

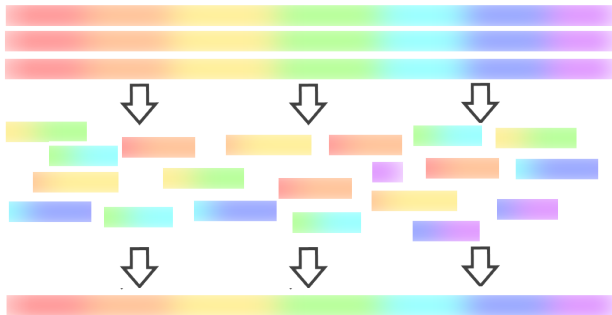
Построение графа связей геномных последовательностей

Черникова Ольга
Руководитель: Пржибельский Андрей

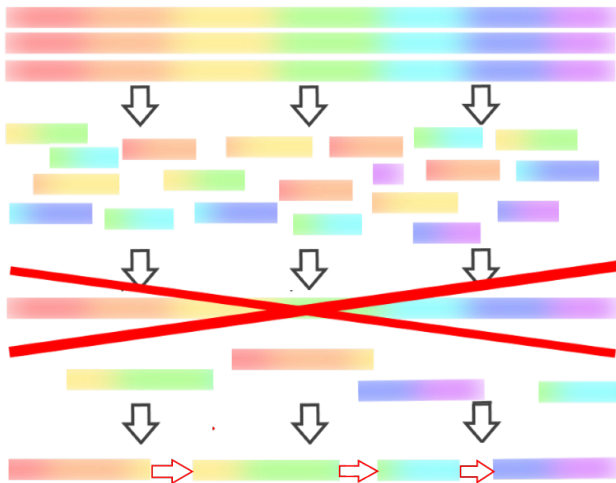
СПб АУ РАН

21 июня 2017

Задача сборки генома



Задача сборки генома

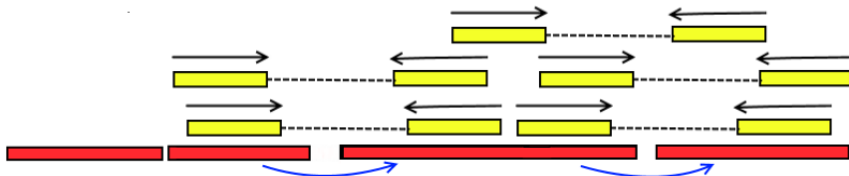


По парным ридам ДНК

■ Парные риды:

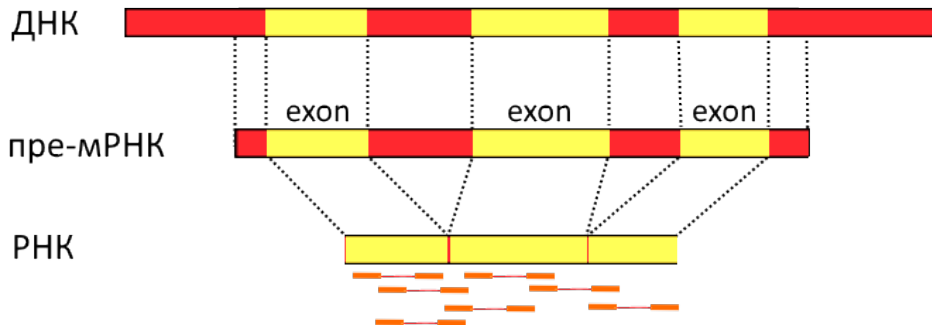


■ Нахождение связей с помощью парных ридов:



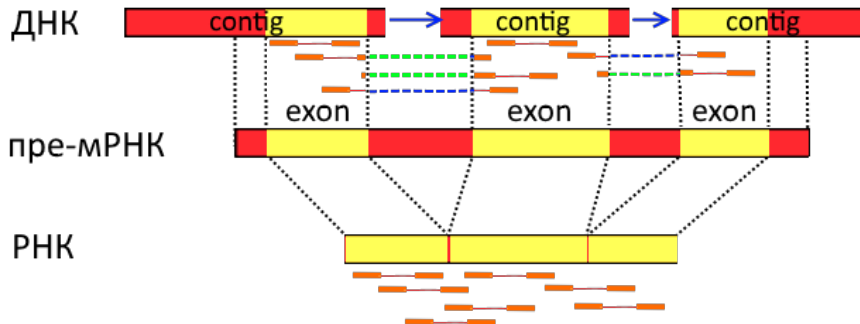
По ридам РНК

Связь между ДНК и РНК:



По ридам РНК

Выравнивание ридов РНК на контиги:



Цель и задачи

Цель

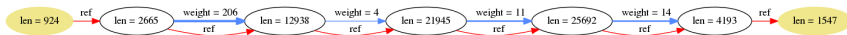
Построения скаффолдов по ридам РНК

Задачи

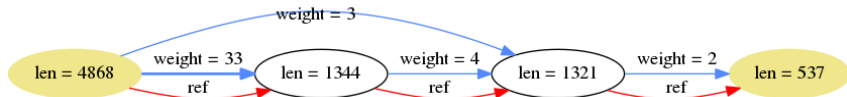
- Построение графа связей
- Построение скаффолдов основываясь на полученных связях
- Создание инструмента для визуализации графа связей между контигами
- Сравнение получившихся результатов с результатами других тулов для построения скаффолодов по ридам РНК

Построение скаффолдов

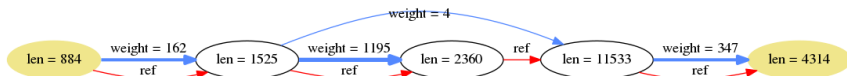
Простые пути



Ребра вдоль пути



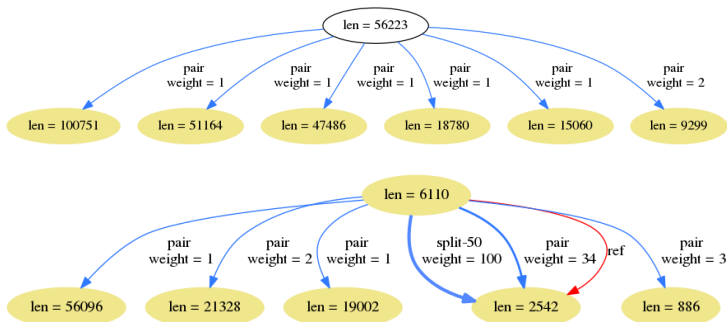
Развилка



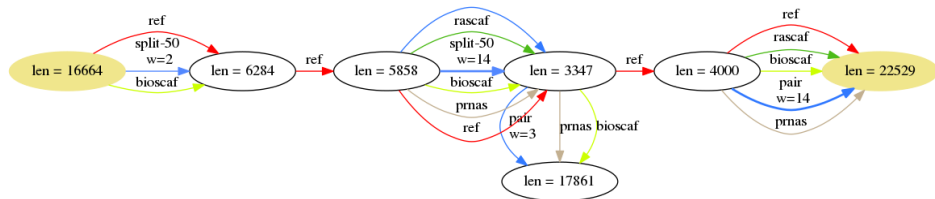
Подбор параметров фильтрации

Проблема

Связи между контигами имеют разную природу.
В каком случае связь между считать достаточно надежной?



Визуализация скаффолдов



Сравнение

	bio_scaffolder	P_RNA_scaffolder	rascaf
NG50	36855	36075	32879
NG75	17299	17188	18395
NGA50	30383	28828	27116
NGA75	12735	12489	11667
LGA50	918	955	995
misassemblies	529	621	521

Результаты и планы

Результаты

- Создание тула для построения скаффолдов по данным РНК
- Создание инструмента для визуализации графа связей, который помогает анализировать корректность построенных соединений

Дальнейшее развитие

- Тестирование и сравнение на большем разнообразии данных
- Ускорение работы приложения
- Реализация новых идей для построения скаффолдов
- Написание документации и удобного интерфейса
- Написать статью

Спасибо за внимание

Репозиторий: https://github.com/olga24912/bio_scaffolder