

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:

ORIGIN

```
1 ttcgtaagtgc ccttcacccat cctatgccat gccgcgcggc tgcctggtag ccctcggccc
61 tgcgcaatca tgaacacgca actaacaagc tcgcgttgta ggacaaggat ggcgatggtc
121 agtactcccc cttcgaactc acttcgcgat actctgccct caaagcagcg ccgcatctcc
181 atccccacgca atcggtaaag gggccggagc gaggcttgct ggctaggggt ccaaaacacc
241 gctcagagct acaagaccac gacatccgcc cgactcagta acaagcaca ctgacgatga
301 tgcgccacag gtcaaaccac caccaaggag ctaggcaccg tcatgcgctc gtcgcccaca
361 aacccacagc agtctgagct ccaggacatg atcaacgagg tcgatgccga caacaacggc
421 accattgact tcccaggtac atcctctcgt acgagtccaa cgtgccacag ctaactctc
481 cagaattcct taccatgatg gcccgcaaga tgaaggacac cgactccgag gaggagatcc
541 gggaagcctt caaggtcttc gaccgcgata acaacggctt catctccgcc gccgaactgc
601 gtcacgtcat gacttctatt ggcgagaaat tgaccgatga cgaggctgac gagatgatcc
661 gggaggctga ccaggacggg gacggccgta tcgactgtag gtcacaact
```

//

Искомая мРНК гриба стемфилиум содержит 685 пар оснований:

ORIGIN

```
1 ccgagttcaa ggaggccttc tccctcttcg taagtagctc cctgtctgct tgacgcgcgg
61 ctgcctagag ccttgtcgtc aacaccatga ccatgttggt gctaacacat gttctctcta
121 taggacaagg atggcgatgg ttagtagtcc ccttcgaac tcacgcccgc gcactctacc
181 cgaagccagc tcgcagtcct cggctctgcac gcgtctcaca ccagacaatc gctgaccatg
241 ttgcgccgca ggtcaaata caaccaagga gctcggaaac gtcacgctc cgctcggcca
301 aaatcccagc gagtctgagc tacaggacat gatcaacgag gtcgacgccc acaacaacgg
361 caccattgac tcccaggta cctcaattct ccgactacca actcaccaaa accctacta
421 accacaaccc agagttctc accatgatgg cccgcaagat gaaggacacc gactctgagg
481 aggagattcg ggaagccttc aaggtctttg accgtgacaa caacggcttc atctccgccg
541 ctgaactgcg tcacgtcatg acttctattg gcgaaaaatt gaccgatgac gaggttgacg
601 agatgatccg ggaggctgac caggacggcg acggccgcac cgactgtagg cgctaaagag
661 ttgctcggtt ccttcaaaaa ctaaa
```

//

Задание 2.

