

ФГБОУ ВО «Российский национальный исследовательский
медицинский университет имени Н.И. Пирогова» Министерства
здравоохранения Российской Федерации
Кафедра молекулярной биологии и медицинской биотехнологии
МБФ

**Анализ результатов
генотипирования, проведенного
различными методами ПЦР.
Анализ и значение полученных
данных**

Занятие II

2018

- Большинство методов анализа генома основано на ПЦР
- Результаты ПЦР анализируют с использованием :
 - Электрофореза
 - Секвенирования
 - Флуоресцентных методов
(ПЦР в реальном времени, real-time PCR)
- Анализ данных

Электрофорез фрагментов ДНК (в том числе продуктов ПЦР, или амплификата) — это аналитический метод, применяемый для разделения, идентификации и очистки фрагментов ДНК по размеру (длине), основанный на движении заряженных макромолекул в постоянном электрическом поле.

- В агарозном геле;
- В полиакриламидном геле (ПААГ).

Электрофорез в агарозном геле:

Чем определяется скорость миграции ДНК в геле?

1. Размер молекул ДНК;

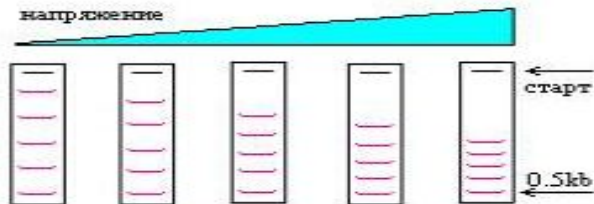
2. Концентрация агарозы

% агарозы	0.3	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.5	2.0
Размер DNA [kbp]	5-60	1-30	1-20	0.8-12	0.6-10	0.5-8	0.5-7	0.4-6	0.2-3	0.1-2

При необходимости разделения фрагментов ДНК, отличающихся на 20 – 50 п.н., используют 2,5-3% агарозный гель.

3. Конформация ДНК (кольцевая, линейная);

4. Напряженность электрического поля;

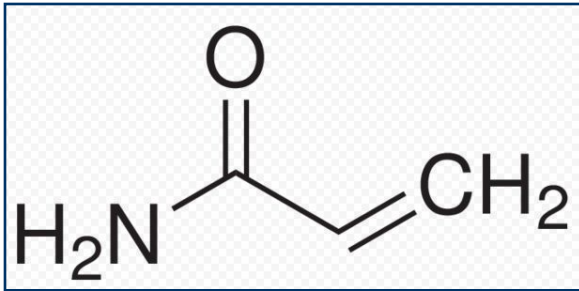


Электрофорез в полиакриламидном геле

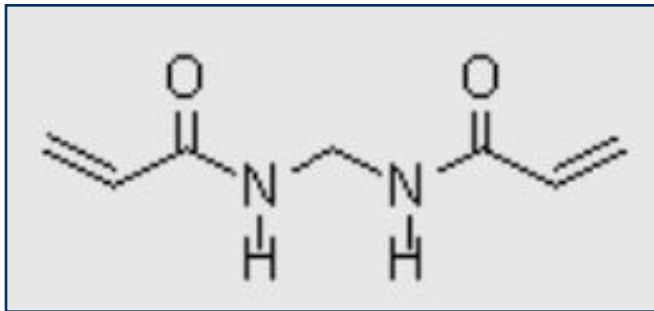
- Для разделения белков в основном используют метод электрофореза в синтетической гелевой среде на основе полиакриламида.

Полиакриламидный гель (ПААГ)

- Для получения полиакриламидного геля используют:
 - акриламид и



- N,N'-метиленабисакриламид (или бисакриламид) (отвечает за образование поперечных сшивок в геле).

