Блок2 Практикум 5 : Парное выравнивание белков

Нуждиной Екатерины, 1 курс ФББ

Задание №1

В этом задании нужно получить несколько коротких фрагментов (по 20 аминокислот) из искусственно смоделированного мутанта своего белка (в моем случае - белка YQGN\_BACSU) при помощи специального скрипта evolve\_protein.pl .

**Таблица №1**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Параметр* | *Опция для скрипта* | *Мутант № 1\_1* | *Мутант № 2\_1* | *Мутант № 3\_1* |
| Вероятность изменения остатка (моделирующая «ошибку» ДНК- полимеразы) | -c | 0.6 | 0.6 | 0.4 |
| Вероятность замены остатка в случае, если данная позиция будет изменена. | -r | 0.6 | 0.8 | 0.8 |

Запустим скрипт evolve\_protein.pl из Putty с нужными нам параметрами (вместо «…» берем соответствующий параметр из таблицы №1 для каждого из мутантов, вместо «///» номер мутанта).

* perl evolve\_protein.pl -i P54491.fasta –с … –r … -o P54491\_result\_///.txt

После получения 3-ех файлов содержащих мутированные последовательности белка из 20 аминокислот создадим файлы формата .fasta, содержащие fasta исходного белка и мутированной последовательности.

Настроим цветовую схему в JalView для аминокислот следующим образом:

1. Положительно заряженные – Lys(K), Arg(R), His(H) - светло-зеленым цветом
2. Отрицательно заряженные – Asp(D), Glu(E) – синим цветом
3. Полярные – Ser(S), Thr(T), Cys(C), Met(M), Asn(N), Gln(Q) – красным цветом
4. Неполярные – Gly(G), Ala(A), Val(V), Leu(L), Ile(I), Pro(P) – желтым
5. Ароматические – Phe(F), Tyr(Y), Trp(W) – темно-зеленым

Для этого зайдем в colour🡪 user defined. В открывшемся окне для аминокислот слева выберем соответственный цвет справа. Сохраним схему «save scheme»

*Мутант № 1\_1*

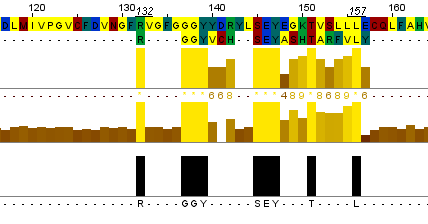
Вероятность изменения остатка - 0,6. Вероятность замены остатка в случае, если данная позиция будет изменена - 0,6.

Процент идентичности: 9/25 (36%)

Процент сходства:12 /25 (48%)

Вес по матрице BLOSUM62: 5 - 4 – 4 - 4 – 4 +6+6+7-1-3+0 - 4 – 4+4+5+7-1+0-1+5+0-1+0+1+4-2 = 17

**Рис.1 (Выравнивание мутированного фрагмента[нижняя последовательность] с исходным белком[верхняя])**



*Мутант № 2\_1*

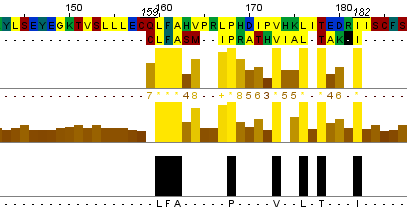
Вероятность изменения остатка - 0,6. Вероятность замены остатка в случае, если данная позиция будет изменена - 0,8.

Процент идентичности: 8/23 (35,8%)

Процент сходства: 11/23 (47,8%)

Вес по матрице BLOSUM62: -3+4+6+4-1+1-4-4+2+7+0-2-1-2+4-3-1+4-4+5-1-1-4+4 = 10

**Рис.2 (Выравнивание мутированного фрагмента[нижняя последовательность] с исходным белком[верхняя])**

****

*Мутант № 3\_1*

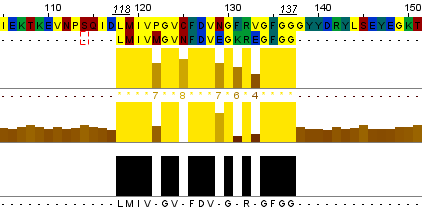
Вероятность изменения остатка - 0,4. Вероятность замены остатка в случае, если данная позиция будет изменена - 0,8.

