

# ПГТ-А МЕТОДОМ NGS

от чего зависит  
разрешение метода?

# Оглавление

- 2 Оглавление
- 3 ПГТ-А**
- 4 Качество исходного биоматериала**
- 5 Биологические причины плохого качества биоматериала
- 6 Технические аспекты эмбриологического этапа
- 7 Технические особенности NGS-платформы для ПГТ-А**
- 9 Равномерность полногеномной амплификации
- 10 Качество сиквенса
- 12 Количество прочтений
- 16 Особенности алгоритма анализа сырых данных
- 20 Влияние интерпретатора на результат**
- 23 Разрешающая способность метода**
- 23 Оценка от производителя платформы
- 24 Исследование геномной ДНК
- 25 Результаты ПГТ-А от партнеров со сбалансированными хромосомными перестройками
- 28 Сравнение с другими платформами
- 29 Клиническая значимость микроделеций и микродупликаций, выявленных в результате ПГТ-А**
- 31 В каком случае стоит искать CNV небольшого размера?
- 32 Резюме
- 33 Список литературы**

# ПГТ-А

**Преимплантационное генетическое тестирование на хромосомные аномалии (ПГТ-А)** — скрининговое исследование, которое используется в цикле вспомогательных репродуктивных технологий и позволяет отобрать для переноса в полость матки эмбрионы с нормальным хромосомным набором.

Ранее для проведения ПГТ-А использовались различные технологии, такие как

- > FISH (флуоресцентная гибридизация *in situ*),
- > qPCR (количественная ПЦР в реальном времени),
- > aCGH (сравнительная геномная гибридизация на микрочипах).

В настоящее время для ПГТ-А чаще всего применяется технология **NGS (секвенирование нового поколения)**, которая позволяет выявлять широкий спектр хромосомных нарушений: анеуплоидии, мозаицизм, некоторые типы полиплоидии, а также сегментарные нарушения.

В технологии NGS существуют платформы для ПГТ-А различных производителей, в связи с чем возникают вопросы о преимуществах и недостатках, а также о точности и разрешающей способности каждой из платформ. Часто при сравнении платформ говорят о количестве прочтений (ридов) на образец, и влиянии этого показателя на итоговый результат. Но важно учитывать, что разные платформы отличаются не только количеством прочтений, они отличаются множеством нюансов практически на всех этапах анализа. Поэтому некорректно сравнивать различные платформы по одному выбранному техническому показателю, вместо того, чтобы оценивать чувствительность и разрешающую способность каждой платформы.

В лаборатории First Genetics для проведения ПГТ-А методом NGS мы используем систему «F-Genetics» и набор «Репролайн» — зарегистрированные в России медицинские изделия, которые прошли валидацию в ходе клинических испытаний. Наша платформа создана в рамках OEM-партнерства с компанией ThermoFisher Scientific на основе прибора IonS5 и набора ReproSeq.

Накопленные данные демонстрируют, что для образцов биопсии трофобласта **возможно обнаружение сегментарных нарушений размером от 2,5 Мб** (миллионов пар оснований). Однако этот показатель нельзя считать гарантированной разрешающей способностью метода, так как разрешение зависит от качества исходного биоматериала.

Мы подготовили этот материал для того, чтобы подробно рассмотреть все факторы, влияющие на разрешение и достоверность результатов ПГТ-А, и рассказать о возможностях и ограничениях метода.

## Факторы, влияющие на качество полученного результата ПГТ-А



### Качество исходного биоматериала

Биологические причины — частичная деградация  
Технические аспекты эбриологического этапа



### Технические особенности NGS-платформы для ПГТ-А

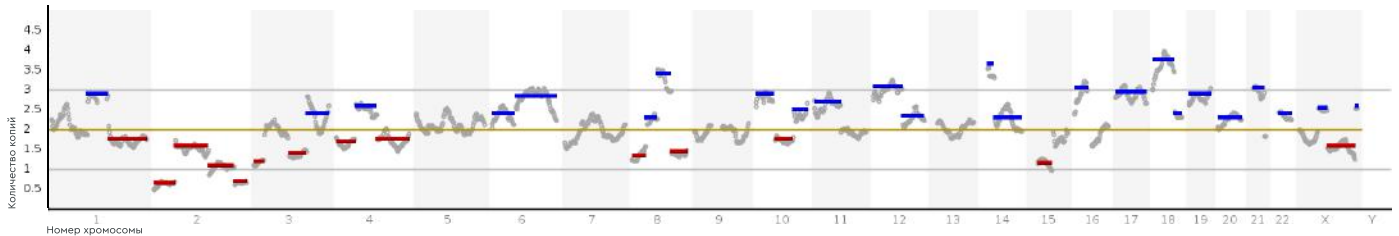
Равномерность полногеномной амплификации  
Количество прочтений  
Качество сиквенса  
Особенности алгоритма анализа сырых данных



### Квалификация эксперта, анализирующего хромосомные профили

# Качество исходного биоматериала

Качество исходного биоматериала напрямую влияет на аккуратность хромосомного профиля, полученного в результате исследования. В самом неблагоприятном случае, лаборатория наблюдает отсутствие амплификации, либо полную деградацию материала ([рисунок 1](#)), при которых интерпретация полученного результата невозможна и требуется повторная биопсия.



**Рисунок 1.** Хромосомный профиль, соответствующий образцу с полной деградацией исходного биоматериала.

Однако часто полученные хромосомные профили доступны для интерпретации, но могут выглядеть более «шумными» из-за частичной деградации материала.

В таком случае достоверность выявления сегментарных нарушений небольшого размера снижается.

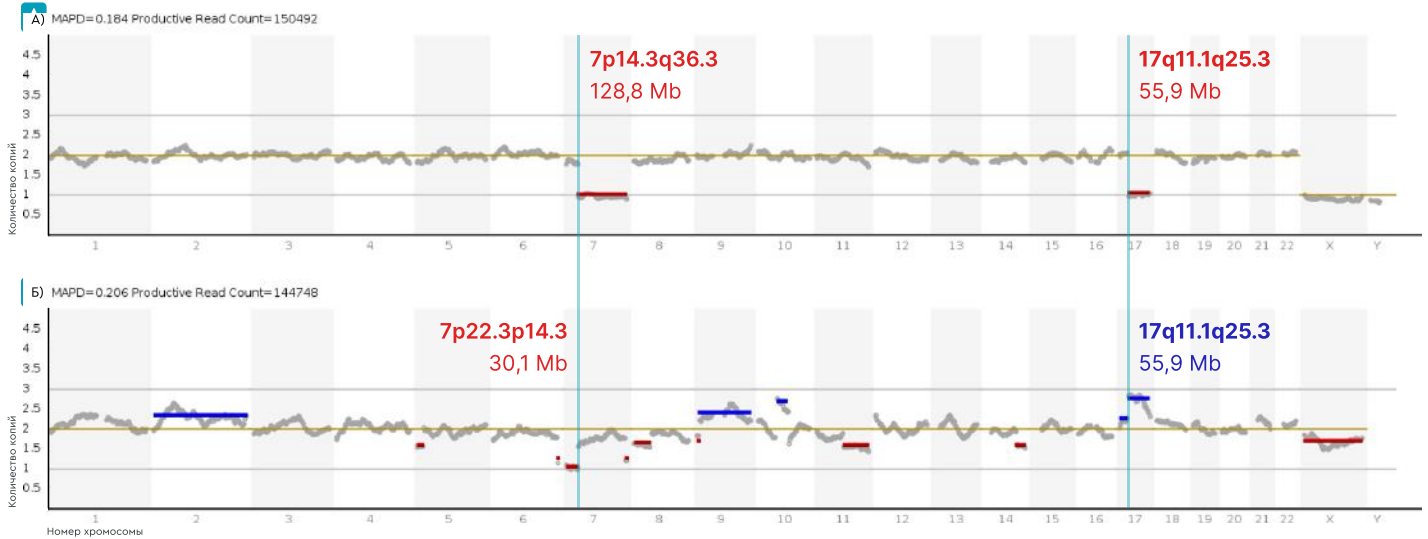
**Частичная деградация биоматериала может быть вызвана множеством причин** — как биологических (гибель клеток и лизис внутриклеточных структур, в том числе ДНК), так и технических (проблемы на этапе биопсии, нарушение условий транспортировки, многократное замораживание-оттаивание образца).

**Устранение технических причин деградации образца — совместная задача генетической и эмбриологической лабораторий.**

Мы внимательно следим за соблюдением температурного режима при транспортировке, а также выстраиваем процесс работы с образцом в лаборатории таким образом, чтобы его размораживание происходило лишь один раз — непосредственно в начале работы.

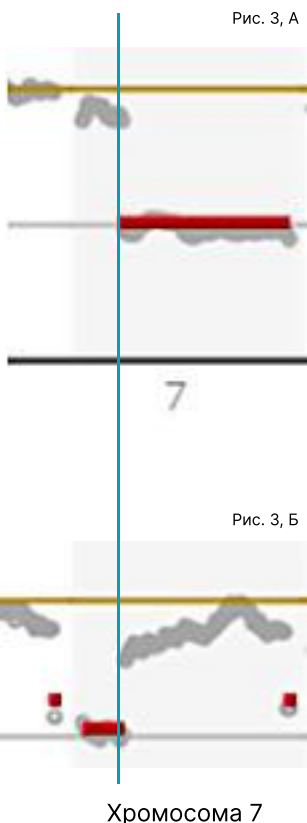
## Биологические причины плохого качества биоматериала

К сожалению, из-за биологических причин не удается полностью избежать наличия образцов плохого качества. На [рисунке 2](#) представлены хромосомные профили, полученные в нашей лаборатории для двух образцов от одного пациента. Биопсия проводилась в один день, одним и тем же эмбриологом, образцы транспортировались в одном штативе и в нашей лаборатории были исследованы в ходе одной и той же постановки. Тем не менее, **качество полученных данных для образцов значительно отличается.**



**Рисунок 2.** Различное качество образцов, полученных и проанализированных в одинаковых условиях. Биопсия трофэктодермы для обоих образцов проведена на 5 день развития эмбрионов. Первый график соответствует эмбриону с характеристикой развития 3ВВ, второй — 4ВВ.

Для первого образца не вызывает сомнения наличие двух сегментарных нарушений — на 7 и 17 хромосомах. В то время как данные, полученные для второго образца, свидетельствуют о частичной деградации биоматериала.



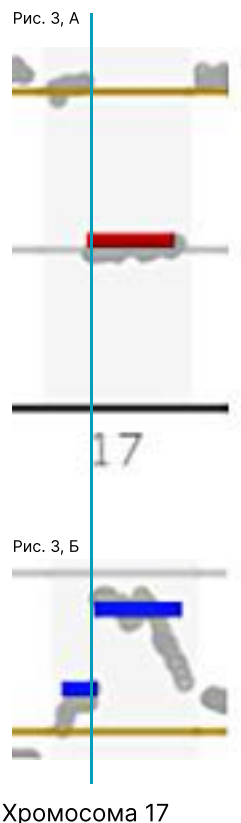
Эта ситуация представляет особый интерес из-за того, что отмеченные на рисунке сегменты на 7 и 17 хромосомах имеют одни и те же координаты в обоих образцах. Благодаря этому можно полагать, что во втором образце действительно присутствуют сегментарные делеция хромосомы 7 и дупликация хромосомы 17, а также то, что один из партнеров является носителем сбалансированной перестройки между этими хромосомами.

Подробнее о наследовании сегментарных нарушений от носителей сбалансированных транслокаций будет сказано в разделе «Результаты ПГТ-А от партнеров со сбалансированными хромосомными перестройками» на странице 24.

Таким образом, эмбрион, соответствующий второму графику, несет хромосомные аномалии и не может быть рекомендован к переносу.

Однако без наличия данных о первом образце, такой вывод сделать было бы невозможно, и для достоверного вывода о хромосомном статусе эмбриона потребовалась бы его повторная биопсия.

В данном случае технические аспекты работы как эмбриологической, так и генетической лаборатории, а также условия транспортировки для обоих образцов были максимально одинаковыми. **Поэтому наиболее вероятной причиной плохого качества второй образца следует считать именно биологические факторы.**



## Технические аспекты эмбриологического этапа

Из-за низкого качества биоматериала мы можем наблюдать:

- > «шумные» хромосомные профили (как на [рисунке 2,Б](#) на стр.5),
- > полную деградацию ДНК ([рисунок 1](#) на стр. 4),
- > отсутствие амплификации.

В каждом конкретном случае крайне сложно определить, что именно стало причиной такого результата — биологические факторы или техническая ошибка на каком-либо этапе исследования.

Однако мы видим **отличия** в частоте таких событий **между различными клиниками**.

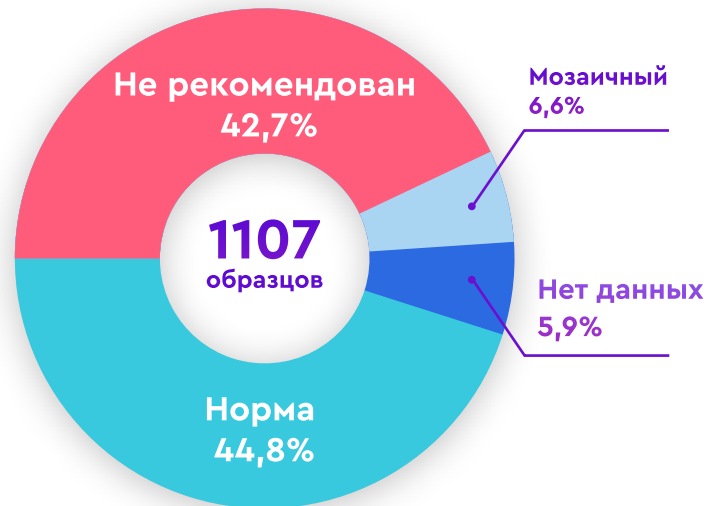
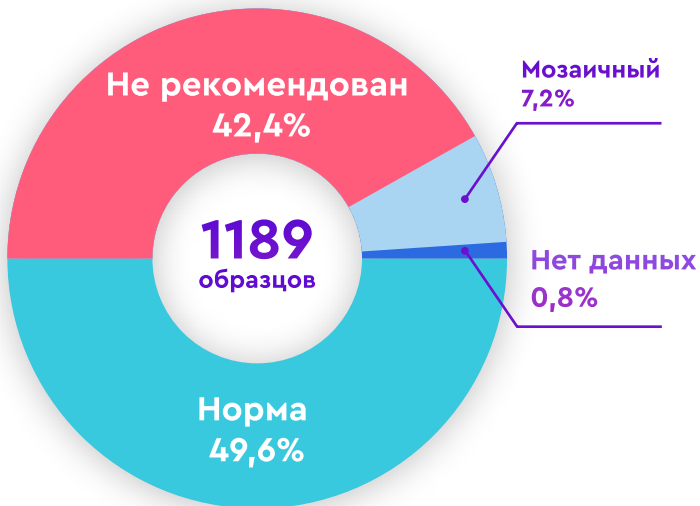
**для 2,5 %** среди всех образцов мы не получаем **достоверный результат и рекомендуем проведение повторной биопсии**.

При этом для отдельных клиник этот показатель составляет менее 1 %, в то время как для других клиник может составлять заметную долю от всех образцов ([рисунок 3](#)). Из-за таких различий можно предполагать, что часто **решающую роль в получении материала плохого качества играют технические аспекты именно эмбриологического этапа**.

**Рисунок 3.** Различие между клиниками по доле образцов, для которых не получено достоверных данных (отсутствие амплификации, либо деградация биоматериала). В группу «Мозаичный» включены все образцы с одним мозаичным событием, которые были рекомендованы к переносу с информированного согласия пациента.

**А** Результаты одной из крупных клиник с минимальной долей неинформативных образцов.

**Б** Суммарные данные для 5 клиник с максимальной долей неинформативных образцов.



Мы понимаем, что генетическая лаборатория не обладает необходимыми знаниями и опытом, чтобы рекомендовать, как именно проводить биопсию и выстраивать внутренние процессы в эмбриологической лаборатории.

Однако большой объем данных позволяет нам выявлять систематические проблемы и рекомендовать эмбриологическим лабораториям-партнерам обратить внимание на технические аспекты своей работы.

# Технические особенности NGS-платформы для ПГТ-А

В [таблице 1](#) представлено сравнение трех наиболее часто используемых NGS-платформ для проведения ПГТ-А:

- > **VeriSeq** от компании Illumina,
- > **ReproSeq** от компании ThermoFisher Scientific (используется в лаборатории First Genetics)
- > **PG-Seq** от компании Perkin Elmer.

Более подробно каждый параметр обсуждается в тексте ниже.

**Сравнение различных платформ построено на документации производителей платформ, доступной в открытых источниках.**

	<b>VeriSeq (Illumina)</b>	<b>ReproSeq (ThermoFisher)</b>	<b>PG-Seq (Perkin Elmer)</b>
Первичное количество ридов на образец	оптимально: 1 000 000 допустимо: 700 000	от 125 000	нет данных
Количество ридов после фильтрации	оптимально: 500 000 допустимо: 250 000	от 100 000	<b>250 000 – 500 000*</b> <small>Увеличение количества прочтений на образец достигается за счет уменьшения количества одновременно секвенируемых образцов, что приводит к увеличению стоимости анализа одного образца. Увеличение количества прочтений в приведенном диапазоне улучшает разрешающую способность метода для данной платформы.</small>
Длина ридов	от 36 п. н.	в среднем 115 п. н.	<b>от 36 до 150 п. н.</b> <small>В зависимости от набора.</small>
Алгоритм анализа данных	настраивается только визуализация данных	настраивается визуализация данных гибкая настройка чувствительности и специфичности возможность использования собственной базовой линии	небольшая настройка чувствительности и специфичности
Разрешающая способность	<b>20 Мб</b> <small>Разрешение для каждой платформы указано в соответствии с документацией производителя. Независимые оценки разрешающей способности платформы VeriSeq указывают на возможность детекции более мелких сегментарных нарушений (Cuman et al., 2018), однако производитель не гарантирует подобное разрешение метода.</small>	<b>8 Мб</b> от 2 Мб при соответствующих настройках алгоритма	<b>5–10 Мб</b> <small>В зависимости от количества прочтений на образец.</small>
Автоматическое выявление сегментарных событий	Нет	Да	Да

**Таблица 1.** Сравнение трех наиболее популярных в настоящее время NGS-платформ для проведения ПГТ-А.

Общая для всех трех платформ схема проведения исследования ПГТ-А представлена на [рисунке 4](#).



Рисунок 4. Общая схема проведения ПГТ-А.

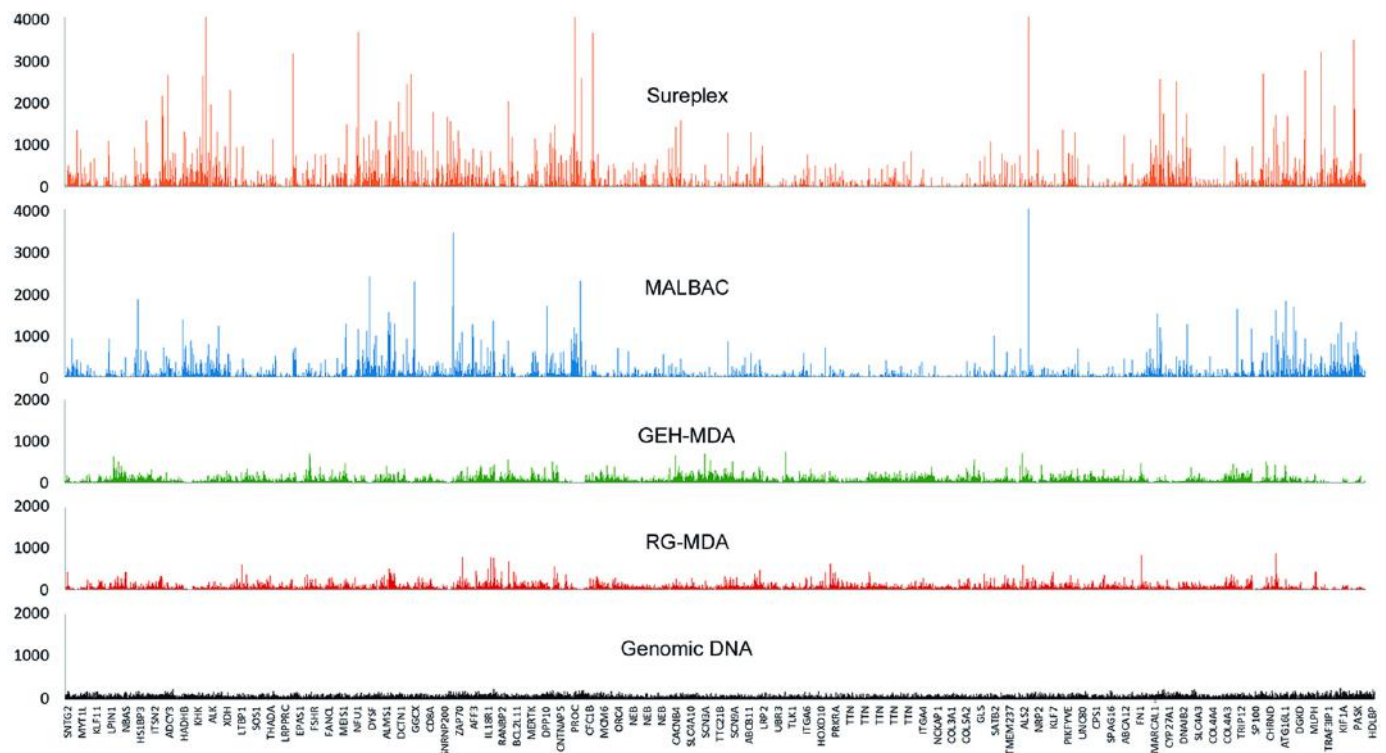
Далее каждый этап будет подробно рассмотрен с точки зрения качества итогового результата.



## Равномерность полногеномной амплификации

При проведении ПГТ-А в качестве начального образца используется всего **5-10 клеток трофэктодермы**. Для того чтобы проанализировать такой образец, на первом этапе необходимо провести **полногеномную амплификацию (WGA, whole genome amplification)** — то есть многократно увеличить количество ДНК за счет синтеза коротких фрагментов — копий исходной ДНК. Существует множество подходов, которые отличаются последовательностью молекулярно-биологических реакций и равномерностью распределения (покрытия) полученных копий по всему геному.

**Неравномерность покрытия возникает из-за того, что одни участки копируются легче и быстрее, а другие медленнее.** В результате продукт реакции содержит больше копий одних фрагментов ДНК, и меньше копий других. И такое неравномерное распределение отражает не исходное строение генома исследуемых клеток, а погрешность полногеномной амплификации. В литературе можно встретить сравнения различных подходов WGA, которые демонстрируют, насколько сильно может отличаться итоговая равномерность покрытия (рисунок 5, Del Rey et al., 2018).



**Рисунок 5.** Из статьи (Del Rey et al., 2018). Графическое представление равномерности покрытия хромосомы 2 для различных систем WGA и неамплифицированной геномной ДНК. Высота столбца отражает количество итоговых фрагментов ДНК, относящихся к каждому участку хромосомы 2.