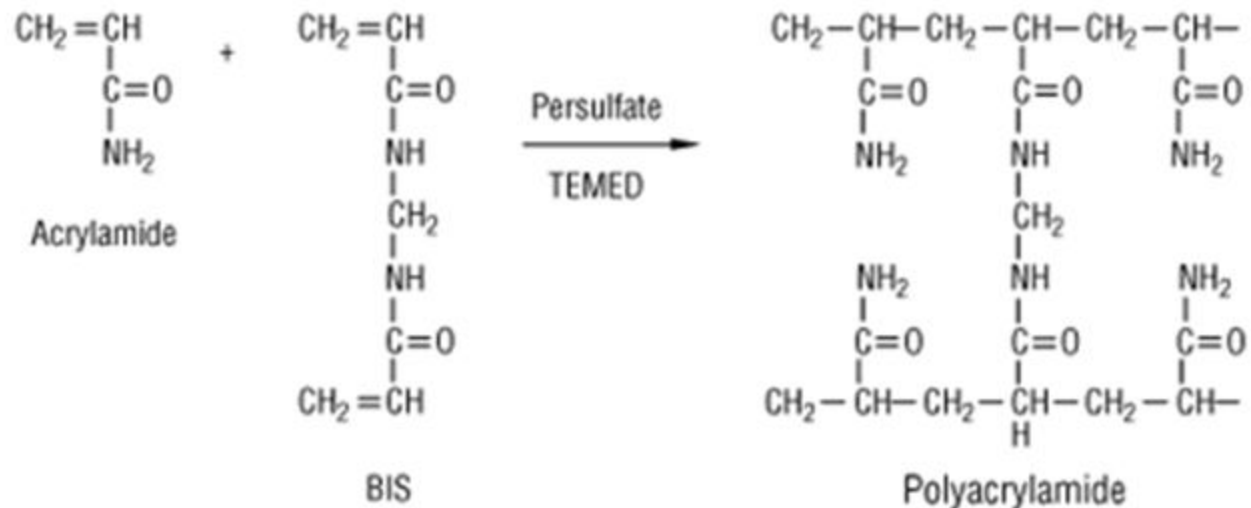


Электрофорез в полиакриламидном геле

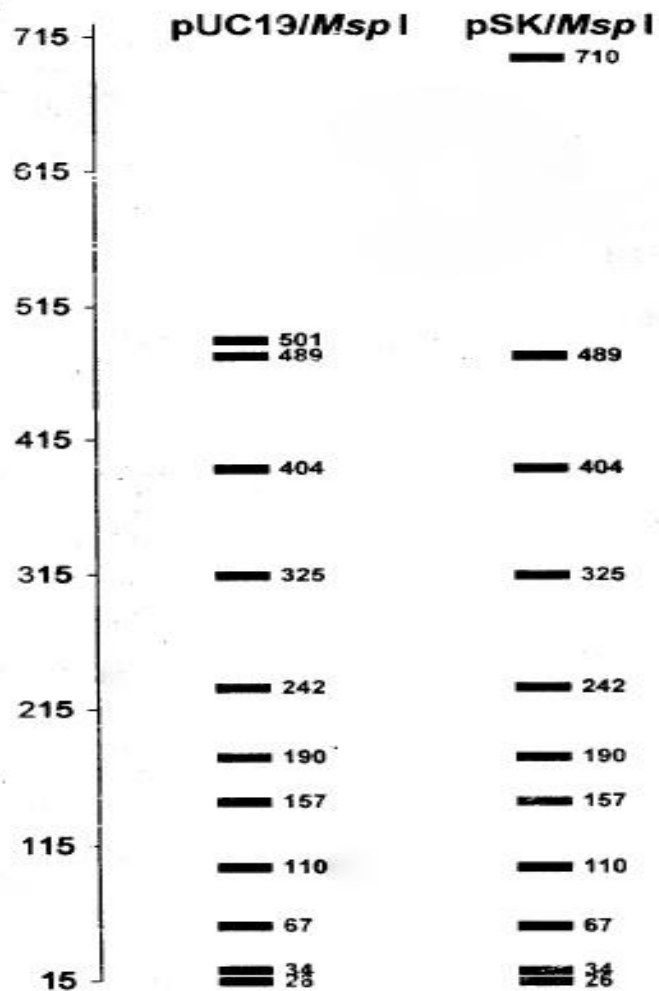
Механизм полимеризации акриламида

- Полимеризация акриламида идет по механизму **свободнорадикальной реакции** и инициируется добавлением или образованием *in situ* свободных радикалов.
- Наиболее распространенным **инициатором полимеризации акриламида** является **персульфат аммония** $((\text{NH}_4)_2\text{S}_2\text{O}_8)$.
- Персульфат может подвергаться одноэлектронному восстановлению:
$$\text{S}_2\text{O}_8^{2-} + 1e \rightarrow \text{SO}_4^{2-} + \text{SO}_4^\bullet$$
- SO_4^\bullet - свободный радикал, запускающий полимеризацию.
- Восстановление персульфата ускоряется в присутствии катализатора - N,N'-тетраметиленамина (**ТЕМЕД** $(\text{H}_3\text{C})_2\text{-N-CH}_2\text{-CH}_2\text{-N-}(\text{CH}_3)_2$)
- Поскольку акриламид имеет только одну реакционную группу, при полимеризации рост цепи идет по типу **«голова – хвост»**. Для образования **поперечных сшивок** необходим **бифункциональный реагент**, такой как бисакриламид.

Когда раствор содержит акриламид, бисакриламид и инициатор полимеризации, происходит одновременное протекание двух реакций – полимеризация акриламида и формирование поперечных сшивок.



Карты маркеров молекулярной массы



Пример генотипирования инсерционно-делеционного (I/D) полиморфизма гена ACE (ангиотензин-превращающий фермент) методом анализа длин продуктов ПЦР

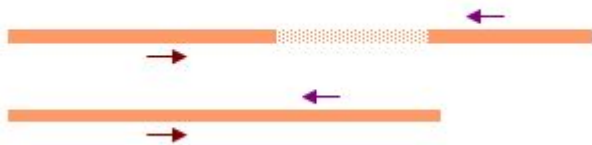


схема расположения праймеров. Область делеции обозначена штриховкой

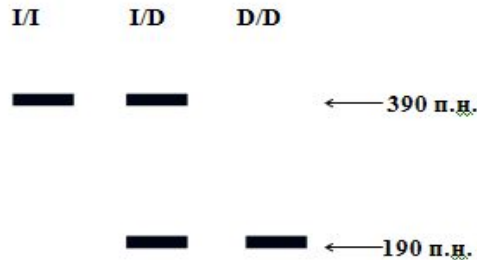
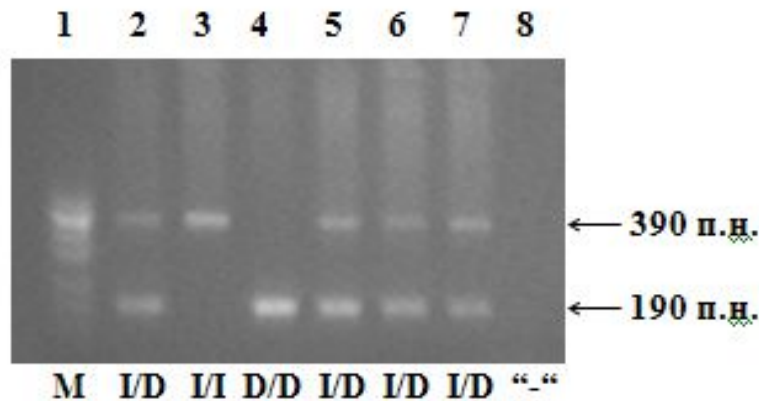
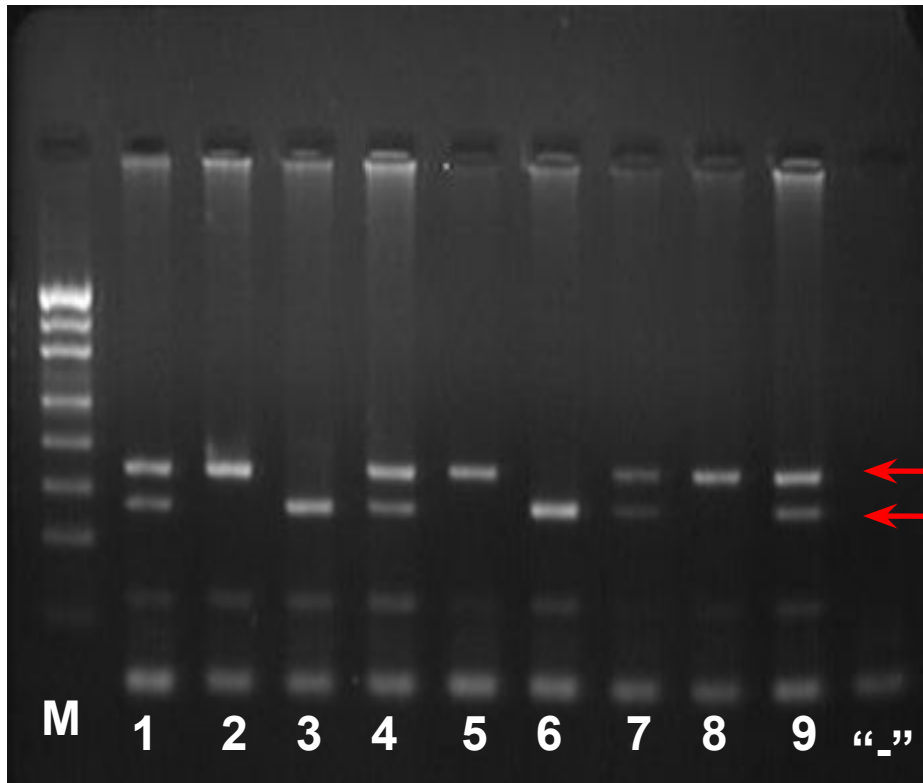


схема набора фрагментов ПЦР, характеризующих рассматриваемый полиморфизм.



анализ фрагментов ПЦР электрофорезом в 2% агарозном геле

Пример генотипирования инсерционно-делеционного полиморфизма в гене *CCR5* методом ПЦР



Анализ длины фрагментов ПЦР электрофорезом в 3% агарозном геле

169 п.н.

137 п.н.

М - маркер молекулярной массы (ДНК рUC19, обработанная эндонуклеазой рестрикции *Msp1*);

1-7 : исследуемые образцы.

8,9:положительные контроли.

Пример генотипирования полиморфизма +49A>G (Thr→Ala) , экзон 1 гена *CTLA4* (ассоциированного с цитотоксическими Т-лимфоцитами белка 4) методом ПЦР – ПДРФ с использованием рестриктазы PspEI.

