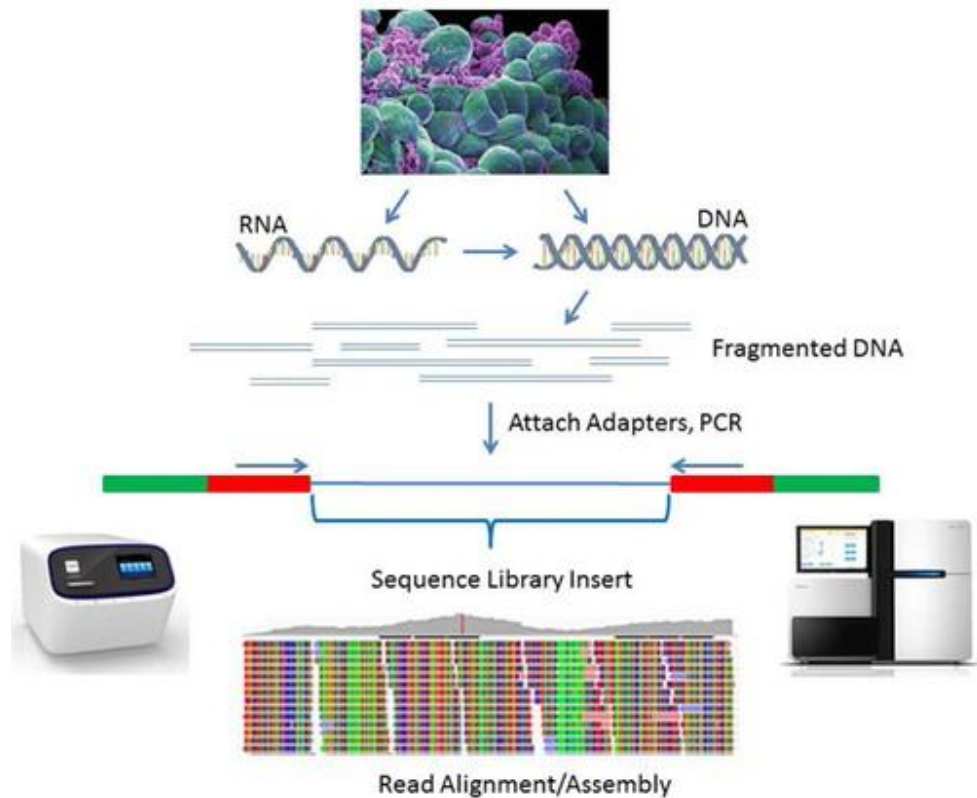


Приготовление библиотек

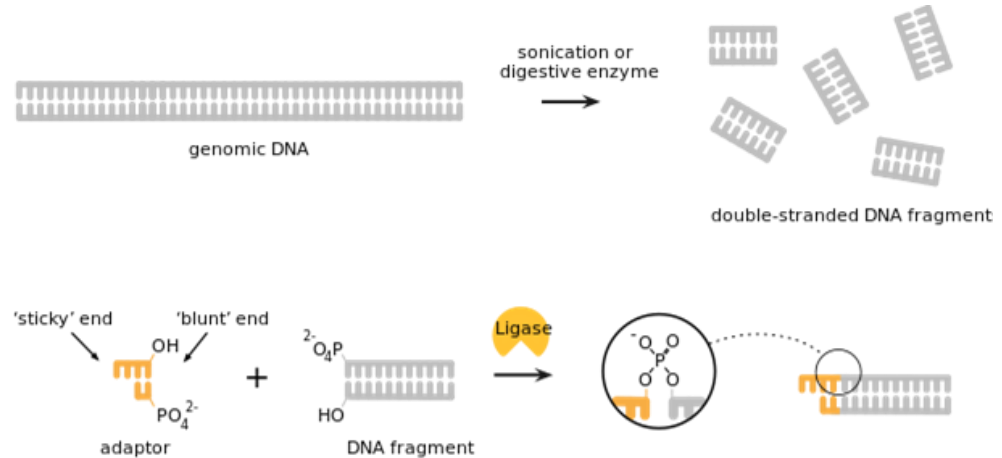
Прикладная биоинформатика: NGS-технологии и Omics-анализ

Basic workflow



Приготовление библиотеки

- 1) Опционально: амплификация таргетных регионов;
- 2) Фрагментация ДНК (необязательно, если ампликоны достаточно коротки);
- 3) Присоединение (ДНК-лигирование) адаптеров;
- 4) Опционально: гибридизационный захват;
- 5) Опционально: универсальная амплификация (праймеры к адаптерам).

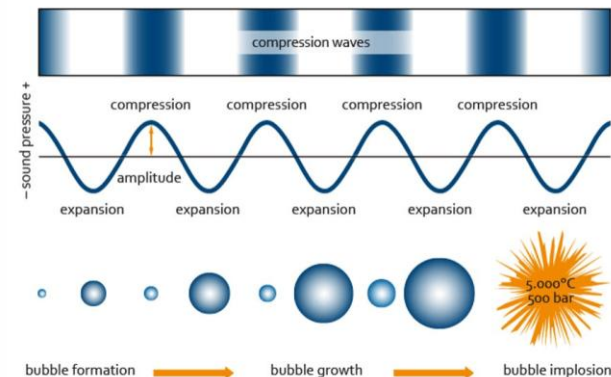


Amplicon sequencing vs. hybrid capture

Фрагментация

1) Физическая:

- a) Acoustic shearing: короткие высокочастотные акустические волны разбивают ДНК на фрагменты длиной от сотен до тысяч п.о.;
- b) Sonication: ультразвуковые волны формируют циклы сжатия и расширения, при низком давлении формируются вакуумные пузыри, которые схлопываются при накоплении энергии (кавитация);
- c) Nebulization, needle shearing...



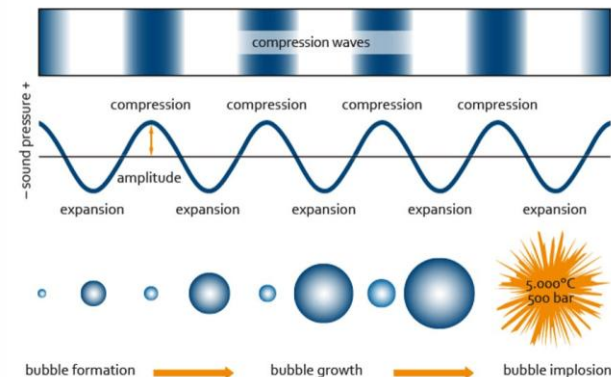
2) Энзиматическая;

3) Химическая.

Фрагментация

1) Физическая:

- a) Acoustic shearing: короткие высокочастотные акустические волны разбивают ДНК на фрагменты длиной от сотен до тысяч п.о.;
- b) Sonication: ультразвуковые волны формируют циклы сжатия и расширения, при низком давлении формируются вакуумные пузыри, которые схлопываются при накоплении энергии (кавитация);
- c) Nebulization, needle shearing...



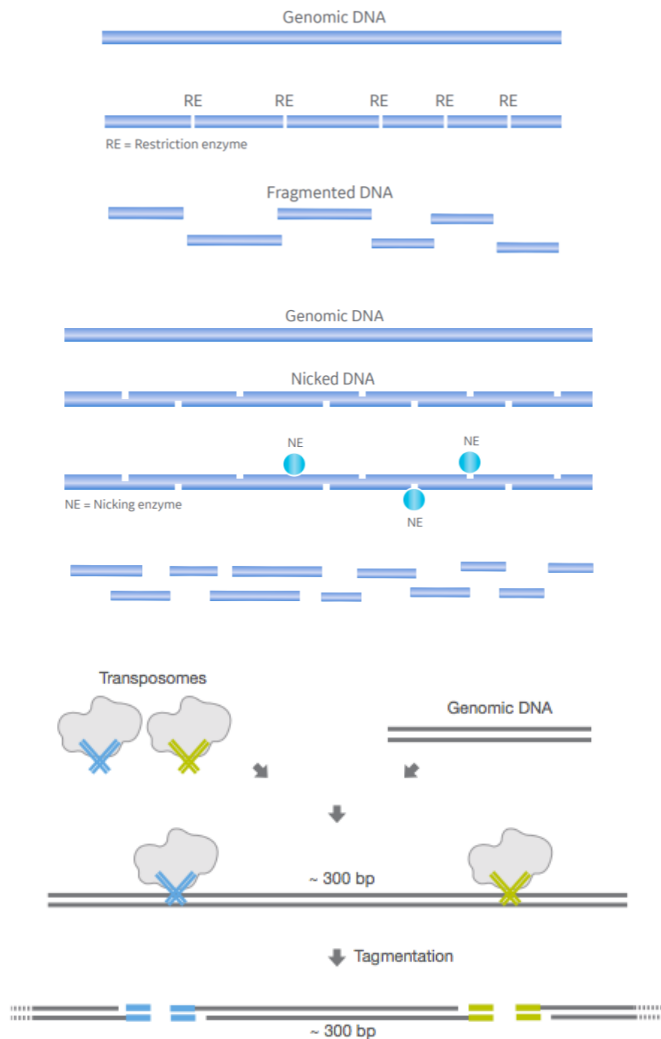
2) Энзиматическая;

3) Химическая.



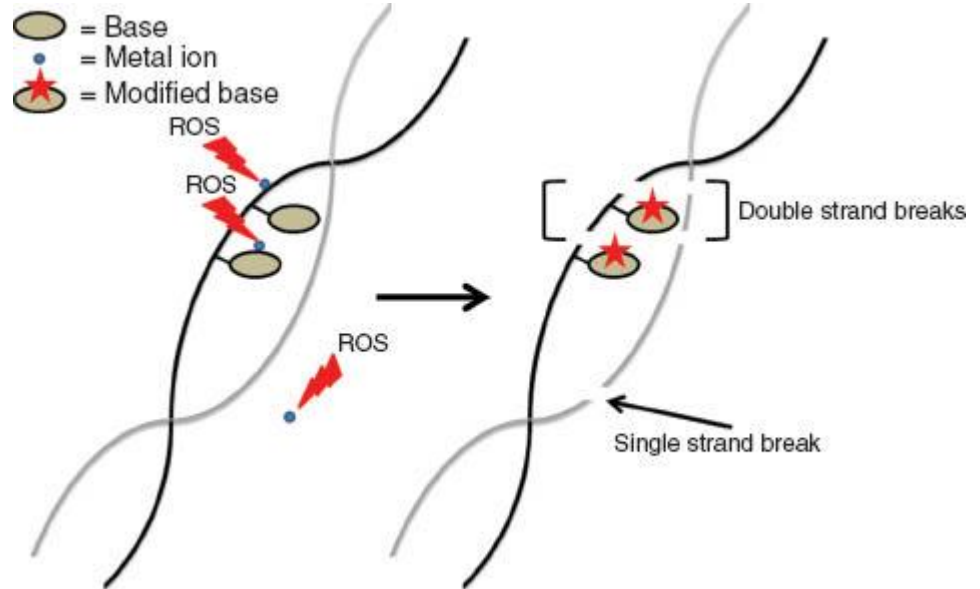
Фрагментация

- 1) Физическая;
- 2) Энзиматическая:
 - а) Рестриктазы (фрагментация будет не случайной, bias!);
 - б) Энзиматические смеси (разрыв одной цепочки ДНКазой I, затем разрыв второй никирующей эндонуклеазой, липкие концы!);
 - с) Тагментация (транспозазы случайно разрезают ДНК, сразу достраивают липкие концы адаптерами);
- 3) Химическая.



Фрагментация

- 1) Физическая;
- 2) Энзиматическая;
- 3) Химическая:
 - а) В основном для фрагментации РНК при помощи бивалентных металлических катионов (магний, цинк) при воздействии высоких температур;
 - б) При помощи комплексообразователей.



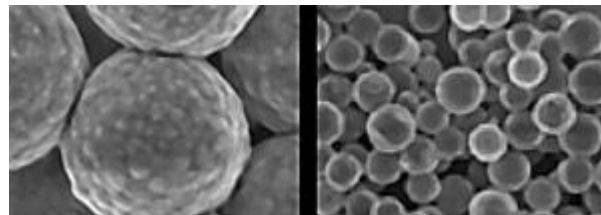
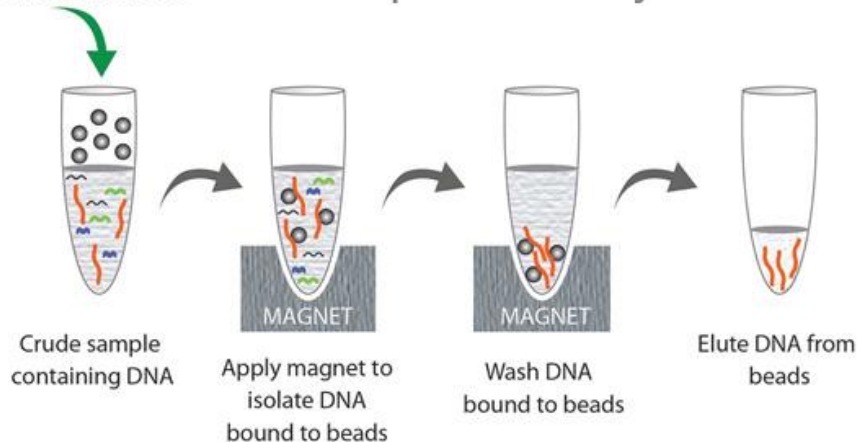
Отбор фрагментов

Магнитные шарики:

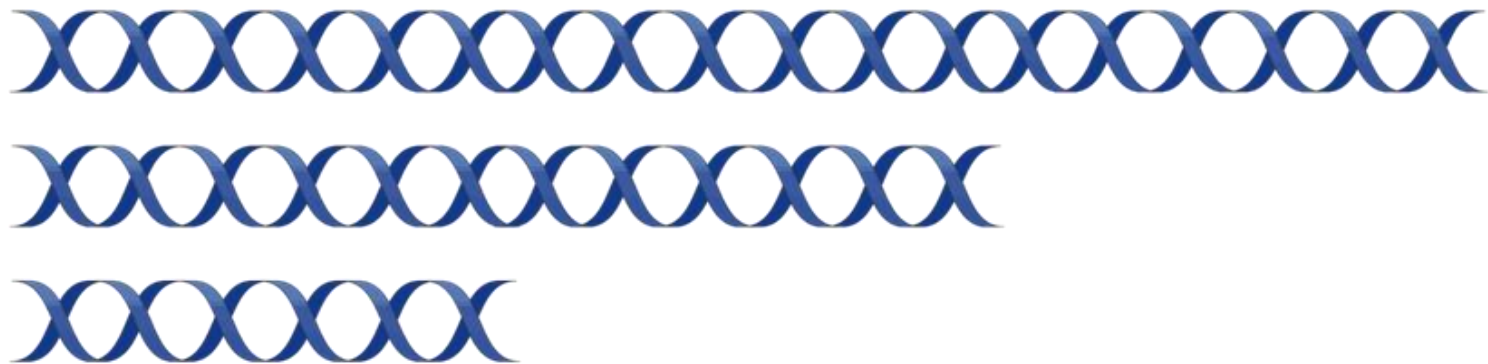
- Соотношение объёмов ДНК и шариков даёт необходимое взаимодействия для прикрепления к шарикам ДНК не ниже определённой длины;
- Первый этап - устранение чересчур длинных фрагментов;
- Второй - коротких (и “мусора”).

Add Si-Mag beads
to lysed cells and tissues

Principle of Assay

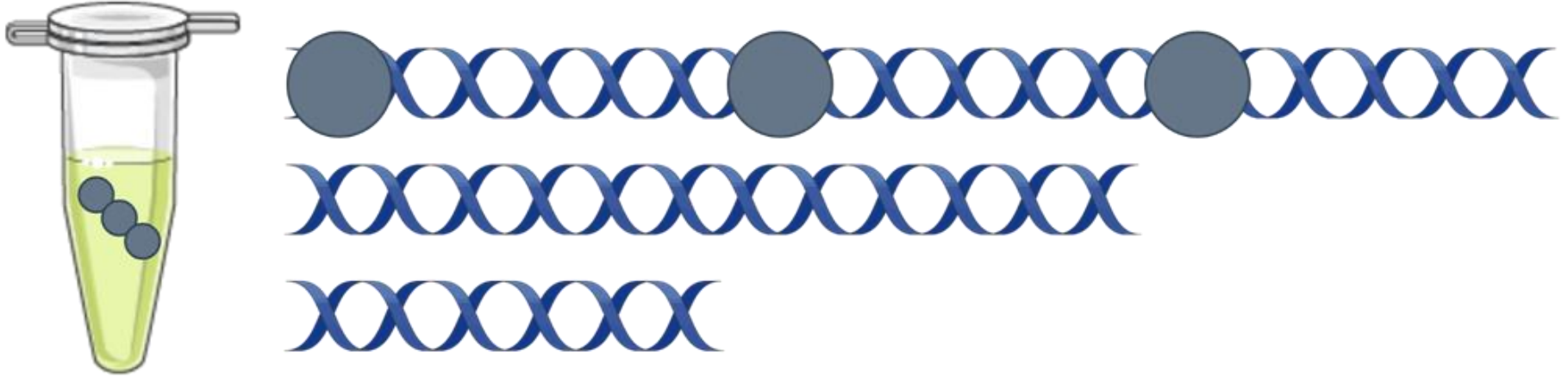


Отбор фрагментов



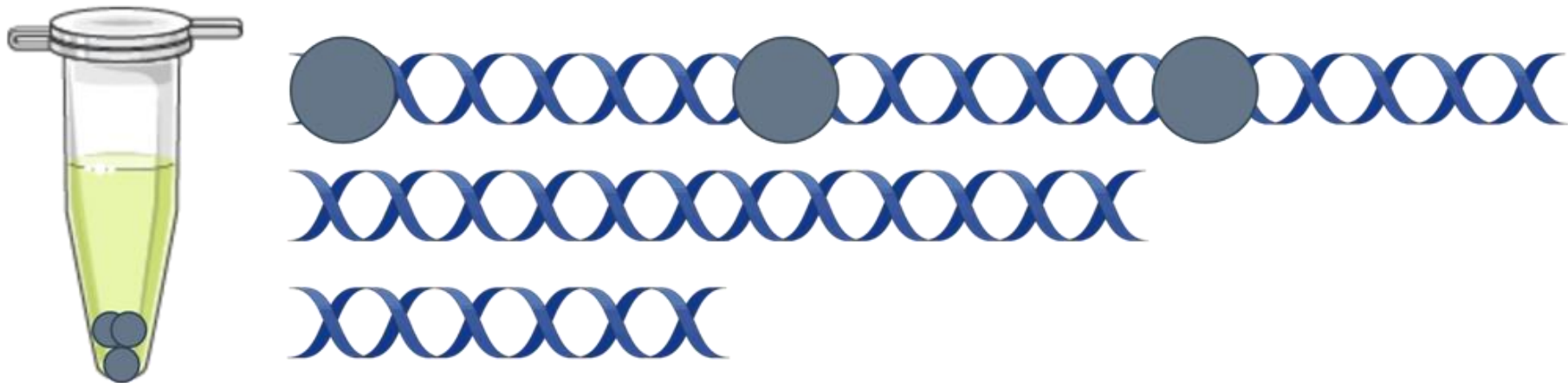
- Условно, в нашем материале есть длинные, средние и короткие фрагменты с частицами, не относящимися к ДНК (их считаем аналогичными коротким).
- Нам нужно оставить только средние.

Отбор фрагментов



- Добавляем шарики, которые в первую очередь присоединяются к самым длинным.

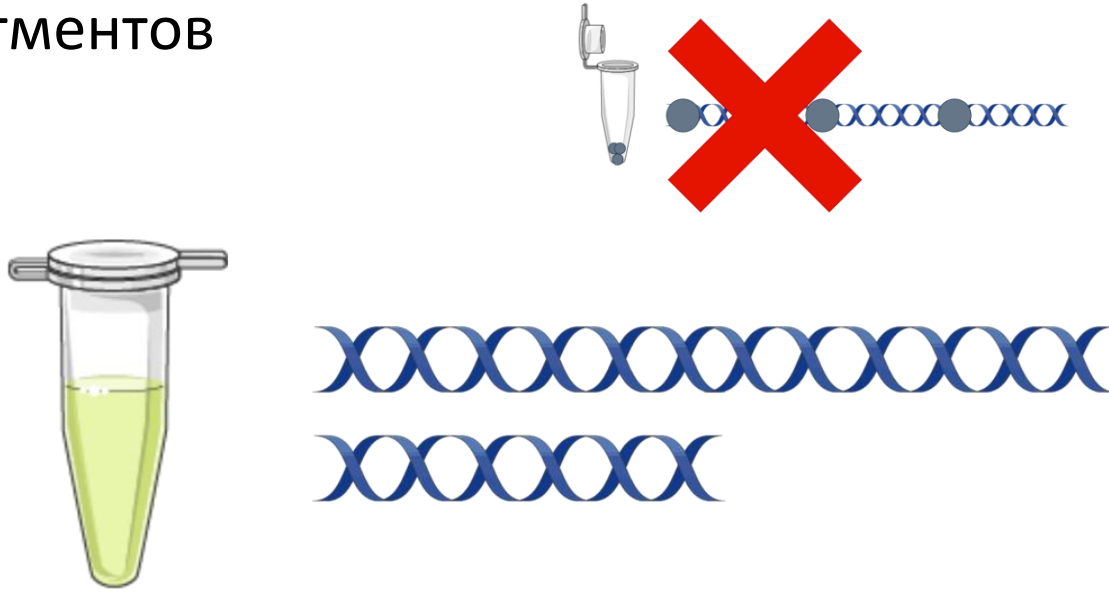
Отбор фрагментов



МАГНИТ

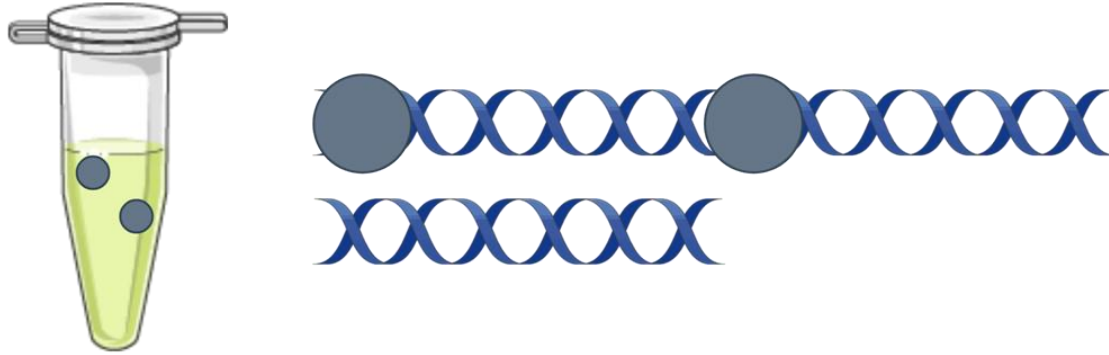
- Магнит снаружи притянет шарики с длинной ДНК, остаётся отобрать супернатант.

Отбор фрагментов



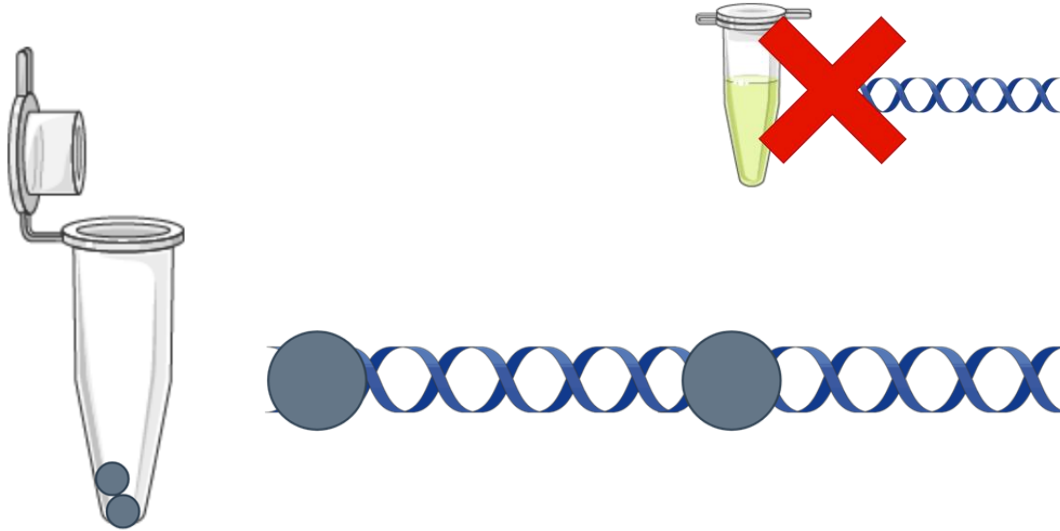
- Супернатант идёт далее, длинные фрагменты удаляем.
- В супернатанте остались средние фрагменты (их берём) и короткие с другими мелкими частицами (это не берём).

Отбор фрагментов



- Вновь добавляем магнитные частицы. Опять связываются самые длинные – только теперь они-то нам и нужны!
- Повторяем процедуру с магнитом, но теперь супернатант выбрасывается, а осадок остаётся.

Отбор фрагментов



- Повторяем процедуру с магнитом, но теперь супернатант выбрасывается, а осадок остаётся. Высушиваем осадок (спирт), отделяем ДНК от частиц (элюент). Частицы удаляются при помощи магнита. ДНК из супернатанта уже будет нужного размера.

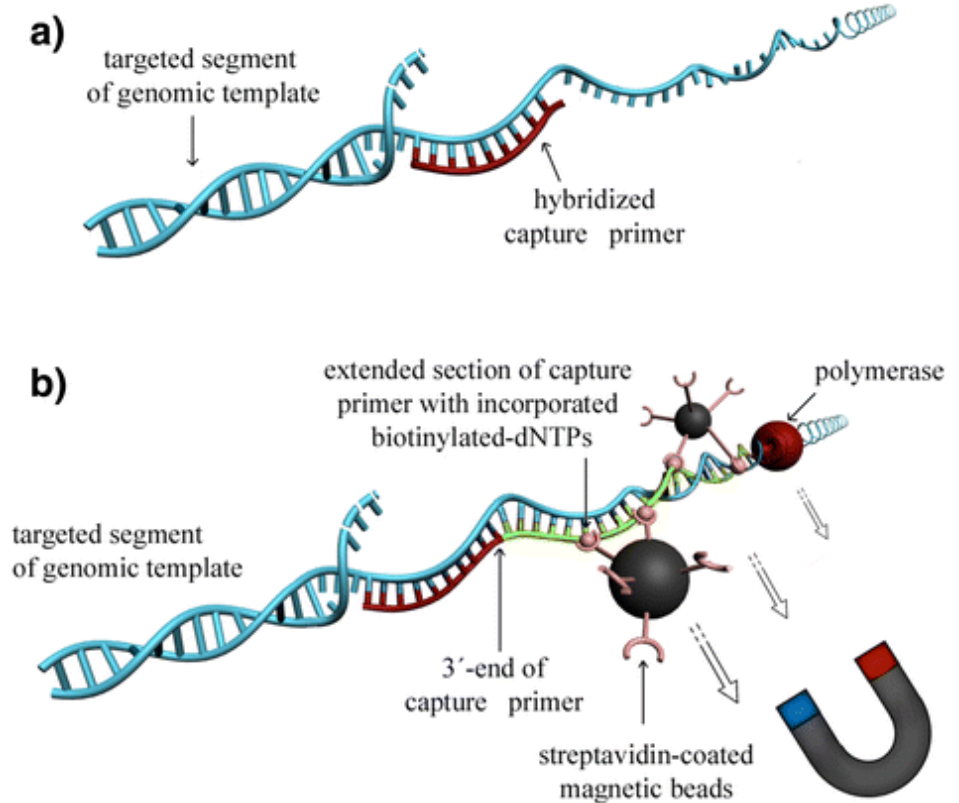
Адаптеры

- Последовательности для прикрепления к ячейке (P5, P7);
- Универсальные праймеры для секвенирования и амплификации (SP1, SP2);
- Индексы проб (barcode: i5, i7);
- Уникальные молекулярные индексы (UMI).