Процент идентичности: 15/20 (75%)

Процент сходства: 16/20 (80%)

Вес по матрице BLOSUM62: 4+5+4+4-2+6+4-3+6+6+4+0+6-3+5-2+6+6+6+6 = 68

**Рис.3 (Выравнивание мутированного фрагмента[нижняя последовательность] с исходным белком[верхняя])**

****

Заданные параметры для мутанта2 дают с вероятностью 0,6\*0,8=0,48 дают замену аминокислотного остатка в определенной позиции.

Последовательность (-c 0.4 -r 0.8) больше всего схожа с изначальной.

Задание №2

Нужно было построить выравнивание своего белка и его предполагаемых ортологов или гомологов.

Были выбраны следующие ортологи:

YQGN\_BACSU , A7Z6Q5\_BACA2 и Q65HC6\_BACLD

При помощи программы Jalview 2.8 с использованием Muscle были получены выравнивания этих последовательностей.

Последовательности попарно анализировались с использованием команды infoalign пакета EMBOSS. Для этого мы зашли через Putty и ввели в командную строку infoalign file.fasta , где file – это путь к файлу с попарными выравниваниями. Получили два соответствующих файла (nuzhdina\_yqgn\_a7z6q5.infoalign и nuzhdina\_yqgn\_q65hc6.infoalign)

Для каждой пары определялся процент идентичности, процент сходства.

Name SeqLen AlignLen Gaps GapLen Ident Similar Differ % Change Weight

YQGN\_BACSU 187 187 0 0 187 0 0 0.000000 1.000000

A7Z6Q5\_BACA2 184 184 0 0 104 27 53 43.478260 1.000000

Соответственно процент идентичности = 104\*100/184 – 56.52%,

Процент сходства = (94+15)\*100/184 – 59.24%

Name SeqLen AlignLen Gaps GapLen Ident Similar Differ % Change Weight

YQGN\_BACSU 187 187 0 0 187 0 0 0.000000 1.000000

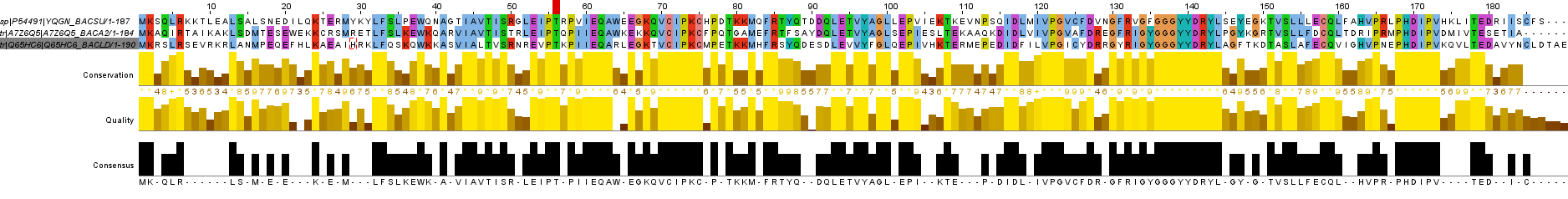
Q65HC6\_BACLD 190 190 0 0 94 38 58 50.526318 1.000000

Соответственно процент идентичности = 94\*100/190 – 49.47%,

Процент сходства = (27+94)\*100/190 – 63.68%

Раскрасим выравнивание по схеме ClustalX (colour🡪ClustalX). Выберем опцию "Above identity ntreshold" для того, чтобы окрашивать только те позиции, в которых превышен процент совпадающих букв (так, чтобы окрашивались позиции, в которых, как минимум, две совпадающие буквы). Для этого порог "identity treshold" установим на 60. (см. рис.2)

**Рис.2 Выравнивание ортологов белка 1YDM, раскрашенное по схеме ClustalX c порогом identity treshold равным 60.**



Сохраним проект JalViw (формат - .jar):

Выберем в главном окне «file» 🡪 «Save Project»