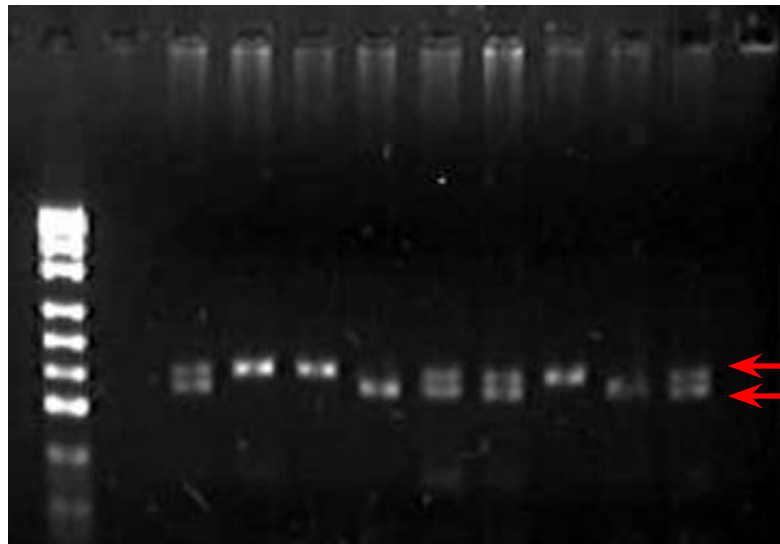
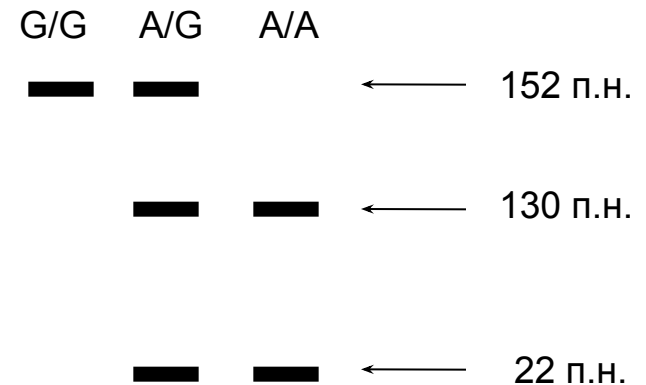


Схема рестрикции:



фрагменты рестрикции после электрофореза в 3%-ом агарозном геле;

М - маркер молекулярной массы (ДНК рUC19, обработанная эндонуклеазой рестрикции Msp1);
 1-6 : исследуемые образцы.
 7-9:положительные контроли.

Сайт узнавания рестриктазы:



Пример генотипирования полиморфизма 4266A>G (Thr312Ala) в гене *FGA* методом ПЦР – ПДРФ с использованием рестриктазы *RsaI*.

Сайт узнавания рестриктазы:

GT↑AC
CA↓TG

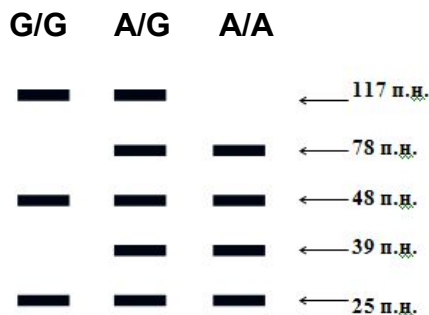
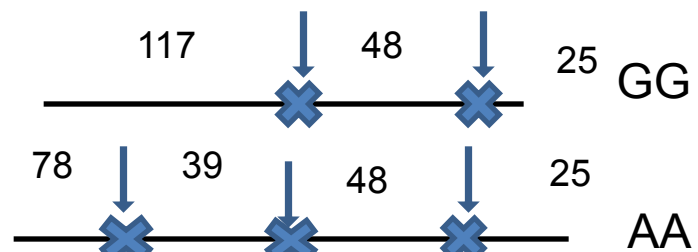
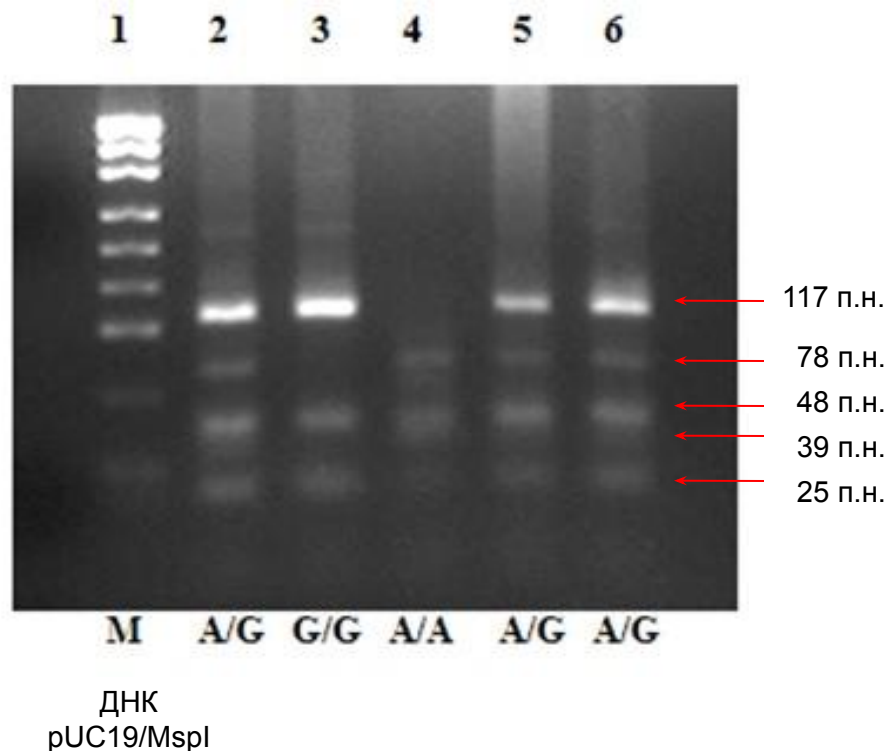
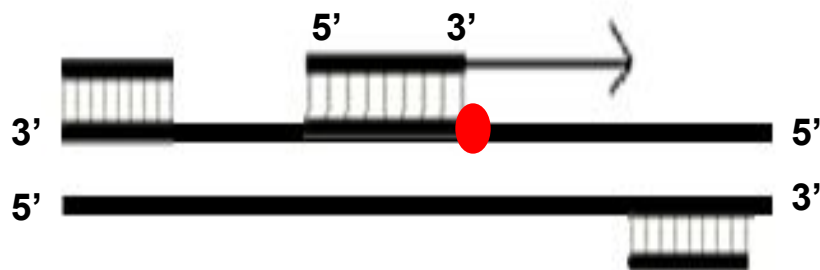


схема набора рестрикционных фрагментов, характеризующих рассматриваемый полиморфизм

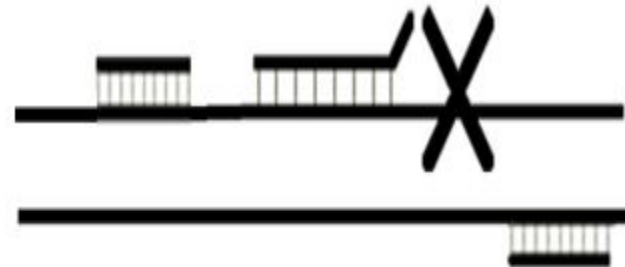


фрагменты рестрикции после электрофореза в 3%-ом агарозном геле;

Общая схема ПЦР с помощью аллелеспецифических праймеров (PCR-SSP)

