2:

ttcgtaagtg ccctcaccat cctatgccat gccgcgcggc tgcctggtag ccctcggccc

tgcgcaatca tgaacacgca actaacaagc tcgcgttgta ggacaaggat ggcgatggtc

agtactcccc cttcgaactc acttccgcat actctgccct caaagcagcg ccgcatctcc

atcccacgca atcggtaaag gggccggagc gaggcttgct ggctaggggt ccaaaacacc

gctcagagct acaagaccac gacatccgcc cgactcagta acaagcacaa ctgacgatga

tgcgccacag gtcaaatcac caccaaggag ctaggcaccg tcatgcgctc gctcggccaa

aaccccagcg agtctgagct ccaggacatg atcaacgagg tcgatgccga caacaacggc

accattgact tcccaggtac atcctctcgt acgagtccaa cgtgccacag ctaacttctc

cagaattcct taccatgatg gcccgcaaga tgaaggacac cgactccgag gaggagatcc

gggaagcctt caaggtcttc gaccgcgata acaacggctt catctccgcc gccgaactgc

gtcacgtcat gacttctatt ggcgagaaat tgaccgatga cgaggtcgac gagatgatcc

gggaggctga ccaggacggt gacggccgta tcgactgtag gtcacaact

ccgagttcaa ggaggccttc tccctcttcg taagtagctc cctgtctgct tgacgcgcgg

ctgcctagag ccttgtcgtc aacaccatga ccatgttgtt gctaacacat gttctctcta

taggacaagg atggcgatgg ttagtagtcc cccttcgaac tcacgcccgc gcactctacc

cgaagccagc tcgcagtcct cggtctgcac gcgtctcaca ccagacaatc gctgaccatg

ttgcgccgca ggtcaaatca ccaccaagga gctcggaacc gtcatgcgct cgctcggcca

aaatcccagc gagtctgagc tacaggacat gatcaacgag gtcgacgccg acaacaacgg

caccattgac ttcccaggta cctcaattct ccgactacca actcaccaaa accctcacta

accacaaccc agagttcctc accatgatgg cccgcaagat gaaggacacc gactctgagg

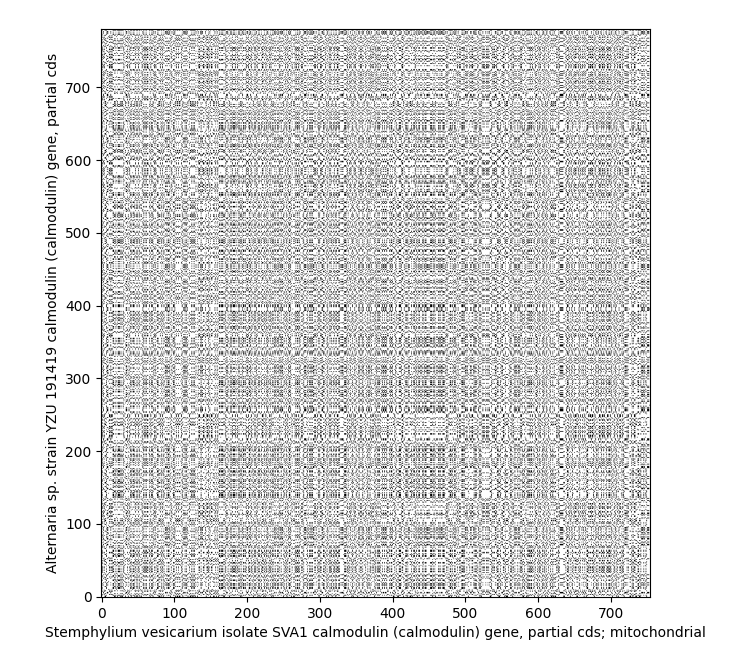
aggagattcg ggaagccttc aaggtctttg accgtgacaa caacggcttc atctccgccg

ctgaactgcg tcacgtcatg acttctattg gcgaaaaatt gaccgatgac gaggttgacg

agatgatccg ggaggctgac caggacggcg acggccgcat cgactgtagg cgctaaagag

ttgctcggtt ccttcaaaca ctaaa

Программа строит вот такую точечную матрицу:



Также программа выдает процент идентичности заданных последовательностей:



Процент идентичности равный 30,37% говорит о том, что исходные последовательности несходственны.

**Задание 4.**

Пользуясь базой данных Nucleotide сайта NCBI для любой выбранной последовательности ДНК организма из варианта записать:

1. Длину последовательности ДНК.
2. Исходный организм, название гена / имя белка.
3. Записать домен, царство, тип, класс, отряд, семейство, род и вид.
4. Авторы последовательности.
5. Дата размещения.
6. Найти и приложить фото исходного организма.

**Вариант 4: *Амбистомы (Ambystoma)***

***Решение:***

1. 1179 пар оснований.
2. Ambystoma ordinarium – Речная амбистома, ген: NADH-дегидрогеназа, субъединица 2 (nad2).
3. Домен – эукариоты, царство – животные, тип – Хордовые, класс – Земноводные, отряд – Хвостатые земноводные, семейство – Амбистомовые, род – Амбистомы, вид – Речная амбистома.
4. Williams,J.S., Niedzwiecki,J.H. and Weisrock,D.W.
5. 09.04.2013
6. Фото:

