

Занятие №7. Секвенирование НК. Секвенирование по Сэнгеру.

Технологические варианты секвенирования

- **Методы секвенирования I поколения** (по Сэнгеру, по Максаму – Гилберту, автоматическое секвенирование по Сэнгеру)
- **Методы секвенирования II поколения**
 - – пиросеквенирование
 - – секвенирование с лигированием
 - – полупроводниковое секвенирование
 - – методы синтеза с обратимым терминированием
- **Методы секвенирования III поколения**
 - – секвенирование с помощью нанопоры

Методические варианты секвенирования

- 1. Секвенирование ДНК (геном) / Секвенирование РНК=>кДНК (транскриптом)
- 2. Ресеквенирование / Секвенирование de novo
- 3. Целевое секвенирование (target sequencing) (напр., экзом, 16S рРНК, позиционные методы) / Полногеномное секвенирование (Whole Genome Sequencing, WGS)
- 4. Секвенирование в пределах одного вида / Метагеномное секвенирование
- 5. Секвенирование ДНК совокупности клеток / Секвенирование ДНК одиночных клеток (single-cell sequence)
- 6. Определение первичной структуры / Определение 3D структуры

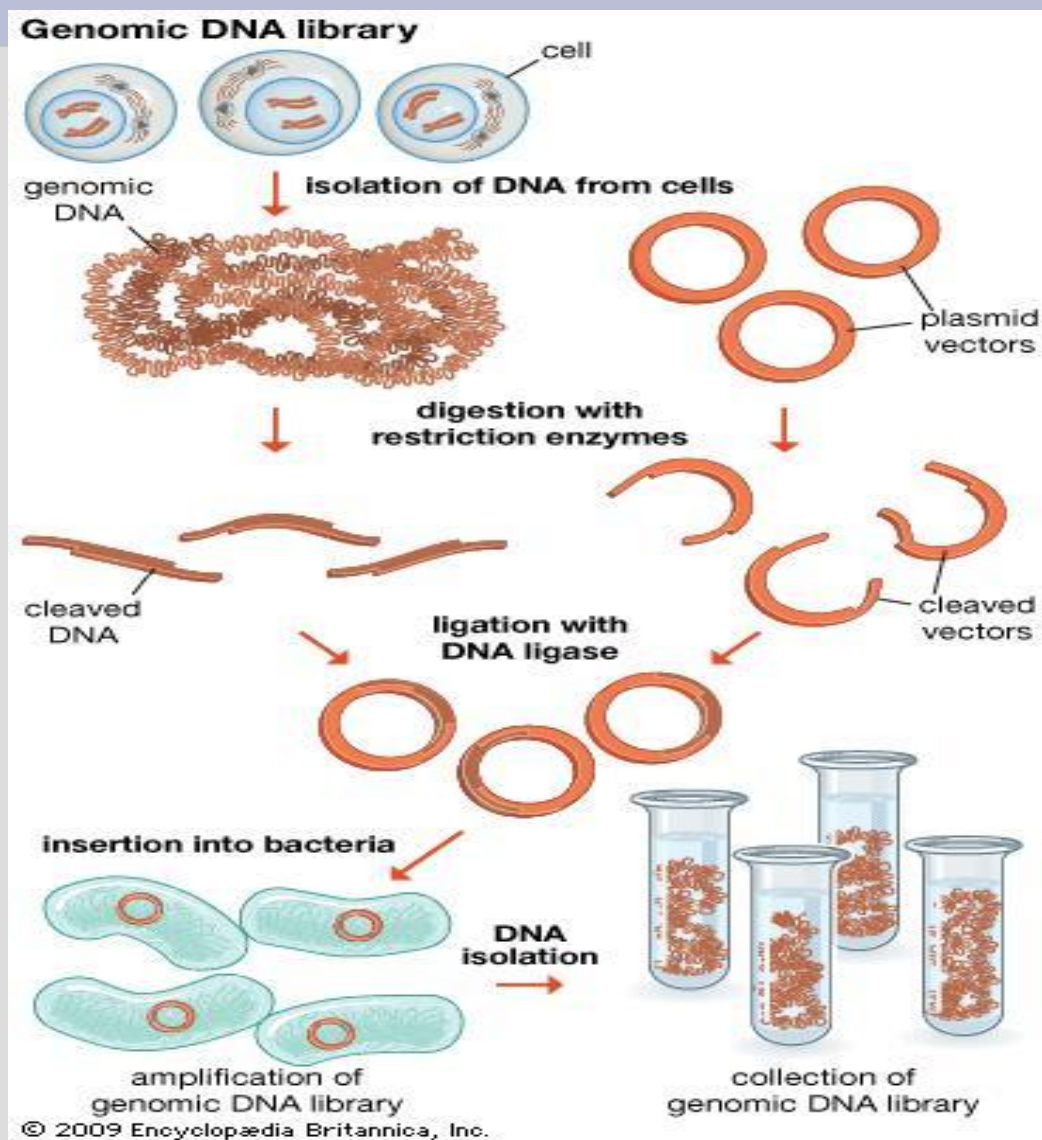
Этапы секвенирования

- 1. Выделение ДНК.
- 2. Создание геномной библиотеки (полногеномное секвенирование) или амплификация участка ДНК (целевое секвенирование).
- 3. Секвенирование.
- 4. Анализ данных.

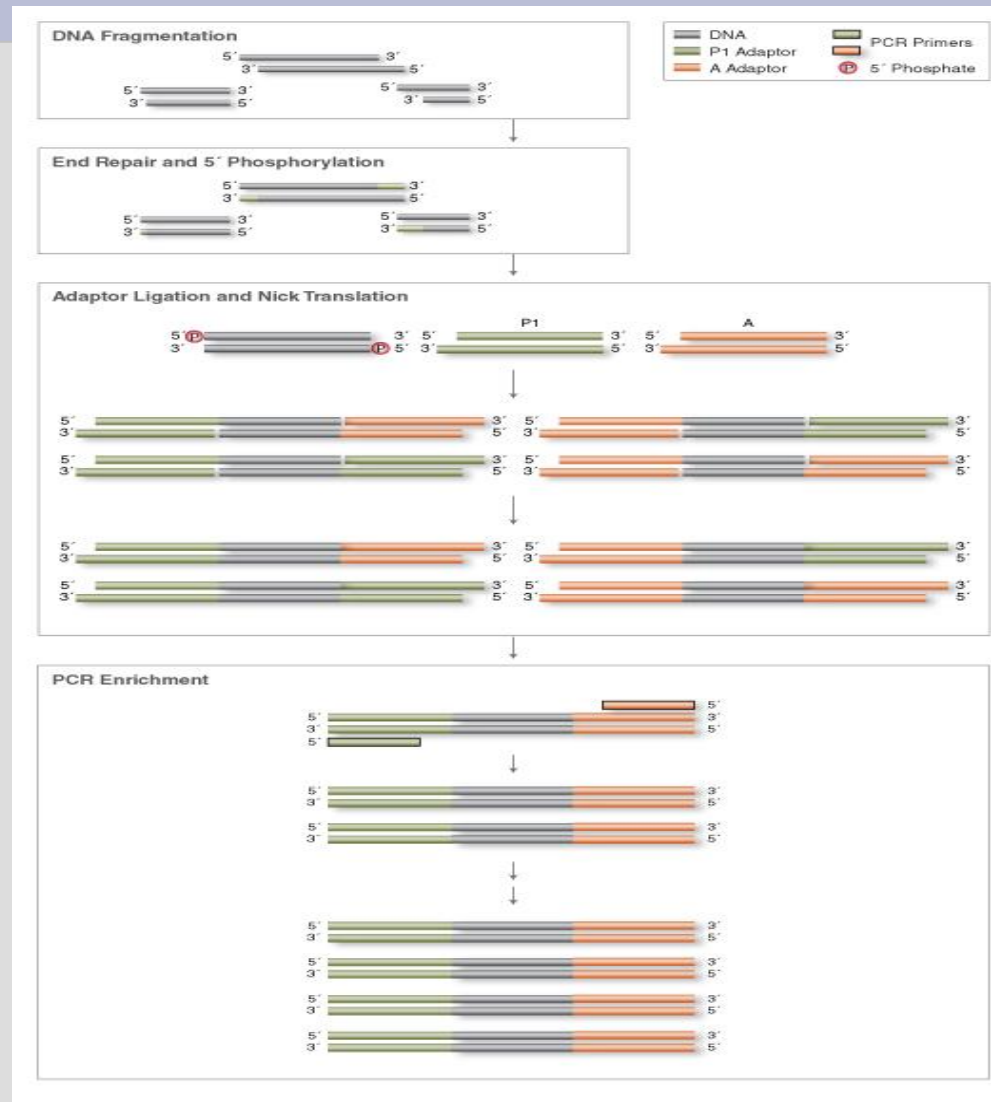
Геномные библиотеки (ДНК-библиотеки)

- **Геномная библиотека** – набор фрагментов одного генома, каждый фрагмент уникален. Для секвенирования необходимо получение сигнала от множества идентичных фрагментов, поэтому из каждого фрагмента геномной библиотеки получают множество копий – клональная библиотека.
- Методы получения геномных библиотек:
- Клонирование в вектор
- ПЦР

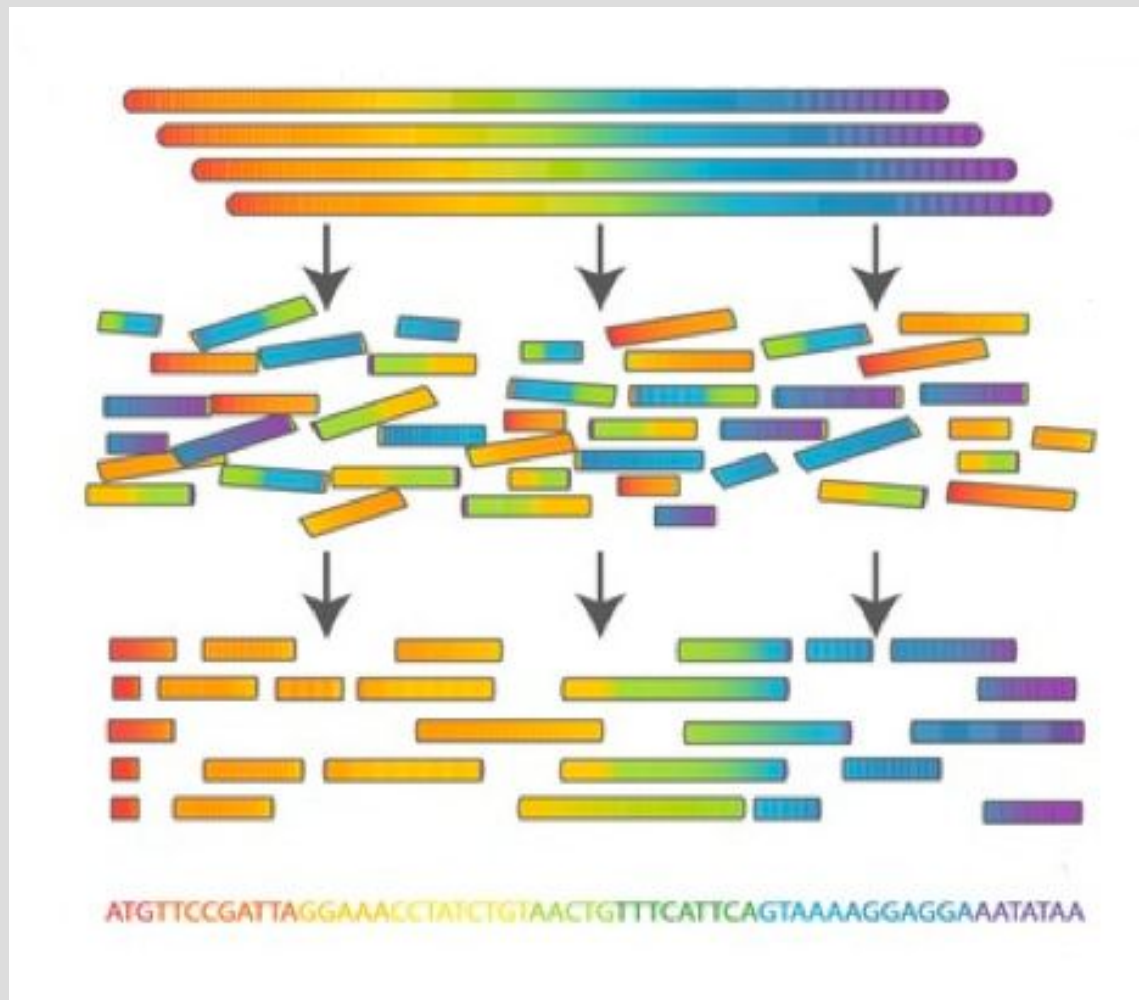
Клонирование в вектор



Создание ДНК-библиотек с помощью ПЦР



Концепция shotgun-секвенирования – получение геномных библиотек



Целевое (таргетное) секвенирование

- В процессе приготовления библиотеки необходимо выделить только целевые участки НК:
- 1). ПЦР со специфическими праймерами.
- 2). Гибридизация со специфическими олигонуклеотидами, меченными биотином – выделение на магнитных частицах, покрытых стрептавидином.
- 3). Выделение на магнитных частицах, покрытых антителами к белкам, связанным с целевой последовательностью (напр., гистон в ChIP-Seq).