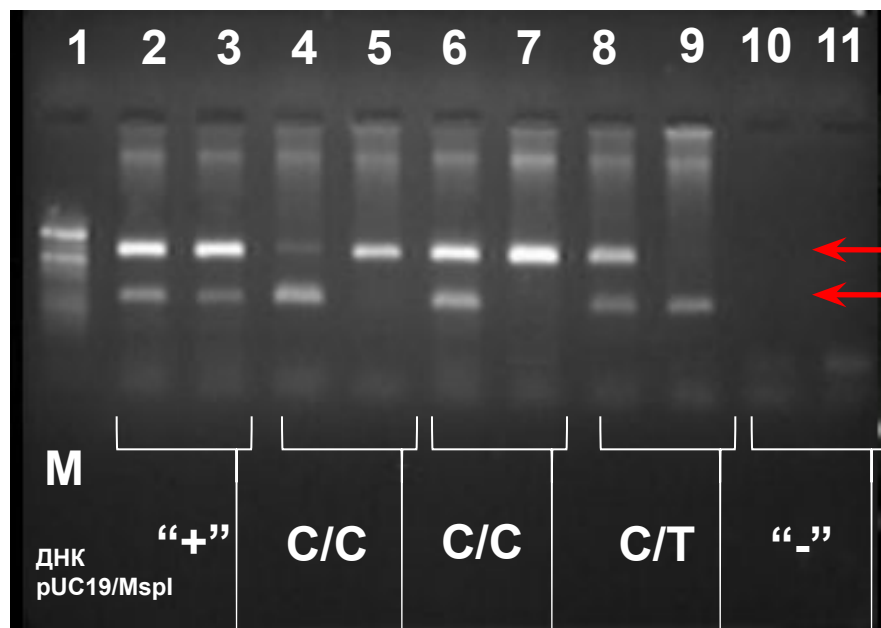


комплементарность праймера аллелю
– амплификация



отсутствие комплементарности
праймера аллелю – нет амплификации

Пример генотипирования полиморфизма С>Т экзон 6 (Ile244Thr) в гене *IL7RA* методом ПЦР с аллелеспецифическими праймерами (пример 1)

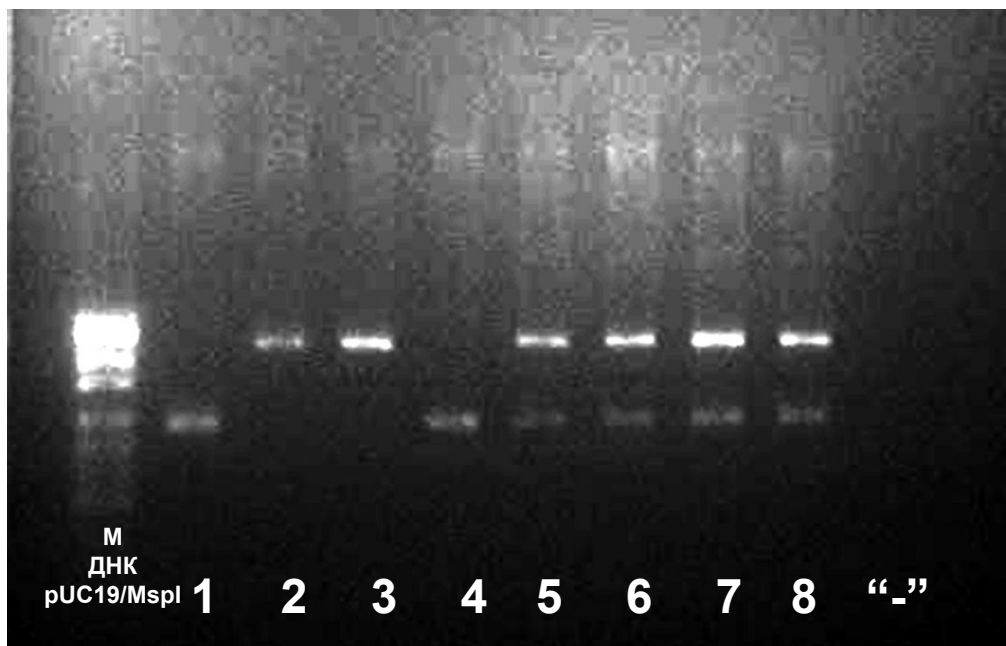


анализ фрагментов ДНК после PCR-SSP
в 2% агарозном геле

В карманы 2,4,6,8 внесены пробы
с аллелеспецифическим
праймером для аллеля С

В карманы 3,5,7,9 внесены пробы
с аллелеспецифическим
праймером для аллеля Т

Пример генотипирования полиморфизма 16725G>C в гене *IFNAR1* методом ПЦР с аллелеспецифическими праймерами (пример 2)



анализ фрагментов ДНК после PCR-SSP
в 2% агарозном геле

← 443 п.н. - внутренний контроль

← 237 п.н. - аллелеспецифический
продукт

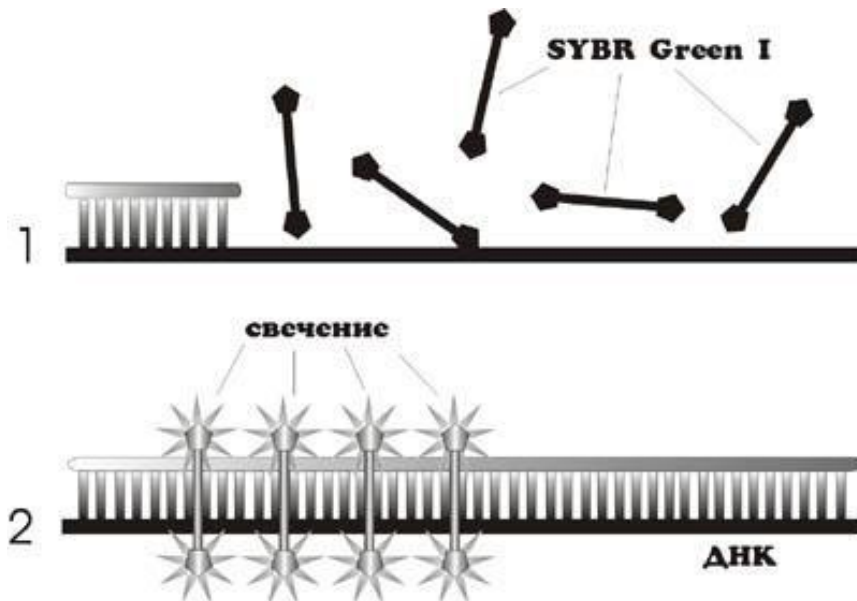
В карманы 1,3,5,7 внесены пробы с
аллелеспецифическим праймером
для аллеля С

В карманы 2,4,6,8 внесены пробы с
аллелеспецифическим праймером
для аллеля G

ПЦР в реальном времени (Real-time PCR) - это семейство методик количественной ПЦР со следующими чертами:

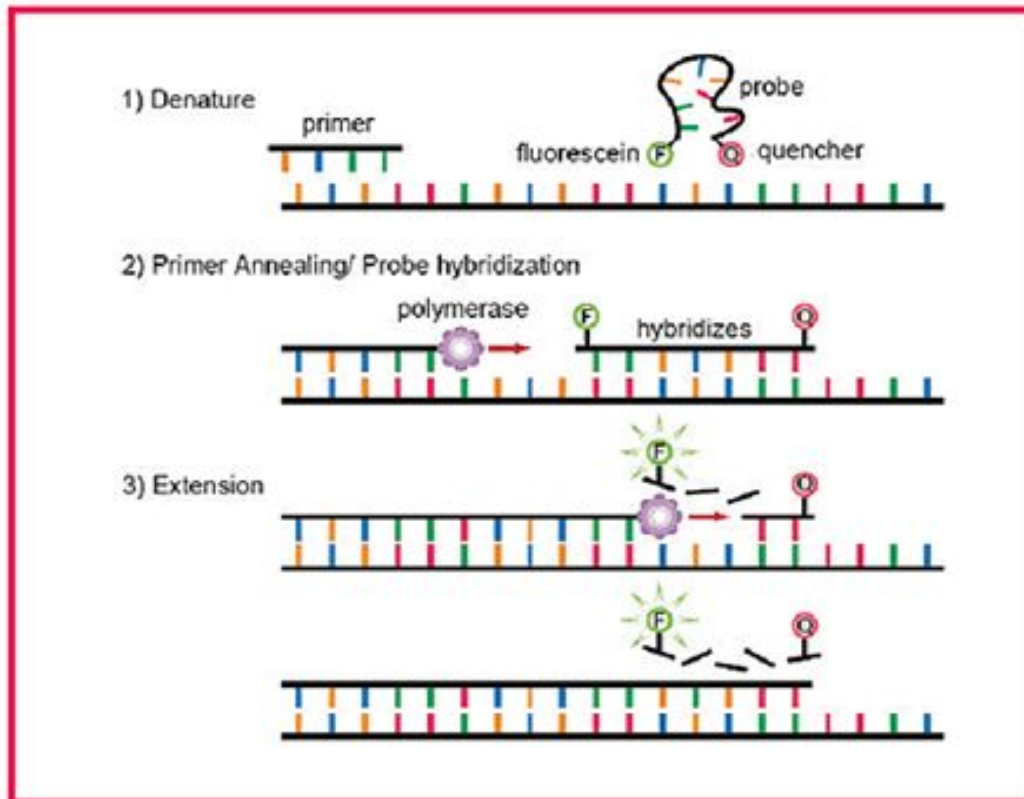
1. Флуоресцентная регистрация накопления вещества;
2. Определение выхода продукта реакции после каждого цикла амплификации;
3. Построение по этим данным кинетической кривой ПЦР;
4. Определение относительной концентрации субстрата на основании анализа этой кривой.

