

Руководство по OMTools v1.4

Олден Леунг , Тинг -Фунг Чан
Китайский университет Гонконга

10 марта 2018 г .

Авторские права 2018, Все права защищены

Содержание

1 Введение	1
2 Доступность и реализация	1
3 Быстрый старт	2
4 Основные процедуры анализа	2
Я картограф	4
5 OMBlastMapper	4
6 OMNAMapper	8
7 OMFMMapper	12
8 Парное выравнивание	16
II Моделирование	19
9 OptMapDataGenerator	19
10 Генератор случайных ссылок	21
III SV-обнаружение	22
11 SVDetection	22
IV Fasta Инструменты	24
12 FastaToOM	24
V Инструменты данных	25
13 DataTools	25
14 Проверка качества данных	27
15 Данные Статистика	28
16 Дублированных Молекул Обнаружение	29
17 Дублированных Молекул Удаление	30
18 FrequentKmerВыделение	31
Инструменты выравнивания VI	32

19 Результаты инструменты	32
20 Результат объединение	36
21 Результат статистика	37
22 PrecisionRecallGraphDataGenerator	38
23 Запрос Обратный	40
24 Выравнивание выделение	41
VII Множественное выравнивание	42
25 Множественное выравнивание	42
VIII Инструменты для множественного выравнивания	49
26 CBLTools	49
27 Анализ производительности множественного выравнивания	50
28 Блок соединение ГрафГенерация	51
IX Филогенетика	52
29 UPGMATreeConstruction	52
X Визуализация	53
30 OMView	53
XI Другие сценарии	69
31 TWINResultRepeatRemover	69
32 Отдельный BNXScan	70

1 Введение

Оптическое картирование — это метод захвата определенных участков фермента на длинной молекуле ДНК. Выс одными данными этой технологии является оптическая карта, представлена в виде кортежем (рисунок 1).



Рисунок 1: Пример оптической карты. Зеленый прямоугольник — это остаток ДНК, а черные столбцы — сайты фермента. Оптическая карта содержит 6 сайтов и 7 сегментов и может быть представлена в виде кортежа: [1518; 15487; 8455; 1350; 25188; 17845; 4948]. Первый и последний сегменты обычно горизонтируются при анализе, поскольку их размеры неточны.

OMTools — программный пакет, обеспечивающий эффективную и интуитивно понятную обработку и визуализацию данных. модули для обработки данных оптического картирования.

2 Документация и реализация

2.1 Документация OMTools

Можно получить по адресу <https://github.com/aldenleung/OMTools> и выпустить по лицензии GPL (подробности см. в дистрибутиве программного обеспечения).

2.2 Минимальные требования

OMTools реализован на Java 1.8. Программное обеспечение было протестировано на Ubuntu 14.10 и Microsoft Windows 7, 10.

2.3 Установка

Все необходимые библиотеки размещены в папке lib/.

1. Скомпилируйте пакет OMTools в папке OMTools:

```
javac -d bin -sourcepath src -cp "lib/*" @classes
```

2. Создайте запускаемый jar-файл для OMTools:

```
jar cvfm OMTools.jar manifest -C bin .
```

3. Запустите OMTools:

```
java -jar OMTools.jar Имя Модуля
```

3 Быстрый старт

```
java -jar OMTools.jar FastaToOM --fastain Ecoli.fa --refmapout Ecoli.ref --enzyme BspQI
```

```
java -jar OMTools.jar Статистика данных --optmapin Ecoli.ref --statout ReferenceStat.txt
```

```
java -jar OMTools.jar OptMapDataGenerator --refmapin Ecoli.ref --optmapout EcoliExample.sdata --cov 100
```

```
java -jar OMTools.jar Статистика данных --refmapin Ecoli.ref --optmapin EcoliExample.sdata --statout Статистика данных .txt
```

```
java -jar OMTools.jar OMBlastMapper --refmapin Ecoli.ref --optmapin EcoliExample.sdata --optresout EcoliExample.omd --thread 2
```

```
java -jar OMTools.jar ResultStatistics --refmapin Ecoli.ref --optmapin EcoliExample.sdata --optresin EcoliExample.omd --statout ResultStat.txt
```

```
java -jar OMTools.jar OMView --viewrefin Ecoli.ref --viewresin EcoliExample.omd --viewregion 1:1-1000000
```

4 Основные процессы анализа

Перед генерацией оптического карттирования пользователям рекомендуется провести моделирование и определить лучший фермент для использования. Если у пользователей есть файлы с борками по ледовательности (например, короткие контиги по ледовательности, длинные каркасы по ледовательности или сущесвующие референтные по ледовательности), они могут выполнить *in silico*-переваривание. Такое преобразование из fasta в данные оптического карттирования можно выполнить с помощью модуля FastaToOM. Пользователи получают две части информации. Во-первых, с помощью модуля DataStatistics пользователи могут генерировать некоторые основные статистические данные переваренной по ледовательности. Одним из наиболее важных чисел, на которые следует обратить внимание, является плотность участков фермента (плотность с игнала). Анализ данных в оптическом карттировании становится очень сложным, если плотность сигнала слишком высока или слишком низка (диапазон «хорошей» плотности сигнала зависит от платформы, используемой для генерации данных оптического карттирования, и пользователям рекомендуется проконсультироваться с поставщиками для получения более подробной информации).

Во-вторых, если конструкция платформы основана на маркировке ферментом никеля, разрывы с айтаницием может стать серьезной проблемой для непрерывности с борками оптического карттирования. Для каждого сайта фермента модуль FastaToOM предоставляет расположение до ближайшего сайта фермента. Пользователи могут предсказать количество разрывов с айтаницием, установив отсечку расположения (то же, расположение между двумя сайтами никеля, которое может привести к разрыву, зависит от платформы, используемой для генерации данных оптического карттирования, и пользователям рекомендуется проконсультироваться с поставщиками для получения более подробной информации). Кроме того, если пользователи хотят завершить с борками с использованием данных оптического карттирования, важно посмотреть на местоположение проигнорируемых разрывов с айтаницием в интересующих контигах. Лучше избегать разрывов с айтаницием вблизи концов контигов.

После определения ис пользованием фермента пользователи могут перейти к эксперименту по оптическому картированию. Когда пользователи получают необработанные данные оптического карттирования с высокой пропускной способностью, рекомендуется сгенерировать статистику по данным оптического карттирования с помощью модуля DataStatistics. Обычно большой интерес представляют некоторые критерии: пропускная способность данных, плотность сигнала и средняя длина молекулы. Затем пользователи могут использовать модуль DataTools для обработки и фильтрации данных. Иногда одни данные дублируются, когда несолько необработанных файлов оптического карттирования объединяются из-за некоторых неожиданных человеческих ошибок. В таком случае пользователи могут захотеть выполнить быстрое сканирование с помощью модуля DuplicatedMoleculesDetection и удалить любые дублированные молекулы с помощью модуля DuplicatedMoleculesRemover.

Далее, одним из анализов данных восходящего потока будет выравнивание. OMTools предоставляет несколько модулей выравнивания, включая модули OMBlastMapper, OMNAMapper и OMFMMapper. Модуль PairwiseAlignment также может помочь выполнить парное выравнивание по нескольким наборам данных.

Затем пользователи могут обрабатывать и фильтровать результаты выравнивания с помощью модуля ResultTools и генерировать

статистика с ис пользованием модуля ResultStatistics. Поскольку несолько методов недавно были опубликованы другими исследователями группами, которые могут выполнять выравнивание данных оптического картирования, пользователи могут рассматривать возможность получения объединения или пересечения результатов выравнивания из нескольких методов с ис пользованием модуля ResultMerger.

Когда дело доходит до разработки новых алгоритмов, связанных с данными оптического картирования, пользователи могут быть заинтересованы в инструментах моделирования. К ним относятся модуль OptMapDataGenerator, который генерирует данные оптического картирования с учетом опорной оптической карты, и модуль RandomReferenceGenerator, который генерирует случайную опорную оптическую карту. Если пользователи разрабатывают алгоритм выравнивания, модуль PrecisionRecallGraphDataGenerator генерирует таблицу, которую можно использовать для генерации графиков точности-воспоминания.

И последнее, но не менее важное: пользователи могут захотеть дополнительного исследования опорных примеров, анализируя данные оптического картирования. Модуль OMView служит для этого многофункциональным визуализатором данных оптического картирования.

Часть 1

Картограф

5 OMBlastMapper

Выполняет выравнивание данных оптического картирования. Алгоритм OMBlast использует подход «seed-and-extend» для выравнивания оптических карт.

5.1 Общие параметры Mapper --minsig Минимальный сигнал

запрос для выравнивания. [По умолчанию: 5]

--minsize Минимальный размер запроса для выравнивания. [По умолчанию: 50000]

--exactmatch Включить точное соответствие запроса с сылке. Отключить эту опцию при выполнении синтаксиса выравнивания. [По умолчанию: true]

5.1.1 Параметры модуля слияния перекрывающихся ся выравниваний

--overlapmergemode Режим 0: Отключить шаг слияния; 1: Объединить одинаковые частичные выравнивания; 2: Объединить перекрывающиеся частичные выравнивания [По умолчанию: 2]

--match Оценка сопоставление сигнала [По умолчанию: 5] --fpr Штраф

за дополнительный сигнал [По умолчанию: 2]

--fpr Штраф за отставание сигнала [По умолчанию: 2] --local Включить

локальное выравнивание [По умолчанию: true]

5.1.2 Параметры фильтра результатов

--filtermode Режим фильтра: 0: Без фильтра; 1: Фильтровать по всем следующим параметрам; 2: Фильтровать по минимуму только оценка [По умолчанию: 0]

--minmatch Минимальное количество совпадений частичного выравнивания [По умолчанию: 3] --maxfp Максимальное

количество дополнительных сигналов частичного выравнивания [По умолчанию: 10000]

--maxfn Максимальное количество пропущенных сигналов частичного выравнивания [По умолчанию: 10000]

--maxfpr Максимальная скорость дополнительных сигналов частичного выравнивания [По умолчанию: 1.0E-4] --maxfmr

Максимальная скорость пропущенных сигналов частичного выравнивания [По умолчанию: 0.5]

--minscore Минимальная оценка частичного выравнивания [По умолчанию: 0,0] --minsubfragratio

Минимальное отношение субфрагментов частичного выравнивания [По умолчанию: 0,0]

--minsigratio Минимальное отношение выровненного сигнала при частичном выравнивании [По умолчанию: 0,0]

--trimmode Режим обрезки: 0: Режим обрезки отключен; 1: Режим обрезки включен [По умолчанию: 0] --maxtrim Максимальное количество шагов обрезки частичного выравнивания [По умолчанию: 5]

--match Оценка сопоставление сигнала [По умолчанию: 5] --fpr Штраф

за дополнительный сигнал [По умолчанию: 2]

--fpr Штраф за отставание сигнала [По умолчанию: 2]

5.1.3 Варианты с оединения выравнивания

--alignmentjoinmode Режим: 0: Без объединения . 1: Стандартное с оединение indel. 2: Стандартное с оединение indel-inv.
3. Стандартное транс локационное прис оединение [По умолчанию : 0]

--closeoref Максимальное расстояние (ссылька) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--closefrag Максимальное расстояние (запрос) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--minmatch Минимальное количество сдвигов соответствия , которые следует считать допустимым частичным спос тавлением. [По умолчанию : 3]

--maxtrim Максимальные шаги обрезки для частичного выравнивания [По умолчанию : 5]

--trimear Допустимая погрешность масштабирования во время обрезки [По умолчанию : 0,1]

--match Оценка споставление сдвига [По умолчанию : 5]

--fpr Штраф за дополнительный сдвиг [По умолчанию : 2]

--fpn Штраф за отставание сдвига [По умолчанию : 2]

--indelp Штраф за прис оединение частичных выравниваний с отношением indel [По умолчанию : 10]

--inpr Штраф за объединение частичных выравниваний с инверсионным отношением [По умолчанию : 30]

--transp Штраф за прис оединение частичных выравниваний с транс локационной связью [По умолчанию : 50]

--localpenalty Включить штраф за локальное выравнивание для окончательного выравнивания (расматривается как глобальное выравнивание)
[По умолчанию : ложно]

--minjoinscore Минимальная оценка объединения окончательного выравнивания [По умолчанию : 30]

--minconf Минимальная достоверность (уникальность) окончательного выравнивания [По умолчанию : 0,4]

--minjoinedfragratio Минимальное отношение выровненной длины к длине запроса [По умолчанию : -1,0]

--minjoinedsigratio Минимальное отношение количества выровненных сдвигов к общему количеству сдвигов запроса [По умолчанию : -1,0]

--overlapalign Разрешить вывод перекрывающихся конечных выравниваний [По умолчанию : true]

--maxalignitem Максимальное количество выводимых окончательных выравниваний. -1: Нет ограничений на количество окончательных выравниваний.
выравнивания [По умолчанию : 1]

5.1.4 Параметры считывающего результата

--optresin Вх одной файл результатов выравнивания для повторного выравнивания

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM
(OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN;
[По умолчанию : -1]

5.2 Параметры OMBlastMapper --local

Включить локальное выравнивание [По умолчанию: true]

--allowequalrefquery Разрешить равную ссылку и запрос в выравнивании. Отключение этой опции запрещает запрос от выравнивания по себе как ссылке [По умолчанию: true]

--allowdiffrefquery Разрешить различные ссылки и запросы в выравнивании. В отличие от allowequalrefquery, отключение этой опции запрещает запросы выравниваться по другим запросам в качестве ссылок [По умолчанию: true]

--meas Ошибка измерения [По умолчанию: 500]

--ear Допустимый диапазон ошибок (допустимая погрешность масштабирования) [По умолчанию: 0,1]

--match Оценка сопоставление с игнора [По умолчанию: 5]

--fpp Штраф за дополнительный сдвиг [По умолчанию: 2]

--fnp Штраф за отставание с игнора [По умолчанию: 2]

--falselimit Максимальное количество последовательных дополнительных отставаний с игнором [По умолчанию: 5]

--maxseedno Максимальное количество первоначальных значений в запросе [По умолчанию: 10]

5.2.1 Параметры заполнения

--seedingmode Режим заполнения : 1: Оптимизировано для длинных k-меров (обычно для k больше 10); 2: Оптимизировано для коротких k-меров (обычно для k меньше или равно 10); -1: Автоматический выбор. [По умолчанию: -1]

--k Длина K-мера. [По умолчанию: 3]

--maxnosignal Максимальная область отставания с игнором между сдвигами для заполнения. [По умолчанию: 10000000]

5.3 Параметры многопоточности

--thread Количество потоков [По умолчанию: 1]

5.4 Параметры чтения с ссылок --refmapin

Вх одн файл карты с ссылок [Обязательно]

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)

(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);

3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX

Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF);

9: Формат OpGen XML; 10: Формат данных Валуева;

11: Формат карт Maligner; [По умолчанию: -1]

5.5 Параметры считывателя данных --

optmapin Вх одной файл оптической карты [обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный с тандемный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OptGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

5.6 Параметры записи результатов --optresout

Вывод файла результата выравнивания [Обязательно]

--optresoutformat Формат файла результата -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

--writeunmap Записать отброшенные или неотображенные молекулы. [По умолчанию: true]

--multiple Записать несколько карт для молекулы. [По умолчанию: true]

--writeinfo Записать информацию о молекуле. [По умолчанию: true]

6 ОМНАMapper

Выполняет выравнивание данных оптического картирования. Алгоритм ОМНА использует эвристический подход для выравнивания оптических карт.