- Платформы **VeriSeq (Illumina)** и **ReproSeq (ThermoFisher)** используют **одинаковый подход** для полногеномной амплификации, поэтому различия в равномерности покрытия на этапе WGA для этих платформ незначительны.
- Платформа **PG-Seq (Perkin Elmer)** использует **иной подход**, который демонстрирует **большую равномерность** покрытия (Deleye et al., 2017).

Тем не менее **для каждой платформы последующие шаги** подготовки библиотек для секвенирования, само секвенирование а также анализ данных, **оптимизированы для выбранного протокола полногеномной амплификации.**

**Поэтому сравнение отдельных подходов WGA скорее несет теоретический, чем практический интерес.**

## Качество сиквенса

После того, как получен продукт многократного копирования ДНК исходных 5-10 клеток, происходит подготовка образца для секвенирования — **создание библиотеки ДНК**.

К фрагментам ДНК присоединяются специальные молекулярные метки — баркоды, уникальные для каждого образца в постановке.

Далее отбираются фрагменты оптимального размера, которые дополнительно амплифицируются и закрепляются на специальной матрице для последующего секвенирования.

**Все эти процессы, а также сам процесс секвенирования проходят по-разному в зависимости от платформы.**

В результате секвенирования получается большое количество прочтений (ридов) — буквенных записей последовательностей нуклеотидов в исходных фрагментах ДНК.

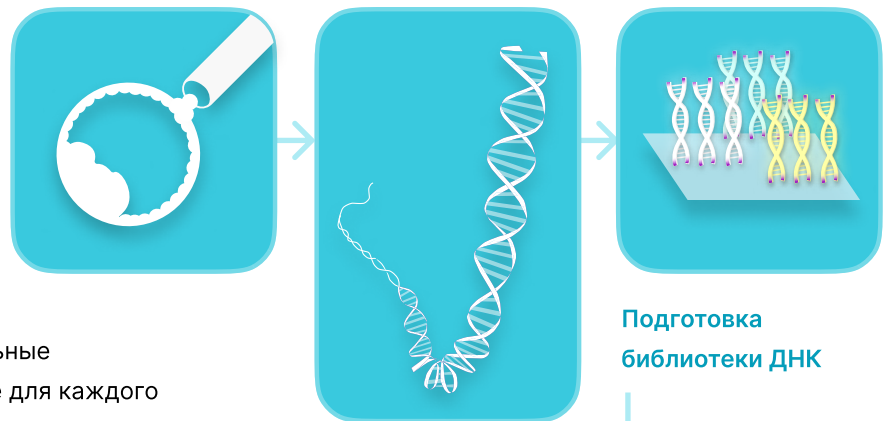
Из всего полученного объема данных **только часть прочтений используется для последующего анализа**. В то время как другая часть исключается из-за низкого качества данных — в том случае, если сигнал, полученный при считывании нуклеотидной последовательности, не прошел контроль качества.

Кроме того, из анализа **исключаются дублицированные** риды — такие прочтения, которые в точности повторяют друг друга. Дублицированные риды появляются из-за копирования фрагментов ДНК в пробирке в процессе постановки и не отражают биологической сути исследуемого образца.

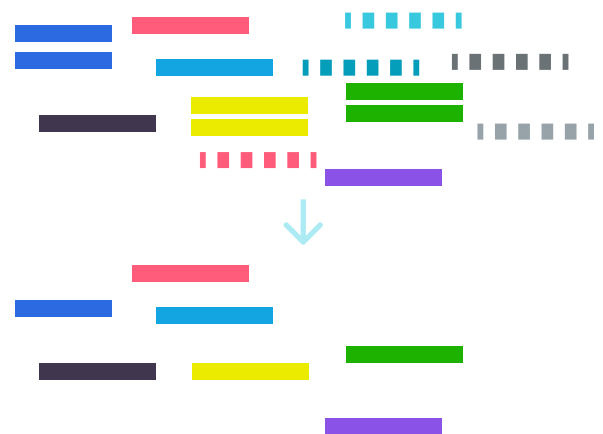
Помимо доли отфильтрованных ридов, второй важный для сравнения параметр — **длина получаемых прочтений**. В дальнейшем при работе алгоритма происходит выравнивание (картирование) полученных ридов — то есть **определяется расположение изначально прочитанного фрагмента ДНК в геноме**.

**Чем длиннее прочтение, тем более однозначно и достоверно определяется его положение.**

Это связано с тем, что короткие прочтения могут одновременно «подходить» по последовательности нескольким различным участкам генома. В таком случае однозначно определить их положение невозможно, и прочтения не используются в итоговом анализе.



Подготовка библиотеки ДНК



Фильтрация ридов



Выравнивание (картирование) ридов

На [рисунке 6](#) изображено, как влияет длина прочтений на возможность их достоверного выравнивания.

По горизонтали отмечены геномные координаты, высота столбца отражает однозначность выравнивания и рассчитывается как  $1/N$ , где  $N$  – количество мест в геноме, к которым подходит такое прочтение.



**Рисунок 6.** Изображение с сайта <https://blog.gene-test.com/sequence-mappability-alignability/>. Влияние длины прочтений на итоговое покрытие генома после выравнивания. На рисунке отражен небольшой фрагмент генома. Горизонтальная ось соответствует координате в геноме, высота столбцов иллюстрирует то, насколько уникально соответствие прочтения указанной длины каждому участку генома. Длины ридов в 36 п. н. часто недостаточно, для того чтобы однозначно определить их положение в геноме, поэтому большие участки генома остаются слабо покрыты после выравнивания. Для ридов в 100 п. н. определение их положения в геноме происходит более достоверно, поэтому непокрытых после выравнивания участков остается значительно меньше.

Таким образом, для всех областей, где зеленый столбец не достигает максимума, однозначно выявить прочтения, соответствующие этим областям невозможно.

Это влияет на последующий алгоритм определения количества копий каждого участка и снижает достоверность полученных данных для таких областей.

Из рисунка видно, что с увеличением длины прочтений таких «проблемных» областей в геноме остается всё меньше.

**Различная длина прочтений в различных платформах приводит к тому, что доля исходно прочитанных ридов, используемая для последующего анализа может значительно отличаться.**

Таким образом при сравнении платформ важно учитывать не первичное количество прочтений, получаемое на один образец, а то количество ридов, которое используется для анализа после фильтрации и выравнивания.

## Количество прочтений

Как уже было сказано, различные платформы помимо исходного количества ридов на образец отличаются друг от друга практически на всех этапах проведения ПГТ-А. Поэтому, для того чтобы отдельно оценить влияние количества прочтений на разрешающую способность ПГТ-А, необходимо **проводить сравнение результатов, полученных на основе различного количества ридов внутри одной и той же платформы.**

Для того чтобы определить, каким образом разрешающая способность нашей платформы зависит от количества прочтений на образец, мы исследовали образцы с известными сегментарными нарушениями.

Подробнее о таких образцах будет сказано в разделе «Разрешающая способность метода» на странице 22.

Было отсеквенировано 8 образцов таким образом, чтобы **количество прочтений на каждый образец составило не менее 300 000 ридов** после фильтрации и выравнивания. Семь образцов были представлены геномной ДНК и еще один – образцом биопсии трофэктодермы, поступившей для проведения ПГТ-А от партнеров с известным носительством сбалансированной транслокации. Далее для всех полученных данных с помощью биоинформатического инструмента SAMtools (Li et al., 2009) были получены «урезанные» варианты файлов с сырыми данными, которые содержали **200 000 и 100 000 прочтений** после фильтрации и выравнивания для каждого образца.



Поскольку файлы с меньшим количеством прочтений были получены из исходных данных секвенирования, а не в результате повторных экспериментов, такой подход позволил исключить влияние различий между постановками и любых других посторонних факторов помимо количества прочтений на качество полученного результата. Анализ данных проводился с использованием различных настроек чувствительности и специфичности алгоритма. **В результате из 15 ожидаемых сегментарных нарушений различного размера было детектировано 14 событий (таблица 2).**

Образец	Сегмент	Размер, Мб	Кол-во прочтений в исходном файле данных	Детекция в исходных данных	при 100 000 прочтений	при 200 000 прочтений
GenomeCNV7	del(3p26.3p26.2)	3,2	403 000	+	+	+
PGT	del(22q11.1q11.21)	3,7	514 000	+	+	+
GenomeCNV6	dup(1p36.33p36.32)	4	294 000	-	-	-
GenomeCNV6	del(17p13.3p13.2)	5,6	294 000	+	+	+
GenomeCNV4	del(2q37.1q37.3)	8,2	450 000	+	+	+
GenomeCNV8	del(10q26.13q26.3)	11,6	349 000	+	+	+
GenomeCNV11	dup(11p15.5p15.3)	11,6	312 000	+	+	+
GenomeCNV4	dup(22q13.1q13.33)	11,7	450 000	+	+	+
GenomeCNV11	del(5p15.33p15.1)	16,9	312 000	+	+	+
GenomeCNV8	dup(5q34q35.3)	17,7	349 000	+	+	+
PGT	del(11q23.3q25)	18,2	514 000	+	+	+
GenomeCNV9	del(13q21.33q32.1)	26	300 000	+	+	+
GenomeCNV5	del(4p16.3p15.2)	26	414 000	+	+	+
GenomeCNV5	dup(1q41q44)	28,2	414 000	+	+	+
GenomeCNV7	dup(18q12.1q23)	48	403 000	+	+	+

**Таблица 2.** Результаты детекции событий различного размера при различном количестве прочтений. **Красным цветом** обозначены размеры делеций, **синим** – дупликаций. Сегментарная дупликация, детектировать которую не удалось ни в одном из вариантов анализа вне зависимости от количества прочтений на образец, отмечена красной заливкой. Все остальные сегментарные нарушения были детектированы при анализе начальных сырых данных, а также при анализе «урезанных» файлов данных с 200 000 и 100 000 прочтений на образец.

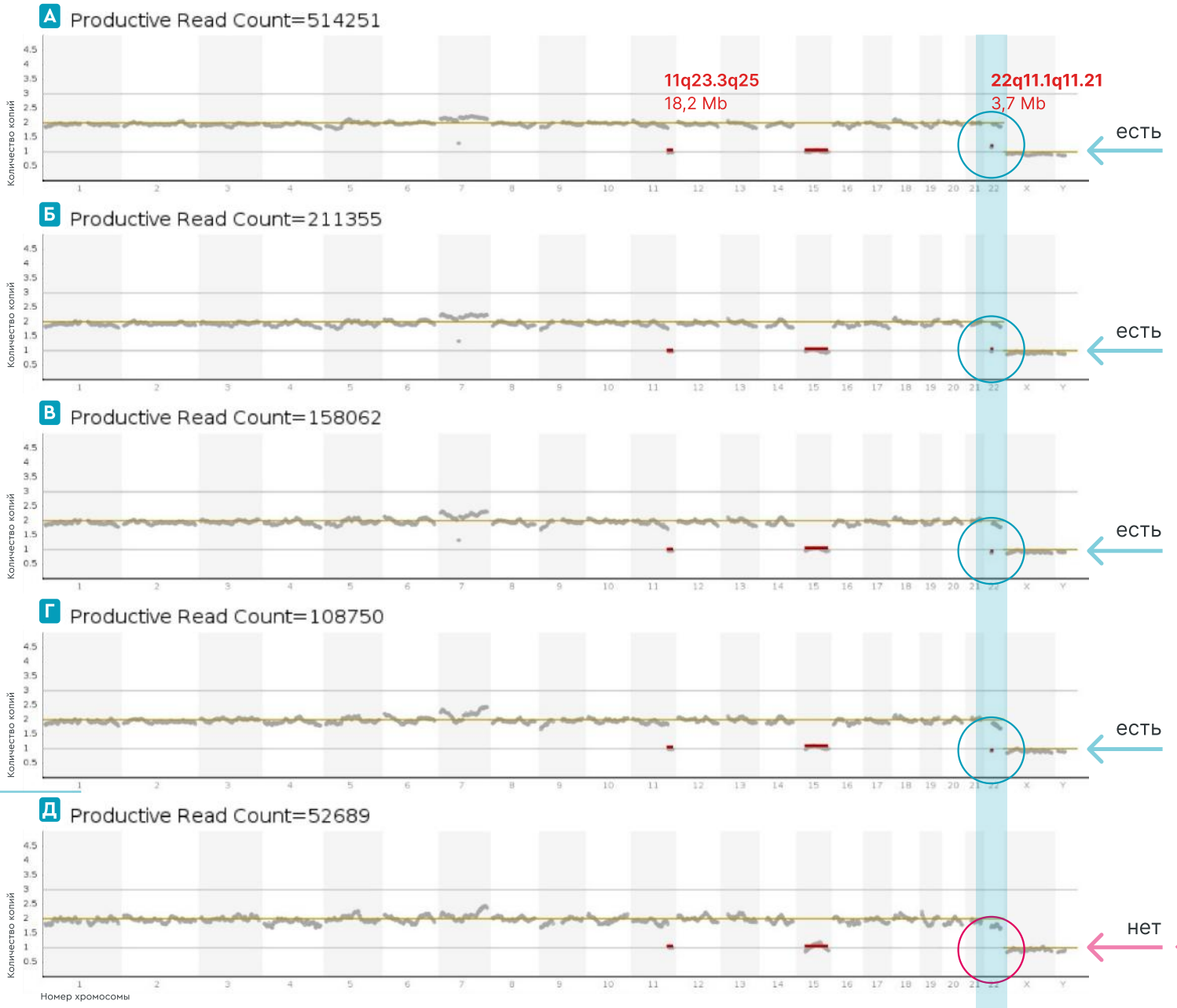
Одно событие размером 4 Мб не удалось детектировать вне зависимости от количества прочтений на образец и использованного алгоритма анализа данных. Подробнее о детекции сегментарных нарушений в образцах геномной ДНК рассказано в разделе «Исследование геномной ДНК» на странице 24.