Использование интеркалирующих агентов (неспецифическая система детекции)



ПЦР в реальном времени со специфичной системой детекции (1)

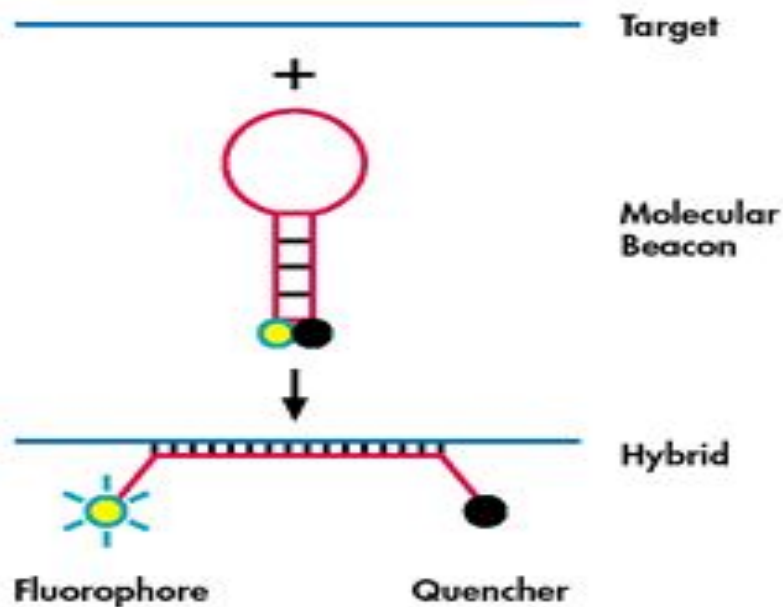
1. Выщепление 5' концевой метки (TaqMan Assay) с помощью 5'-экзонуклеазной активности ДНК-полимеразы.



TaqMan® Probe Method

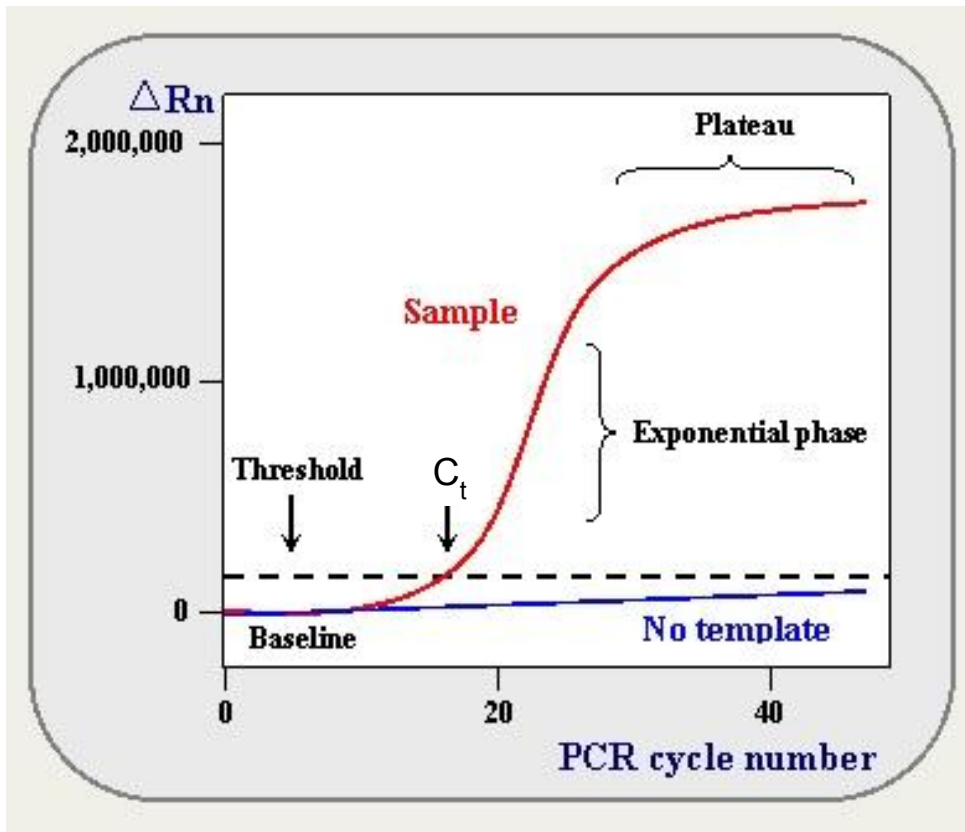
ПЦР в реальном времени со специфичной системой детекции (2)

2. Использование зондов с комплементарными концевыми последовательностями (molecular beacons).



- наличие комплементарных концевых последовательностей приводит к образованию шпильки и тушению флуоресценции;
- при связывании зонда происходит расхождение флуоресцентной метки и гасителя.

Модель графика накопления ДНК в ходе ПЦР в реальном времени



Baseline – базовый уровень флуоресценции: флуоресцентный сигнал накапливается, но он ниже пороговой величины, детектируемой прибором

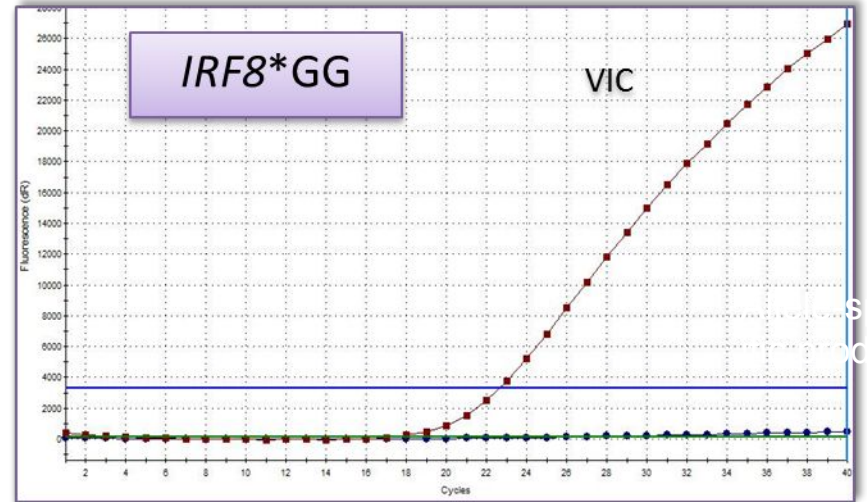
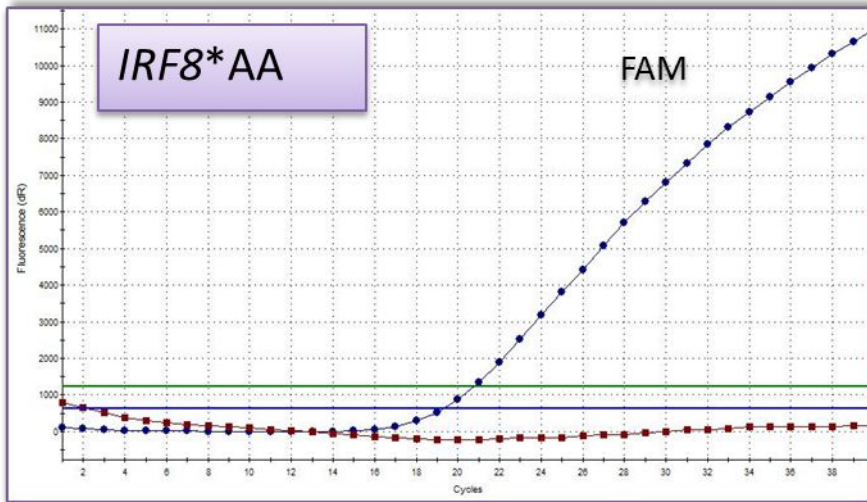
ΔR_n – приращение (изменение) флуоресцентного сигнала в каждый момент времени