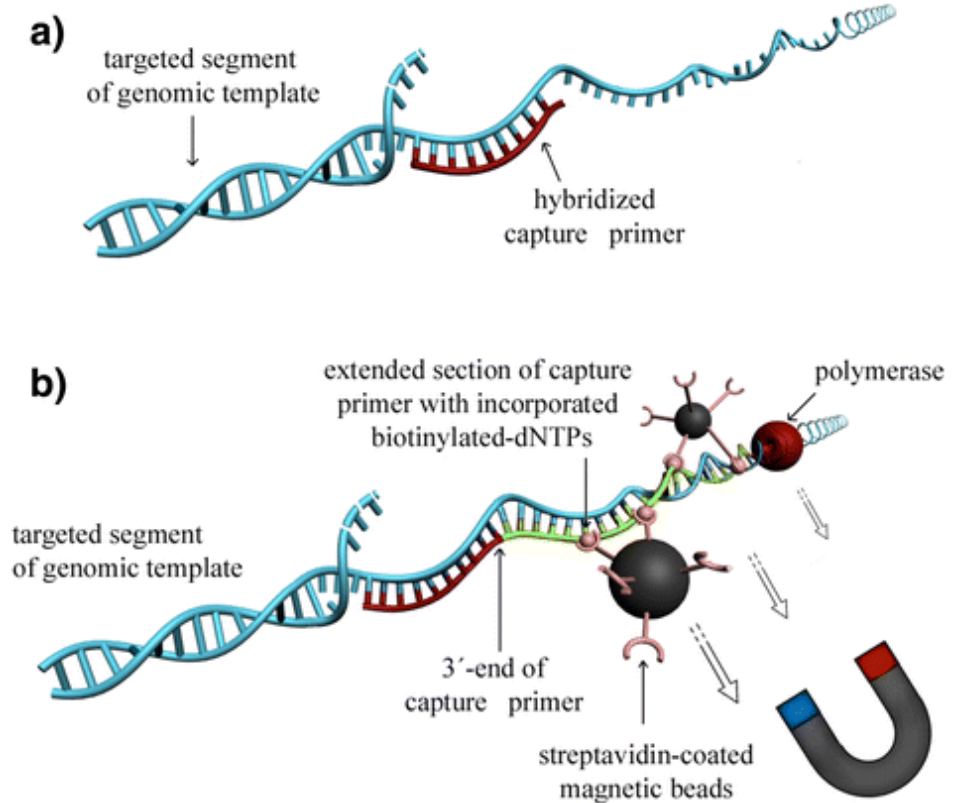
Гибридизационный захват

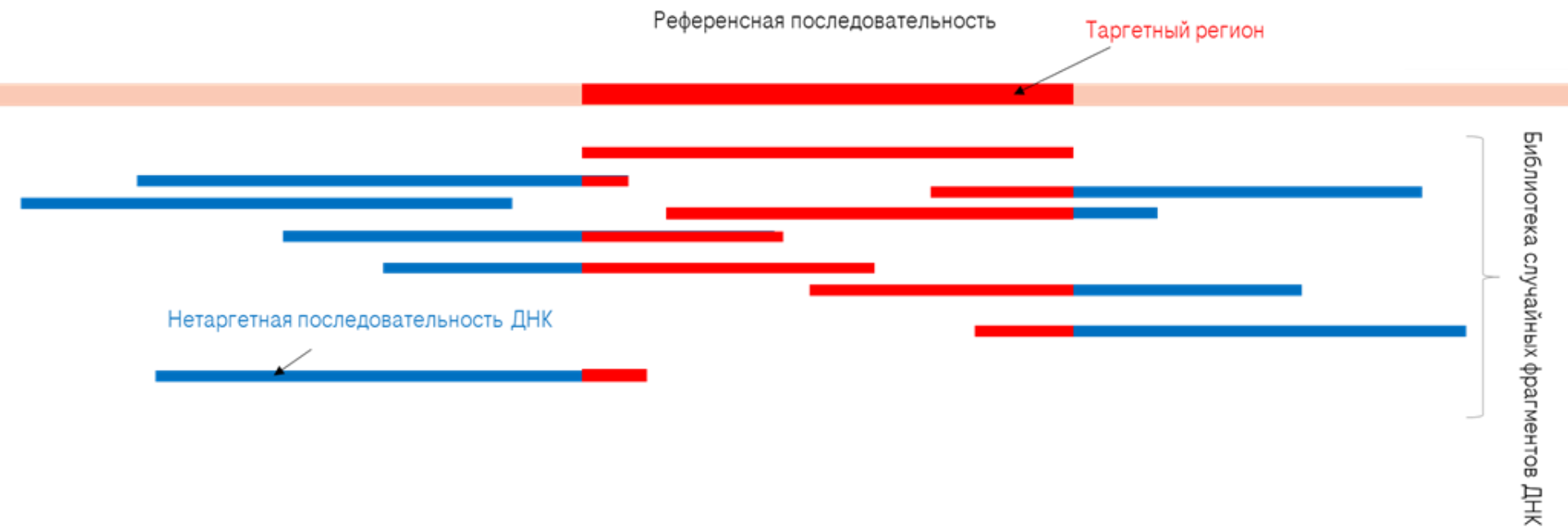
- Гибридизация с зондом, комплементарным мишени;
- Достройка зонда при помощи обычных и биотинилированных дНТФ;
- Добавление магнитных шариков со стрептавидином, прикрепление мишеней к ним;
- Отделение шариков от мишеней при нагреве.

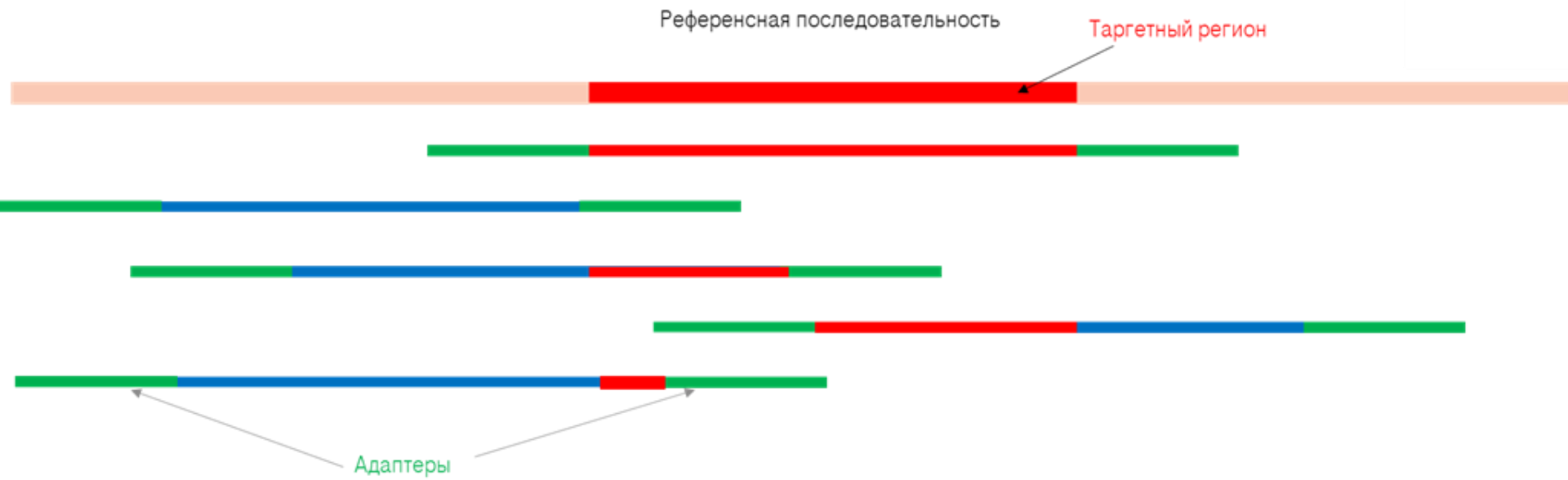


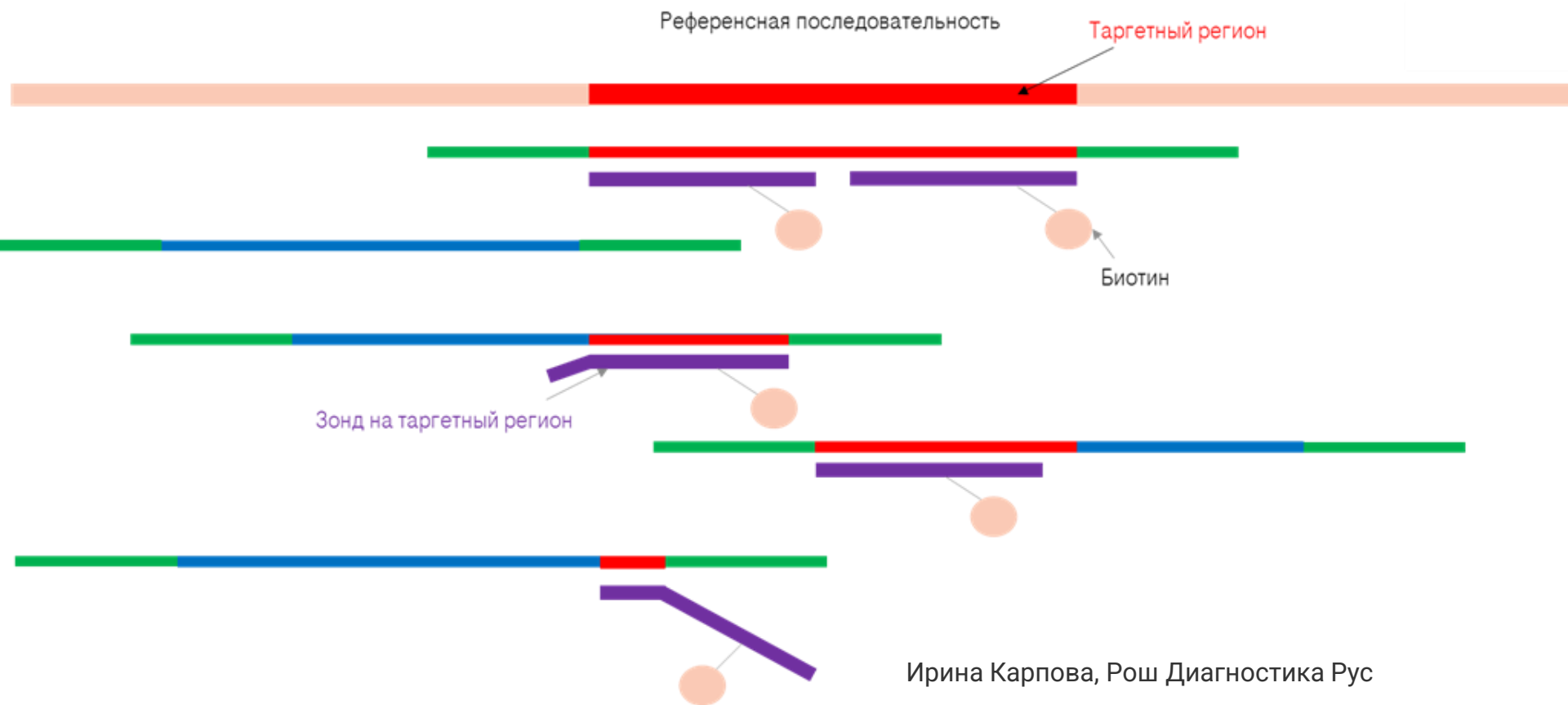
Гибридизационный захват

- Для присоединения ДНК к шарикам их необходимо лишить гидратной оболочки, для чего используется этанол.
- В дальнейшем шарики высушивают от этанола.
- Пересушивание и недостаточная сушка могут привести к повреждению ДНК и снижению объёма библиотеки.



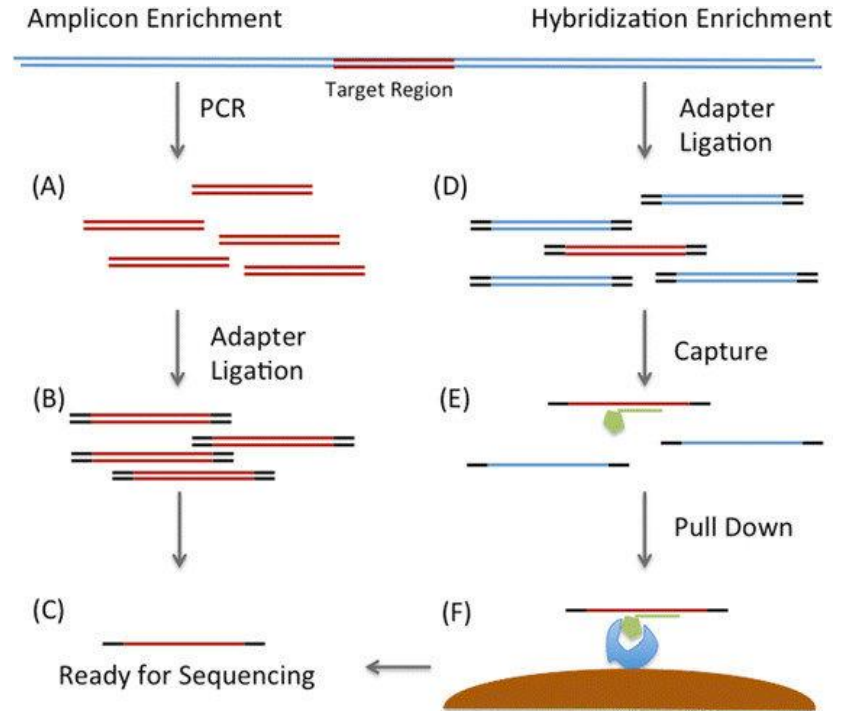






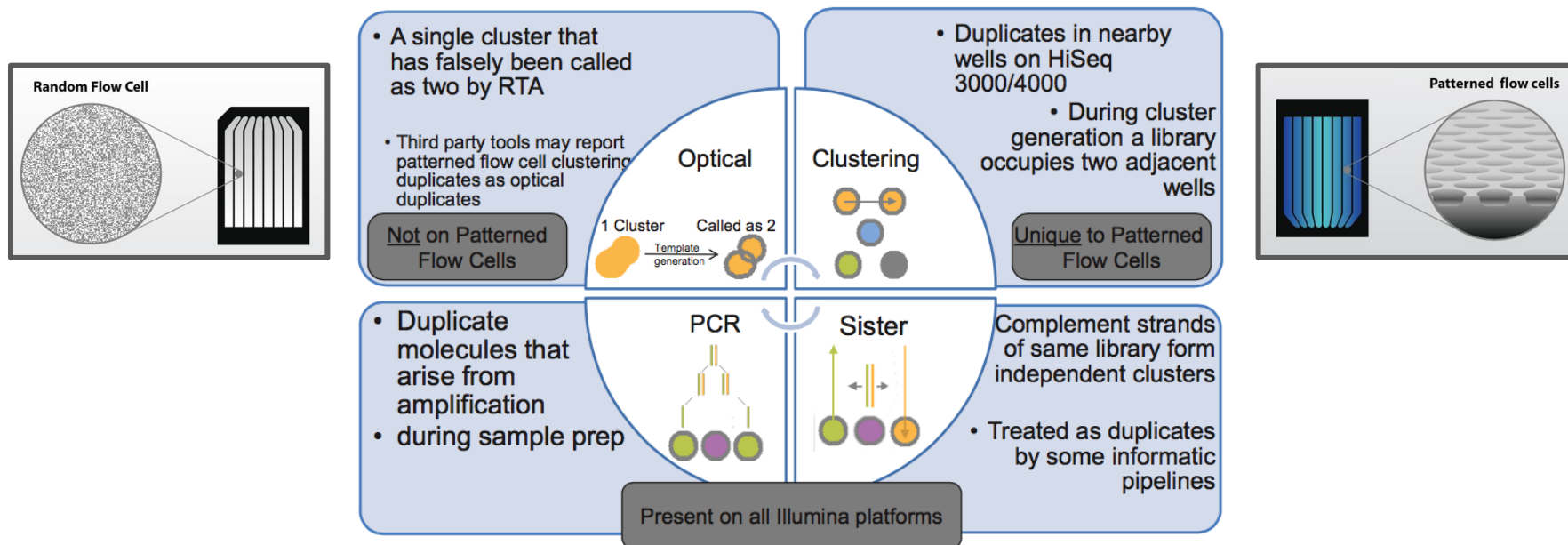
Наработка ампликонов

- Альтернативой гибридационному захвату является наработка ампликонов при помощи ПЦР.
- Праймеры обеспечивают специфическую амплификацию целевых регионов.
- Часто для этого пулируется несколько реакций.



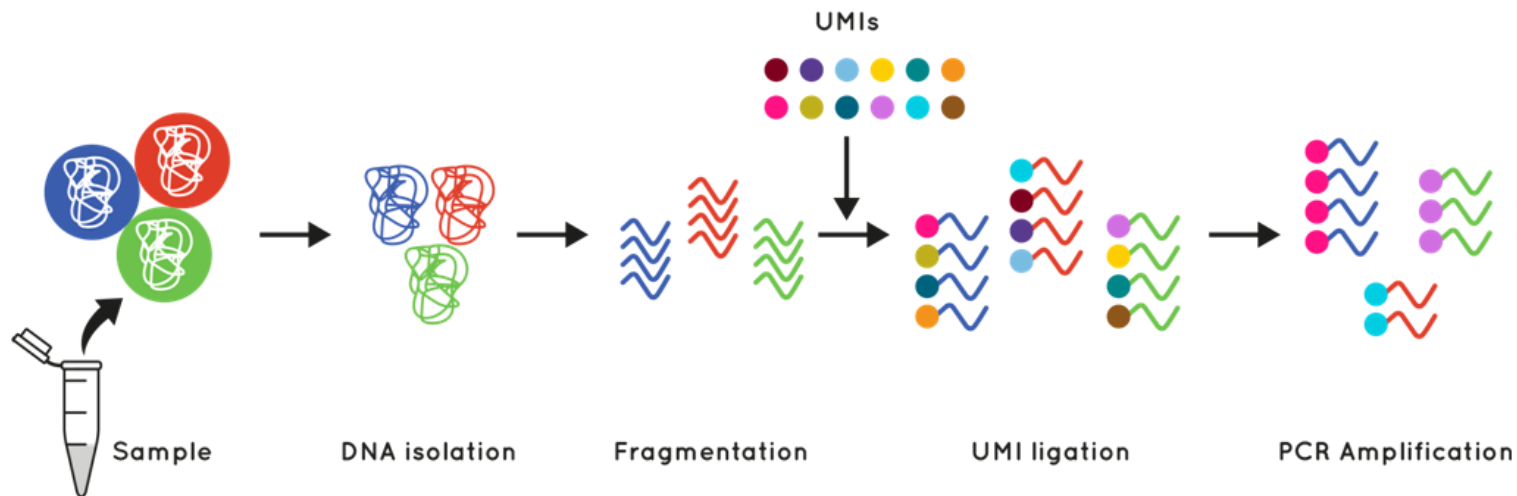
Дубликаты в секвенировании

Дубликаты – это одинаковые риды или такие риды, начальные нуклеотиды которых совпадают при выравнивании с референсным геномом.



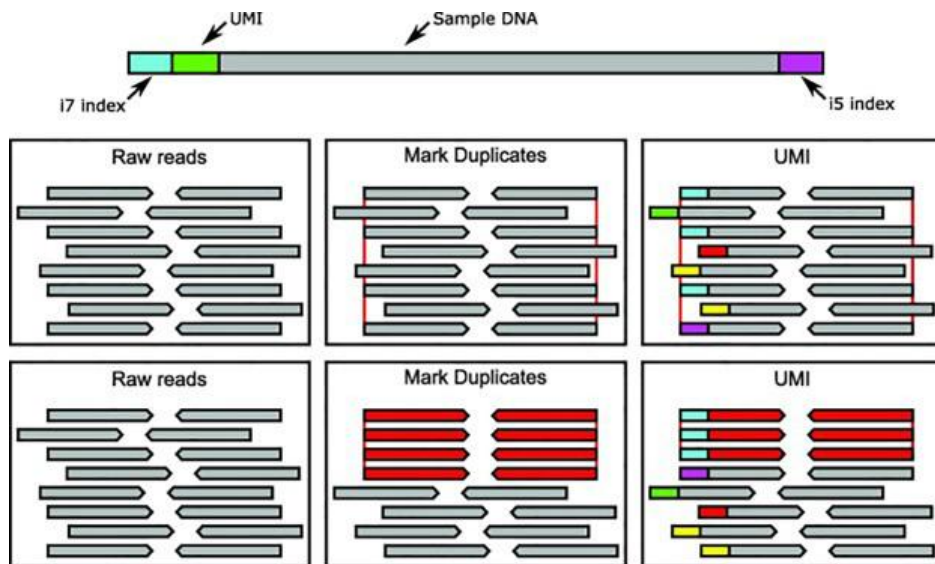
Важность UMI технологии

UMI – короткие ДНК-последовательности (заранее известные), которые лигируются с фрагментами ДНК исследуемой пробы. Все UMI – разные, поэтому каждая молекула помечена уникальным молекулярным индексом.



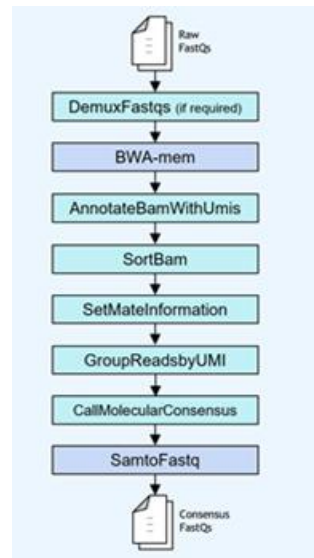
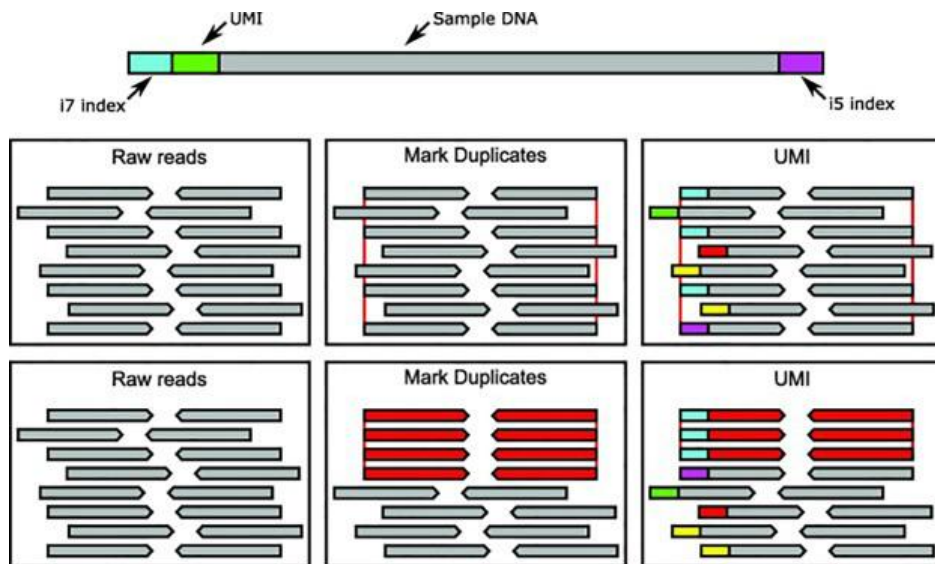
Важность UMI технологии

При анализе UMI позволяют более корректно отмечать дубликаты, появившиеся в эксперименте.



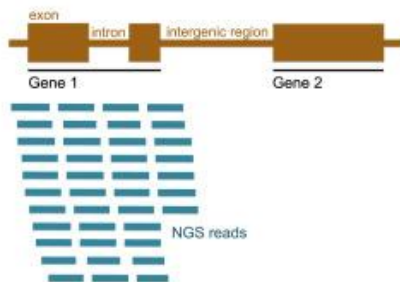
Важность UMI технологии

При анализе UMI позволяют более корректно отмечать дубликаты, появившиеся в эксперименте.



Таргетное, полноэкзомное и полногеномное секвенирование

Targeted Sequencing (TS)

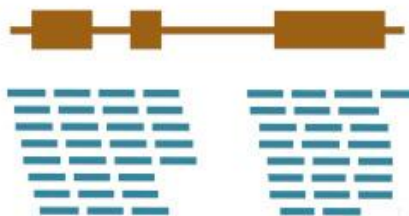


Selection (capture):
Panel of genes of interest
(typically 20-200 predefined genes)

- greatest sequencing depth*
- most cost effective

Fixed selection of genes,
analysis cannot be expanded

Whole Exome Sequencing (WES)

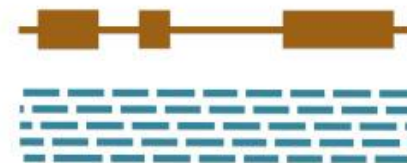


Selection (capture):
All exons of all known genes
(1.5-2% of all human DNA)

- variable read depth at boundaries
- greater sequencing depth*
- more cost effective

Subsequent *in silico* selection of genes/gene list(s) of interest
(but entire dataset generated: analysis can be expanded to other genes)

Whole Genome Sequencing (WGS)



No selection:
Entire human DNA analyzed
Including introns, RNA genes, etc.

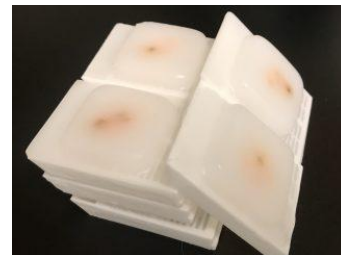
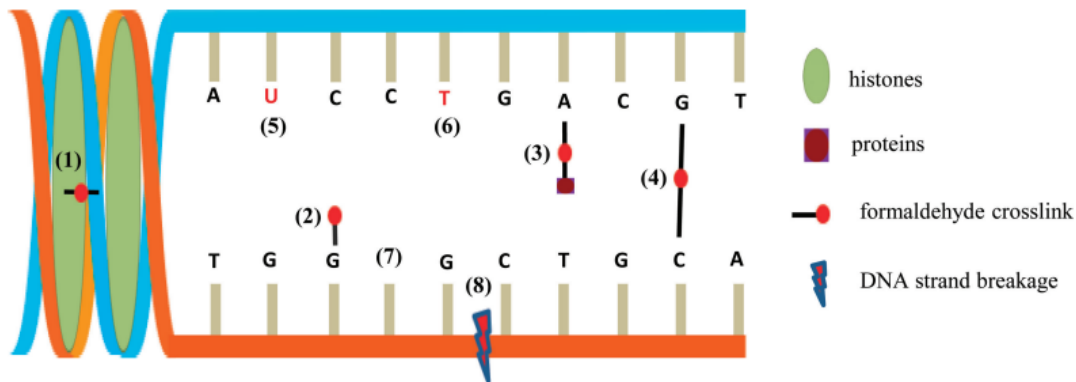
- moderate read depth*
- similar read depth throughout
- may also detect copy number variants and repeat expansions
- higher price

Механизм повреждения ДНК при фиксации

Формальдегид (составная часть формалина) может формировать кросслинки между ДНК и другими макромолекулами, нарушая стабильность двухцепочечной структуры (например, метиленовые мостики с аминокислотами).

Фрагментация усиливается при пониженном pH и увеличенной длительности взаимодействия формалина с материалом.

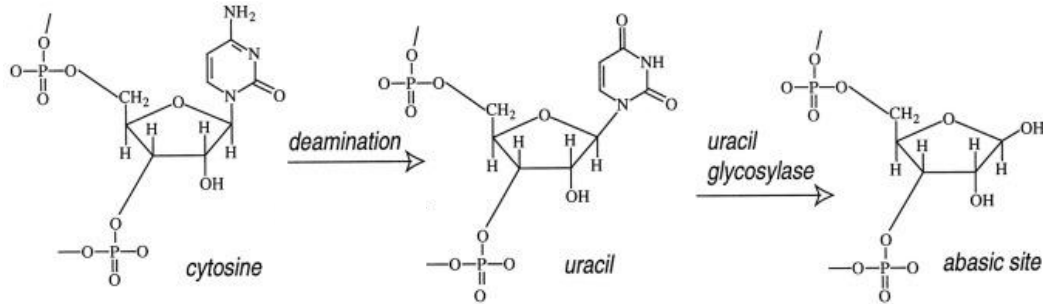
Histone–DNA crosslinks (1), formaldehyde–DNA adducts (2), DNA–protein crosslinks (3), and DNA–DNA crosslinks (4). Uracil (5) and thymine (6). Abasic sites (7). Fragmentation of DNA (8).



Механизм повреждения ДНК при фиксации

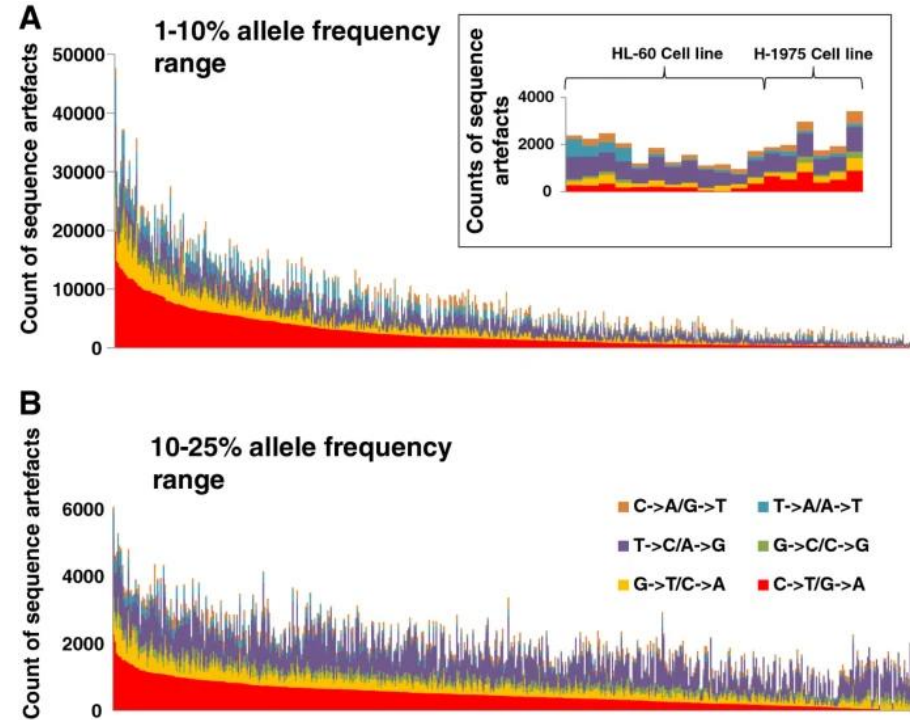
Гидролиз, приводящий к отделению азотистого основания от дезоксирибозы, приводит к формированию дезокси-бесосновных сайтов.

Гидролитическое деаминирование может превращать цитозин в урацил



“Formalin-fixed paraffin-embedded” материал

- Чаще всего в ДНК из FFPE-материала наблюдаются ложные замены C>T/G>A.
- Например, таким образом может обнаруживаться ложноположительная замена NRAS G12D, приводящая к меланоме.
- В рамке на графике – структура артефактов секвенирования ДНК, полученной из не зафиксированного материала.



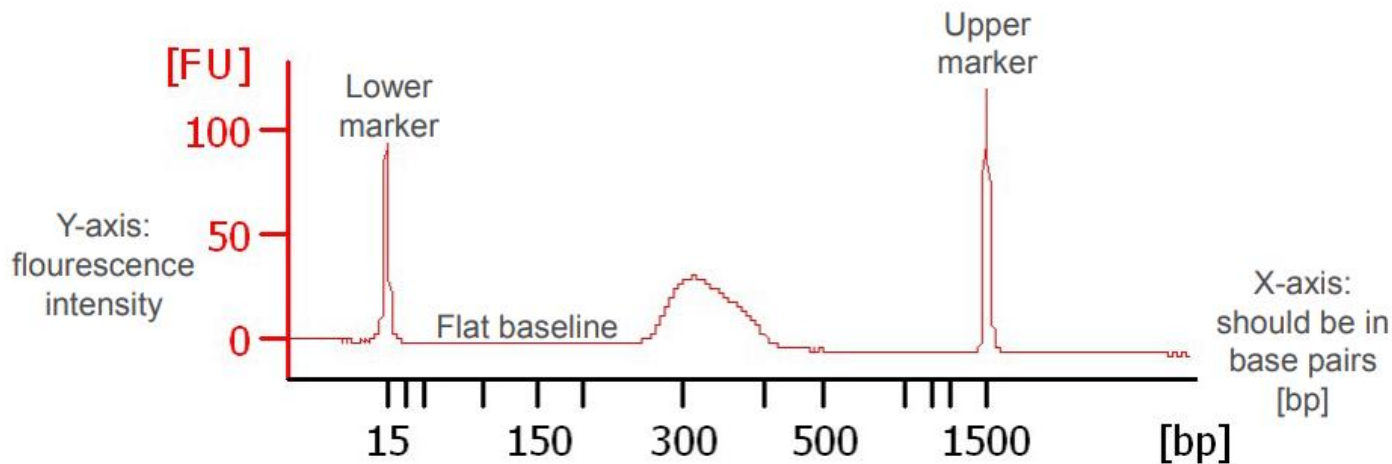
Амплификация перед загрузкой в прибор

- В случае, когда имеется достаточное количество ДНК для загрузки в секвенатор, дополнительная амплификация не требуется.
- В противном случае, проводится (дополнительная) амплификация, при этом используются универсальные последовательности адаптеров.



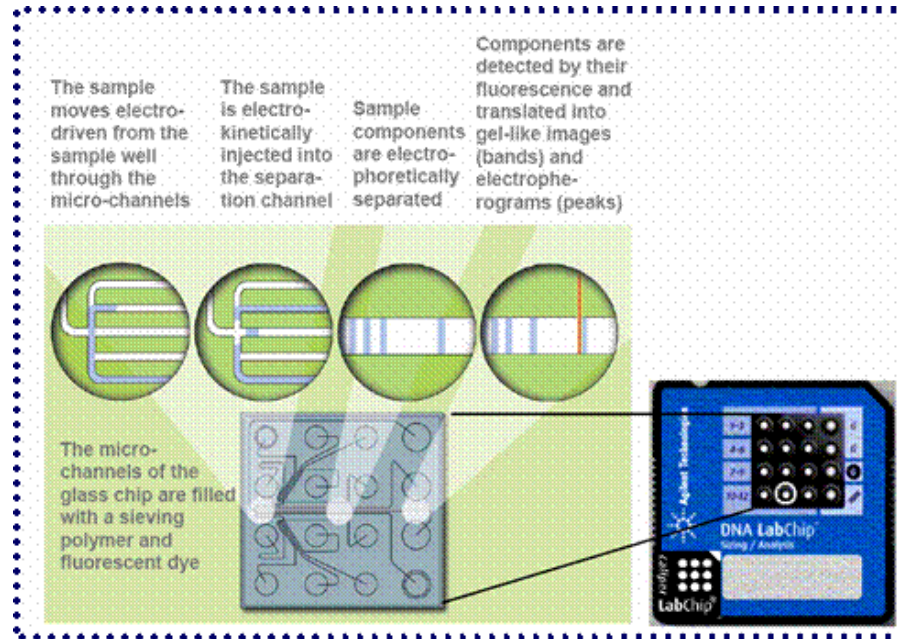
Bioanalyzer

Для оценки корректности приготовления библиотеки необходимо проводить контроль концентрации ДНК и оценивать распределение фрагментов по длине.



Bioanalyzer

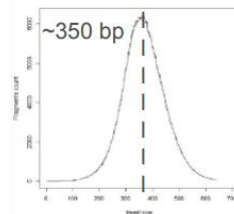
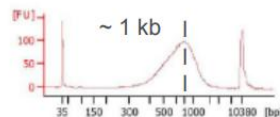
Принцип метода состоит в электрофорезе молекул в микрокапиллярах на чипе.



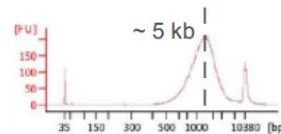
Нормальный размер библиотеки

- Размеры библиотеки на электрофорезе отличаются в большую сторону от размера вставки (как минимум, в составе библиотеки вставку окружают адаптеры, на фореze будут димеры и проч.), но пропорциональны ей.

TruSeq PCR-Free 350 bp insert

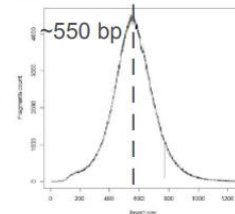


TruSeq PCR-Free 550 bp insert



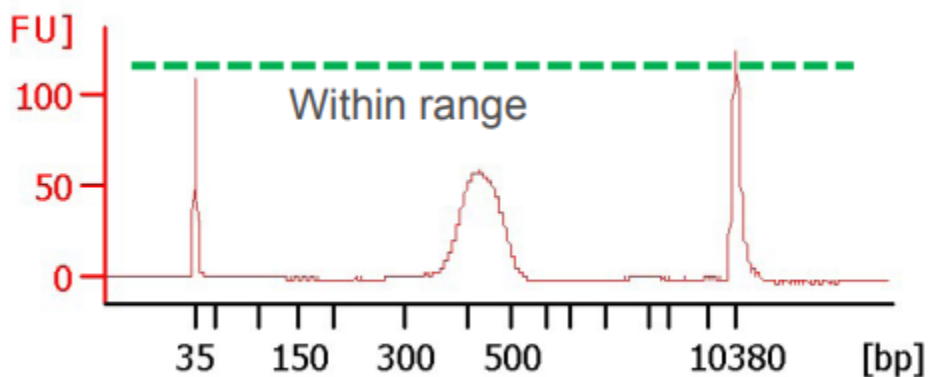
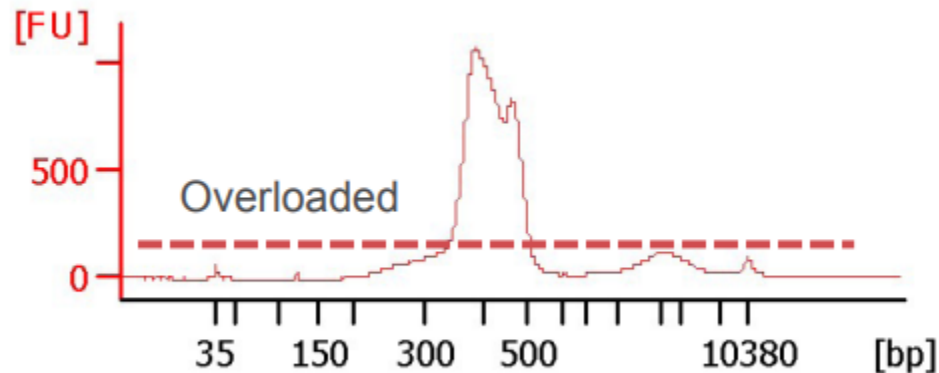
Final library
Bioanalyzer traces

Post-run analysis
average insert sizes



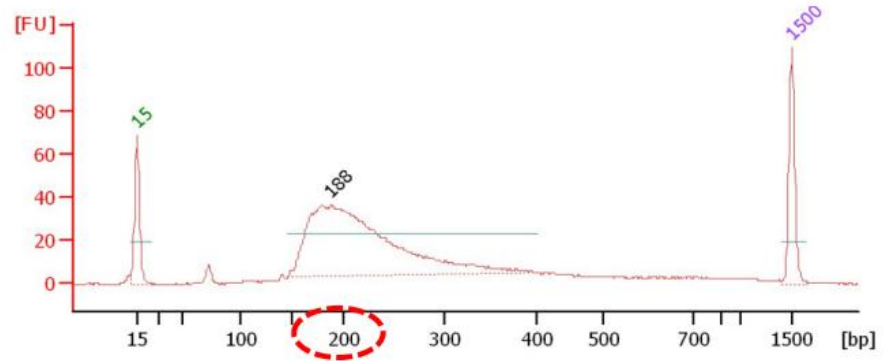
Слишком высокая концентрация библиотеки

Решение: развести библиотеку и перегнать форец.



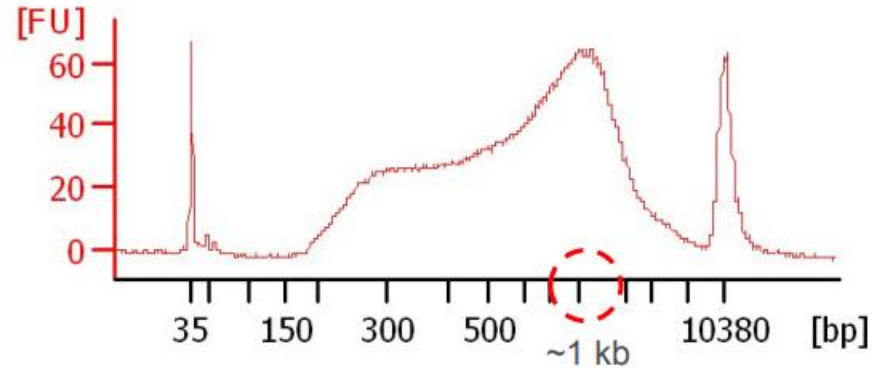
Недостаточный размер библиотеки

- В случае ДНК низкого качества, малого исходного количества ДНК, или избыточной фрагментации, размер библиотеки может получиться слишком малым.
- Для выяснения причин необходимо проводить контроль качества исходного материала (исходная фрагментация материала (FFPE?), концентрация)), а также контролировать фрагментацию ДНК в ходе приготовления библиотеки.



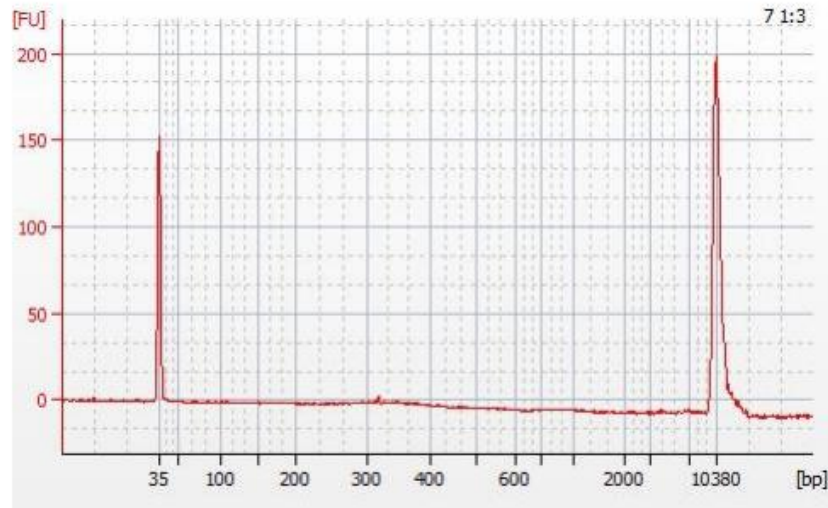
Избыточный размер библиотеки

- В случае слишком высокого исходного количества ДНК, недостаточной фрагментации, может сформироваться библиотека избыточного размера.
- Для выяснения причин этого необходимо контролировать концентрацию ДНК в исходном материале, проводить фрагментацию согласно протоколу, а также не допускать контаминации ингибиторами фрагментации.



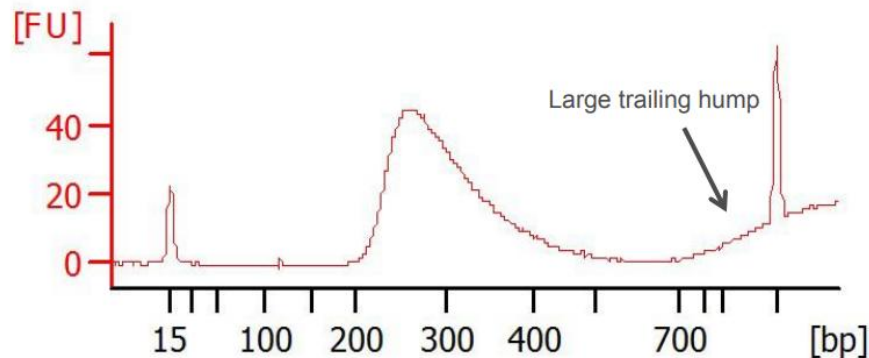
Отсутствие пика библиотеки

- В подобных случаях секвенирование не должно проводиться.
- Во всех остальных случаях оно может быть проведено, и вероятность адекватного результата будет ненулевой.



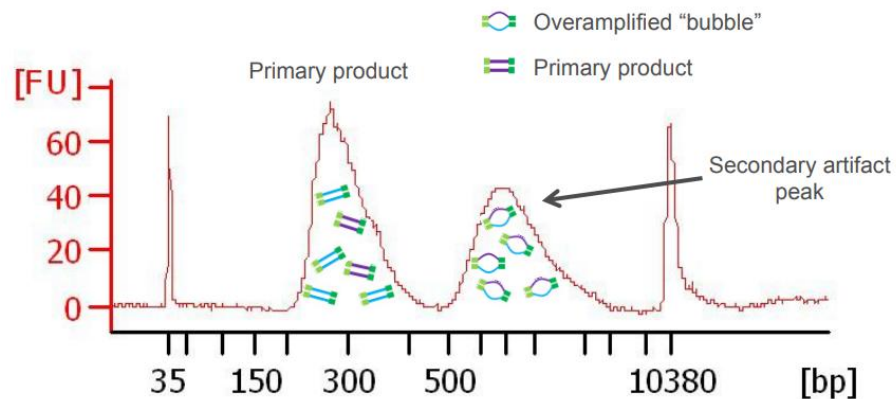
Дополнительные пики: магнитные шарики

- Существует вероятность занесения в библиотеку магнитных шариков.
- В таком случае следует удалить их при помощи магнитов согласно инструкции.



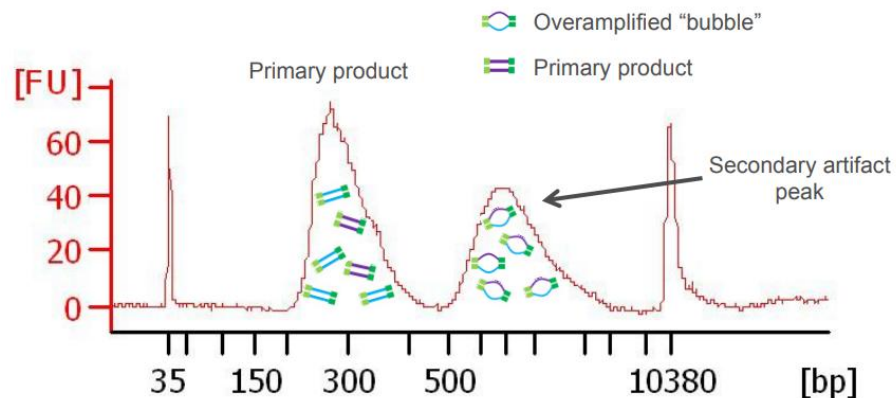
Дополнительные пики: bubble products

- В случае избыточной амплификации при помощи универсальных праймеров может происходить гибридизация некомплементарных участков.
- Они не будут амплифицироваться, однако, могут создавать структуры, которые из-за своей формы (в виде пузыря, когда некомплементарные вставки лежат некомпактно) медленнее перемещаются в геле, формируя дополнительный пик.



Дополнительные пики: bubble products

- Для проверки: провести денатурацию (95C), дать медленно остыть в термостате, переставить электрофорез. Дополнительный пик должен слиться с основным пиком библиотеки; в противном случае причину следует искать в другом.



Дополнительные пики: димеры адаптеров

- Малое количество ДНК, низкое качество исходного материала и недостаточная очистка могут приводить к попаданию значительного количества адаптеров в библиотеку. Адаптеры формируют устойчивые димеры.
- Димеры адаптеров могут занимать место на ячейке и секвенироваться, поэтому в случае обнаружения димеров необходимо удалить их при помощи дополнительной очистки (например, при помощи магнитных шариков).