Секвенирование по Максаму – Гилберту

- Основа метода – химическое расщепление множества копий одной целевой молекулы ДНК в позициях определённых нуклеотидов.

Этапы секвенирования по Максому – Гилберту

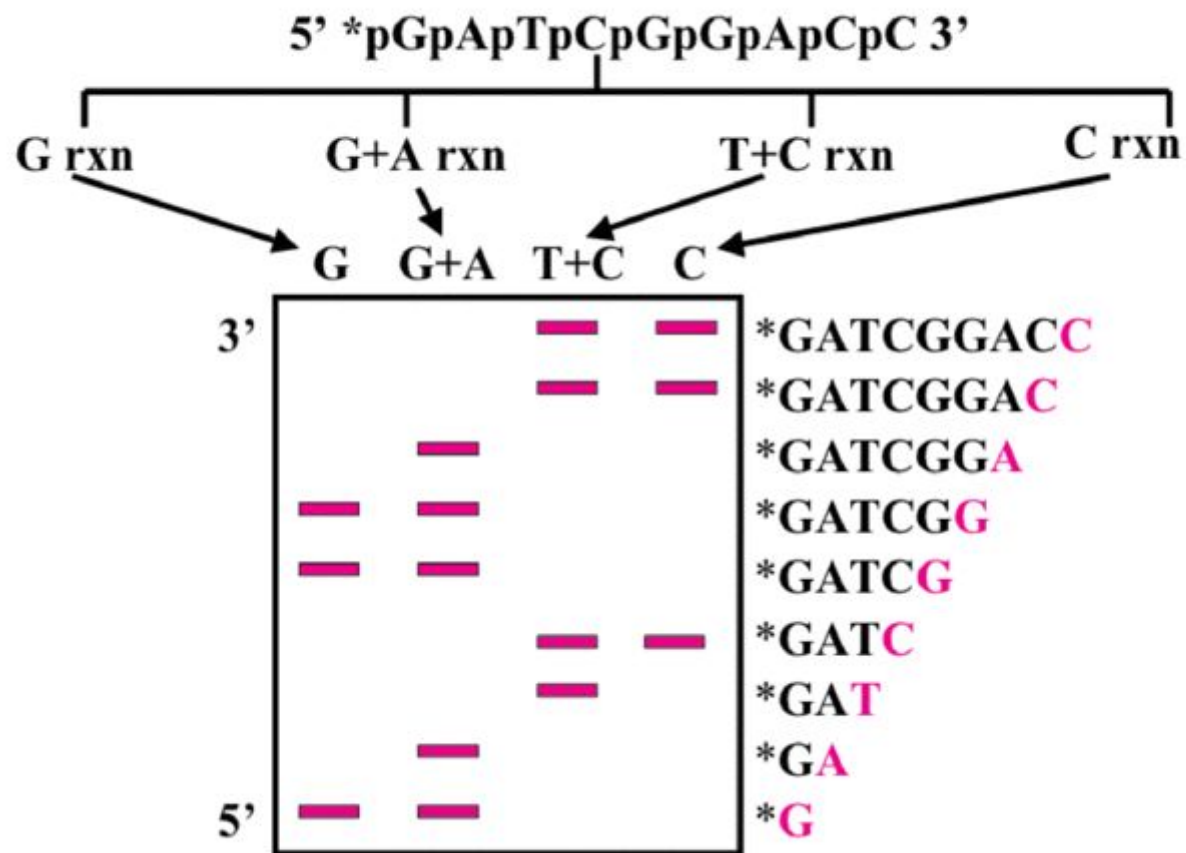
- 1). Получение множества копий целевой последовательности (100-1000 п.о.).
- 2). Мечение молекул ДНК на 5'-конце с помощью ^{32}P -фосфата.
- 3). Разделение смеси на 4 части.
- 4). Двухстадийный гидролиз молекул в каждой части (4 реакции) по определённому основанию или двум химически близким основаниям.
- 5). Электрофорез с радиографической детекцией.