Threshold – пороговая линия, обозначающая пороговый уровень флуоресценции. Выбирается произвольно исходя из базового уровня флуоресценции. Сигнал, обнаруживаемый над пороговой линией считают истинным сигналом.

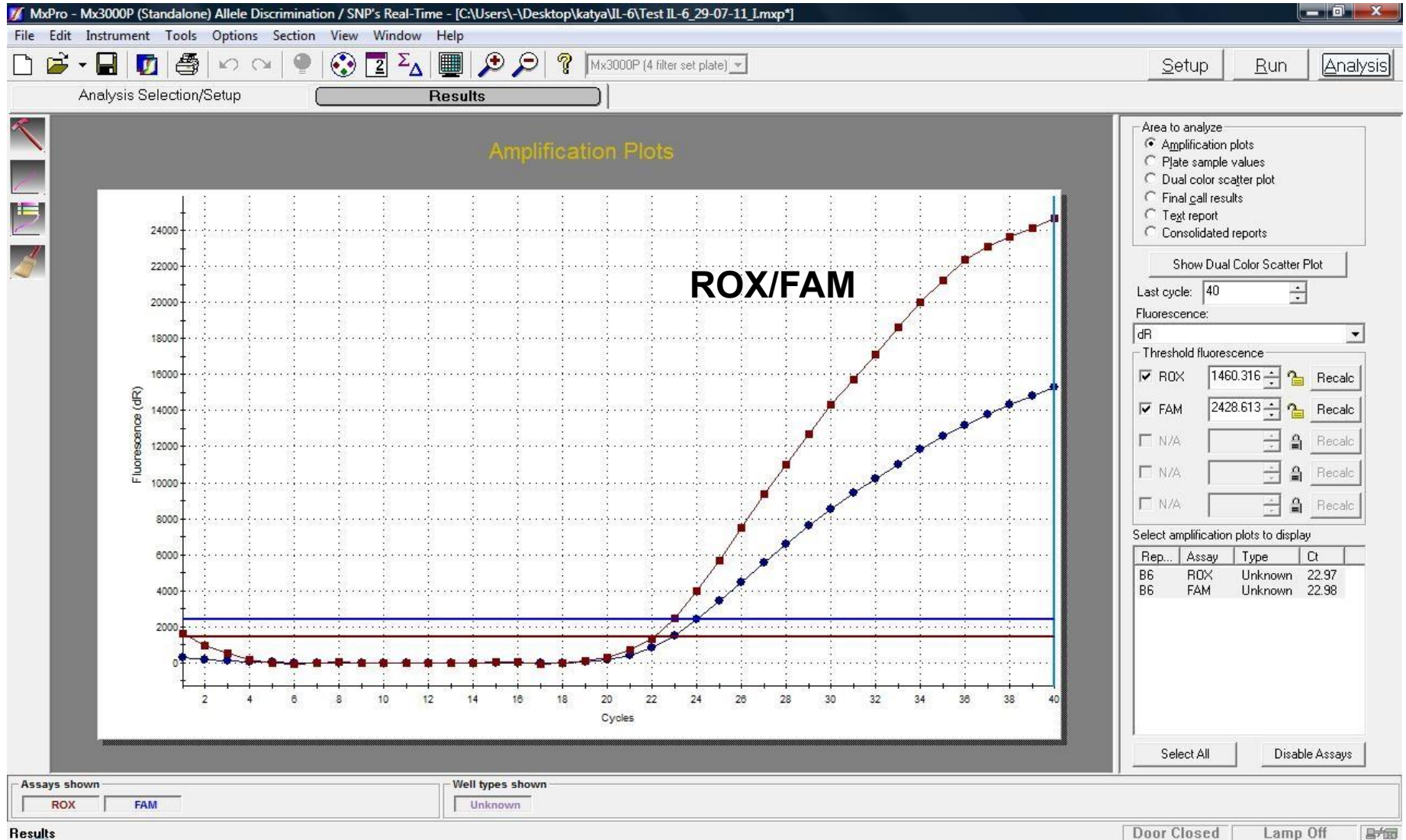
C_t – пороговый цикл реакции, при котором уровень репортерной флуоресценции превышает ее пороговое значение.

Результаты генотипирования полиморфизма rs17445836 в гене *IRF8*

гомозиготные образцы

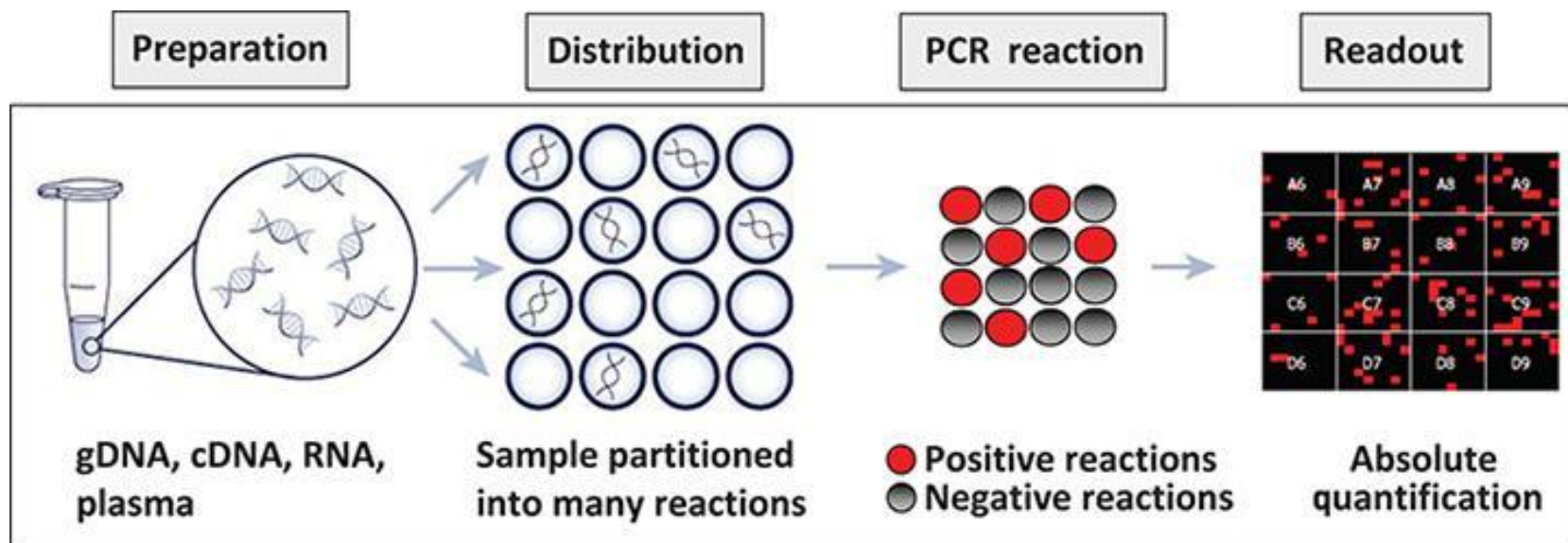


Гетерозиготный образец (ген IL6)



Digital PCR

- Позволяет определять абсолютное количество копий ДНК-мишени в образце.
- Обладает высокой чувствительностью и специфичностью.
- Проводится множество параллельных реакций на малом количестве материала каждая.
- Для коррекции искажений, возникающих в результате возможного попадания в лунку более 1 молекулы ДНК применяется поправка, учитывающая распределение Пуассона.



Digital PCR (2)

Решаемые задачи:

- **Количественная оценка биомаркеров рака.** Мутации, ассоциированные с онкологией, часто не удается детектировать из-за их низкой концентрации по сравнению с фоновой ДНК дикого типа в том же образце.
- **Точная количественная оценка вирусной нагрузки.**
- **Определение количества копий гена (CNV).**
- **Валидация и количественная оценка библиотек NGS.** Используя технологии цифровой ПЦР, можно качественно и количественно оценить созданную библиотеку для наиболее эффективного использования секвенаторов.
- **Анализ экспрессии генов.** Капельная цифровая ПЦР позволяет количественно определить уровень экспрессии гена: возможна оценка минимального 10% изменения экспрессии, в том числе и при низких концентрациях.
- **Стандартизация экспериментов и сравнение результатов, полученных в разных лабораториях.**

Под секвенированием ДНК понимают определение её нуклеотидной последовательности

Компоненты успеха:

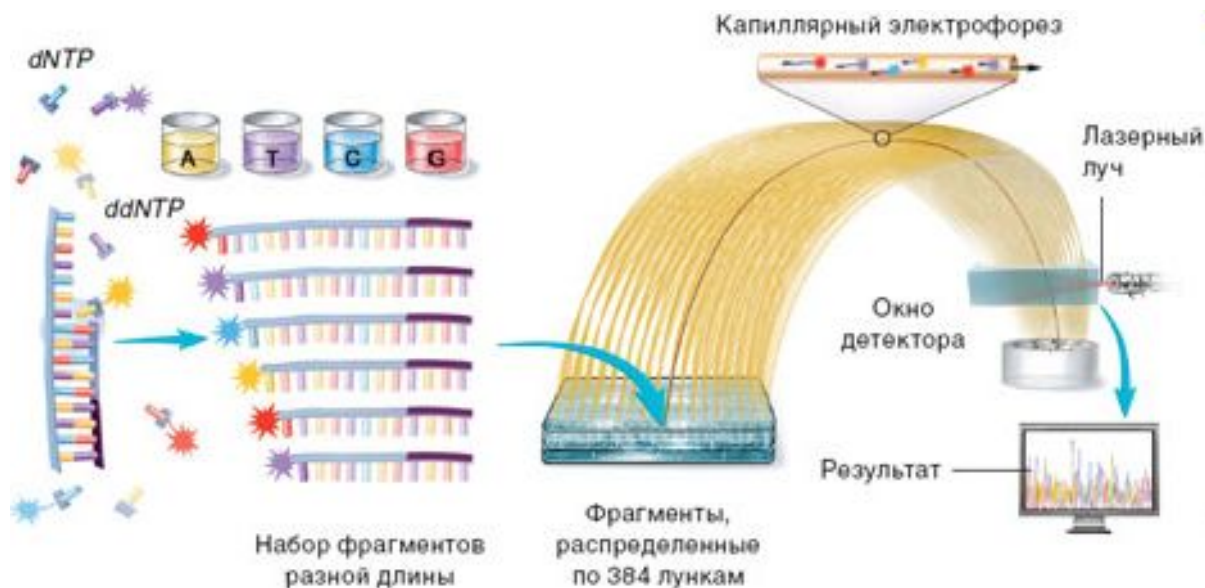
- **Результат секвенирования определяется качеством выделения матрицы, ее количеством, ее сложностью, выбором праймера, с которого начинается секвенирование;**
- **При соблюдении оптимальных условий в ходе проведения одной реакции секвенирования можно прочесть до 1000 нуклеотидов. Регулярно получаемая дальность чтения - 600-700 нуклеотидов;**
- **Контролем прохождения реакции служит проведение реакции на контрольной матрице pGEM-3Zf(+).**

Секвенирование ДНК по Сэнгеру

Раствор с одноцепочечными фрагментами и праймерами распределяют по четырём пробиркам, в каждую из которых добавлены четыре разные dNTP и один из флуоресцентно меченных дидезоксинуклеозидтрифосфатов (ddNTP). Удлинение гибридизовавшегося с ДНК-фрагментом праймера происходит до тех пор, пока в цепь не включится ddNTP. В этом месте синтез останавливается, и в результате в каждой из пробирок образуется уникальный набор отрицательно заряженных фрагментов разной длины, оканчивающихся одним из меченых ddNTP.

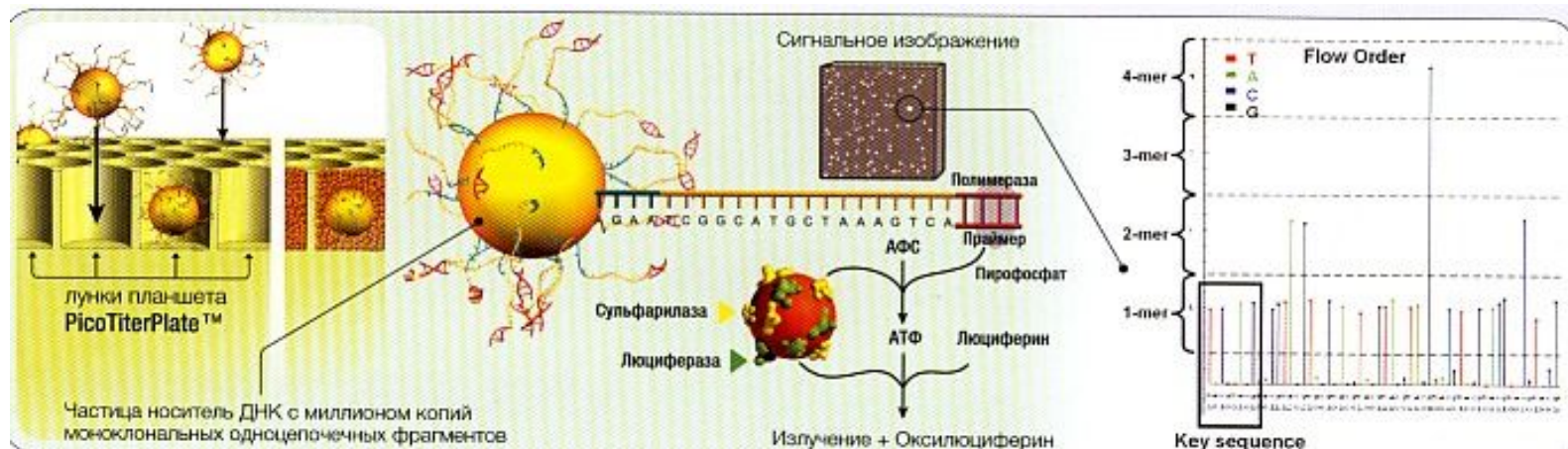
Секвенирование ДНК по Сэнгеру

- Фрагменты разделяют по размеру с помощью капиллярного электрофореза. Когда фрагменты определённой длины проходят через окно детектора, освещаемое лазерным лучом, ddNTP начинают флуоресцировать. Длина волны флуоресценции зависит от того, какой именно ddNTP находится у них на конце, так что на выходе получается цветная картинка, которую можно трансформировать в нуклеотидную последовательность.



Технология пиросеквенирования

- При прохождении нуклеотидов происходит одновременное секвенирование уникальных одноцепочечных ДНК на каждой частице, в каждой лунке планшета.
- Если через лунку проходит нуклеотид, комплементарный матрице, полимераза удлиняет цепь, встраивая этот нуклеотид.
- Добавление нуклеотида приводит к высвобождению пирофосфата и далее к реакции, генерирующей световой сигнал.
- Этот сигнал регистрируется ССD-камерой прибора.
- Интенсивность сигнала пропорциональна количеству нуклеотидов, встроенных в цепь ДНК за время одного прохода нуклеотидов.



Научная задача:

1. Провести генотипирование полиморфного участка **SNP 16725G>C** в гене *IFNAR1* (субъединица I рецептора интерферонов типа I) для 131 здорового человека и 226 больных рассеянным склерозом русского происхождения;
2. Определить аллельную частоту, частоту встречаемости аллелей и генотипов;
3. Проверить выборку больных и здоровых на соблюдение закона Харди-Вайнберга;
4. Сравнить аллельную частоту, частоту встречаемости аллелей и генотипов в группах больных и здоровых людей (метод случай-контроль).

Аллельные частоты, частоты встречаемости аллелей и генотипов по полиморфному участку C16725G гена IFNAR1 у больных рассеянным склерозом и здоровых индивидов

Аллели и генотипы	Больные РС 226 человека	Здоровые 131 человек
Частота аллелей, число (%)		
C	172 (38.1%)	93 (35.5%)
G	280 (61.9%)	169 (64.5%)
Частота встречаемости аллелей, число (%) носителей		
C	134 (59.3%)	80 (61.1%)
G	188 (83.2%)	118 (90.1%)
Частота генотипов, число (%) носителей		
C/C	38 (16.8%)	13 (10.0%)
G/G	92 (40.7%)	51 (39.0%)
C/G	96 (42.5%)	67 (51.0%)

закон Харди—Вайнберга (закон популяционного равновесия)

утверждает, что в теоретической идеальной популяции распределение генотипов будет оставаться постоянным из поколения в поколение и соответствовать уравнению:

$$p^2 + 2pq + q^2 = 1, \text{ где}$$

p^2 — доля гомозигот по одному из аллелей;

p — частота этого аллеля;

q^2 — доля гомозигот по альтернативному аллелю;

q — частота соответствующего аллеля;

$2pq$ — доля гетерозигот.

Соблюдение закона популяционного равновесия Харди-Вайнберга

Нулевая гипотеза: **популяция находится в равновесии**

$$\chi^2 = \frac{(N_{11} - \hat{p}^2 N)^2}{\hat{p}^2 N} + \frac{(N_{12} - 2\hat{p}\hat{q}N)^2}{2\hat{p}\hat{q}N} + \frac{(N_{22} - \hat{q}^2 N)^2}{\hat{q}^2 N}$$

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

O - наблюдаемая численность (OBSERVED)
E – ожидаемая численность (EXPECTED)

The Hardy–Weinberg expectation is:

$$\text{Exp}(AA) = p^2 n$$

$$\text{Exp}(Aa) = 2pqn$$

$$\text{Exp}(aa) = q^2 n$$

n – численность популяции

- 1) Степеней свободы: 1 (количество степеней свободы при проверке на соблюдение закона Х-В = количество генотипов - количество аллелей);
- 2) 5% уровень значимости для 1 степени свободы составляет 3.84;
- 3) **Т.к. величина χ^2 меньше 3.84, нулевая гипотеза не отвергается.**

- **p-value** – это вероятность справедливости нулевой гипотезы.
 $p=0.05$ (5%) показывает, что сделанный при анализе некоторой группы вывод является лишь случайной особенностью этих объектов с вероятностью только 5%. Другими словами, с очень большой вероятностью (95%) вывод можно распространить на все объекты.
- **Отношение шансов** (ОШ, Odds ratio, OR) - это отношение шансов события для первой группы объектов к шансам события для второй группы объектов.
Шанс – это отношение вероятности того, что события произойдёт к вероятности того, что событие не произойдёт.

Если ОШ=1, то шанс для первой группы равен шансу для второй группы

Если ОШ>1, то шанс для первой группы больше шанса для второй группы

Если ОШ<1, то шанс для первой группы меньше шанса для второй группы

- **Доверительный интервал** (ДИ, confidence interval, CI) для некоторой величины - это диапазон вокруг значения величины, в котором

Аллельные частоты, частоты встречаемости аллелей и генотипов по полиморфному участку 16725C>G гена *IFNAR1* у больных рассеянным склерозом и здоровых индивидов

Аллели и генотипы	Больные РС 226 человека	Здоровые 131 человек
Частота аллелей, число (%)		
C	172 (38.1%)	93 (35.5%)
G	280 (61.9%)	169 (64.5%)
Частота встречаемости аллелей, число (%) носителей		
C	134 (59.3%)	80 (61.1%)
G	188 (83.2%)	118 (90.1%)
Частота генотипов, число (%) носителей		
C/C	38 (16.8%)	13 (10.0%)
G/G	92 (40.7%)	51 (39.0%)
C/G	96 (42.5%)	67 (51.0%)

Аллельные частоты, частоты встречаемости аллелей и генотипов по полиморфному участку 16725C>G гена *IFNAR1* у больных рассеянным склерозом и здоровых индивидов

Аллели и генотипы	Больные РС 226 человека	Здоровые 131 человек	Значение <i>p</i> при сравнении частот	OR, 95%CI
Частота аллелей, число (%)				
C	172 (38.1%)	93 (35.5%)	Н.з.	
G	280 (61.9%)	169 (64.5%)	Н.з.	
Частота встречаемости аллелей, число (%) носителей				
C	134 (59.3%)	80 (61.1%)	Н.з.	
G	188 (83.2%)	118 (90.1%)	<i>p</i> =0.048	0.5 (0.3-1.1) Н.з.
Частота генотипов, число (%) носителей				
C/C	38 (16.8%)	13 (10.0%)	<i>p</i> =0.048	1.8 (0.9-3.6) Н.з.
G/G	92 (40.7%)	51 (39.0%)	Н.з.	
C/G	96 (42.5%)	67 (51.0%)	Н.з.	

Н.з. - не обнаружено значимых различий.

Аллельные частоты, частоты встречаемости аллелей и генотипов по полиморфному участку 1444С>Т гена *CRP* у пациентов с инфарктом миокарда и здоровых индивидов

Аллели и генотипы	Пациенты с ИМ 218 человека	Здоровые индивиды 153 человека	Значение p	OR (95%CI)
Аллели, число (%) аллелей				
С	301 (69.5%)	225 (73.5%)	н.з.	
Т	135 (31%)	81 (26.5%)	н.з.	
Аллели, число (%) носителей				
С	193(88.5%)	146 (95.4%)	0.01	0.4 (0.2-0.9)
Т	110 (50.5%)	74 (48.4%)	н.з.	
Генотипы, число (%) носителей				
С/С	108 (49.5%)	79 (51.6%)	н.з.	
С/Т	85 (39%)	67 (43.8%)	н.з.	
Т/Т	25 (11.5%)	7 (4.6%)	0.01	2.7 (1.1-6.4)