Крайне важно, что для всех ожидаемых сегментарных событий их детекция происходила одинаково вне зависимости от количества ридов на образец в диапазоне от 100 до 300 тысяч.

Пример полученных хромосомных профилей для файлов с разным количеством прочтений представлен на [рисунке 7](#).

Несмотря на небольшие различия в визуализации, во всех вариантах от 100 до 500 тысяч ридов на образец были обнаружены ожидаемые сегментарные нарушения.

В то же время 50 тысяч ридов на образец позволили детектировать только более крупные события, в то время как делеция размером 3,7 Мб выявлена не была.



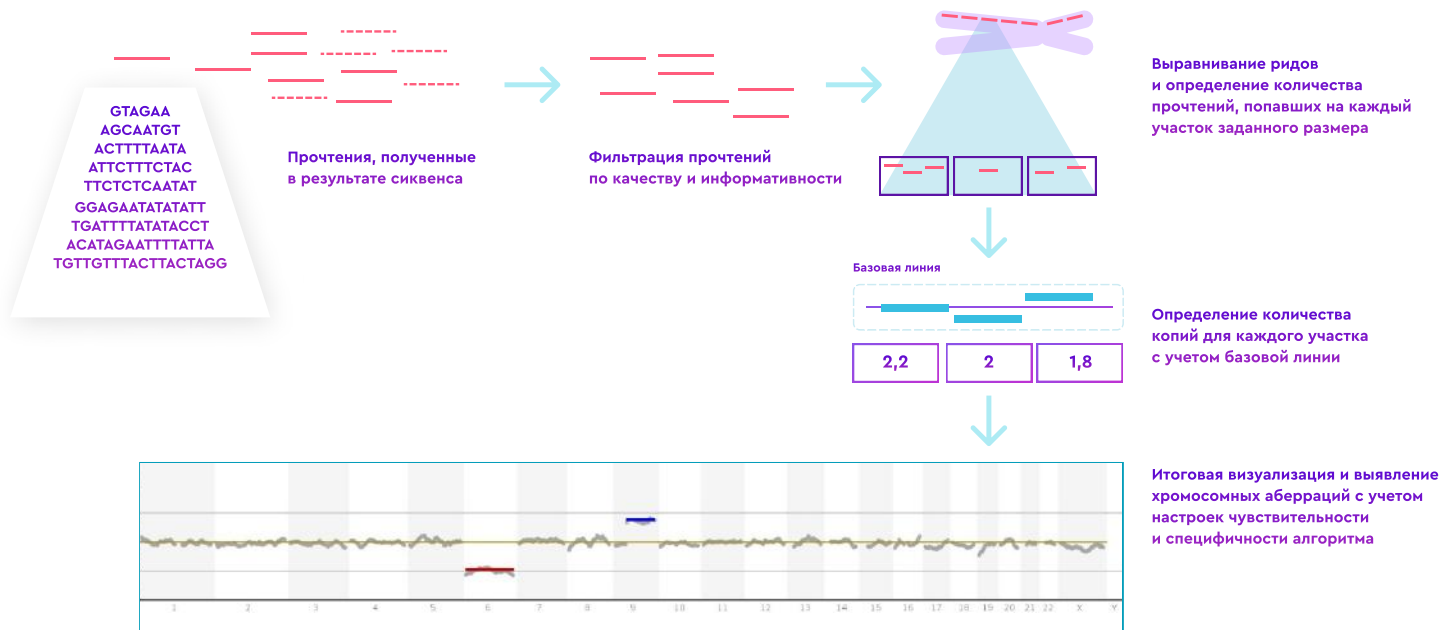
**Рисунок 7.** Результаты биоинформатического уменьшения количества ридов, проведенного для образца биопсии трофанодермы, от пациентов с носительством транслокации 46,XX,t(11;22)(q24;q12) (образец PGT в таблице 2). На рисунке указано итоговое количество прочтений на образец после фильтрации и выравнивания. Несмотря на небольшие различия в визуализации, во всех вариантах от 100 до 500 тысяч ридов на образец были детектированы сегментарные делеции размером 18,2 Мб и 3,7 Мб, соответствующие регионам, участвующим в родительской хромосомной перестройке. В то же время 50 тысяч ридов на образец оказалось достаточно для детекции сегментарного нарушения в 18,2 Мб, но недостаточно для детекции сегментарного нарушения в 3,7 Мб на хромосоме 22.

Исследование демонстрирует, что **100 000 прочтений** на образец после фильтрации и выравнивания, рекомендованное производителем платформы, **оптимально** для последующего анализа и реализации разрешающей способности метода менее 5 Мб.

**Увеличение количества прочтений** избыточно и **не приводит к улучшению качества результата**, при условии хорошего качества биоматериала.

## Особенности алгоритма анализа сырых данных

Схема работы алгоритма поиска хромосомных нарушений при проведении ПГТ-А методом NGS изображена на [рисунке 8](#).



**Рисунок 8.** Схема работы алгоритма, используемого для поиска хромосомных нарушений при проведении ПГТ-А методом NGS.

### Алгоритм состоит из следующих основных шагов:

1

**Фильтрация прочтений** — на этом этапе из всех полученных данных отсеиваются риды плохого качества и дублицированные риды. Подробнее об этом этапе рассказано в разделе «Качество сиквенса» на странице 10.

2

**Выравнивание (картирование) ридов** — то есть определение участка в геноме, которому соответствует прочитанный фрагмент ДНК. На данном этапе дополнительно отфильтровываются все прочтения, для которых не удастся однозначно определить их положение в геноме.

3

Весь геном разбивается на отрезки заданного размера, после чего происходит подсчет количества прочтений, попавших на каждый участок. Разные платформы используют различный размер таких отрезков.

Так, алгоритм BlueFuse (Illumina) использует отрезки длиной 1 Мб и не позволяет изменять их размер.

В то время как программное обеспечение IonReporter (ThermoFisher) позволяет изменять размер отрезков от 0,5 Мб до 2 Мб и более.

**Чем меньше размер отрезков** используется в алгоритме при анализе данных, тем более короткие сегментарные события могут быть детектированы и **тем выше чувствительность метода**. Увеличение чувствительности позволяет детектировать сегментарные события меньшего размера, но в то же время приводит к увеличению количества ложноположительных результатов – избыточной детекции артефактных сегментарных событий.



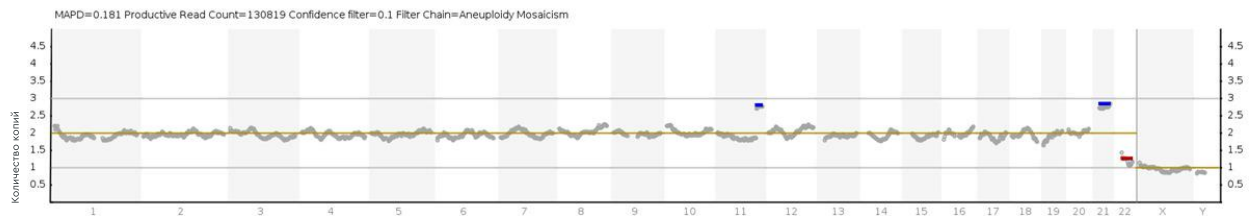
На **рисунке 9** показано, как **различные настройки алгоритма влияют на итоговую визуализацию** одних и тех же исходных данных.

Данные хромосомные профили представлены как крайние варианты настроек и иллюстрируют важность понимания работы алгоритма для корректной последующей интерпретации графиков экспертом.

Алгоритмы Б и В представлены для иллюстрации значимости настроек алгоритма и не используются в нашей лаборатории.

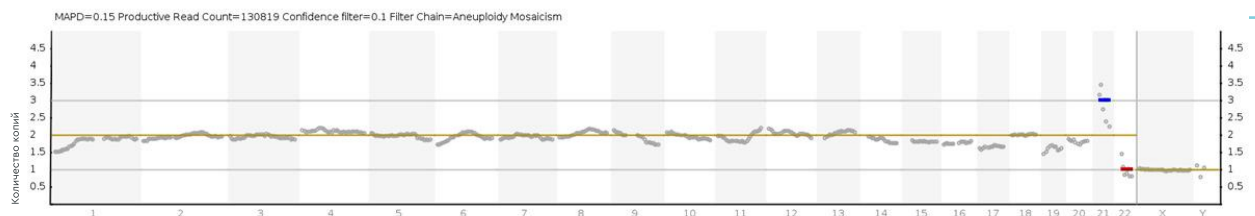
**А** Оптимальный алгоритм:

позволяет выявить сегментарную дупликацию размером 20 Мб на хромосоме 11.



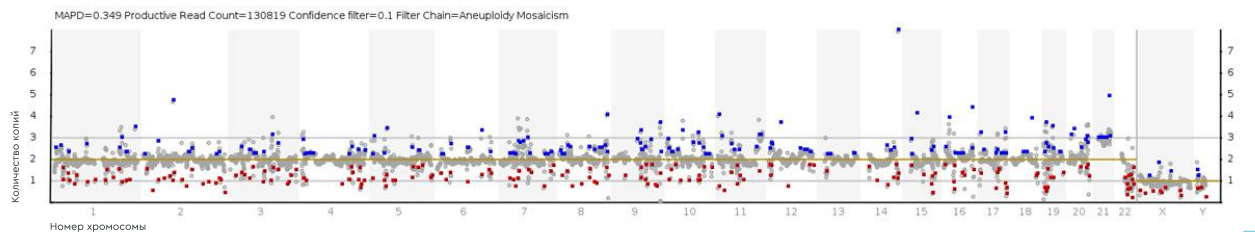
**Б** Слишком нечувствительный алгоритм – не подходит для использования в работе:

сглаживает дупликацию и делает её незаметной для интерпретатора.



**В** Излишне чувствительный алгоритм – не подходит для использования в работе:

слишком сильно разбивает профиль на отдельные события и не позволяет эксперту проводить адекватную интерпретацию.