Описать какие последовательности вы нашли, за что они отвечают и каким организмам 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()
```

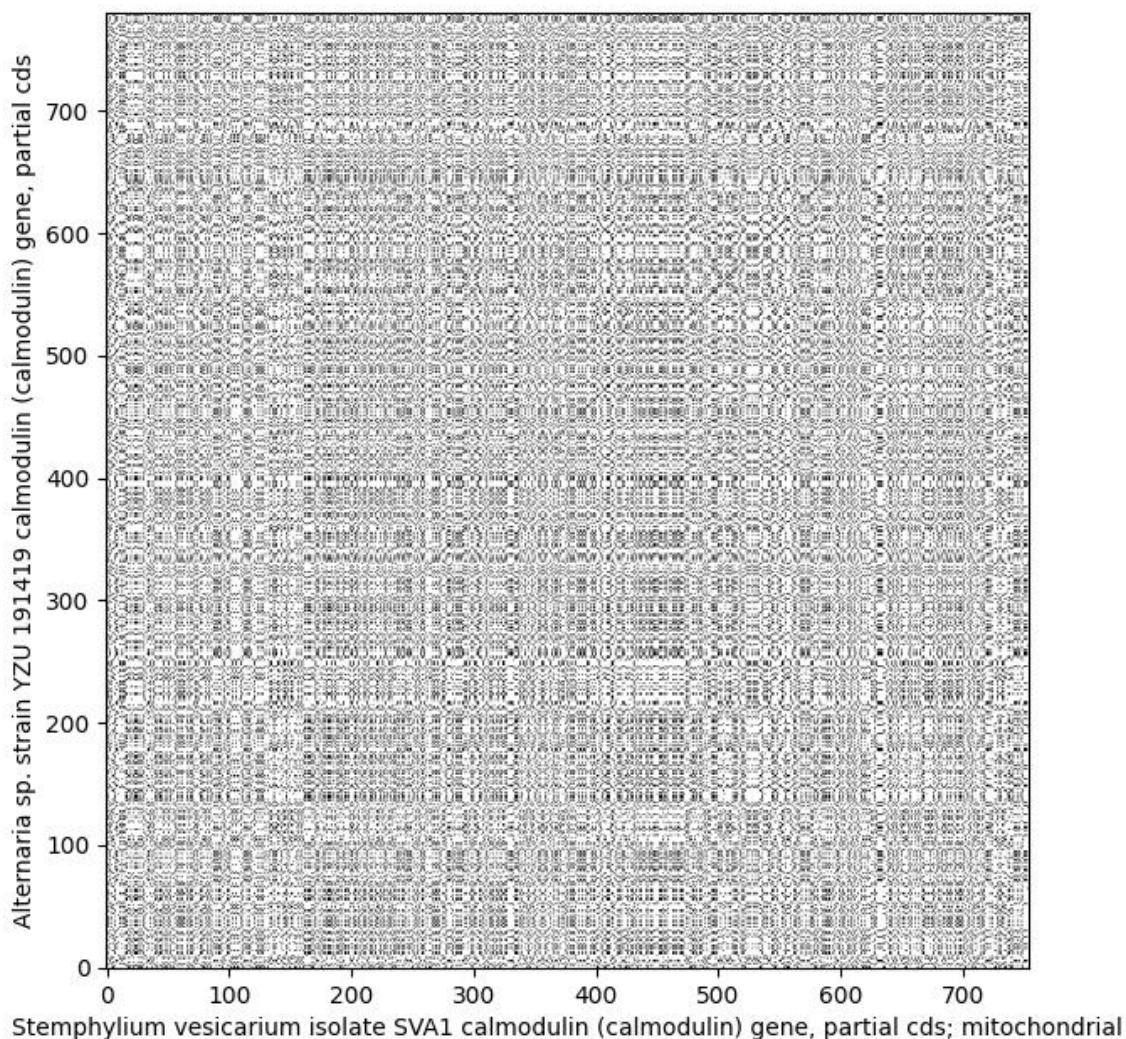
```
# вычисляем процент идентичности последовательностей
matches = sum(1 for a, b in zip(seq1, seq2) if a == b)
identity = (matches / min(len(seq1), len(seq2))) * 100
print(f"Identity: {identity:.2f}%")
```

Рассматриваемые последовательности записаны в текстовые файлы seq1 и seq2:

```
ttcgtaagtg ccctcacat cctatgccat gccgcgcggc tgcctggtag ccctcggccc
tgcgcaatca tgaacacgca actaacaagc tcgcgttgta ggacaaggat ggcgatggtc
agtactcccc cttcgaactc acttccgcat actctgccct caaagcagcg ccgcatctcc
atccacgca atcggtaaag gggccggagc gaggccttgct ggctaggggt ccaaaacacc
gctcagagct acaagaccac gacatccgcc cgactcagta acaagcaca ctgacgatga
tgcgccacag gtcaaatac caccaaggag ctaggcaccg tcatgcgctc gtcgggcaa
aaccacagcg agtctgagct ccaggacatg atcaacgagg tcgatgccga caacaacggc
accattgact tcccaggtac atcctctcgt acgagtccaa cgtgccacag ctaacttctc
cagaattcct taccatgatg gcccgcaaga tgaaggacac cgactccgag gaggagatcc
gggaagcctt caaggtcttc gaccgcgata acaacggcct catctccgcc gccgaactgc
gtcacgtcat gacttctatt ggcgagaaat tgaccgatga cgaggctgac gagatgatcc
gggaggctga ccaggacggt gacggccgta tcgactgtag gtcacaact
```

```
ccgagttcaa ggaggccttc tccctcttcg taagtagctc cctgtctgct tgacgcgcgg
ctgcctagag cttgtcgtc aacaccatga ccatgttggt gctaacacat gttctctcta
taggacaagg atggcgatgg ttagtagtcc cccttcgaac tcacgcccgc gactctacc
cgaagccagc tcgcagtcct cggctcgcac gcgtctcaca ccagacaatc gctgaccatg
ttgcgccgca ggtcaaata ccaccaagg gctcggaacc gtcattgcgt cgctcggcca
aaatcccagc gagtctgagc tacaggacat gatcaacgag gtcgacgccg acaacaacgg
caccattgac ttcccaggta cctcaattct ccgactacca actcaccaa accctcacta
accacaaccc agagttcctc accatgatgg cccgcaagat gaaggacacc gactctgagg
aggagattcg ggaagccttc aaggctttg accgtgacaa caacggcttc atctccgccg
ctgaactgcg tcacgtcatg acttctattg gcgaaaaatt gaccgatgac gaggttgacg
agatgatccg ggaggctgac caggacggcg acggccgcat cgactgtagg cgctaaagag
ttgctcggtt cttcaaaca ctaaa
```

Программа строит вот такую точечную матрицу:



Также программа выдает процент идентичности заданных последовательностей:

```
PS E:\Лабы\Биоинф  
Identity: 30.37%
```

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:

