

Практические занятия по пробоподготовке, секвенированию и первичному анализу данных NGS с последующим анализом рисков при их использовании в клинической генодиагностике

Научный руководитель :
Александр Павлов, Sequoia genetics

Задачи проекта:

- Акцептировать знания о практических аспектах проведения пробоподготовки и секвенирования для полногеномного анализа.
- Научиться разбираться в метриках качества данных и форме их представления.
- Научиться понимать варианты построения pipeline'ов для анализов данных таргетного и полногеномного анализа для клинической генодиагностики.
- Провести анализ рисков при разработке решений в области генетического анализа на базе технологии NGS.

Практические занятия по пробоподготовке, секвенированию и первичному анализу данных NGS с последующим анализом рисков при их использовании в клинической генодиагностике

Научный руководитель :
Александр Павлов, Sequoia genetics

Задачи проекта:

- Акцептировать знания о практических аспектах проведения пробоподготовки и секвенирования для полногеномного анализа.
- Научиться разбираться в метриках качества данных и форме их представления.
- Научиться понимать варианты построения pipeline'ов для анализов данных таргетного и полногеномного анализа для клинической генодиагностики.
- Провести анализ рисков при разработке решений в области генетического анализа на базе технологии NGS.

Принципиальные отличия клинического и академического применений технологий NGS

	Академическое применение	Клиническое применение
Цели	Выявление общих закономерностей и взаимосвязей	Постановка диагноза и выбор лечения для конкретного пациента
Возможность набора статистики	Практически всегда присутствует и может компенсировать неточность пайплайнов и инструментов	отсутствует
Необходимая чувствительность	Зависит от эксперимента	Всегда высокая
Необходимая скорость анализа	Не критична	Всегда высокая
Критичность ошибок	Низкая	Чрезвычайно высокая
Квалификация персонала в области биоинформатики	Высокая	отсутствует

Этап I: Пробоподготовка

Сбор проб

Выделение геномной ДНК
($C \geq 3.3 \text{ ng}/\mu$)

Амплификация (PCR) регионов интереса

Отрезание праймеров

Лигирование адапторов

Очистка

Вторичная амплификация

Очистка от геномной ДНК и коротких фрагментов

Квантификация

Библиотека

Смесь библиотек

Эмульсионная PCR на сферах

Очистка

Контрольные сферы

Загрузка чипа



Риски на этапе пробоподготовки:

- 1) риск неравномерности покрытия при амплификации методом PCR
- 2) риск плохой очистки (этапы "сбор проб" → ... → "библиотека" осуществляется вручную)

- ## Риски на этапе пробоподготовки:
- 3) риск плохой загрузки чипа (также осуществляется вручную)

Базовые проекты для практических занятий

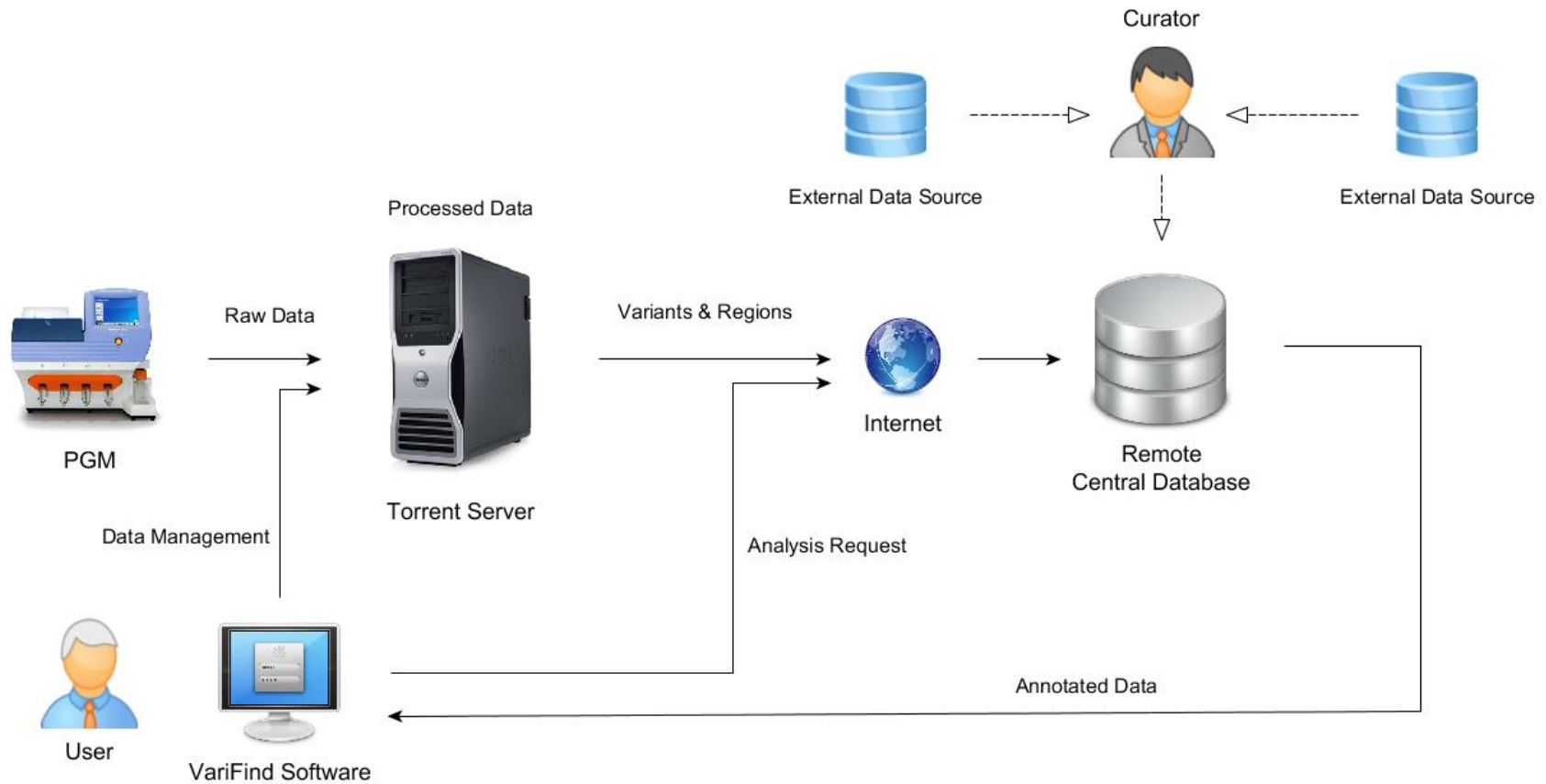
Таргетная неонатальная NGS диагностика

- Секвенирование 3 протяженных генов новорожденных с множеством горячих точек
 - CFTR (муковисцидоз)
 - PAH (фенилкетонурия)
 - GALT (галактоземия)
- Амплификация со 192 пар праймеров (20,73 kb)
- Платформа Ion PGM

Пренатальная неинвазивная NGS диагностика

- В крови матери присутствует ДНК плода (фетальная ДНК)
- Доля фетальной ДНК составляет 5-10% от всей внеклеточной ДНК крови беременной
- Теоретическая возможность диагностики анеуплоидий
- Секвенирование внеклеточной ДНК беременной
- Платформа Ion Proton

Таргетная неонатальная NGS диагностика



Этап II: Секвенирование

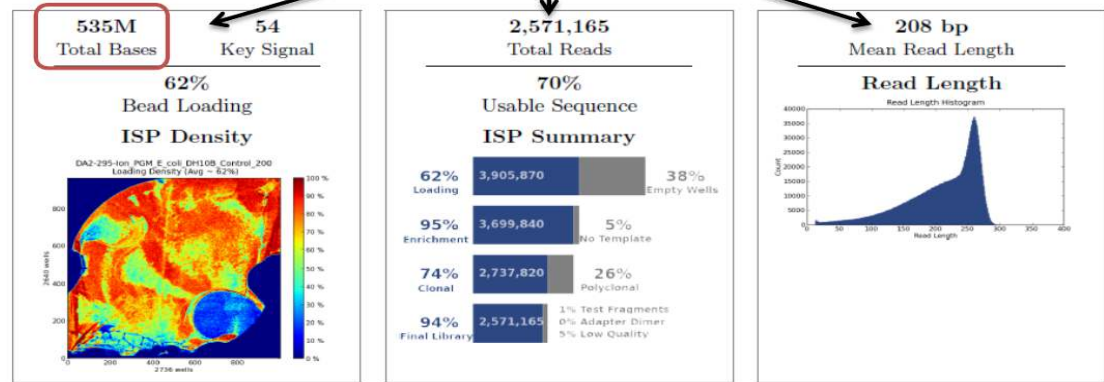


Ion 316™ Chip
300-600 Mb (~350-400 Mb на практике)
48 образцов 20.73kb с 350x покрытием



Ion Torrent Ion PGM™ Sequencer

Total bases or throughput will depend on the **loading density**, **total number of reads (usable sequence)** and **read length**; each of which is affected by several factors described below.

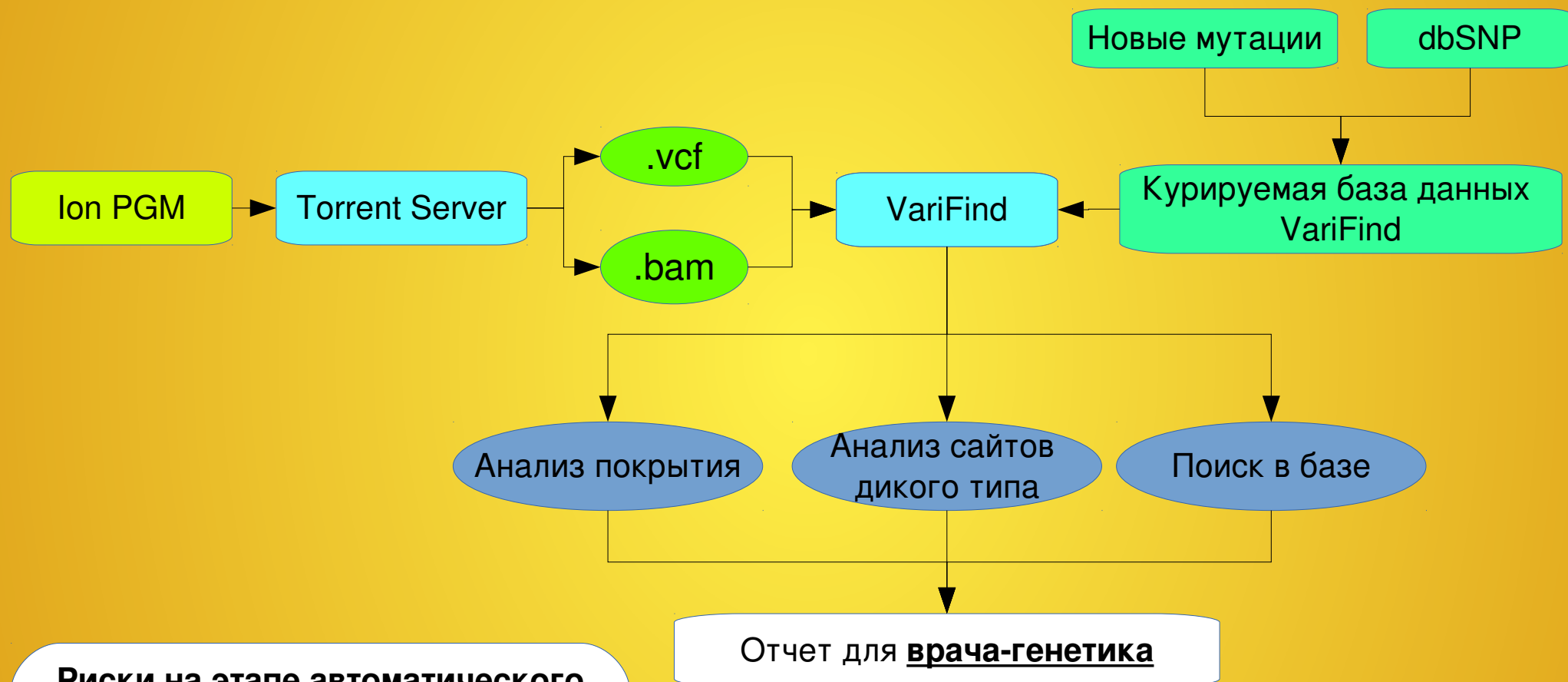


Метрики качества запуска:

1. Доля загруженных ячеек чипа
2. Доля сфер с ампликонами
3. Доля моноклональных сфер
4. Доля ридов требуемого качества

Риски на этапе секвенирования:
ошибки секвенирования

Этап III: Анализ данных



Риски на этапе автоматического анализа данных:

- 1) невыявление патогенного варианта
- 2) ложное выявление патогенного варианта
- 3) отсутствие выявленного варианта в базе

Риски на этапе финального (ручного) анализа отчета:

- 1) игнорирование метрик качества
- 2) нарушение правил использования

Учет выявленных рисков

I этап - пробоподготовка:

- 1) принципиально неустранимая неравномерность амплификации методом PCR, обусловленная анализируемой последовательностью ДНК
- 2) отклонения от методики

II этап - секвенирование:

- 1) случайные ошибки секвенатора
- 2) систематические ошибки секвенатора в районах гомополимеров

III этап - автоматический анализ данных:

- 1) невыявление патогенного варианта
- 2) ложное выявление патогенного варианта
- 3) отсутствие выявленного варианта в базе

IV этап - финальный анализ отчета врачом-генетиком:

- 1) игнорирование метрик качества
- 2) нарушение правил использования

Выводы

- 1) использование решений на базе технологий NGS в качестве единственного диагностического инструмента преждевременно
- 2) подобные решения являются хорошими кандидатами на роль дополнительных средств диагностики, но необходимо с осторожностью относиться к получаемым результатам