Номер хромосомы

**Рисунок 9.** Влияние параметров алгоритма обработки данных на итоговую визуализацию, интерпретируемую экспертом. Алгоритмы с различными настройками чувствительности применены к одному и тому же файлу сырых данных.

4

Для каждого полученного **отрезка** определяется **количество его копий в образце**.

При этом учитывается так называемая **базовая линия**. Как уже было сказано в разделе «Равномерность полногеномной амплификации» на странице 9, весь **геном покрыт прочтениями неравномерно** из-за того, что различные участки нарабатываются в процессе полногеномной амплификации с разной эффективностью. Такая **неравномерность** связана с особенностями последовательности ДНК и **повторяется от образца к образцу**. Это **систематическое искажение** и **учитывает базовая линия**.

То есть при определении количества копий каждого участка, сравнивается количество ридов, попавших на данный участок, с тем количеством прочтений, которое в среднем приходится на этот участок в других образцах с нормальным кариотипом.

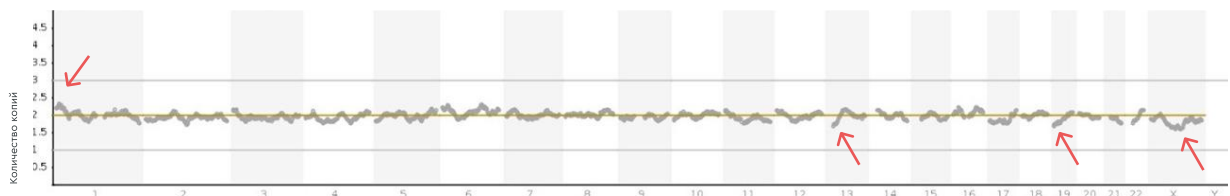
Неравномерность полногеномной амплификации **зависит в том числе и от оборудования**, на котором проводят ПГТ-А, и поэтому может слабо варьироваться от одной лаборатории к другой.

Программное обеспечение IonReporter (ThermoFisher), в отличие от других платформ, позволяет создавать свою собственную базовую линию и таким образом корректировать систематические ошибки, повторяющиеся от образца к образцу.

В нашей лаборатории используется собственная базовая линия, для построения которой использовались образцы с известным исходом переноса эмбриона – рождением здорового ребенка.

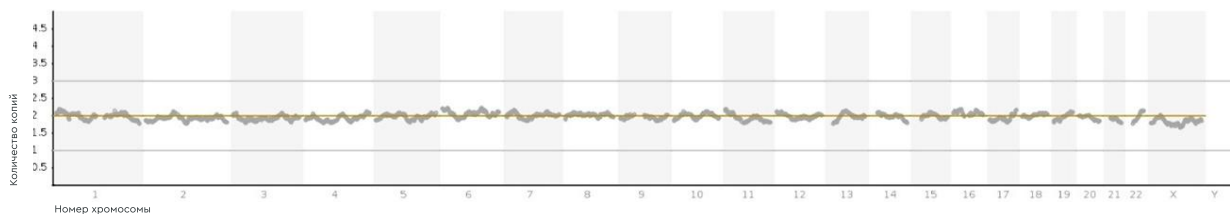
На [рисунке 10](#) показано, как влияет использование собственной базовой линии на полученный хромосомный профиль. Для большинства образцов такая корректировка оказывается незначительной и лишь слегка выравнивает полученный профиль, однако в некоторых случаях наличие систематической ошибки может приводить к ложноположительному выявлению хромосомных нарушений.

**A** Стандартная базовая линия



Использование собственной базовой линии выправляет небольшие отклонения (показаны стрелками), которые регулярно повторяются в различных образцах и являются артефактом.

**B** Собственная базовая линия:



**Рисунок 10.** Влияние собственной базовой линии на конечный хромосомный профиль. Результаты анализа данных с помощью алгоритма с использованием: А) стандартной базовой линии; Б) собственной базовой линии.

5

На заключительном этапе происходит **визуализация данных** – построение хромосомного профиля (каждому проанализированному отрезку соответствует одна точка) и определение хромосомных aberrаций.

В отличие от программы BlueFuse (Illumina), которая в автоматическом режиме выявляет и отмечает только нарушения, затрагивающие целую хромосому, IonReporter (ThermoFisher) также автоматически обнаруживает сегментарные нарушения и рассчитывает для них **статистические характеристики достоверности**.

Автоматическое выявление сегментарных событий происходит в соответствии с настройками чувствительности и специфичности алгоритма и **уменьшает риск незамеченных событий** из-за ошибки интерпретатора.

Программное обеспечение IonReporter, используемое в нашей лаборатории, имеет ряд важных преимуществ:

- позволяет использовать гибкие настройки чувствительности и специфичности алгоритма,
- выявляет и отмечает сегментарные события в автоматическом режиме
- рассчитывает статистическую достоверность для таких событий.

Это значительно уменьшает риск ошибки интерпретатора и позволяет уменьшить субъективность при анализе данных.

Еще один важный момент — возможность использовать собственную базовую линию, которая корректирует систематические ошибки, связанные с индивидуальными особенностями работы оборудования в каждой лаборатории.

# Влияние интерпретатора на результат

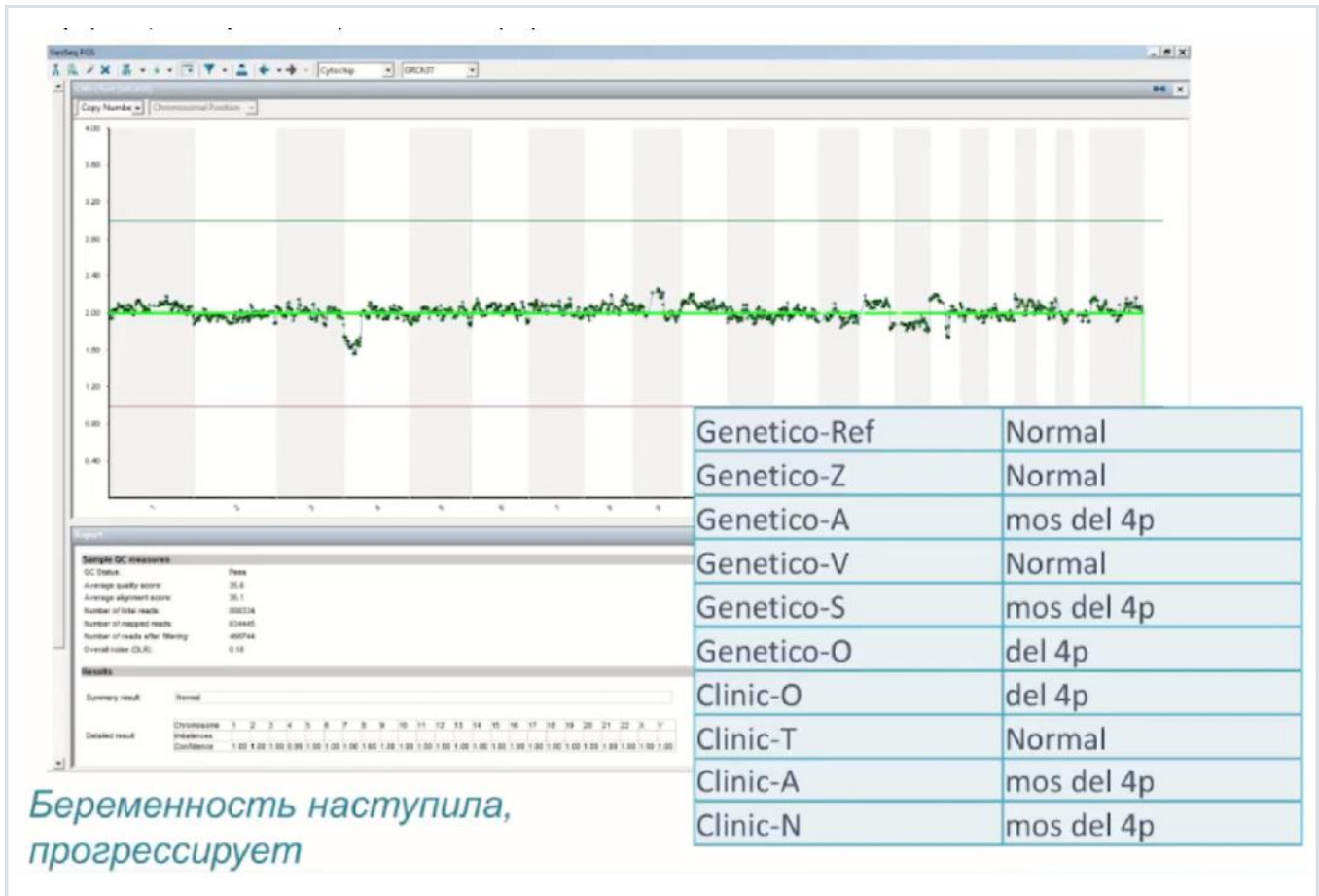
Как уже было показано выше, на вид получаемого в результате ПГТ-А хромосомного профиля может влиять множество факторов, связанных как с самим материалом, так и с особенностями проведения исследования и анализа данных.

Поэтому **ключевую роль в получении достоверного заключения играет эксперт**, интерпретирующий полученные данные. Именно эксперт принимает решение о том, какие события считать истинными, формирует запись итогового кариотипа и принимает решение о необходимости проведения повторного анализа при получении данных неудовлетворительного качества.

Таким образом в заключение неизбежно вносится **субъективность**, и возникает проблема зависимости полученного результата от конкретного человека, работающего с хромосомными профилями. Влияние человеческого фактора в интерпретации результатов ПГТ-А уже не раз обсуждалось. Так на [рисунке 11](#) представлен слайд из презентации лаборатории «Генетико», иллюстрирующий, что в зависимости от мнения эксперта, судьба эмбриона может быть прямо противоположной.

В то время, как одни эксперты рекомендуют эмбрион к переносу как эуплоидный, другие выявляют такой кариотип, который не допускает эмбриотрансфер.

**Рисунок 11.** В зависимости от эксперта, интерпретирующего полученные данные, судьба эмбриона может быть прямо противоположной. Слайд из презентации лаборатории «Генетико» на тему «Клиническая интерпретация результатов преимплантационного тестирования на хромосомные аномалии», доступном в YouTube на канале Российской Ассоциации Репродукции Человека.



Крайне важно, чтобы интерпретатор не просто смотрел на полученную визуализацию хромосомного профиля, а глубоко понимал весь процесс работы с образцом и детали работы алгоритма анализа данных. Только такой подход может позволить отличать истинные события от ложноположительного результата.

Ранее на [рисунке 9](#) (стр. 17) было показано что применение различных настроек алгоритма анализа данных может значительно влиять на получаемый хромосомный профиль.

Использование оптимальных настроек алгоритма и выбор адекватной чувствительности для каждого конкретного образца крайне важны для корректной интерпретации экспертом полученного хромосомного профиля.

Лаборатория First Genetics является партнером компании ThermoFisher, благодаря чему наши сотрудники прошли **подробные тренинги** по работе с данными и их интерпретации, глубоко **понимают алгоритм анализа данных и возможности его настроек**. Это позволяет нам более гибко анализировать каждый образец и улучшает качество интерпретации получаемых хромосомных профилей.

**Уменьшить субъективность** интерпретации позволяют некоторые преимущества программного обеспечения IonReporter (TheromFisher). Среди них — автоматическое выявление сегментарных нарушений и определение статистических характеристик достоверности, на которые ориентируется эксперт помимо общего визуального впечатления от полученного хромосомного профиля.

О преимуществах программного обеспечения IonReporter (TheromFisher) было рассказано ранее в разделе «Особенности алгоритма анализа сырых данных» на страницах 16-19.

Но помимо этого, для уменьшения зависимости результата от мнения конкретного сотрудника лаборатории, а также для уменьшения риска человеческой ошибки,

- каждое заключение в нашей лаборатории после оформления **проходит проверку двумя сотрудниками**, которые интерпретируют данные независимо друг от друга.
- **При расхождении интерпретации** образец выносится на обсуждение всеми сотрудниками лаборатории, в результате которого **принимается коллегиальное решение**.

- Наконец, в настоящее время мы занимаемся разработкой **инструмента интерпретации получаемых данных на основе методов машинного обучения**. Ожидается, что за счет обобщения и усреднения данных о вынесенных ранее вердиктах такой инструмент будет подсказывать, какое решение уже было принято в подобных случаях. Таким образом можно добиться более стабильной интерпретации сложных случаев и уменьшить субъективность в каждом конкретном случае. Кроме того, первичное автоматическое формирование заключения позволит дополнительно снизить риск ложного результата из-за возможной ошибки при записи кариотипа интерпретатором.

Таким образом в нашей лаборатории для уменьшения влияния человеческого фактора на качество интерпретации данных предприняты следующие шаги:

1

Сотрудники лаборатории обучаются непосредственно в компании ThermoFisher (производитель платформы ПГТ-А).

2

Используемые алгоритмы автоматически выявляют не только события, затрагивающие целую хромосому, но и сегментарные события.

3

При интерпретации кроме общего вида хромосомного профиля учитываются статистические показатели достоверности для каждого события, рассчитанные алгоритмом.

4

Мы проводим разработку инструмента для интерпретации данных, построенного на методах машинного обучения. Это позволит внести стабильность в процесс интерпретации и уменьшить вероятность человеческой ошибки при формировании итогового кариотипа.

5

Каждый образец проходит проверку как минимум двумя сотрудниками лаборатории, которые интерпретируют данные независимо друг от друга и при расхождении мнений выносят коллегиальное решение.

# Разрешающая способность метода

Теперь, понимая все факторы, которые влияют на полученные данные, попробуем определить разрешающую способность метода.

Как правило, хромосомные нарушения в эмбриональных клетках — это случайные события, которые не подчиняются какой-то конкретной закономерности. Кроме того, в ходе проведения исследования клетки полностью разрушаются и не подлежат повторному независимому анализу. Из-за этих факторов для большинства образцов, поступающих на ПГТ-А, не могут быть получены данные об «истинном кариотипе» исследованных клеток, который был бы подтвержден независимыми методами. Учитывая уже продемонстрированную разнородность исходного биоматериала по качеству, а также проблему мозаицизма в эмбриональных клетках, можно представить, насколько нетривиальной становится задача определения разрешающей способности метода для ПГТ-А.

## Оценка от производителя платформы

Самый простой для лаборатории способ определить разрешающую способность метода — обратиться к характеристикам, заявленным производителем платформы. На [рисунке 12](#) представлены отрывки из документации к трем наиболее часто используемым платформам для проведения ПГТ-А методом NGS.

Сравнение данных платформ уже было представлено в таблице 1 на стр. 7.

**Ion ReproSeq PGS Kits**  
Detection of aneuploidies, segmental aneuploidies, and mosaicism using Ion ReproSeq kits

- Default high-sensitivity workflows can consistently call subchromosomal CNV events of  $\geq 8$  Mb, whereas CNV events of 2–8 Mb can be detected using adjustable custom sensitivity parameters for automated calling

**VeriSeq™ PGS Kit - MiSeq® System**  
Rapid aneuploidy screening using the MiSeq System.

BlueFuse Multi Analysis Software processes and displays the data to provide genomic profiles of each sample in a run. Whole chromosome aneuploidy is called automatically. The effective resolution of the assay is 20 Mb. Reports are generated automatically for each sample and for the cycle. Views provided are similar to those of the 24sure® array solution for PGS.

**PG-Find™ Software**  
User Guide

**NOTE:** The default settings are optimised for detecting copy number events of 5Mb or larger, so adjustment of target bin width should not be necessary unless desired minimum resolution is below 5Mb, or the required target width is not available.

Target Resolution	Number of Mapped Reads	Target Bin Width
25Mb	1,000,000	500,000, 1,000,000
15Mb	2,000,000	500,000, 1,000,000
10Mb	3,000,000	500,000, 1,000,000
5Mb	7,000,000	500,000, 1,000,000
2.5Mb	14,000,000	500,000, 1,000,000

**NOTE:** Some copy number events in the range of 5–10 Mb may not be detected using the default settings. To investigate these events, provided the requisite mapped reads are available, we recommend reanalysing the data using a sig. threshold of 1.0E-6 (no adjustment to other settings), but caution will be needed when interpreting the results as the likelihood of false positives will increase.

**Рисунок 12.** Фрагменты документации, описывающие разрешающую способность трех различных NGS-платформ для ПГТ-А.

Важно отметить, что приведенные данные **не говорят о том, как интерпретировать** выявляемые события меньшего размера, а также о том, к какому типу образцов и для какого качества исходного биоматериала применимы такие характеристики. Поэтому многие лаборатории проводят дополнительную валидацию используемых методов.

Ниже приведены некоторые способы экспериментальной оценки разрешающей способности NGS-платформ для ПГТ-А.

## Исследование геномной ДНК

Один из способов оценки разрешающей способности метода — использование геномной ДНК из крови пациентов с хромосомными нарушениями небольшого размера, подтвержденными независимыми методами исследования (например, хромосомным микроматричным анализом, ХМА).

В нашей лаборатории было проанализировано 10 таких образцов геномной ДНК, несущих 19 различных сегментарных нарушений.

Из них успешно было детектировано 18 событий.

Семь из исследованных образцов были отсекувенированы в рамках оценки влияния количества прочтений на качество полученных результатов (раздел «Количество прочтений» на странице 12).

Остальные три образца были отсекувенированы по стандартному протоколу, количество прочтений после фильтрации для них составило от 108 до 133 тысяч на образец.

**Результаты детекции всех событий при ≈ 100 000 прочтений на образец.**

Регион	Размер, Мб	Тип нарушения	Детекция
3p26.3p26.2	3,2	Делеция	+
17q25.3	3,5	Дупликация	+
1p36.33p36.32	4	Дупликация	-
17p13.3p13.2	5,6	Делеция	+
7p22.3p22.1	6,3	Делеция	+
8p23.3p23.1	6,7	Дупликация	+
2q37.1q37.3	8,2	Делеция	+
4p16.3p16.1	8,8	Делеция	+
10q26.13q26.3	11,6	Делеция	+
11p15.5p15.3	11,6	Дупликация	+
22q13.1q13.33	11,7	Дупликация	+
5p15.33p15.1	16,9	Делеция	+
5q34q35.3	17,7	Дупликация	+
13q21.33q32.1	26	Делеция	+
4p16.3p15.2	26	Делеция	+
7p14.3p11.2	27,6	Дупликация	+
1q41q44	28,2	Дупликация	+
7q11.21q21.2	29,1	Делеция	+
18q12.1q23	48	Дупликация	+

Таблица 3. Результаты детекции 19 сегментарных нарушений в образцах геномной ДНК, наличие которых подтверждено методом ХМА.

На рисунке 13 показан хромосомный профиль с наименьшим детектированным событием — сегментарной делецией хромосомы 3 размером 3,2 Мб.

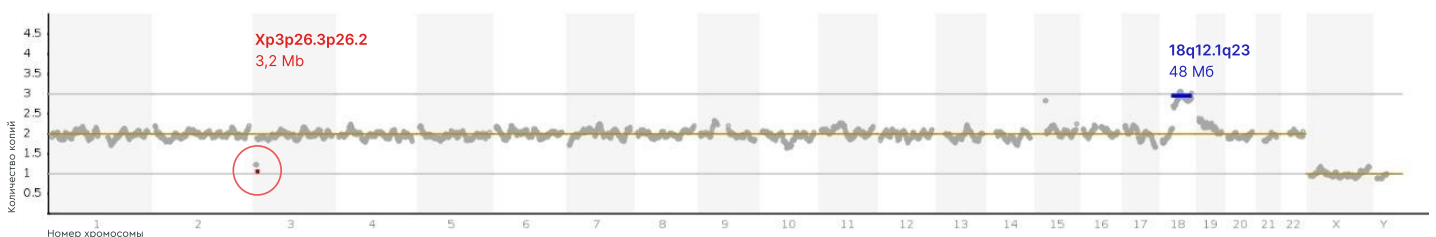


Рисунок 13. Делеция 3,2 Мб на хромосоме 3 и дупликация 48 Мб на хромосоме 18, выявленные в образце геномной ДНК в результате протокола ПГТ-А и подтвержденные методом ХМА

С учетом небольшого количества образцов полученные данные позволяют оценить разрешающую способность нашей платформы в 5 Мб, но демонстрируют возможность обнаружения и сегментарных нарушений меньшего размера.

Важно отметить, что, как правило, качество геномной ДНК лучше, чем качество ДНК в образце биопсии трофэктодермы. Поэтому полученная таким способом разрешающая способность показывает лишь теоретические возможности метода и не гарантирует подобное разрешение на клинических образцах.

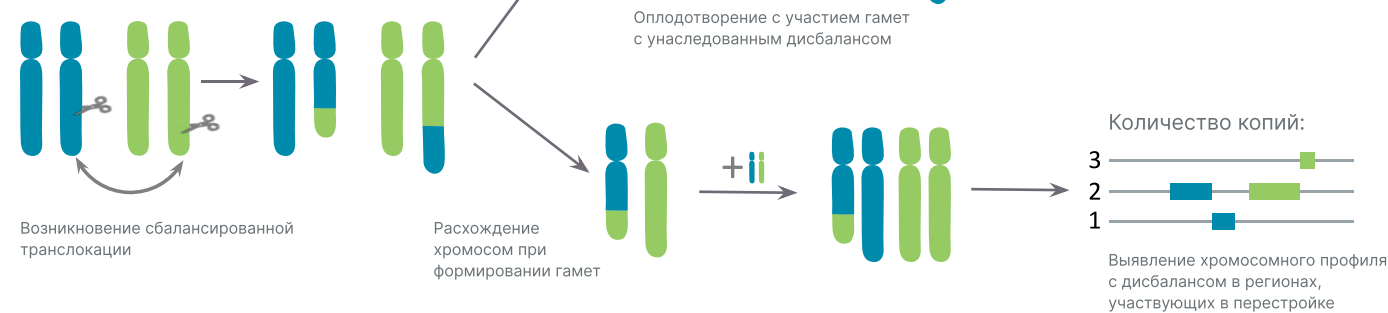


## Результаты ПГТ-А от партнеров со сбалансированными хромосомными перестройками

Как уже было сказано, в большинстве случаев для образцов биопсии трофэктодермы невозможно получить информацию о том, какой хромосомный набор несут клетки на самом деле, чтобы проверить кариотип, выявленный платформой для ПГТ-А. Однако **в некоторых случаях сегментарные нарушения могут быть унаследованными** — при наличии сбалансированных хромосомных перестроек у родителей (**рисунок 14**). В таком случае при описанном кариотипе родителей оказываются **заранее известны конкретные регионы**, где ожидается появление сегментарных нарушений — последствий родительских хромосомных перестроек.

**Рисунок 14.** Механизм наследования сегментарных нарушений при носительстве сбалансированной хромосомной перестройки одним из партнеров.

На рисунке изображены только варианты с наследованием хромосомного дисбаланса. Наследование сбалансированных вариантов не изображено для краткости.



Главный **недостаток использования таких образцов** для оценки разрешающей способности метода состоит в том, что образцы с ожидаемыми сегментарными нарушениями небольшого размера **встречаются** достаточно **редко**. Поэтому получить значимую выборку для сегментарных нарушений заданного размера оказывается невозможно.

В то же время такой подход имеет важное **преимущество**. В данном случае определение разрешения метода происходит на клинических образцах биопсии трофэктодермы, а значит позволяет **оценить реальную, а не теоретическую разрешающую способность метода**.

Мы проанализировали

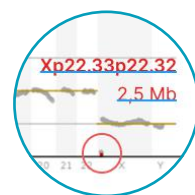
**157 образцов**,  
полученных

**от 44 семей**

с известным носительством сбалансированных хромосомных перестроек.

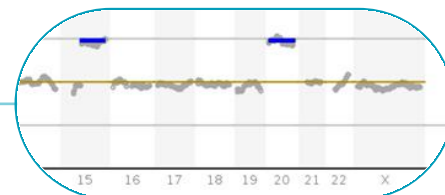
**Из них 66 образцов**

показали сегментарные нарушения в ожидаемых регионах.



Наименьший размер выявленного сегментарного нарушения составил 2,5 Мб.

Фрагмент рисунка 15, А



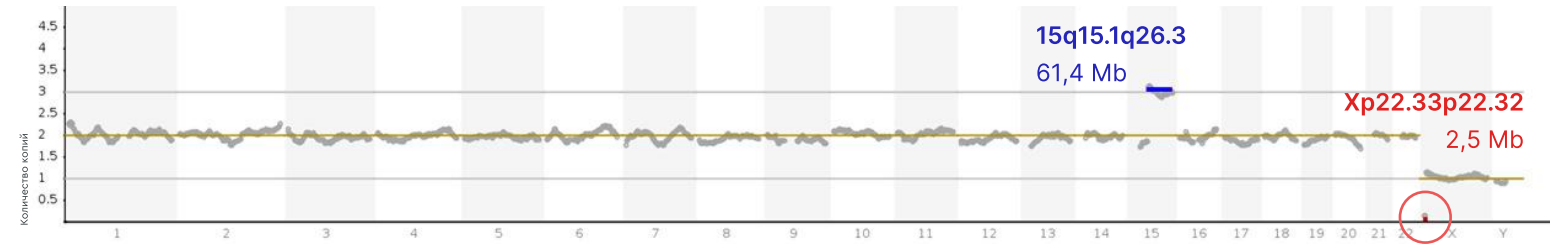
Однако для второго образца, полученного от тех же партнеров, была выявлена лишь дупликация хромосомы 15, но не выявлено сегментарное нарушение хромосомы X.

Фрагмент рисунка 15, Б

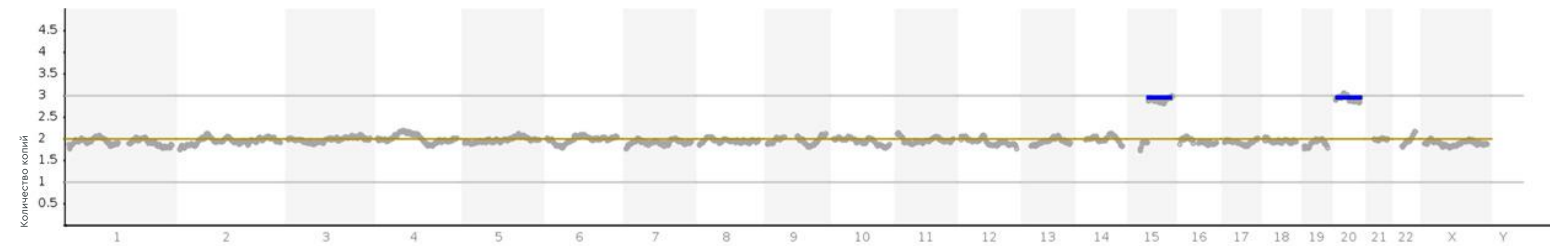
Из **рисунка 14** видно, что при наследовании несбалансированного варианта чаще всего сегментарные нарушения будут наблюдаться сразу в обоих регионах, участвующих в родительской перестройке. Таким образом, по наличию или отсутствию сегментарного нарушения в более крупном регионе можно с высокой вероятностью предполагать, должно ли наблюдаться событие на более коротком участке.

Так, например, для образца на **рисунке 15, В** нормальный профиль хромосомы 15 говорит о том, что эмбрион унаследовал сбалансированный вариант, и с большой вероятностью сегментарное нарушение хромосомы X из-за родительской транслокации в данном образце действительно отсутствует.

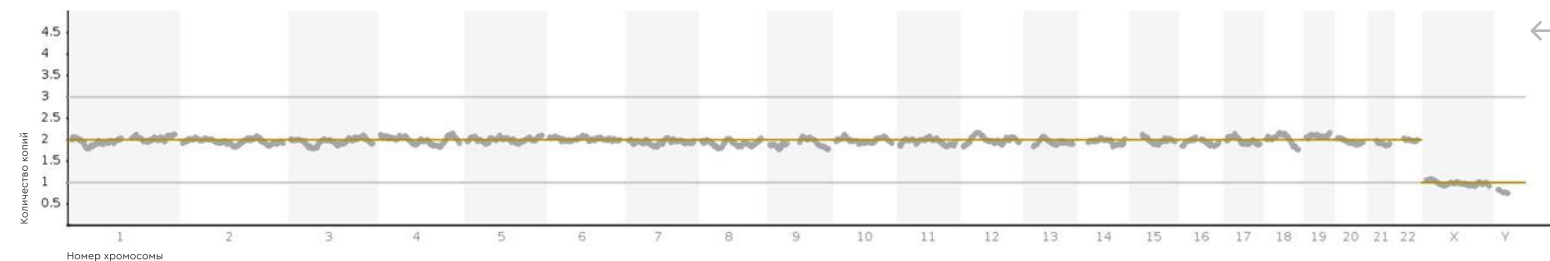
**А** Сегментарная делеция хромосомы X размера 2,5 Мб соответствует региону, участвующему в сбалансированной перестройке.



**Б** Наличие сегментарной дупликации хромосомы 15 говорит о том, что эмбрион унаследовал несбалансированный вариант. А значит, с большой вероятностью ожидается сегментарное нарушение на хромосоме X, которое детектировать не удалось.



**В** Отсутствие сегментарных нарушений на хромосоме 15 говорит о том, что эмбрион унаследовал сбалансированный вариант хромосом, участвующих в транслокации.



**Рисунок 15.** Последствия сбалансированной транслокации 46,XX,t(15;X)(q15;p22), выявленные в ходе ПГТ-А.

Такая ситуация говорит о том, что

**2,5 Мб** —

**достижимое разрешение** метода, которое, однако, может не реализовываться в зависимости от качества исходного материала.

Для всех проанализированных образцов, полученных от носителей сбалансированных транслокаций, размер хотя бы одного из регионов, участвующих в перестройке, составил не менее 18 Мб. При этом качество полученных данных для всех образцов оказалось достаточно хорошим, чтобы считать детекцию сегментов такого размера достоверной.

Таким образом, для всех образцов наличие или отсутствие сегментарного нарушения хотя бы в одном из регионов, участвующих в перестройке, позволило определить нам, унаследовал ли исследуемый эмбрион несбалансированный вариант.

Благодаря этому мы сделали вывод об отсутствии детекции мелких сегментарных нарушений в тех случаях, когда они ожидалось.



Результаты наличия или отсутствия детекции для **20 наиболее коротких сегментарных нарушений** представлены в [таблице 4](#).

Из таблицы видно, что среди всех ожидаемых событий нам не удалось зафиксировать наличие **лишь в двух случаях**, при этом ожидаемый размер нарушений составил 2,5 и 4,5 Мб. Все ожидаемые события большего размера были успешно выявлены.

**Таблица 4.** Результаты детекции 20 наиболее мелких из ожидаемых сегментарных нарушений. Все события большего размера были детектированы, для краткости эти результаты в таблицу не включены. Для анализа детекции сегментарных нарушений все образцы были проанализированы в рамках стандартного протокола секвенирования, во всех случаях количество прочтений составило 100 000 – 150 000 на образец.

Регион	Размер, Мб	Тип нарушения	Детекция
Xp22.33p22.32	2,5	Делеция	+
Xp22.33p22.32	2,5	-	-
22q11.1q11.21	3,6	Делеция	+
22q11.1q11.21	4,5	Делеция	+
22q11.1q11.21	4,5	Дупликация	+
22q11.1q11.21	4,5	Делеция	+
22q11.1q11.21	4,5	Дупликация	+
22q11.21q13.33	4,5	-	-
Yq11.223q12	4,9	Делеция	+
Yq11.223q12	4,9	Делеция	+
2q37.1q37.3	9	Дупликация	+
2q37.1q37.3	9	Дупликация	+
Yq11.221q12	10,8	Делеция	+
8p23.3p22	12,1	Делеция	+
18p11.32p11.1	15,4	Дупликация	+
18p11.32p11.1	15,4	Делеция	+
7q35q36.3	15,7	Делеция	+
7q35q36.3	15,7	Дупликация	+
11q23.3q25	18,2	Делеция	+
13q31.3q34	19,6	Делеция	+

Таким образом, в результате проведенного анализа показано, что **2,5 Мб** может достигать разрешающая способность нашей платформы для ПГТ-А.

А для событий **от 5–10 Мб** их детекция стабильна.

Тем не менее размер выборки, а также продемонстрированная ранее разнородность образцов по качеству, требуют осторожного отношения к сегментарным событиям на границе разрешающей способности метода.

## Сравнение с другими платформами

Еще один способ определения разрешающей способности метода — сравнение результатов, полученных с помощью различных платформ.

Основной недостаток такого подхода состоит в том, что при расхождении результатов, чаще всего нет возможности определить, какая платформа «ошиблась», поскольку каждое исследование может давать как ложноотрицательный, так и ложноположительный результат.

Тем не менее, в открытых источниках встречается большое количество сравнений разрешающей способности двух наиболее популярных платформ для ПГТ-А — **ReproSeq** (ThermoFisher Scientific) и **VeriSeq** (Illumina). Как правило они показывают **полное совпадение результатов обеих платформ и их разрешающую способность около 5 Мб**.



Результаты одного из подобных независимых сравнений представлены на [рисунке 16](#).

Assessment of resolution

Abnormality Size	Chromosome	ReproSeq-PGS	VeriSeq-PGS
Whole Chromosome	21 (trisomy)	Detected	Detected
Whole Chromosome	X (monosomy)	Detected	Detected
Whole Chromosome	X (trisomy)	Detected	Detected
+33.3Mb	21	Detected	Detected
+28.2Mb	8	Detected	Detected
+26.2Mb	20	Detected	Detected
-7.8Mb	8	Detected	Detected
-5.5Mb	13	Detected	Detected
+5.1Mb	15	Not Detected	Not Detected
+4.6Mb	8	Detected	Detected
-3.6Mb	17	Not Detected	Not Detected
-3.5Mb	22	Not Detected	Not Detected
-2.5Mb	10	Not Detected	Not Detected
+2.2Mb	15	Not Detected	Not Detected
-1.5Mb	17	Not Detected	Not Detected

\* Both platforms saw all of chromosome 21 as trisomic

\*\* Mosaic gain of chromosome 10 seen on both platforms

**Рисунок 16.** Результаты независимого сравнения разрешающей способности платформ ReproSeq (ThermoFisher Scientific) и VeriSeq (Illumina), проведенного лабораторией NxGen MDx. Презентация «Independent Evaluation of Reproseq PGS» доступна на сайте <https://www.genomeweb.com/>.

# Клиническая значимость микроделеций и микродупликаций, выявленных в результате ПГТ-А

Любой диагностический или скрининговый метод неизбежно дает ложноположительные и ложноотрицательные результаты. Для того чтобы найти баланс чувствительности и специфичности, обсуждая возможность детекции небольших сегментарных нарушений при ПГТ-А, важно учитывать распространенность и клиническую значимость таких нарушений.