6.1 Общие параметры картографа

```
--minsig Минимальный сигнал запроса для выравнивания. [По умолчанию: 5]
--minsize Минимальный размер запроса для выравнивания. [По умолчанию: 50000]
--exactmatch Включить точное соответствие запроса с сылкой. Отключить эту опцию при выполнении синтаксиса выравнивания. [По умолчанию: true]
```

6.1.1 Параметры модуля слияния перекрывающихся ся выравниваний

```
--overlapmergemode Режим: 0: Отключить шаг слияния; 1: Объединить одинаковые частичные выравнивания; 2: Объединить перекрывающиеся частичные выравнивания [По умолчанию: 2]
--match Оценка сопоставление сигнала [По умолчанию: 5]
--fpp Штраф за дополнительный сигнал [По умолчанию: 2]
--fnp Штраф за отсутствие сигнала [По умолчанию: 2]
--local Включить локальное выравнивание [По умолчанию: true]
```

6.1.2 Параметры фильтра результатов

```
--filtermode Режим фильтра: 0: Без фильтра; 1: Фильтровать по всем следующим параметрам; 2: Фильтровать по минимуму только оценка [По умолчанию: 0]
--minmatch Минимальное количество совпадений частичного выравнивания [По умолчанию: 3]
--maxfp Максимальное количество дополнительных сигналов частичного выравнивания [По умолчанию: 10000]
--maxfn Максимальное количество пропущенных сигналов частичного выравнивания [По умолчанию: 10000]
--maxfpr Максимальная скорость дополнительных сигналов частичного выравнивания [По умолчанию: 1.0E-4]
--maxfnr Максимальная скорость пропуска сигналов частичного выравнивания [По умолчанию: 0,5]
--minscore Минимальная оценка частичного выравнивания [По умолчанию: 0,0]
--minsubfragratio Минимальное отношение субфрагментов частичного выравнивания [По умолчанию: 0,0]
--minsigratio Минимальное отношение выровненного сигнала при частичном выравнивании [По умолчанию: 0,0]
--trimmode Режим обрезки: 0: Режим обрезки отключен; 1: Режим обрезки включен [По умолчанию: 0]
--maxtrim Максимальное количество шагов обрезки частичного выравнивания [По умолчанию: 5]
--match Оценка сопоставление сигнала [По умолчанию: 5]
--fpp Штраф за дополнительный сигнал [По умолчанию: 2]
--fnp Штраф за отсутствие сигнала [По умолчанию: 2]
```

6.1.3 Варианты с оединения выравнивания

--alignmentjoinmode Режим: 0: Без объединения . 1: Стандартное с оединение indel. 2: Стандартное с оединение indel-inv.
3. Стандартное транс локационное прис оединение [По умолчанию : 0]

--closeoref Максимальное расстояние (ссылька) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--closefrag Максимальное расстояние (запрос) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--minmatch Минимальное количество сдвигов соответствия , которые следует считать допустимым частичным спос тавлением. [По умолчанию : 3]

--maxtrim Максимальные шаги обрезки для частичного выравнивания [По умолчанию : 5]

--trimear Допустимая погрешность масштабирования во время обрезки [По умолчанию : 0,1]

--match Оценка споставление сдвига [По умолчанию : 5]

--fpr Штраф за дополнительный сдвиг [По умолчанию : 2]

--fpn Штраф за отставание сдвига [По умолчанию : 2]

--indelp Штраф за прис оединение частичных выравниваний с отношением indel [По умолчанию : 10]

--inpr Штраф за объединение частичных выравниваний с инверсионным отношением [По умолчанию : 30]

--transp Штраф за прис оединение частичных выравниваний с транс локационной связью [По умолчанию : 50]

--localpenalty Включить штраф за локальное выравнивание для окончательного выравнивания (расматривается как глобальное выравнивание)
[По умолчанию : ложно]

--minjoinscore Минимальная оценка объединения окончательного выравнивания [По умолчанию : 30]

--minconf Минимальная достоверность (уникальность) окончательного выравнивания [По умолчанию : 0,4]

--minjoinedfragratio Минимальное отношение выровненной длины к длине запроса [По умолчанию : -1,0]

--minjoinedsigratio Минимальное отношение количества выровненных сдвигов к общему количеству сдвигов запроса [По умолчанию : -1,0]

--overlapalign Разрешить вывод перекрывающихся конечных выравниваний [По умолчанию : true]

--maxalignitem Максимальное количество выводимых окончательных выравниваний. -1: Нет ограничений на количество окончательных выравниваний.
выравнивания [По умолчанию : 1]

6.1.4 Параметры считывателя результатов

--optresin Вх одный файл результатов выравнивания для повторного выравнивания

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM
(OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN;
[По умолчанию : -1]

6.2 Параметры OMNAMapper --local Включить

локальное выравнивание [По умолчанию : true]

--localstart Локальная начальная позиция для выравнивания , 0: начинается с каждого сегмента (исчерпывающий), x: начинается с первого сегмента, -x: запускается без последних x сегментов. [По умолчанию : 0]

--scorefilter Первичный фильтр очков во время выравнивания [По умолчанию : 30]

--deg Вырождение сегментов закрытия для обработки ошибок разрешения . [По умолчанию : 1500]

--meas Ошибка измерения [По умолчанию : 500]

--ear Допустимый диапазон ошибок (допустимость решения масштабирования) [По умолчанию : 0,1]

--match Оценка сопоставление сегмента [По умолчанию : 5]

--fpr Штраф за дополнительный сегмент [По умолчанию : 2]

--fnr Штраф за отсутствие сегмента [По умолчанию : 2]

--falselimit Максимальное количество последовательных ложных сегментов [по умолчанию : 5]

6.3 Параметры многопоточности

--thread Количество потоков [По умолчанию : 1]

6.4 Параметры чтения с сывороток --refmapin Вх одной

файл карты с сывороток [Обязательно]

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию : -1]

6.5 Параметры считывателя данных --

optmapin Вх одной файл оптической карты [обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию : -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию : 3.0]

6.6 Параметры записи результатов --

optresout Вывод файла результатов выравнивания [Обязательно]

--optresoutformat Формат файла результата -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

--writeunmap Записать отброшенные или неотображенные молекулы. [По умолчанию: true]

--multiple Записать несколько карт для молекулы. [По умолчанию: true]

--writeinfo Записать информацию о молекуле. [По умолчанию: true]

7 OMFMapper

Выполняет выравнивание данных оптического картирования. Алгоритм OMFMapper использует индексный подход для выравнивания оптических карт.

7.1 Общие параметры картографа

```
--minsig Минимальный сигнал запроса для выравнивания. [По умолчанию: 5]
--minsize Минимальный размер запроса для выравнивания. [По умолчанию: 50000]
--exactmatch Включить точное соответствие запроса с сылкой. Отключить эту опцию при выполнении симметричного выравнивания. [По умолчанию: true]
```

7.1.1 Параметры модуля слияния перекрывающихся ся выравниваний

```
--overlapmergemode Режим: 0: Отключить шаг слияния; 1: Объединить одинаковые частичные выравнивания; 2: Объединить перекрывающиеся частичные выравнивания [По умолчанию: 2]
--match Оценка сопоставления сигнала [По умолчанию: 5]
--fpp Штраф за дополнительный сигнал [По умолчанию: 2]
--fnp Штраф за отсутствие сигнала [По умолчанию: 2]
--local Включить локальное выравнивание [По умолчанию: true]
```

7.1.2 Параметры фильтра результатов

```
--filtermode Режим фильтра: 0: Без фильтра; 1: Фильтровать по всем следующим параметрам; 2: Фильтровать по минимуму только оценка [По умолчанию: 0]
--minmatch Минимальное количество совпадений частичного выравнивания [По умолчанию: 3]
--maxfp Максимальное количество дополнительных сигналов частичного выравнивания [По умолчанию: 10000]
--maxfn Максимальное количество пропущенных сигналов частичного выравнивания [По умолчанию: 10000]
--maxfpr Максимальная скорость дополнительных сигналов частичного выравнивания [По умолчанию: 1.0E-4]
--maxfnr Максимальная скорость пропуска сигналов частичного выравнивания [По умолчанию: 0,5]
--minscore Минимальная оценка частичного выравнивания [По умолчанию: 0,0]
--minsubfragratio Минимальное отношение субфрагментов частичного выравнивания [По умолчанию: 0,0]
--minsigratio Минимальное отношение выравненного сигнала при частичном выравнивании [По умолчанию: 0,0]
--trimmode Режим обрезки: 0: Режим обрезки отключен; 1: Режим обрезки включен [По умолчанию: 0]
--maxtrim Максимальное количество шагов обрезки частичного выравнивания [По умолчанию: 5]
--match Оценка сопоставления сигнала [По умолчанию: 5]
--fpp Штраф за дополнительный сигнал [По умолчанию: 2]
--fnp Штраф за отсутствие сигнала [По умолчанию: 2]
```

7.1.3 Варианты с оединения выравнивания

--alignmentjoinmode Режим: 0: Без объединения . 1: Стандартное с оединение indel. 2: Стандартное с оединение indel-inv.
3. Стандартное транс локационное прис оединение [По умолчанию : 0]

--closeoref Максимальное расстояние (ссылька) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--closefrag Максимальное расстояние (запрос) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--minmatch Минимальное количество сдвигов соответствия , которые следует считать допустимым частичным спос тавлением. [По умолчанию : 3]

--maxtrim Максимальные шаги обрезки для частичного выравнивания [По умолчанию : 5]

--trimear Допустимая погрешность масштабирования во время обрезки [По умолчанию : 0,1]

--match Оценка споставление сдвига [По умолчанию : 5]

--fpr Штраф за дополнительный сдвиг [По умолчанию : 2]

--fpn Штраф за отставание сдвига [По умолчанию : 2]

--indelp Штраф за прис оединение частичных выравниваний с отношением indel [По умолчанию : 10]

--inpr Штраф за объединение частичных выравниваний с инверсионным отношением [По умолчанию : 30]

--transp Штраф за прис оединение частичных выравниваний с транс локационной связью [По умолчанию : 50]

--localpenalty Включить штраф за локальное выравнивание для окончательного выравнивания (расматривается как глобальное выравнивание)
[По умолчанию : ложно]

--minjoinscore Минимальная оценка объединения окончательного выравнивания [По умолчанию : 30]

--minconf Минимальная достоверность (уникальность) окончательного выравнивания [По умолчанию : 0,4]

--minjoinedfragratio Минимальное отношение выровненной длины к длине запроса [По умолчанию : -1,0]

--minjoinedsigratio Минимальное отношение количества выровненных сдвигов к общему количеству сдвигов запроса [По умолчанию : -1,0]

--overlapalign Разрешить вывод перекрывающихся конечных выравниваний [По умолчанию : true]

--maxalignitem Максимальное количество выводимых окончательных выравниваний. -1: Нет ограничений на количество окончательных выравниваний.
выравнивания [По умолчанию : 1]

7.1.4 Параметры считывателя результатов

--optresin Вх одной файл результатов выравнивания для повторного выравнивания

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM
(OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN;
[По умолчанию : -1]

7.2 Параметры OMFMMapper --meas Ошибка

измерения [По умолчанию : 500]

--ear Допустимый диапазон ошибок (допустимая погрешность масштабирования) [По умолчанию : 0,1]

--match Оценка сопоставление с игнора [По умолчанию : 5]

--fpr Штраф за дополнительный сдвиг [По умолчанию : 2]

--fnr Штраф за отсутствие сдвига [По умолчанию : 2]

--rfalselimit Максимальное количество последовательных ложных сдвигов на ссылку [По умолчанию : 5]

--qfalselimit Максимальное количество последовательных ложных сдвигов на запрос [По умолчанию : 5]

--cfalselimit Максимальное количество последовательных ложных сдвигов для ссылки и запроса [По умолчанию : 5]

--malignscore Минимальная оценка на этапе выравнивания [По умолчанию : 20]

7.2.1 Параметры заполнения

--seedingmode Режим заполнения : 1: Оптимизировано для длинных k-меров (обычно для k больше 10); 2: Оптимизировано для коротких k-меров (обычно для k меньше или равно 10); -1: Автоматический выбор. [По умолчанию : -1]

--k Длина k-мера. [По умолчанию : 3]

--maxnosignal Максимальная область отсутствия сдвига между сдвигами для заполнения. [По умолчанию : 10000000]

7.3 Параметры многопоточности --thread

Количество потоков [По умолчанию : 1]

7.4 Параметры считывателя ссылок --refmapin

Вх одн файл карты с ссылок [Обязательно]

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию : -1]

7.5 Параметры считывателя данных --

optmapin Вх одн файл оптической карты [Обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию : -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию : 3.0]

7.6 Параметры записи и результатов --

optresout Вывод файла результатов выравнивания [Обязательно]

--optresoutformat Формат файла результата -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

--writeunmap Записать отброшенные или неотображенные молекулы. [По умолчанию: true]

--multiple Записать несколько карт для молекулы. [По умолчанию: true]

--writeinfo Записать информацию о молекуле. [По умолчанию: true]

8 Парное выравнивание

Выполняет парное выравнивание файлов данных на основе OMBlastMapper. Введите несколько файлов данных для парного выравнивания между каждой парой из них.

8.1 Общие параметры картографа

```
--minsig Минимальный сдвиг запроса для выравнивания. [По умолчанию: 5]
--minsize Минимальный размер запроса для выравнивания. [По умолчанию: 50000]
--exactmatch Включить точное соответствие запроса с сылкой. Отключить эту опцию при выполнении с амовыравнивания. [По умолчанию: true]
```

8.1.1 Параметры модуля слияния перекрывающихся ся выравниваний

```
--overlapmergemode Режим: 0: Отключить шаг слияния; 1: Объединить одинаковые частичные выравнивания; 2: Объединить
перекрывающиеся частичные выравнивания [По умолчанию: 2]
--match Оценка сопоставление с игнором [По умолчанию: 5]
--fpp Штраф за дополнительный сдвиг [По умолчанию: 2]
--fnp Штраф за отставание с игнором [По умолчанию: 2]
--local Включить локальное выравнивание [По умолчанию: true]
```

8.1.2 Параметры фильтра результатов

```
--filtermode Режим фильтра: 0: Без фильтра; 1: Фильтровать по всем следующим параметрам; 2: Фильтровать по минимуму
только оценка [По умолчанию: 0]
--minmatch Минимальное количество совпадений частичного выравнивания [По умолчанию: 3]
--maxfp Максимальное количество дополнительных сдвигов частичного выравнивания [По умолчанию: 10000]
--maxfn Максимальное количество пропущенных сдвигов частичного выравнивания [По умолчанию: 10000]
--maxfpr Максимальная скорость дополнительных сдвигов частичного выравнивания [По умолчанию: 1.0E-4]
--maxfnr Максимальная скорость пропуска сдвигов частичного выравнивания [По умолчанию: 0,5]
--minscore Минимальная оценка частичного выравнивания [По умолчанию: 0,0]
--minsubfragratio Минимальное отношение субфрагментов частичного выравнивания [По умолчанию: 0,0]
--minsigratio Минимальное отношение выровненного сдвига при частичном выравнивании [По умолчанию: 0,0]
--trimmode Режим обрезки: 0: Режим обрезки отключен; 1: Режим обрезки включен [По умолчанию: 0]
--maxtrim Максимальное количество шагов обрезки частичного выравнивания [По умолчанию: 5]
--match Оценка сопоставление с игнором [По умолчанию: 5]
--fpp Штраф за дополнительный сдвиг [По умолчанию: 2]
--fnp Штраф за отставание с игнором [По умолчанию: 2]
```

8.1.3 Варианты с оединения выравнивания

--alignmentjoinmode Режим: 0: Без объединения . 1: Стандартное с оединение indel. 2: Стандартное с оединение indel-inv.
3. Стандартное транс локационное прис оединение [По умолчанию : 0]

--closeoref Максимальное расстояние (ссылка) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--closefrag Максимальное расстояние (запрос) между двумя частичными выравниваниями, которые необх одимо объединить [По умолчанию : 250000]

--minmatch Минимальное количество сдвигов соответствия , которые следует считать допустимым частичным спос тавлением. [По умолчанию : 3]

--maxtrim Максимальные шаги обрезки для частичного выравнивания [По умолчанию : 5]

--trimear Допустимая погрешность масштабирования во время обрезки [По умолчанию : 0,1]

--match Оценка споставление сдвига [По умолчанию : 5]

--fpr Штраф за дополнительный сдвиг [По умолчанию : 2]

--fpn Штраф за отставание сдвига [По умолчанию : 2]

--indelp Штраф за прис оединение частичных выравниваний с отношением indel [По умолчанию : 10]

--inpr Штраф за объединение частичных выравниваний с инверсионным отношением [По умолчанию : 30]

--transp Штраф за прис оединение частичных выравниваний с транс локационной связью [По умолчанию : 50]

--localpenalty Включить штраф за локальное выравнивание для окончательного выравнивания (расматривается как глобальное выравнивание)
[По умолчанию : ложно]

--minjoinscore Минимальная оценка объединения окончательного выравнивания [По умолчанию : 30]

--minconf Минимальная достоверность (уникальность) окончательного выравнивания [По умолчанию : 0,4]

--minjoinedfragratio Минимальное отношение выровненной длины к длине запроса [По умолчанию : -1,0]

--minjoinedsigratio Минимальное отношение количества выровненных сдвигов к общему количеству сдвигов запроса [По умолчанию : -1,0]

--overlapalign Разрешить вывод перекрывающихся конечных выравниваний [По умолчанию : true]

--maxalignitem Максимальное количество выводимых окончательных выравниваний. -1: Нет ограничений на количество окончательных выравниваний.
выравнивания [По умолчанию : 1]

8.1.4 Параметры считывателя результатов

--optresin Вх одный файл результатов выравнивания для повторного выравнивания

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM
(OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN;
[По умолчанию : -1]

8.2 Параметры OMBlastMapper --local

Включить локальное выравнивание [По умолчанию: true]

--allowequalrefquery Разрешить равную ссылку из запроса в выравнивании. Отключение этой опции запрещает запрос от выравнивания по себе как ссылке [По умолчанию: true]

--allowdiffrefquery Разрешить различные ссылки из запросов в выравнивании. В отличие от allowequalrefquery, отключение этой опции запрещает запросы выравниваться по другим запросам в качестве ссылок [По умолчанию: true]

--meas Ошибка измерения [По умолчанию: 500]

--ear Допустимый диапазон ошибок (допустимая погрешность масштабирования) [По умолчанию: 0,1]

--match Оценка сопоставления с идентификатором [По умолчанию: 5]

--fpr Штраф за дополнительный сдвиг [По умолчанию: 2]

--fnr Штраф за отставание сдвигом [По умолчанию: 2]

--falselimit Максимальное количество последовательных дополнительных отставаний сдвигом [По умолчанию: 5]

--maxseedno Максимальное количество первоначальных значений в запросе [По умолчанию: 10]

8.2.1 Параметры заполнения

--seedingmode Режим заполнения: 1: Оптимизировано для длинных k-меров (обычно для k больше 10); 2: Оптимизировано для коротких k-меров (обычно для k меньше или равно 10); -1: Автоматический выбор. [По умолчанию: -1]

--k Длина K-мера. [По умолчанию: 3]

--maxnosignal Максимальная область отставания сдвигом между сдвигами для заполнения. [По умолчанию: 10000000]

8.3 Параметры многопоточности

--thread Количество потоков [По умолчанию: 1]

8.4 Параметры считывателя данных

--optmapin Входной файл оптической карты [обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF);
9: Формат OpGen XML; 10: Формат данных Валуев; 11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

8.5 Варианты парного выравнивания

--output префикс вывода [Обязательно]

--rerun Повторно запустить, даже если файл результатов существует [По умолчанию: false]

Часть 2

Моделирование

9 OptMapDataGenerator

Генерирует моделированные данные из эталона.

9.1 Параметры чтения с правочной литературы

--refmapin Вх одной файл с правочной карты
--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию : -1]

9.2 Несколько вариантов чтения с правочных материалов

--refmaplistin Вх одной список файлов с правочной карты с сопротивлением с топонимами
--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию : -1]

9.3 Параметры генератора данных --rsIn Ошибки

разрешения [По умолчанию : 1200]
--meas Ошибка измерения [По умолчанию : 500]
--fsize Средний размер фрагмента [По умолчанию : 200000]
--fubound Верхняя граница размера, включительно [По умолчанию : 1000000]
--flbound Нижняя граница, включительно [По умолчанию : 100000]
--median Медиана для шкалы [По умолчанию : 1.0]
--scalesd SD для масштаба [по умолчанию : 0,04]
--subound Верхняя граница шкалы, включительно [По умолчанию : 1.3]
--silbound Нижняя граница шкалы, включительно [По умолчанию : 0,7]
--fpr час tota ложных срабатываний [По умолчанию : 1.0E-5]
--fnr час tota ложных отрицательных результатов [По умолчанию : 0,1]
--seed Случайное семя
--indelsize Случайный размер вставки/удаления [По умолчанию : 0]

```
--inversionmode Режим инверсии. 0: без инверсии. 1: инверсия второй половины [По умолчанию: 0]
--cov Oxят вывода данных [По умолчанию: 10.0]
--molen Количествомолекул, которые будут сгенерированы. Переопределяет параметр покрытия, если установлено положительное число
[По умолчанию: -1]
```

9.4 Параметры записи данных --optmapout

Выходной файл оптической карты [обязательно]

```
--optmapoutformat -1: Автоматически определяется по расширению файла: 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: Формат файла Spots (SPOTS); 3:
Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO)
(эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]
```

10 Генератор с лучайных с с ылок

Генерирует случайные с правочными карты путем перемешивания порядка сегментов во всех одных с правочными картах.

10.1 Параметры чтения с правочной литературы

--refmapin Вх одной файл с правочной карты [обязательно]

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

10.2 Параметры референтного автора

--refmapout Вывод файла с правочной карты [обязательно]

--refmapoutformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: Формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

Часть 3 Обнаружение СВ

11 SVDetection

Представляет базовый модуль обнаружения SV для оптической карты выравнивания

11.1 Параметры чтения с правочной литературы

--refmapin Вх одной файл с правочной карты [обязательно]

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный с стандартный формат (REF)
(эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF);
9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

11.2 Параметры считывающего устройства данных

--optmapin Вх одной файл оптической карты

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный с стандартный формат (REF)
(эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF);
9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

11.3 Параметры считывающего устройства результатов

--optresin Вх одной файл результатов выравнивания [Обязательно]

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного о выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

11.4 Стандартные параметры SV Writer

--svout Вых одной SV-файл

11.5 Обнаружение SV

--svmode Режим SV. 1: подходит с разделением карты (экспериментальный); 2: подходит с новой картой; 3: подходит с разделением карты и подходит с новой картой [По умолчанию: 3]

--minsupport Минимальная поддержка молекул для SV [По умолчанию: 5]

--flanksig Минимальное количество фланговых сигналов для выравнивания [По умолчанию: 5]

11.5.1 Параметры обнаружения на основе разделенной карты

-closeref Максимальное расстояние (ссылька) между двумя результатами, которые следуют расматривать в одном классе. [Определение]
ошибка: 250000]

-closefrag Максимальное расстояние (фрагмент) между двумя результатами, которые следуют расматривать в одном классе.
[По умолчанию: 250000]

-closesv Закрыть SV (SV объединяются на основе равенства bp1bp2. [По умолчанию: 10000]

-mergesv Объединить SV (SV объединяются на основе простоя с равенства регионов). [По умолчанию: true]

-flanksize Минимальный размер фланга в выравнивании (используется только при обнаружении разделенной карты) [По умолчанию: 0]

-minindelsize Минимальный размер вставки для SV [по умолчанию: 1500]

-maxindelsize Максимальный размер вставки для SV [по умолчанию: 1000000]

-maxsupport Максимальная поддержка молекул для SV [По умолчанию: 100]

-minsvscore Минимальный балл SV (с отношение поддержки/оппозиции [По умолчанию: 0,5]

11.5.2 Параметры обнаружения на основе сигнала

-mininvsig Минимальный сигнал, участвующий в инверсии [По умолчанию: 4]

-maxinvsig Максимальный размер инверсии [По умолчанию: 100000]

-meas Ошибка измерения [По умолчанию: 500]

-ear Допустимый диапазон ошибок (допустимая погрешность масштабирования) [По умолчанию: 0,1]

-deg Вырождение сигналов закрытия для обработки ошибок разрешения. [По умолчанию: 1500]

Часть 4

Инструменты Fasta

12 FastaToOM

Выполняет *in silico*-расщепление последовательности ДНК.

12.1 Параметры потокового Fasta Reader

--fastain вх одной файл fasta [обязательно]

12.2 Варианты ввода фермента

--enzyme Вс троенные ферменты [BspQI, BbvCI, AlwI, BsmAI, BstNBI, BsmI, BsrDI, BssSI, BtsI] (поддержка ввода нескольких ферментов)

--enzymestring Последовательность фермента (например, GCTCTTC)

12.3 Параметры автора ссылок

--refmapout Выход файлов с правочной карты [обязательно]

--refmapoutformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: Формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

12.4 Параметры программирования разрыва сайтинга

--nsbout Потенциальный сайт разрыва прерывает вывод (программа полезна для новых ферментов разрыва только)

Часть 5

Инструменты данных

13 DataTools

Представляет базовые функции для фильтрации и обработки данных оптического картирования.

13.1 Параметры инструментов данных

```
--idprefix Добавить префикс к всем идентификаторам

--idmodify Преобразовать все идентификаторы в x..x + n - 1 (x: вх одное значение, n: количество оптических карт в данных
файл). Отрицательное значение отключает эту функцию. [По умолчанию: -1]

--idmodifylog Файл журнала, содержащий преобразования идентификаторов

--staticid Использовать статический идентификатор для всех оптических карт (переделять все остальные параметры, связанные с идентификатором)

--fix Исправить данные (отрицательная коррекция расположения сигнала и т. д.) [По умолчанию: true]

--condense Объединить несколько сигналов, расположенных ближе, чем параметр, в один сигнал [По умолчанию: 0]

--removeseg Удалить элементы, меньшие, чем параметр [По умолчанию: -1]

--minsize Данные с минимальным размером для сокращения [По умолчанию: 0]

--minsig Данные с минимальным сигналом для сокращения [По умолчанию: 0]

--dataid Список идентификаторов данных, которые необязательно извлечь

--region Список регионов для извлечения.

--shift Сдвиг вперед(вправо) x bp (предположим, круговой) [По умолчанию: 0]

--randdata Количества случайных данных, которые необязательно извлечь

--seed Начальное число, используемое при извлечении случайных данных.

--concat Объединить все записи данных в одну запись. -1: не активировано; Неотрицательное значение: пробел (сегмент без сигнала) между каждой
записью данных. Игнорировать любые функции модификации данных [По умолчанию: -1]
```

13.1.1 Параметры считывателя ConcatInfo

```
--concatin Вх однай файл ConcatInfo.
```

13.1.2 Параметры записи ConcatInfo

```
--concatout Выход файла ConcatInfo.
```

13.1.3 Фильтрация низкой сложности

```
--lowcom Сохранить/удалить молекулы с низкой сложностью -1: Сохранить низкую сложность; 0: Ничего не делать;
1: Сохранить высокую сложность [По умолчанию: 0]

--maxdensity Максимальная плотность на 100 кбит/с для фильтрации [по умолчанию: 25,0]

--maxseed Максимальное семя для фильтрации [по умолчанию: 5]
```

13.2 Параметры считывающего устройства данных

--optmapin Вх однай файл оптической карты [обязательно]
--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]
--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

13.3 Параметры записи и данных --

optmapout Вых однай файл оптической карты [обязательно]

--optmapoutformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: Формат файла Spots (SPOTS); 3:
Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

14 Проверка качества данных

Выполняет базовую проверку качества данных, включая распределение длины молекул, плотности сигнала и длины сегмента.

14.1 Параметры чтения с правочной литературы

--refmapin Вх одной файл с правочной карты

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла СМАР (СМАР); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

14.2 Параметры считывающего устройства

--optmapin Вх одной файл оптической карты [обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла СМАР (СМАР); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

14.3 Параметры проверки качества данных

--name Использовать имя набора данных вместо имени файла.

--prefix Префикс вывода статистики [По умолчанию: вывод]

--gradcolor Использовать градиентный цвет [По умолчанию: false]

--imageformat Форматы сокращения изображения . [svg; png; jpg;] [По умолчанию: png]

15 Данные Статистика

Формирует статистику файла данных .

15.1 Параметры статистики данных

```
--winsize Размер окна для размера молекулы [По умолчанию: 25000]
--maxsize Максимальный размер. -1 для автоматической настройки [По умолчанию: -1]
--maxsignal Максимальный сигнал. -1 для автоматической настройки [По умолчанию: -1]
--statout Вывод статистики [Обязательно]
```

15.2 Параметры считывателя данных

```
--optmapin Вх одной файл оптической карты [обязательно]
--optmapinformат -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла СМАР (СМАР); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]
```

15.3 Параметры чтения с правочной литературы

```
--refmapin Вх одной файл с правочной карты
--refmapinformат -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла СМАР (СМАР); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]
```

16 Дублированных МолекулОбнаружение

Обнаруживает дублированные молекулы в наборе данных оптической карты. Дублированные молекулы с одержат одинаковое количество сегментов, а разница между размерами каждого сегмента очень мала (обычно менее 100 пар оснований)

16.1 Параметры многопоточного заполнения

--meas Ошибки измерения . Обычно это намного меньше обычных ошибок измерения при обнаружении дублированные молекулы [По умолчанию : 100]

--eag Допустимый диапазон ошибок [По умолчанию : 0,0]

--thread Количество потоков [По умолчанию : 1]

16.2 Параметры заполнения

--seedemode Режим заполнения : 1: Оптимизировано для длинных k-меров (обычно для k больше 10); 2: Оптимизировано для коротких k-меров (обычно для k меньше или равно 10); -1: Автоматический выбор. [По умолчанию : -1]

--k Длина K-мера. [По умолчанию : 3]

--maxnosignal Максимальная область отсутствия сигналов между сигналами для заполнения . [По умолчанию : 10000000]

16.3 Параметры обнаружения дублированных молекул

--dipout Файлы, содержащие дублированные молекулы [Обязательно]

--minseg Минимальное количество сегментов, которые следует считывать дублированными [По умолчанию : 15]

16.4 Параметры считывателя данных

--optmapin Вх одной файл оптической карты [обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Стандартный формат (REF) (Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию : -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию : 3.0]

17 Дублированных Молекул Удаление

Удаляется обнаруженные дублированные молекулы из файла данных

17.1 Параметры удаления дублированных молекул

--dupin Файлы, с одержащие дублированные молекулы [Обязательно]

17.2 Параметры считывателя данных

--optmapin Вх одной файл оптической карты [обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3:
Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

17.3 Параметры записи данных --optmapout

Вых одной файл оптической карты [обязательно]

--optmapoutformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: Формат файла Spots (SPOTS); 3:
Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

18 FrequentKmerВыделение

Возвращает идентификаторы сегментов в километрах с числом сдвигов, превышающим минимальное значение, по одному и тому же запросу.

18.1 Параметры считывателя данных

--optmapin Вх однофайл оптической карты [обязательно]
--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF);
9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]
--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

18.2 Параметры записи идентификатора сегмента

--sfout Вых одинфайл идентификатора сегмента

18.3 Параметры заполнения

--seedingmode Режим заполнения : 1: Оптимизировано для длинных k-меров (обычно для k больше 10); 2: Оптимизировано для коротких k-меров (обычно для k меньше или равно 10); -1: Автоматический выбор. [По умолчанию: -1]
--k Длина K-мера. [По умолчанию: 3]
--maxnosignal Максимальная область отсечения сигнала между сигналами для заполнения. [По умолчанию: 10000000]

Часть 6

Инструменты для выравнивания

19 Результаты инструменты

Представляет базовые функции для фильтрации и обработки результатов выравнивания

19.1 Параметры считывателя результатов

--optresin Вх одной файл результатов выравнивания [Обязательно]

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

19.2 Параметры записи результатов

--optresout Вывести файл результатов выравнивания

--optresoutformat Формат файла результата-1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

--writeunmap Записать отброшенные или неотображенные молекулы. [По умолчанию: true]

--multiple Записать несколько карт для молекулы. [По умолчанию: true]

--writeinfo Записать информацию о молекуле. [По умолчанию: true]

19.3 Параметры чтения с правочной литературы

--refmapin Вх одной файл с правочной карты

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

19.4 Параметры считывателя данных

--optmapin Вх одной файл оптической карты

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

--bnxsng Значение фильтра BNX SNR [по умолчанию: 3.0]

19.5 Параметры вывода данных

--mapout Вывод картированных молекул
--unmapout Вывод неотображенных молекул
--optmapoutformat -1: Автоматически определяется по расширению файла; 0: Стандартный с тандартный формат (REF)
 (Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: Формат файла Spots (SPOTS); 3:
 Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
 Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
 Формат OpGen XML; 10: Формат данных Валуева;
 11: Формат карт Maligner; [По умолчанию: 1]

19.6 Параметры инструментов результатов

--qprefix Добавить префикс к имени запроса.
--rprefix Добавить префикс к имени ссылки.
--disinvalid Отменить недействительные результаты. [По умолчанию: true]
--conf Пере считать доказательности результата [По умолчанию: false]
--dataid Список идентификаторов данных, которые необходимо извлечь
--region Регион в формате chrN:start-end или chrN:start
--refnamemodify Изменить имя ссылки в соответствии с целевым файлом в формате: src\!Target
--dataremoval Удалить результат с именами запросов в файле
--joinresult Представить частичные выравнивания в одном выравнивании (пробел заполняется дополнительными и отсутствующими символами). Работает только с частичными выравниваниями с отношением indel [По умолчанию: false]

19.6.1 Опции прерывания результатов

--breakermode Режим 0: Отключить функцию прерывания; 1: Разбить выравнивание на segmente запроса/ссылки, размер которых слишком сильно отличается, нанес несколько частичных выравниваний [По умолчанию: 0]
--meas Ошибка измерения [По умолчанию: 500]
--ear Допустимый диапазон ошибок (допустимая погрешность масштабирования) [По умолчанию: 0,1]
--match Оценка за соответствие с идентификатором [По умолчанию: 5]
--fpr Штраф за дополнительный с идентификатором [По умолчанию: 2]
--fnr Штраф за отсутствие с идентификатором [По умолчанию: 2]

19.6.2 Параметры фильтра результатов

--filtermode Режим фильтра: 0: Без фильтра; 1: Фильтровать по всем следующим параметрам; 2: Фильтровать по минимуму только оценка [По умолчанию: 0]
--minmatch Минимальное количество совпадений частичного выравнивания [По умолчанию: 3]
--maxfp Максимальное количество дополнительных с идентификатором частичного выравнивания [По умолчанию: 10000]
--maxfn Максимальное количество пропущенных с идентификатором частичного выравнивания [По умолчанию: 10000]

--maxfpr Максимальная скорость дополнительных сдвигов частичного выравнивания [По умолчанию: 1.0E-4] --maxfnr
Максимальная скорость пропущенных сдвигов частичного выравнивания [По умолчанию: 0.5]
--minscore Минимальная оценка частичного выравнивания [По умолчанию: 0,0] --minsubfragratio
Минимальное отношение субфрагментов частичного выравнивания [По умолчанию: 0,0]
--minsigratio Минимальное отношение выровненной длины к длине запроса [По умолчанию: 0,0] --trimmode Режим подстройки: 0: Режим подстроки отключен; 1: Режим подстроки включен [По умолчанию: 0] --maxtrim Максимальное количество шагов подстроки частичного выравнивания [По умолчанию: 5]
--match Оценка сопоставления сдвигов [По умолчанию: 5] --fpp Штраф за дополнительный сдвиг [По умолчанию: 2]
--fnp Штраф за отсутствие сдвигов [По умолчанию: 2]

19.6.3 Параметры соединения выравнивания

--alignmentjoinmode Режим: 0: Без объединения . 1: Стандартное соединение indel. 2: Стандартное соединение indel-inv.
3. Стандартное транслокационное соединение [По умолчанию: 0]
--closeoref Максимальное расстояние (ссылка) между двумя частичными выравниваниями, которые необходимо объединить [По умолчанию: 250000]
--closefrag Максимальное расстояние (запрос) между двумя частичными выравниваниями, которые необходимо объединить [По умолчанию: 250000]
--minmatch Минимальное количество сдвигов с соответствием, которые следует считать допустимым сопоставлением. [По умолчанию: 3]
--maxtrim Максимальные шаги обрезки для частичного выравнивания [По умолчанию: 5] --trimear Допустимая погрешность масштабирования во время обрезки [По умолчанию: 0,1] --match Оценка сопадающей сдвигов [По умолчанию: 5]
--fpp Штраф залишний сдвиг [По умолчанию: 2] --fnp Штраф за отсутствие сдвигов [По умолчанию: 2]
--indelp Штраф за объединение частичных выравниваний с отношением indel [По умолчанию: 10] --invp Штраф за объединение частичных выравниваний с отношением инверсией [По умолчанию: 30] --transp Штраф за объединение частичных выравниваний с отношением транслокации [По умолчанию: 50]
--localpenalty Включить штраф за локальное выравнивание для окончательного выравнивания (расматривается как глобальное выравнивание) [По умолчанию: ложно]
--minjoinscore Минимальная оценка объединенного окончательного выравнивания [По умолчанию: 0]
--minconf Минимальная достоверность (уникальность) окончательного выравнивания [По умолчанию: 0.0]
--minjoinedfragratio Минимальное отношение выровненной длины к длине запроса [По умолчанию: -1,0]
--minjoinedsigratio Минимальное отношение количества выровненных сдвигов к общему количеству сдвигов запроса [По умолчанию: -1,0]
--overlapalign Разрешить вывод перекрывающихся конечных выравниваний [По умолчанию: true] --maxalignitem
Максимальное количество выводимых конечных выравниваний. -1: Нет ограничений на количество конечных выравниваний [По умолчанию: 1]

19.6.4 Варианты подъема

--liftoverin Вх одной файл liftOver, Формат: хромосомы координат размер\n

19.7 Варианты анализа результатов моделирования

--rocout Вывести таблицу для построения ROC-кривой

20 Результат Объединение

Объединя ет результаты выравнивания , полученные разными методами

20.1 Параметры считывателя результатов --optresin

Вх одног файла результатов выравнивания [Обязательно]

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного о выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию : -1]

20.2 Параметры слияния результатов --resultkey

Ключи (имена) для представления файлов результатов [Обязательно]

--gapallowed Разрешены пробелы между результатами [По умолчанию : 0]

--analyzeall Анализировать только если запрос присутствует во всех результатах [По умолчанию : false]

--prefix Префикс выходного файла [Обязательно]

--outtype Тип выходного файла [По умолчанию : .omd]

21 Результат Статистика

Формирует статистику результатов выравнивания

21.1 Параметры чтения с сылок --refmapin Вх одной

файл карты с сылок [Обязательно]

--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент
формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

21.2 Параметры считывателя данных

--optmapin Вх одной файл оптической карты [обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент
формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

21.3 Параметры считывателя результатов

--optresin Вх одной файл результатов выравнивания [Обязательно]

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат
подробного выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального
сответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

21.4 Параметры статистики результата --

--checkstrand Проверка цепочки на корректность [По умолчанию: true]

--covout Вывод покрытия (в разработке)

--statout Вывод статистики

22 PrecisionRecallGraphDataGenerator

Генерирует таблицу данных для графиков точного отзыва. Этот модуль предполагает одно выравнивание (может содержать не сколько частичных выравниваний) на один запрос. Вам необходимо использовать те же параметры модуля присоединения выравнивания, если файл выравнивания генерируется с помощью карты OMTools. Если вы используете другие инструменты выравнивания, установите alignmentjoinmode на 0.

22.1 Параметры чтения с правочной литературы

```
--refmapin Вх одной файл с правочной карты [обязательно]
--refmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный с стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева
11: Формат карт Maligner; [По умолчанию: -1]
```

22.2 Параметры считывающего устройства данных

```
--optmapin Вх одной файл оптической карты [обязательно]
--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный с стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева
11: Формат карт Maligner; [По умолчанию: -1]
```

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

22.3 Параметры считывающего устройства результатов

```
--optresin Вх одного файла результатов выравнивания [Обязательно]
--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного о
выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат
Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]
```

22.4 Варианты соединения выравнивания

```
--alignmentjoinmode Режим: 0: Без объединения . 1: Стандартное соединение indel. 2: Стандартное соединение indel-inv.
3. Стандартное транслокационное соединение [По умолчанию: 0]

--closeref Максимальное расстояние (ссылька) между двумя частичными выравниваниями, которые необходимо объединить [По умолчанию:
250000]

--closefrag Максимальное расстояние (запрос) между двумя частичными выравниваниями, которые необходимо объединить [По умолчанию:
250000]

--minmatch Минимальное количество соглашений, которые следуют с читать допустимым частичным спосаблением. [По умолчанию: 3]

--maxtrim Максимальные шаги обрезки для частичного выравнивания [По умолчанию: 5]

--trimear Допустимая погрешность масштабирования во время обрезки [По умолчанию: 0,1]
```

```
--match Оценка за сопоставление с именем [По умолчанию: 5]
--fpp Штраф за дополнительный с именем [По умолчанию: 2]
--fnp Штраф за отсутствие с именем [По умолчанию: 2]
--indelp Штраф за присоединение частичных выравниваний с отношением indel [По умолчанию: 10]
--inpr Штраф за объединение частичных выравниваний с инверсационным отношением [По умолчанию: 30]
--transp Штраф за присоединение частичных выравниваний с транслокационной связью [По умолчанию: 50]
--localpenalty Включить штраф за локальное выравнивание для окончательного выравнивания (расматривается как глобальное выравнивание)
[По умолчанию: ложно]
--minjoinscore Минимальная оценка объединенного окончательного выравнивания [По умолчанию: 30]
--minconf Минимальная достоверность (уникальность) окончательного выравнивания [По умолчанию: 0,4]
--minjoinedfragratio Минимальное отношение выровненной длины к длине запроса [По умолчанию: -1,0]
--minjoinedsigratio Минимальное отношение количества выровненных с именем к общему количеству с именем запроса [По умолчанию: -1,0]

--overlapalign Разрешить вывод перекрывающихся ся конечных выравниваний [По умолчанию: true]
--maxalignitem Максимальное количество выводимых окончательных выравниваний. -1: Нет ограничений на количество окончательных выравниваний.
выравнивания [По умолчанию: 1]

22.5 Параметры графика точности отзыва --prgout Вывод таблицы графика
точности отзыва [Обязательно]
--checkstrand Проверка цепочки на корректность [По умолчанию: true]
--sortstrat Сортировать по «оценке» или «уверенности» [По умолчанию: оценка]
```

23 Запрос Обратный

Изменяется запрос в соответствии с результатами выравнивания так, чтобы они имели ту же цепочку, что и ссылка. В результатах выравнивания для каждого запроса допускается только одна ссылка и одна цепочка (иначе будет выдано исключение). Инверсионные не поддерживаются.

