

Обработка данных секвенирования Sanger

Мелешко Дмитрий, Малыгина Татьяна

СПБАУ, Институт биоинформатики

Декабрь 21, 2013

Полная автоматизация процесса обработки данных секвенирования:

1. Обработка выходных файлов секвенатора Sanger
2. Выравнивание полученных ридов с различными параметрами
3. Построение консенсусных последовательностей и хроматограмм

Зачем все это нужно?

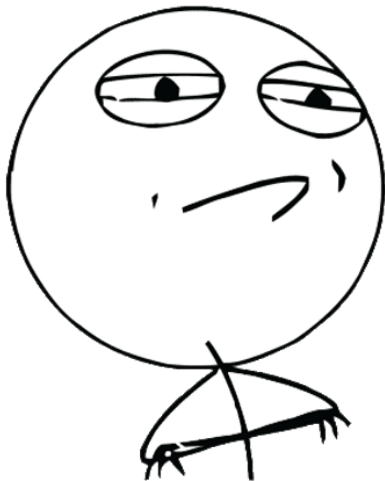
Существующие программные средства не предназначены для массовой обработки результатов секвенирования Sanger, хороших средств для построения консенсусных хроматограмм вовсе не существует

Полная автоматизация процесса обработки данных секвенирования:

1. Обработка выходных файлов секвенатора Sanger
2. Выравнивание полученных ридов с различными параметрами
3. Построение консенсусных последовательностей и хроматограмм
4. Заставить работать вместе Ruby и C++

Challenge accepted!

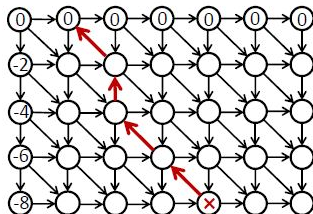
CHALLENGE ACCEPTED



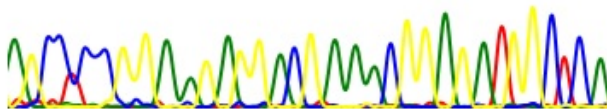
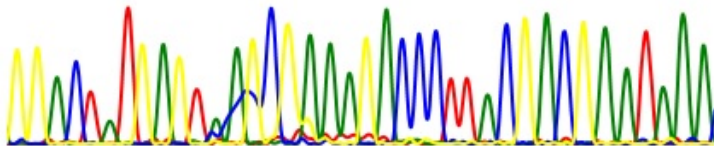
Semi-Global Alignment

Biological problem: Find approximate matches to a given pattern within a large sequence. For example, seek promoters within a DNA sequence, or a copies of a domain within a protein sequence.

Solution: *Semi-global alignment.* Needleman-Wunsch algorithm with two modifications: 1) Penalize end gaps in the pattern, but not in the long sequence; 2) Trace back from the highest scoring node along the long edge of the path graph.



Построение хроматограмм



I SHOULD BE:



I AM:



Спасибо за внимание!