Микроделеционные и микродупликационные синдромы, связанные с сегментарными хромосомными нарушениями размером менее 5-10 Мб — это очень обширная и разнородная группа состояний, и однозначной оценки распространенности этих состояний в популяции нет.

## Синдром Ди Джорджи

от 1 случая на 1 000 до 1 на 6 000

Например, самый распространенный и изученный микроделеционный синдром — синдром Ди Джорджи встречается, по разным оценкам, от 1 случая на 1 000 до 1 на 6 000 (Weise et al., 2012). Этот синдром вызван потерей участка хромосомы 22 размером от 1,5 до 3 Мб (регион 22q11.2) и может иметь различные клинические проявления даже внутри одной семьи.

Так, при тяжелых формах синдром проявляется множественными нарушениями со стороны сердечно-сосудистой и иммунной систем, дефектами нёба, приводящими к нарушению питания и речи, тяжелыми нарушениями развития и патологией со стороны практически всех систем органов.

В то же время описаны случаи, при которых синдром проявлялся лишь в виде незначительных интеллектуальных или социальных нарушений и был диагностирован у родителей только после рождения ребенка с выраженной патологией (McDonald-McGinn and Sullivan, 2011).

## Синдромы Вильямса, Ангельмана и Прадера-Вилли

1 на 10 000 или реже

Другие известные синдромы, которые могут быть вызваны микроделециями, такие как синдромы Вильямса, Ангельмана и Прадера-Вилли встречаются в популяции с частотой 1 на 10 000 или реже (Goldenberg, 2018; Gurrieri and Accadia, 2009).

## Интерпретация CNV

Отдельную сложность представляет интерпретация выявляемых CNV (copy number variations, вариации количества копий участков ДНК).

Как правило, задачу определения клинической значимости CNV решают при проведении хромосомного микроматричного анализа пациентам с подозрением на хромосомные нарушения и наличием выраженной клинической картины заболевания. При анализе CNV в геноме человека было выявлено, что ДНК двух здоровых людей может различаться наличием или отсутствием участков размером в несколько миллионов пар оснований (Sebat et al., 2004).

Это означает, что наличие CNV размером даже в несколько Мб не обязательно приводит к развитию заболевания

(Bateman et al., 2018; Filges et al., 2009; Maurin et al., 2017).

Для определения клинической значимости выявляемых CNV учитываются многие критерии, такие как:

- > размер,
- > тип (делеция или дупликация),
- > генный состав,
- > унаследован ли вариант от родителей или возник *de novo*.

Микродупликации часто могут иметь более мягкие проявления и труднее поддаются интерпретации, чем микроделеции. Более крупные нарушения, как правило, чаще оказываются патогенными, в то время как для интерпретации более мелких событий необходим более подробный анализ. Однако даже очень небольшие варианты (размером менее 1 Мб) могут оказаться патогенными (Watson et al., 2014).

Таким образом, суммарная частота всех заболеваний, связанных с сегментарными хромосомными нарушениями небольшого размера, **не превышает 1 на 1 000**.

При этом для многих из них невозможно достоверное определение клинической значимости без сопоставления литературных данных и клинических проявлений, которые в принципе недоступны для эмбрионов, исследуемых при ПГТ-А.

## В каком случае стоит искать CNV небольшого размера?

Мы убеждены, что в первую очередь ПГТ-А должно помогать в решении репродуктивных задач.

Основные цели проведения исследования:

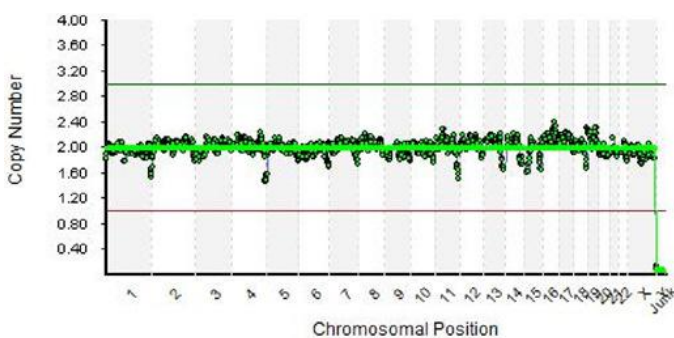
- повышение эффективности переноса эмбрионов в цикле ЭКО,
- сокращение времени до наступления беременности,
- уменьшение частоты потери беременности.

ПГТ-А позволяет снизить вероятность рождения ребенка с хромосомными аномалиями, но не позволяет полностью исключить её, в том числе из-за мозаицизма. Поэтому к беременности, полученной после проведения ПГТ-А, необходимо относиться с той же степенью осторожности, что и к спонтанной беременности.

При повышении чувствительности метода в стремлении детектировать как можно более мелкие сегментарные нарушения, неизбежно увеличивается доля ложноположительных результатов. Помимо уже продемонстрированных проблем с образцами низкого качества (рисунок 2, стр. 5), а также выбором излишне чувствительного алгоритма обработки данных (рисунок 9, стр. 17), даже при оптимизации этих условий, неизбежно могут возникать артефакты анализа. К сожалению, такие артефакты достаточно частые и даже упоминаются в руководствах по интерпретации хромосомных профилей (рисунок 17).

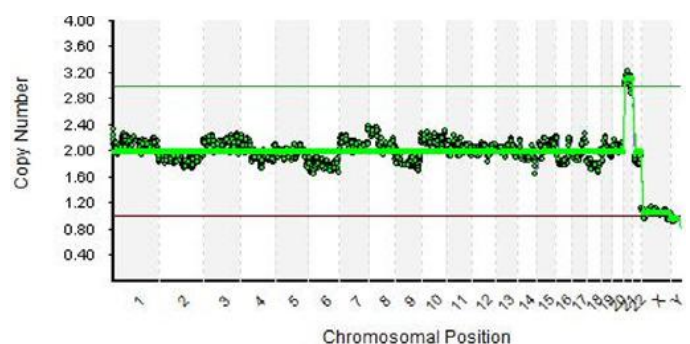
**Рисунок 17.** Примеры ложноположительных результатов, приведенных в руководстве по интерпретации данных A Technical Guide to Aneuploidy Calling with VeriSeq PGS (Illumina).

**А** Теломерный эффект, который может быть ошибочно принят за микроделеции на концах хромосом.



Учитывая относительно небольшую частоту микроделеционных и микродупликационных синдромов, сложность интерпретации выявляемых CNV небольшого размера, а также возможность ложноположительной детекции небольших хромосомных нарушений, важно правильно выбрать критерии интерпретации и включения в итоговое заключение подобных событий.

**Б** «Волновой эффект», который может быть ошибочно принят за низкоуровневый мозаицизм.



Чрезмерное повышение чувствительности метода приводит к гипердиагностике и **неоправданному уменьшению количества эмбрионов, рекомендуемых к переносу**. В результате такого подхода ПГТ-А может уже не помогать, а только мешать решению репродуктивных задач в циклах ВРТ.

Поэтому мы считаем важным использовать чувствительные алгоритмы анализа сырых данных только в тех случаях, когда в образце ожидаются хромосомные нарушения небольшого размера (например, при носительстве родителями сбалансированных хромосомных перестроек), и доверять только тем событиям менее 5-10 Мб, которые наблюдаются в ожидаемых регионах.

## Резюме:

1

Технические характеристики платформы ПГТ-А, используемой в лаборатории First Genetics, позволяют детектировать события размером **от 2,5 Мб** при оптимальном качестве исходного биоматериала.

2

Разрешающая способность и точность метода ПГТ-А во многом зависит от качества исходного биоматериала.

3

Для большинства образцов разрешающая способность нашего метода ПГТ-А составляет **5–10 Мб**.

4

Сверхчувствительный поиск сегментарных нарушений небольшого размера должен проводиться только для тех образцов и в тех областях, где такие нарушения ожидаются исходя из клинической информации о пациентах.

5

К беременности, полученной после проведения ПГТ-А, необходимо относиться с такой же осторожностью, как к любой другой беременности. Проблема возможных хромосомных нарушений у плода после проведения ПГТ-А должна решаться правильным консультированием пациентов, своевременными скринингами и при необходимости проведением инвазивной диагностики.

# Список литературы

Bateman, M.S., Collinson, M.N., Bunyan, D.J., Collins, A.L., Duncan, P., Firth, R., Harrison, V., Homfray, T., Huang, S., Kirk, B., et al. (2018).

**Incomplete penetrance, variable expressivity, or dosage insensitivity in four families with directly transmitted unbalanced chromosome abnormalities.**

Am. J. Med. Genet. A 176, 319–329.

Cuman, C., Beyer, C.E., Brodie, D., Fullston, T., Lin, J.I., Willats, E., Zander-Fox, D., and Mullen, J. (2018).

**Defining the limits of detection for chromosome rearrangements in the preimplantation embryo using next generation sequencing.**

Hum. Reprod. 33, 1566–1576.

Deleye, L., Tilleman, L., Vander Plaetsen, A.-S., Cornelis, S., Deforce, D., and Van Nieuwerburgh, F. (2017).

**Performance of four modern whole genome amplification methods for copy number variant detection in single cells.**

Sci. Rep. 7, 3422.

Filges, I., Röthlisberger, B., Noppen, C., Boesch, N., Wenzel, F., Necker, J., Binkert, F., Huber, A.R., Heinimann, K., and Miny, P. (2009).

**Familial 14.5 Mb interstitial deletion 13q21.1-13q21.33: clinical and array-CGH study of a benign phenotype in a three-generation family.**

Am. J. Med. Genet. A 149A, 237–241.

Goldenberg, P. (2018).

**An Update on Common Chromosome Microdeletion and Microduplication Syndromes.**

Pediatr. Ann. 47, e198–e203.

Gurrieri, F., and Accadia, M. (2009).

**Genetic imprinting: the paradigm of Prader-Willi and Angelman syndromes.**

Endocr. Dev. 14, 20–28.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009).

**The Sequence Alignment/Map format and SAMtools.**

Bioinformatics 25, 2078–2079.

Maurin, M.-L., Arfeuille, C., Sonigo, P., Rondeau, S., Vekemans, M., Turleau, C., Ville, Y., and Malan, V. (2017).

**Large Duplications Can Be Benign Copy Number Variants: A Case of a 3.6-Mb Xq21.33 Duplication.**

Cytogenet. Genome Res. 151, 115–118.

McDonald-McGinn, D.M., and Sullivan, K.E. (2011).

**Chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome).**

Medicine (Baltimore). 90, 1–18.

Del Rey, J., Vidal, F., Ramírez, L., Borràs, N., Corrales, I., Garcia, I., Martinez-Pasarell, O., Fernandez, S.F., Garcia-Cruz, R., Pujol, A., et al. (2018).

**Novel Double Factor PGT strategy analyzing blastocyst stage embryos in a single NGS procedure.**

PLoS One 13, e0205692.

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Månér, S., Massa, H., Walker, M., Chi, M., et al. (2004).

**Large-scale copy number polymorphism in the human genome.**

Science 305, 525–528.

Watson, C.T., Marques-Bonet, T., Sharp, A.J., and Mefford, H.C. (2014).

**The genetics of microdeletion and microduplication syndromes: an update.**

Annu. Rev. Genomics Hum. Genet. 15, 215–244.

Weise, A., Mrasek, K., Klein, E., Mulatinho, M., Llerena, J.C., Hardekopf, D., Pekova, S., Bhatt, S., Kosyakova, N., and Liehr, T. (2012).

**Microdeletion and microduplication syndromes.**

J. Histochem. Cytochem. 60, 346–358.



## АО «Ферст Генетикс»

г. Москва, Инновационный центр Сколково, Большой бульвар, 42с1

Лицензия № 008 от 14 декабря 2018 г.

**+7 (499) 673-02-80**

Звонки принимаем с 09:00 до 24:00 каждый день

**info@f-genetics.com**

Отвечаем в течение двух часов