23.1 Параметры считывателя результатов

```
--optresin Вх одной файл результатов выравнивания [Обязательно]
--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]
```

23.2 Параметры считывателя данных --

```
optmapin Вх одной файл оптической карты [Обязательно]
--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]
```

23.3 Параметры записи данных

```
--optmapout Вых одной файл оптической карты [обязательно]
--optmapoutformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: Формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]
```

24 Выравнивание Выделение

Выделяет сегменты, используемые в выравнивании х в качестве идентификаторов сегментов.

24.1 Параметры считывателя данных --optmapin

Вх однай файл оптической карты [Обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF);
9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

24.2 Параметры считывателя результатов

--optresin Вх однай файл результатов выравнивания [Обязательно]

--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания OM (OMA); 1: Формат подробного о выравнивания OM (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]

24.3 Параметры записи идентификатора сегмента

--siout Вых однай файл идентификатора сегмента

24.4 Параметры BED-писатель

--bedout Вых однай BED-файл

24.5 Параметры подсчетки выравнивания

--mode Выделить сегменты в -1: Только ссылка; 2: Только запрос; 3: И запрос, и ссылка [По умолчанию: 3]

--overlap Требовать перекрывающийся запрос и ссылку (по крайней мере с одним сигналом) [По умолчанию: false]

--strict Выделять сегменты только в последовательном сопадении сигналов (без дополнительных или отсутствующих сигналов между ними) [По умолчанию: false]

Часть VII

Множественное выравнивание

25 Множественное выравнивание

Выполняет множественное выравнивание, используя несколько оптических карт в качестве запросов

25.1 Параметры считывателя данных

```
--optmapin Вх одной файл оптической карты [обязательно]
--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(Эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла СMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]
```

25.2 Несколько вариантов выравнивания

```
--maconfig Файл конфигурации множественного выравнивания
--minlinksize Минимальный размер сегмента в процессе вызова [По умолчанию: 1000]
--maref Ссылки для множественного выравнивания
--maorderby 1: Первое вложение блока; 2: кластеризация по ходу [По умолчанию: 1]
```

25.3 Параметры коллинеарного блок-писца

```
--cblout Вывод нескольких коллинеарных блоков выравнивания.
```

25.4 Коллинеарный блок-писатель портатива

```
--cboout Несколько коллинеарных блоков выравнивания упорядочиваются вывод
```

25.5 Конфигурация множественного выравнивания

Весь процесс множественного выравнивания можно разбить на три основных подмодуля — вызов, построение и слияние/ближение слияния.

Первый подмодуль «link» обрабатывает вх одные файлы выравнивания , класс теризаци ии и множественное выравнивания для создания потенциальных «сегментных связей» между сегментами из разных запросов. Эти связи необязательны для стратегий группировки сегментов в коллинеарные блоки. Связи, созданные для этого подмодуля , должны быть высоконадежными, поскольку они определяют глобальную структуру множественного выравнивания .

Второй подмодуль «build» компилирует сегментные связи, созданные из предыдущего подмодуля , в коллинеарные блоки.

Третий подмодуль «merge» и «mergeproximity» объединяет похожие коллинеарные блоки в один коллинеарный блок, не нарушая глобальной структуры множественного выравнивания . Поскольку консервативная стратегия обычно используется для обмена чувствительности на более высокую с практической точки зрения этапе «link», некоторые сегменты, принадлежащие одному и тому же коллинеарному блоку, могут быть помещены в несколько коллинеарных блоков. Объединение основано на(1) менее надежных связях («merge») или (2) длине проксиимальных блоков («mergeproximity»).

Альтернативный подмодуль «mask» маркирует и предотвращает вызов и слияние сегментов, которые приводят к несекционному множественному выравниванию , в указанных выше подмодулях . Этот модуль используется только в том случае, если запросы состоят из множества повторяющихся сигнальных шаблонов или срезовых сегментных дубликатов, которые нарушают процесс множественного выравнивания .

25.5.1 Файл конфигурации множественного выравнивания

В файле конфигурации каждого строка начинается с команды, соответствующей модулю в нескольких процессах выравнивания , за которой следуют параметры и вх одные файлы.

`mask(Необязательно) mask[Вх одной файл]`

Запретить обработку сегментов, используемых в файле выравнивания , на этапах «вызова» и «объединения».

`сылка ссылка[Вх одной файл] [k=? (класс теризаци и); maxnosignal=? (класс теризаци и)]`

Создать сегментные ссылки из файлов выравнивания , класс теризаци ии и множественное выравнивания . Для ввода файла класс теризаци и требуется два параметра k и maxnosignal.

`построить построить [allowrearrangement=true]`

Объедините эти сегментные связи в коллинеарные блоки.

`mergeproximity (Необязательно) mergeproximity [meas=? (ошибка измерения); ear=? (ошибка масштабирования)]`

Объединить соседние блоки, если их длины достаточно близки с учетом допустимого уровня ошибок.

`mergeproximitysimple (Необязательно) mergeproximitysimple [meas=? (погрешность измерения); ear=? (масштабирование ошибки)]`

Объединить блоки в близости, если их длины блоков достаточно близки с учетом допустимого уровня ошибок. Эта команда использует упрощенную версию mergeproximity для сокращения времени обработки памяти, но объединяется только близкие блоки с одинаковым уникальным родительским блоком.

`merge (Необязательно) merge [Вх одной файл] [mergechain=true; k=? (класс теризаци и); maxnosignal=? (класс теризаци и)]`

Объединить коллинеарные блоки с помощью сегментных связей. Вх одные файлы и параметры аналогичны тем, что в команде link.

Файл конфигурации должен содержать как минимум одну команду «link», за которой следует одна команда «build».

Команды «mask», «mergeproximity», «mergeproximitysimple» и «merge» являются необязательными.

Пример файла конфигурации

```
маска SegmentMaskFile.si

ссылка AlignmentFile1.oma AlignmentFile2.oma

ссылка ClusteringFile1.kc k=7;maxnosignal=1000000

ссылка MultipleAlignmentFile1.cbl

сборка allowrearrangement=true

mergeproximitysimple изм=500; ух о=0.1

мера близости слияния = 500; ух о = 0,1

объединить AlignmentFile3.oma mergechain=true
```

25.5.2 Страгегия оптимизации конфигурации

Поскольку оптимизация параметров во многом зависит от характера запросов, с должно иметь единственную конфигурацию, которая подходит для всех случаев. В этом разделе обобщена и представлена стратегия оптимизации конфигурации. Нетерпеливые пользователи могут пропустить подробности здесь и прочитать разделы ниже для шаблонов конфигурации для конкретных случаев.

Основные усилия по оптимизации направлены на то, какие файлы должны быть предоставлены на этапах «связать» и «слиять», где пользователи должны решить, что предоставлять в качестве подсказок для ссылок сегментов. Другие этапы менее сложны и не требуют общей оптимизации.

маска Сегменты, которые приводят к неспецифическому выравниванию, но с высоким баллом, должны быть замаскированы, поскольку они приводят к неправильным связям сегментов для первоначального построения коллинеарного блока и нарушают структуру множественного выравнивания. Обычно эти сегменты происходят из повторяющихся сигнальных схем или сегментарных дубликатов. Пользователи могут найти эти сегменты путем самоуравнивания каждого запроса:

Маска

```
OMBlastMapper -refmapin query.data -optmapin query.data -alignmentjoinmode 0 -filtermode 1 -minconf 0 -allowdiffrefquery
false -exactmatch false -maxalignitem -1 -minjoinscore 50 -fpp 10 -fnr 10 AlignmentHighlight -mode 3 -optmapin query.data
-optresin
queryselfalign.oma -siout query-mask.si
```

Пользователи могут уменьшить параметры minjoinscore, fpp и fnr для менее строгого выравнивания, чтобы исключить больше потенциальных сегментов, вовлеченные в сегментарные дубликаты. Этот шаг не является необходимым, если сегментарные дубликаты отсутствуют в запросах.

link [«link»](#) является наиболее важной частью оптимизации. Ссылки сегментов должны быть построены с использованием высоконадежных доказательств, поскольку каждая ссылка сегмента определяет окончательную глобальную структуру множественного выравнивания. Пользователи должны быть сконсервативны и использовать строгие параметры для выравнивания. Основная цель оптимизации — связать как можно больше запросов, сократив при этом корректную структуру.

Оптимизация может быть упрощена как процесс выбора правильных файлов. Во-первых, пользователи могут импортировать файлы выравнивания от высокой до низкой достоверности. Для неоднозначной общей структуры множественного выравнивания пользователи должны импортировать меньше файлов выравнивания с низкой достоверностью. Для множественного выравнивания с большим количеством похожих шаблонов, не объединенных в одинаковые коллинеарные блоки, вместе с этим пользователи должны импортировать больше файлов выравнивания с низкой достоверностью.

Файл множественного выравнивания также может быть исключен для ссыльного вызова. Обычно это происходит при множественном выравнивании крупные эукариотические гены, где запросы были частично выровнены в скользящих окнах.

Обратите внимание, что во время оптимизационного шага «link» пользователи могут отключить команды «merge» и «mergeproximity».

Пользователям не нужно оптимизировать этап «сборки».

merge На основе сегментных связей шаг «merge» объединяет два коллинеарных блока только в том случае, если слияние не нарушает глобальную структуру множественного выравнивания (т.е. не создает перегруппировки). Процесс оптимизации «merge» по окончании с «link», за исключением того, что файлы для сегментных связей могут иметь более низкую достоверность в целом. Этот шаг настоятельно не рекомендуется использовать в регионе с вариацией числа копий. Обратите внимание, что этот шаг требует очень много времени и памяти и не применим к эукариотическим геномам.

mergeproximity Пользователям необходимо учитывать только один фактор при оптимизационном шаге «mergeproximity» — толерантность к ошибкам (ошибка измерения и ошибка масштабирования). Поэтому, более высокая толерантность к ошибкам должна применяться при множественном выравнивании необработанных молекул оптического картирования. Напротив, более низкая толерантность к ошибкам может использоваться при множественном выравнивании с обрывами контигов оптического картирования или последовательностей, переваренных *in-silico*. Пользователи найдут шаг «mergeproximity» полезным практически во всех областях применения. Шаг требует много времени и памяти, а упрощенная, но менее мощная версия «mergeproximitysimple» может использоваться в более крупных эукариотических генах.

«mergeproximitysimple» — это упрощенная версия «mergeproximity».

Он ограничивает объединение с единими блоками только в том случае, если у них есть один родительский блок.

25.5.3 Шаблон конфигурации

Геномы прокариот

Выравнивание

```
OMBlastMapper -refmapin query.data -optmapin query.data -optresout query1.oma -filtermode 1 -alignmentjoinmode
0 -maxalignitem -1 -minconf 0 -thread 2 - allowequalrefquery false -exactmatch false -writeinfo false -writeunmap false
-fpp 15 -fnp 15 -ear 0.05 -meas 500 -minjoinscore 50 OMBlastMapper -refmapin query.data -optmapin query.data -
optresout query2.oma -filtermode 1 -alignmentjoinmode 0 -
maxalignitem -1 -minconf 0 -thread 2 - allowequalrefquery false -exactmatch false -writeinfo false -writeunmap false -
fpp 10 -fnp 10 -ear 0.05 -meas 500 -minjoinscore 50 OMBlastMapper -refmapin query.data -optmapin query.data -
optresout query3.oma -filtermode 1 -alignmentjoinmode 0 -maxalignitem -1 -minconf 0 -thread 2 - allowequalrefquery
false -exactmatch false -writeinfo false -writeunmap false -fpp
5 -fnp 5 -ear 0.05 -meas 500 -minjoinscore 50 OMBlastMapper -refmapin query.data -optmapin query.data -optresout
query4.oma -filtermode 1 -alignmentjoinmode 0 -maxalignitem -1 -minconf 0 -thread 2 - allowequalrefquery false -
exactmatch false -writeinfo false -writeunmap false -fpp 50 -fnp 50 -ear 0.05 -meas 500 -minjoinscore 30 OMBlastMapper
-refmapin query.data -optmapin query.data -optresout
query5.oma -filtermode 1 -alignmentjoinmode 0 -maxalignitem -1 -minconf 0 -thread 2 - allowequalrefquery false -
exactmatch false -writeinfo false -writeunmap false -fpp 50 -fnp 50 -ear 0.05 -meas 500 -minjoinscore 20 OMBlastMapper
-refmapin query.data -optmapin query.data -optresout query6.oma -filtermode 1 -alignmentjoinmode 0 -maxalignitem
-1 -minconf 0 -thread 2 -allowequalrefquery false -exactmatch
false -writeinfo false -writeunmap false -fpp 2 -fnp 2 -ear 0.05 -meas 500 -minjoinscore 50
```

Файл конфигурации множественного выравнивания

```
ссылка запрос 1.ома
ссылка запрос 2.ома
ссылка запрос 3.ома
сборка allowrearrangement=true
mergeproximity изм.=500;ух о=0,05 объединение
запрос 1.ома объединение
запрос 2.ома объединение
запрос 3.ома объединение
запрос 4.ома объединение
запрос 5.ома объединение
запрос 6.ома
mergeproximity изм.=500;ух о=0,05
```

Варианты номеров копий

Выравнивание

```
OMBlastMapper -refmapin query.data -optmapin query.data -optresout query1.oma -filtermode 1 -alignmentjoinmode  
0 -maxalignitem -1 -minconf 0 -thread 2 - allowequalrefquery false -exactmatch false -writeinfo false -writeunmap false  
-fpp 10 -fnp 10 -ear 0.05 -meas 500 -minjoinscore 100 OMBlastMapper -refmapin query.data -optmapin query.data -  
optresout query1.oma -filtermode 1 -alignmentjoinmode 0 -  
maxalignitem -1 -minconf 0 -thread 2 - allowequalrefquery false -exactmatch false -writeinfo false -writeunmap false -  
fpp 5 -fnp 5 -ear 0.05 -meas 500 -minjoinscore 50 OMBlastMapper -refmapin query.data -optmapin query.data -optresout  
query1.oma -filtermode 1 -alignmentjoinmode 0 -maxalignitem -1 -minconf 0 -thread 2 -allowequalrefquery false -  
exactmatch false -writeinfo false -writeunmap false -fpp 50 -  
fnp 50 -ear 0.05 -meas 500 -minjoinscore 30
```

Конфигурация множественного выравнивания

```
ссылка query1.oma  
ссылка query2.oma  
ссылка query3.oma  
сборка allowrearrangement=true  
mergeproximity meas=500;ear=0.05
```

Часть VIII

Несколько инструментов выравнивания

26 CBLTools

Предоставляет базовые функции для фильтрации и обработки результатов множественного выравнивания.

26.1 Параметры считывателя данных

--optmapin Вх одн файл оптической карты

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS); 3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9: Формат OpGen XML; 10: Формат данных Валуева; 11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

26.2 Параметры коллинеарного считывателя блоков

--cblin Ввод нескольких коллинеарных блоков выравнивания.

26.3 Параметры коллинеарного блок-писца

--cblout Вывод нескольких коллинеарных блоков выравнивания.

26.4 Параметры инструментов CBL

--cboout Вывод коллинеарного порядка блоков

26.4.1 Параметры фильтрации

--allblock Извлечь запросы, содержащие все блоки в этом списке

--reservequery Резервировать запросы, даже если они не соответствуют другим критериям

26.4.2 Параметры сортировки

--sortblock Сортировать запросы по наличию блока

--sortdesflanklen Сортировать запросы по длине между двумя фланговыми блоками (FB1a FB1b) по убыванию. порядок

27 Анализ производительности множественного выравнивания

Анализирует эффективность множественного выравнивания ОМ на основе множественного выравнивания последовательностей.

27.1 Параметры считывателя данных --

optmapin Вх одной файл оптической карты [Обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент
формата REF); 9: Формат OptGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

27.2 Считыватель MAF

--mafif Ввод формата множественного выравнивания .

27.3 Параметры коллинеарного считывателя блоков

--cblin Ввод нескольких коллинеарных блоков выравнивания .

27.4 Множественные варианты выравнивания

--statout Вывод статистики

28 Блок С оединение ГрафГ енераци я

Создает файл точечной диаграммы, отображающей последовательность коллинеарных блоков.

28.1 Параметры коллинеарного считывателя блоков

--cblin Ввод нескольких коллинеарных блоков выравнивания .

28.2 Варианты генерации графа с оединений блоков

--maref Ссылки для множественного выравнивания

--reverse Изменить направление просчета коллинеарных блоков на противоположное [По умолчанию: false]

--minedgeweight Фильтр с минимальным весом ребра [По умолчанию: 0]

--displayall Вывести все блоки (Установите false, чтобы программа не выводила блоки без какой-либо связи с другими блоками) [По умолчанию: true]

--dotout Вывод файла с точкой [обязательно]

Часть 9

Филогенетика

29 UPGMA Tree Construction

Реконструирует филогенетическое дерево на основе результатов множественного выравнивания с использованием подхода UPGMA

29.1 Варианты построения UPGMA Tree

- matrixout Вывести матрицу расстояний [обязательно]
- treeout Выводить деревья в формате newick [Обязательно]
- startblock Стартовый блок (Экспериментальные параметры)
- stopblock Стоп-блок (Экспериментальные параметры)

29.2 Параметры считывателя данных

- optmapin Вх одн файл оптической карты [обязательно]
- optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF)
(эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла СМАР (СМАР); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата
REF); 9: Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]
- bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

29.3 Параметры коллинеарного считывателя блоков

- cblin Ввод нескольких коллинеарных блоков выравнивания .

Часть X

Визуализация

30 ОМВид

Визуализирует данные оптического картирования . OMView предоставляет графический интерфейс для визуализации данных оптического картирования для различных целей.

30.1 Загрузка данных

```
--viewrefin Загрузить ссылки
--viewmapin Загрузить молекулы
--viewresin Загрузить результаты выравнивания
--viewcblin Загрузить коллинеарные блоки
--viewcboin Загрузить коллинеарные блоки (порядок)
--viewcbcin Загрузить коллинеарные блоки (цвет)
--viewannoин Загрузить аннотации
--viewin Автоматический ввод файлов
```

30.2 Просмотр открытия

```
--viewregion Показать определенный регион в региональном представлении
--viewanchor Показать определенный якорь в представлении якорей
--viewanchorregion Укажите регион для якоря в представлении якоря .
--viewalignment Показать определенное выравнивание
--viewma Автоматически открывать представление с несколькими выравниваниями [По умолчанию: false]
--viewmb Автоматически открывать представление блока множественного выравнивания [По умолчанию: false]
--viewmolecule Автоматически открывать вид молекулы [По умолчанию: false]
--viewsave Сохранять представления в определенном месте вместе запуска OMView
--viewsaveformat Форматы сохранения изображения . [svg; png; jpg; ] [По умолчанию: png]
```

30.3 Настройки просмотра --dnaratio

```
Соотношение ДНК по умолчанию [По умолчанию: 400,0]
--zoom Уровень масштабирования по умолчанию [По умолчанию: 1,0]
--viewbreakresult Включить прерыватель результатов [по умолчанию: false]
--viewunmap Включить неотображенную часть [по умолчанию: false]
--viewsettingin Вх одноименное настройки OMView
```

30.4 Помощь

--help Отобразить меню с правки

30.5 Процедуры визуализации

30.5.1 Макет OMView

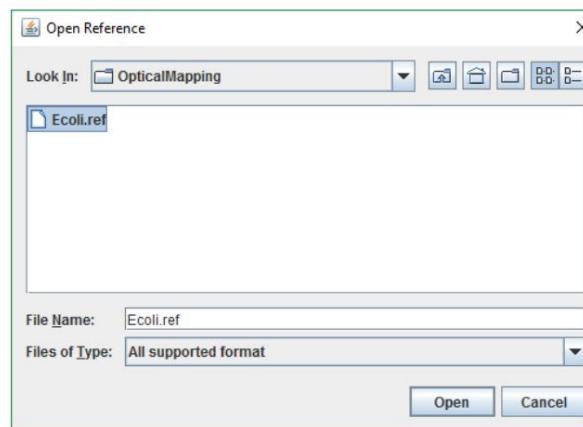


30.5.2 Загрузка необходимых данных

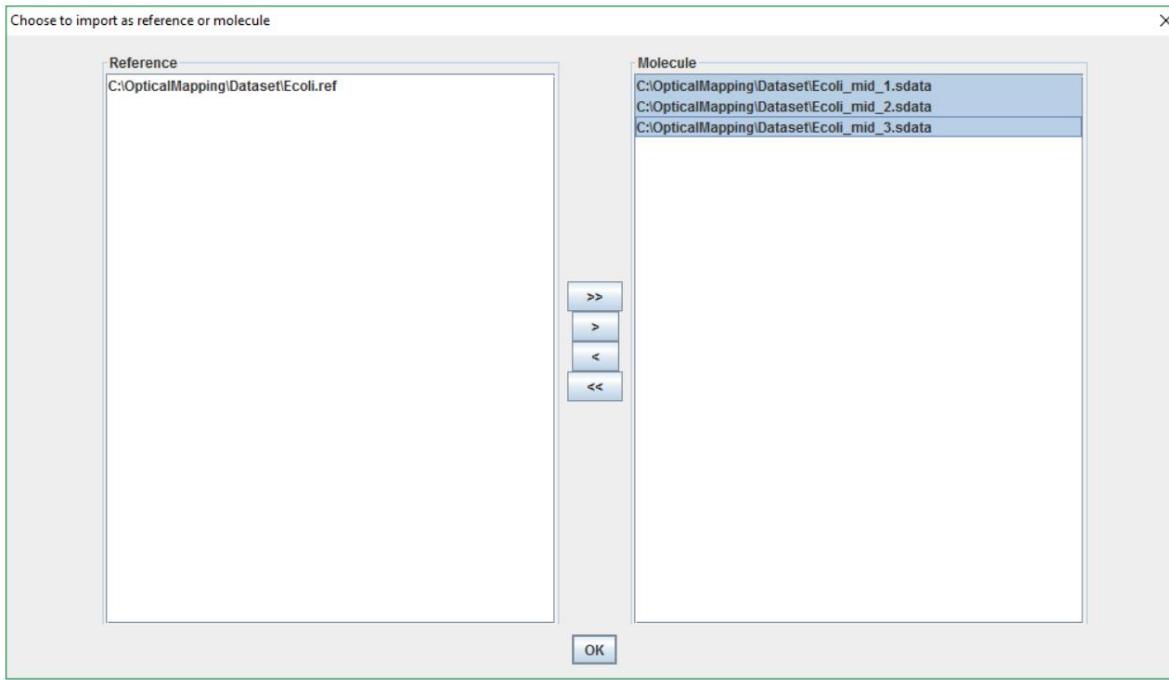
Существует два способа загрузки данных в OMView: (1) загрузить наборы данных из меню и (2) перетащить наборы данных в программу.

Примечание: пользователи должны сначала загрузить файлы с сылкой и молекул, прежде чем загружать другие файлы (выравнивание, файлы аннотаций и множественных выравниваний)

Выберите файлы данных из меню. После выбора файла данных для загрузки (FileLoad) выберите целиевой файл и нажмите «Открыть».



Перетаскивание наборов данных. В программу можно одновременно перетаскивать несколько файлов.



Поскольку форматы файлов ссылок и молекул одинаковы, пользователям необходимо указать, файлы загружаются как ссылки или молекулы.

Зависимость файла

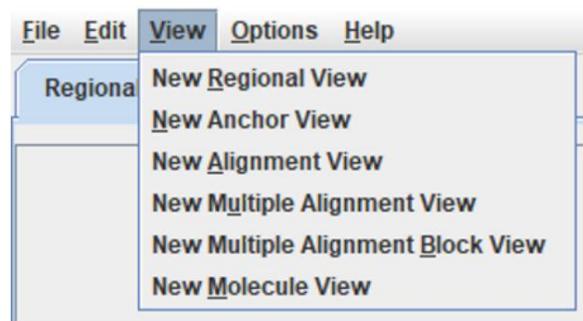
Ссылка и молекулы: Нет требований к другим файлам данных. Обратите внимание, что идентификатор ссылки не должен дублироваться среди всех файлов ссылок. Аналогично, идентификатор молекулы не должен дублироваться.

Выравнивания: требуется ссылки и молекулы (если файл выравнивания содержит информацию о молекулах, загрузка файла молекул не требуется)

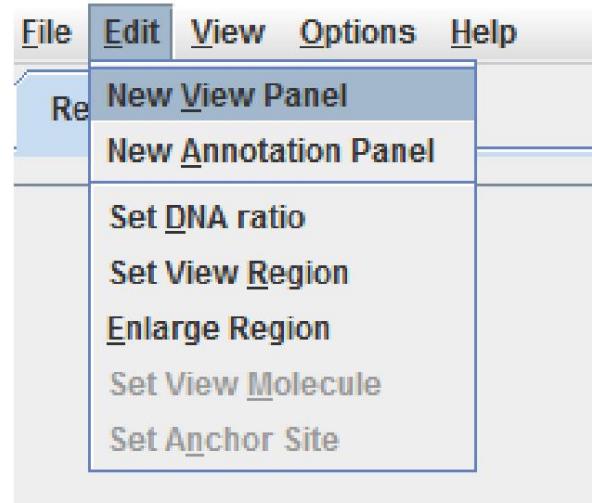
Аннотации: Требуется ссылка

Множественные выравнивания: требуется молекулы

30.5.3 Запуск представления



Пользователи могут открывать новые вкладки просмотра в меню Вид. По умолчанию OMView инициализируется с новым рабочим региональным видом.

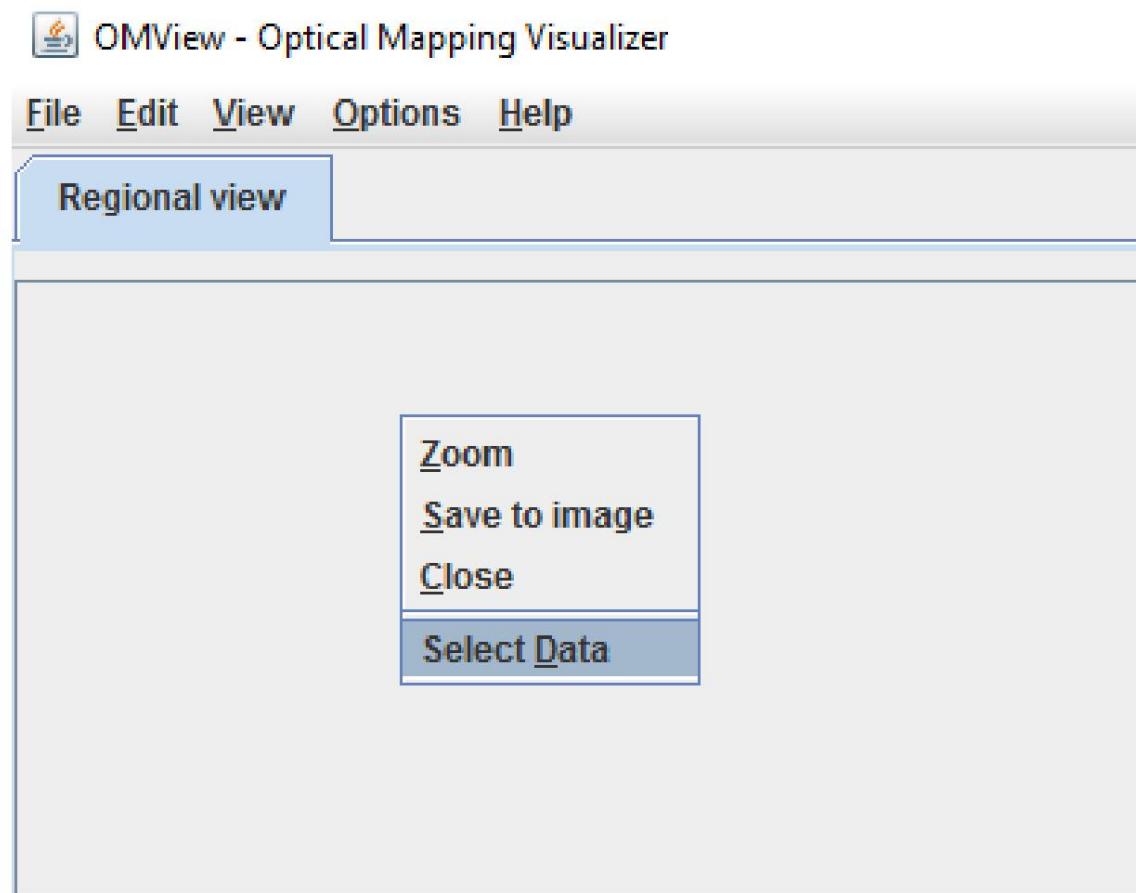


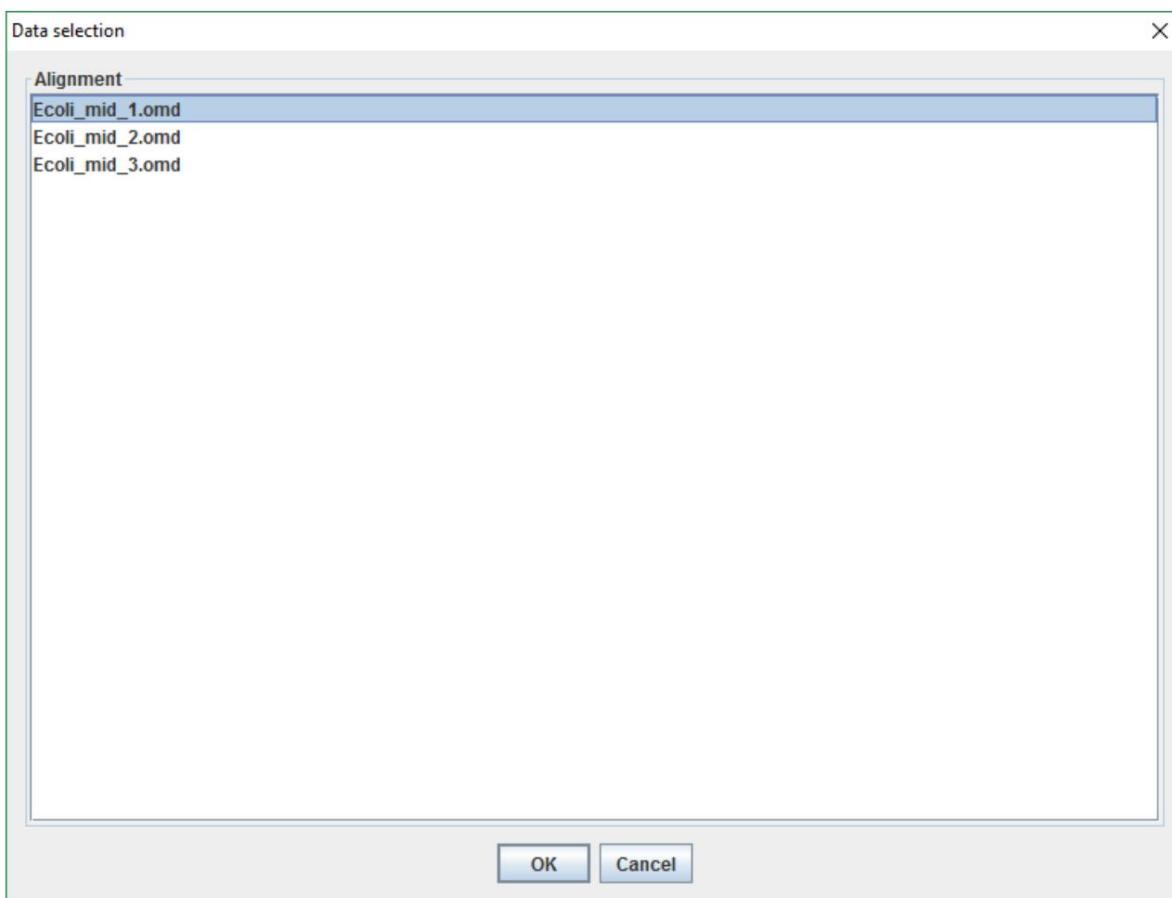
Каждая вкладка по умолчанию содержит панель просмотра. Некоторые представления поддерживают визуализацию более чем одной панели просмотра (регистриональный вид, вид коря и вид молекулы) и панелей аннотаций (регистриональный вид и вид коря) на одной вкладке. Пользователи могут добавлять новые панели в меню «Изменить».

30.5.4 Выбор данных на панели просмотра

В контекстном меню каждой панели просмотра выберите «Выбрать данные», чтобы выбрать данные для отображения.

Открывается диалоговое окно, в котором пользователи могут выбрать данные. Обратите внимание, что если результаты выравнивания более одного набора загружены на одной панели просмотра, убедитесь, что их соответствующие идентификаторы молекул различны.





30.5.5 Процедуры по конкретным видам

Региональный вид Региональный вид отображает трассы в виде обзора в выбранном регионе



После открытия вкладки регионального представления установите целевой регион как X:NNNNNN-NNNNNN (EditSetRegion). Должна появиться языковая панель (красный правомоугольник) с языками (черные вертикальные полосы).

Примечание: Имя ссылается на блок (крайний правый столбик) с именем файла вертикальной полосы.

Фильм снят на киностудии «СоюзМедиа» в городе Чебоксары. Премьера фильма состоится 11 декабря в кинотеатре «Колизей».

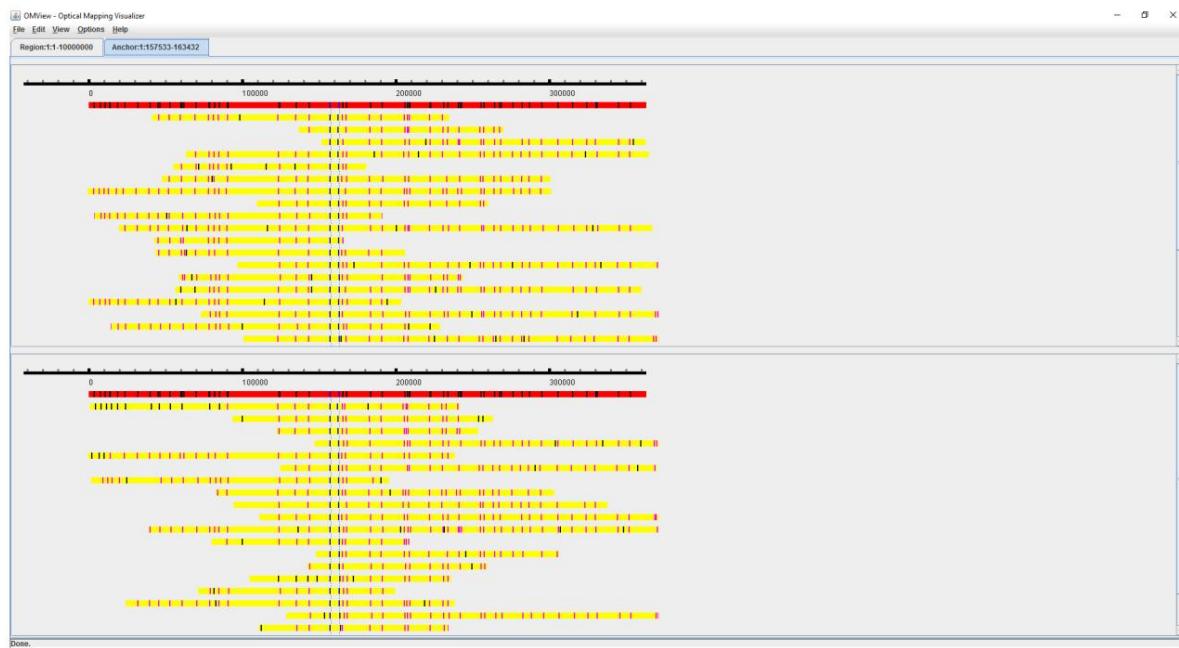
настрижка гармонично подчеркивает имя с ее звуками «спит» или «т».

После выбора выравниваний и аннотаций результаты будут отображены на панелях просмотра.
Выровненные части молекул показаны в цветовом спектре (от зеленого до красного) в зависимости от коэффициента масштабирования от 0,5 до 1,5, где желтый цвет означает коэффициент масштабирования 1) с розовыми и черными сигналами, указывающими на отображенные или неотображенные сигналы.

Наведите курсор на сдвиг, чтобы отобразить его информацию. Чтобы отобразить детали выравнивания определенной молекулы, просто щелкните на цепи молекулы, и будет создана новая вкладка представления выравнивания. Чтобы просмотреть все молекулы, выровненные с определенным опорным сдвигом, щелкните на этом сдвиге, чтобы создать новое представление привязки.

Пример: (Ссылка) Ecoli.ref, Ecoli mid 1.omd, Ecoli mid 2.omd, Ecoli mid 3.omd, Ecoli.gff

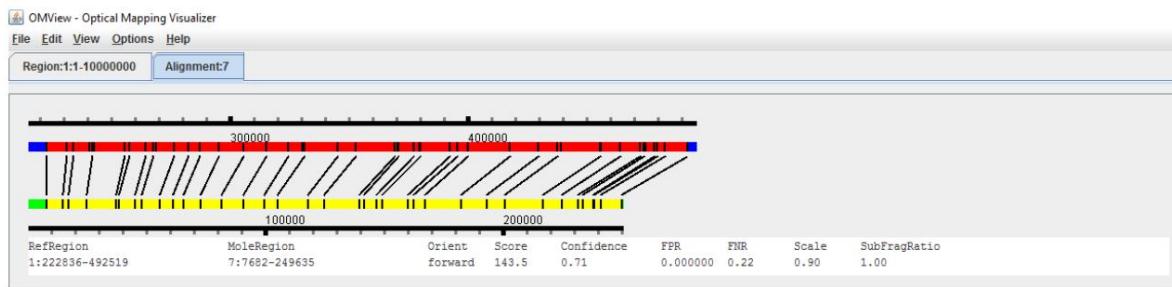
Представление «Якорь» Представление «Якорь» отображает выравнивания , с соответствующими выбранным с игналам, для проверки структурных изменений.



Процедура настройки аналогична процедуре в региональном представлении. После открытия вкладки представления якоря установите якорный сайт как X>NNNNNNN>NNNNNNN (EditSet Anchor Site). Обратите внимание, что якорный сайт должен представлять положение одного или двух сайтов. Пользователи могут установить регион как X>NNNNNNN>NNNNNNN (EditSet Region). По умолчанию регион установлен на 200 кбит/с от якорных сайтов.

Пример: (Ссылка) Ecoli.ref, Ecoli mid 1.omd, Ecoli mid 2.omd, Ecoli mid 3.omd, Ecoli.gff - - -

Вид выравнивания Вид выравнивания отображает детали выравнивания отдельной молекулы.

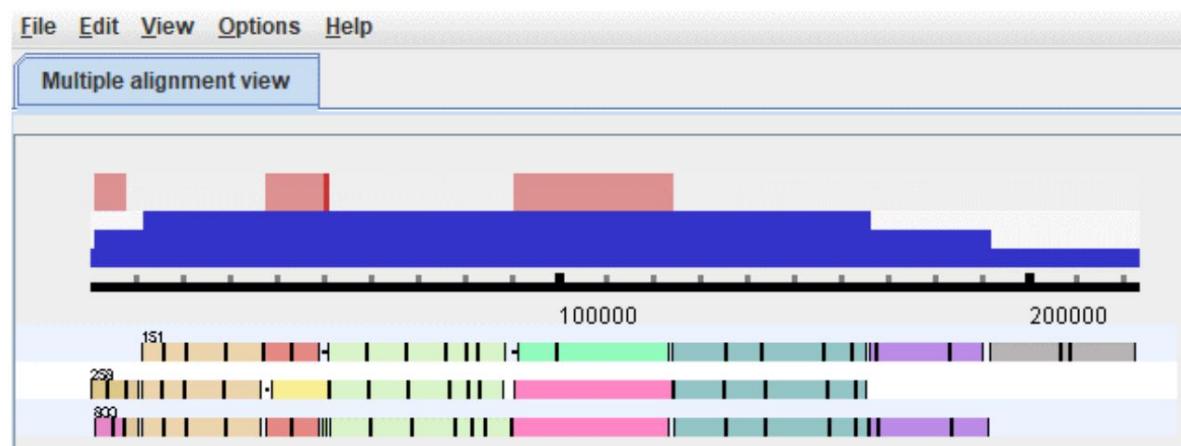


В представлении выравнивания установите идентификатор молекулы, чтобы просмотреть детали выравнивания (EditSet View Molecule). Пользователи получат предупреждение с сообщением, если для выбранной молекулы выравнивание не существует.

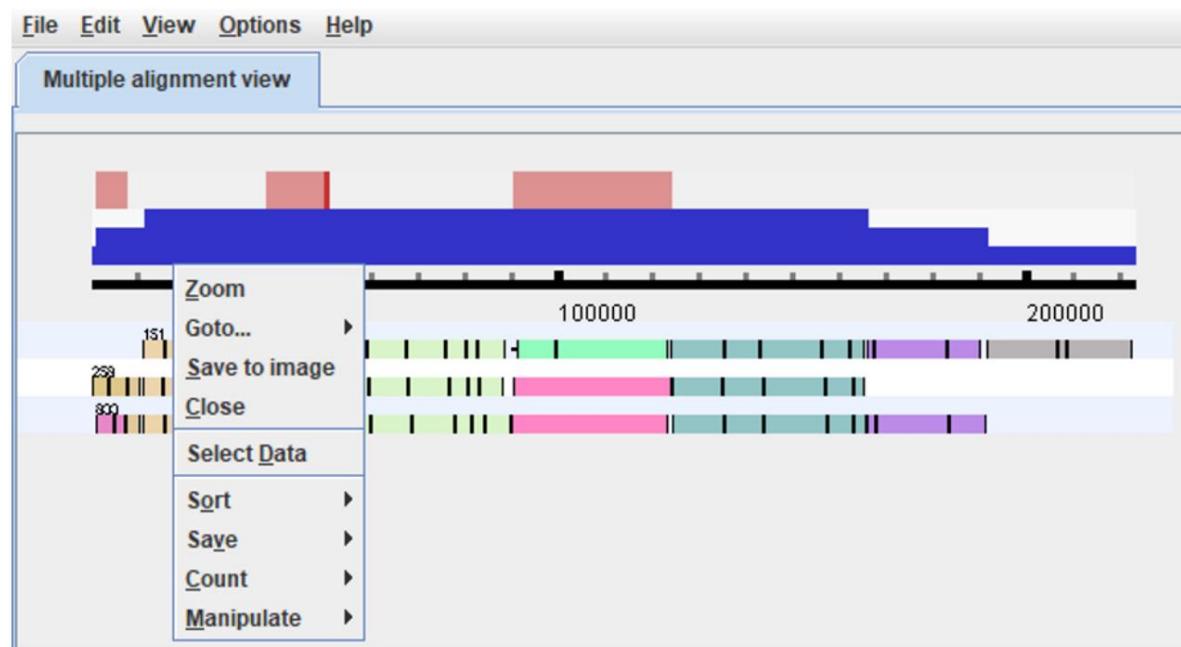
Верхний и нижний прямые линии представляют собой и молекулу соответственно. Два набора линеек указывают координаты относительно начала эталонной хромосомы и начала оптической молекулы. Обратите внимание, что молекула не масштабируется и остается в прямой ориентации, в то время как ссылька может быть масштабирована и перевернута для лучшей визуализации. Пользователи могут просмотреть текстовое поле ниже, чтобы получить дополнительные сведения о выравнивании.

Пример: (Ссылка) Ecoli.ref, Ecoli mid 1.omd

Представление множественного выравнивания Представление множественного выравнивания отображает множественные выравнивания всех запросов для геномного сравнения.



Вид множественного выравнивания отображается автоматически после установки данных. Если файл цвета (.cbc) не при этом коллинеарным блокам будет назначен случайный цвет.

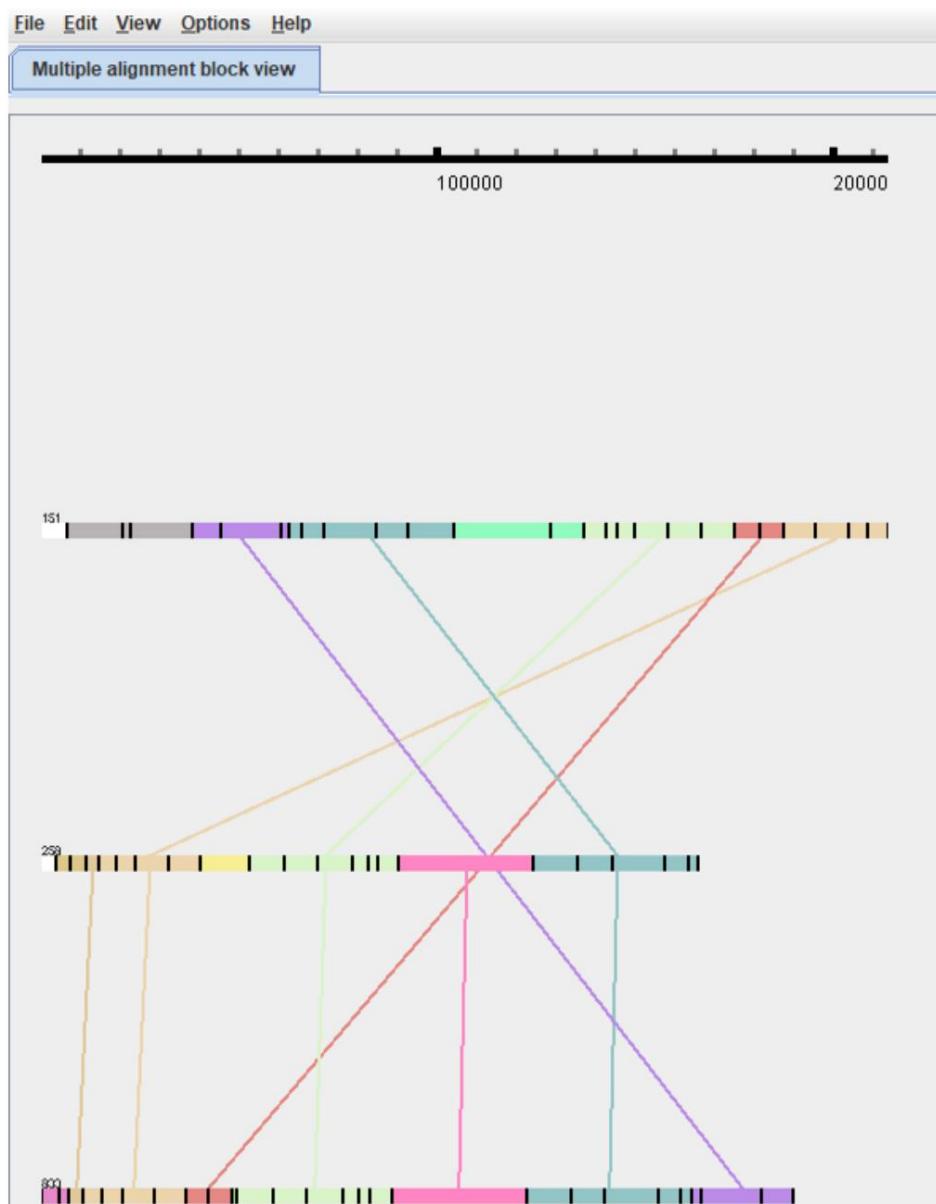


Щелкните правой кнопкой мыши, чтобы сортировать и управлять множественными выравниваниями. Пользователи могут сортировать текущее множественное выравнивание после манипуляции. Также доступна функция подсчета, а статистика выводится в консоли.

Пример: (Молекула) Ecoli MA.sdata, Ecoli MA.cbl, Ecoli MA.cbo, Ecoli MA.cbe

-

Вид блока множественного выравнивания отображает альтернативную визуализацию множественного выравнивания. Запросы остаются в исходном виде без изменений. Сегменты из одного коллинеарного блока соединены линиями одного цвета.



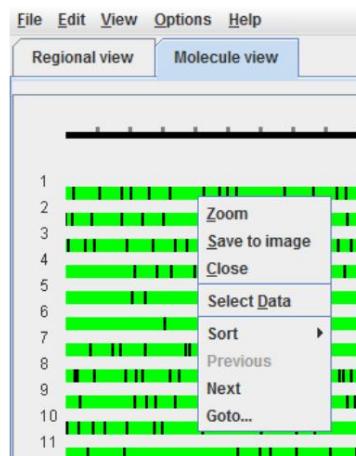
Вид блока множественного выравнивания отображается автоматически после установки данных. Если файл цвета (.cbc) не указан, коллинеарным блокам будет назначен случайный цвет.

Пример: (Молекула) Ecoli MA.sdata, Ecoli MA.cbl, Ecoli MA.cbo, Ecoli MA.cbc

-



Вид молекул В виде молекул панель отображает с транзиц у, с одержащую 100 молекул. Перейдите на другую транзицу с помощью пунктов меню правой кнопки мыши Предыдущий, Следующий и Перейти. Также доступна функция сортировки для сортировки молекул по размеру молекулы, количеству сигналов в молекуле или имени молекулы.



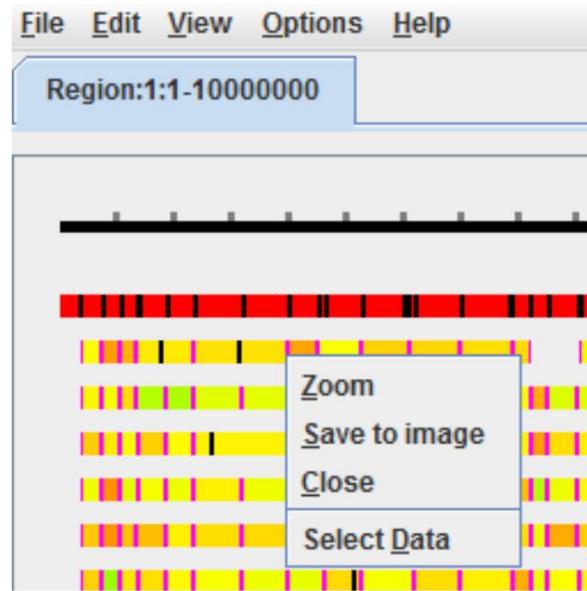
Щелкните правой кнопкой мыши, чтобы отсортировать молекулы.

Пример: (Молекула) Ecoli mid 1.sdata -

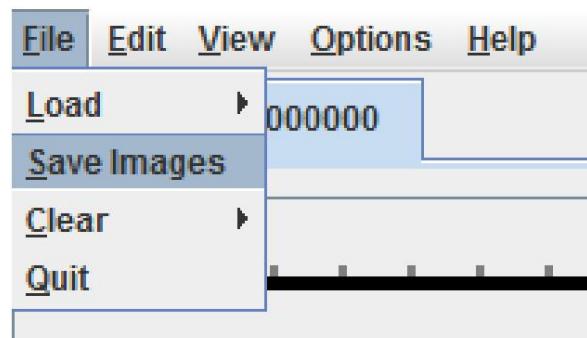
30.5.6 Экспорт изображения

Изображения можно экспортить на отдельных панелях или на нескольких панелях в форматах SVG, PNG и JPG.

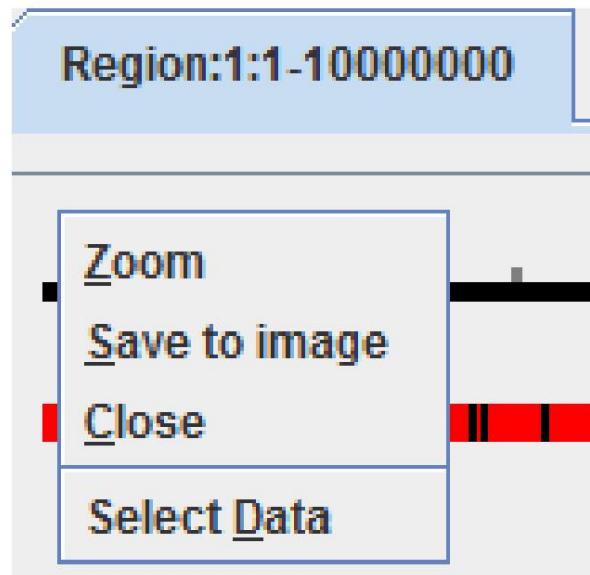
Отдельная панель В контекстном меню пользователя могут выбрать «Сохранить как изображение», чтобы сохранить текущую панель просмотра.



Все панели на одной вкладке Пользователи могут выбрать меню Файл Сохранить изображения, чтобы сохранить все панели просмотра на текущей вкладке в одно изображение.



30.5.7 Закрытие вкладок и панелей



Чтобы закрыть вкладку, щелкните по ней средней кнопкой мыши. Чтобы закрыть панель просмотра, щелкните правой кнопкой мыши панель и выберите «Закрыть».

30.6 Часто задаваемые вопросы (FAQ)

В: Существует ли пакетный режим генерации изображений?

А: Да. Вы можете использовать опцию `-viewsave` для пакетной генерации изображений.

В: Могу ли я выбрать несколько элементов одновременно для импорта с правочных или выборочных данных?

А: Да. Удерживайте клавишу `Control` при выборе нескольких элементов.

В: Я загрузил все файлы, включая скрипты, молекулы и результаты выравнивания, но ничего не вижу.

В чем дело?

А: Чтобы визуализировать скрипты и результаты выравнивания, (1) необходимо выбрать данные на панели просмотра (Выбрать данные в контекстном меню); и (2) необходимо задать область просмотра с помощью `EditSet View Region`.

В: Я установил область просмотра или якорь сайта для визуализации своих данных, но получаю ошибку «Скрипта не найдена».

Что не так? А: Наиболее распространенной причиной ошибки является неправильный ввод имени скрипта. Обратите внимание на разницу между «`chr1`», «`Chr1`», «`CHR1`» и «`1`».

В: Почему скорость загрузки файла очень низкая после загрузки половины моих данных?

А: Убедитесь, что у вас достаточно памяти для хранения всех данных. Используйте параметр `Xmx`, чтобы выделить больше памяти для Java-машины. Не загружайте с слишком много данных в экземпляре `OMView`, если на вашей машине недостаточно памяти.

В: Скорость загрузки регионального вида низкая.

А: Попробуйте ограничить диапазон региона. Обычно региону больше 1 Мбит/с требуется некоторое время, чтобы полностью загрузка.

В: Какие форматы принимает `OMView`?

А: Форматы файлов скриптов и молекул: REF, FA01, SPOTS, DATA, SDATA, BNX, CMAP, OPT, SILICO и OpGen XML. Форматы результатов выравнивания: OMA, OMD, XMAP, Valouev et al., SOMA v2 Unique Match и Twin PSL. Форматы аннотаций: BED, GVF и OSV, AGP

Часть XI

Другие сценарии

31 TWINResultRepeatRemover

Удаление повторных результатов выравнивания из TWIN.

31.1 Параметры считывателя результатов

```
--optresin Вх одног файла результатов выравнивания [Обязательно]  
--optresinformat -1: Автоматически определяется по расширению файла; 0: Формат выравнивания ОМ (OMA); 1: Формат подробного выравнивания ОМ (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]
```

31.2 Параметры записи результатов

```
--optresout Вывести файл результатов выравнивания [обязательно]  
--optresoutformat Формат файла результата -1: Автоматически определяется по расширению файла; 0: Формат выравнивания ОМ (OMA); 1: Формат подробного выравнивания ОМ (OMD); 2: Формат XMAP (XMAP); 3: Формат Valouev et al.; 4: Формат уникального соответствия SOMA v2; 5: Формат Twin PSL; 6: Формат Maligner ALN; [По умолчанию: -1]  
  
--writeunmap Записать отброшенные или неотображенные молекулы. [По умолчанию: true]  
--multiple Записать несколько карт для молекулы. [По умолчанию: true]  
--writeinfo Записать информацию о молекуле. [По умолчанию: true]
```

32 Отдельный BNXScan

Разделяет файл bnx на несколько файлов bnx в соответствии с глобальным номером сканирования.

32.1 Параметры считывателя данных --optmapin

Вх однай файл оптической карты [Обязательно]

--optmapinformat -1: Автоматически определяется по расширению файла; 0: Справочный стандартный формат (REF) (эквивалент формата SILICO); 1: формат fasta-01 (FA01); 2: формат файла Spots (SPOTS);
3: Стандартный формат молекулы (DATA); 4: Формат моделирования молекулы (SDATA); 5: Файл BNX
Формат (BNX); 6: Формат файла CMAP (CMAP); 7: Формат SOMA opt (OPT); 8: Формат SOMA silico (SILICO) (эквивалент формата REF); 9:
Формат OpGen XML; 10: Формат данных Валуева;
11: Формат карт Maligner; [По умолчанию: -1]

--bnxsnr Значение фильтра BNX SNR [по умолчанию: 3.0]

32.2 Отдельные параметры BNXScan

--prefix Префикс вывода [По умолчанию: сканирование]