Химический гидролиз по основаниям

Реакция (нуклеотиды)		

Схема метода Максама – Гилберта

Maxam-Gilbert sequencing

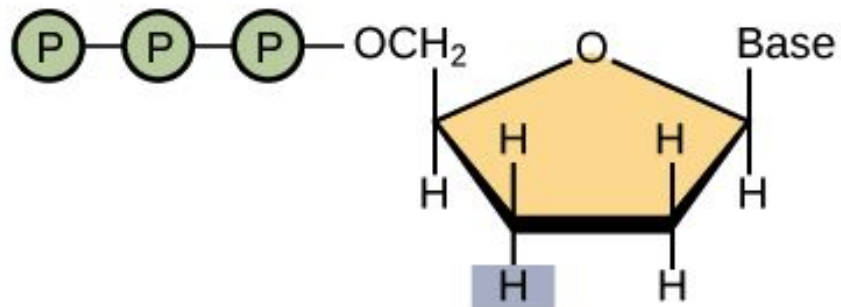


Характеристика секвенирования по Максаму – Гилберту

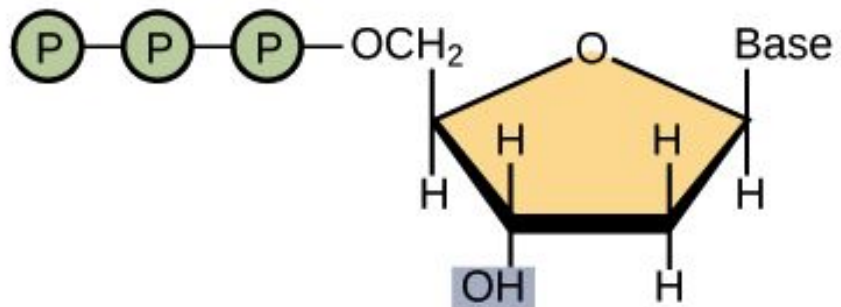
- **Преимущества:**
- Высокая точность
- Большая длина прочтения (300-1000 п.о.)
- Не требует знания концевых нуклеотидов целевой последовательности (не нужны праймеры)
- “Не боится” вторичных структур ДНК
- **Недостатки:**
- Низкая производительность
- Высокая стоимость в пересчёте на производительность
- Высокая трудоёмкость и опасность работы

Практически не используется (“преигрывает другим методам”)

Основа метода Сэнгера: терминация синтеза



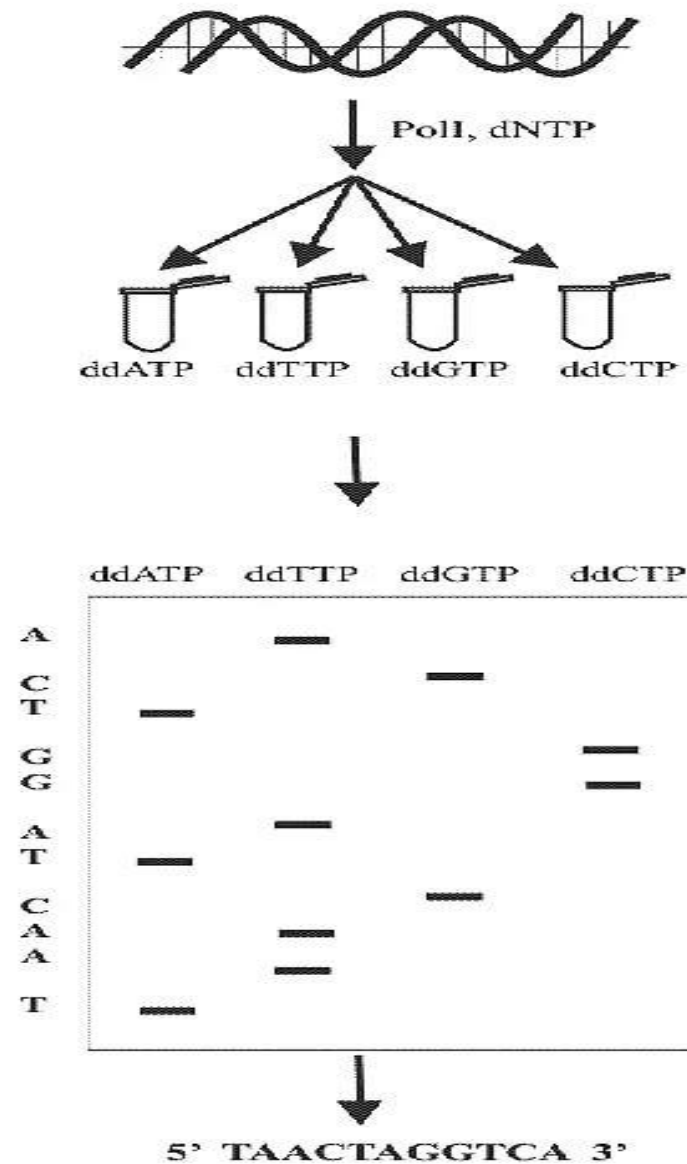
Dideoxynucleotide (ddNTP)



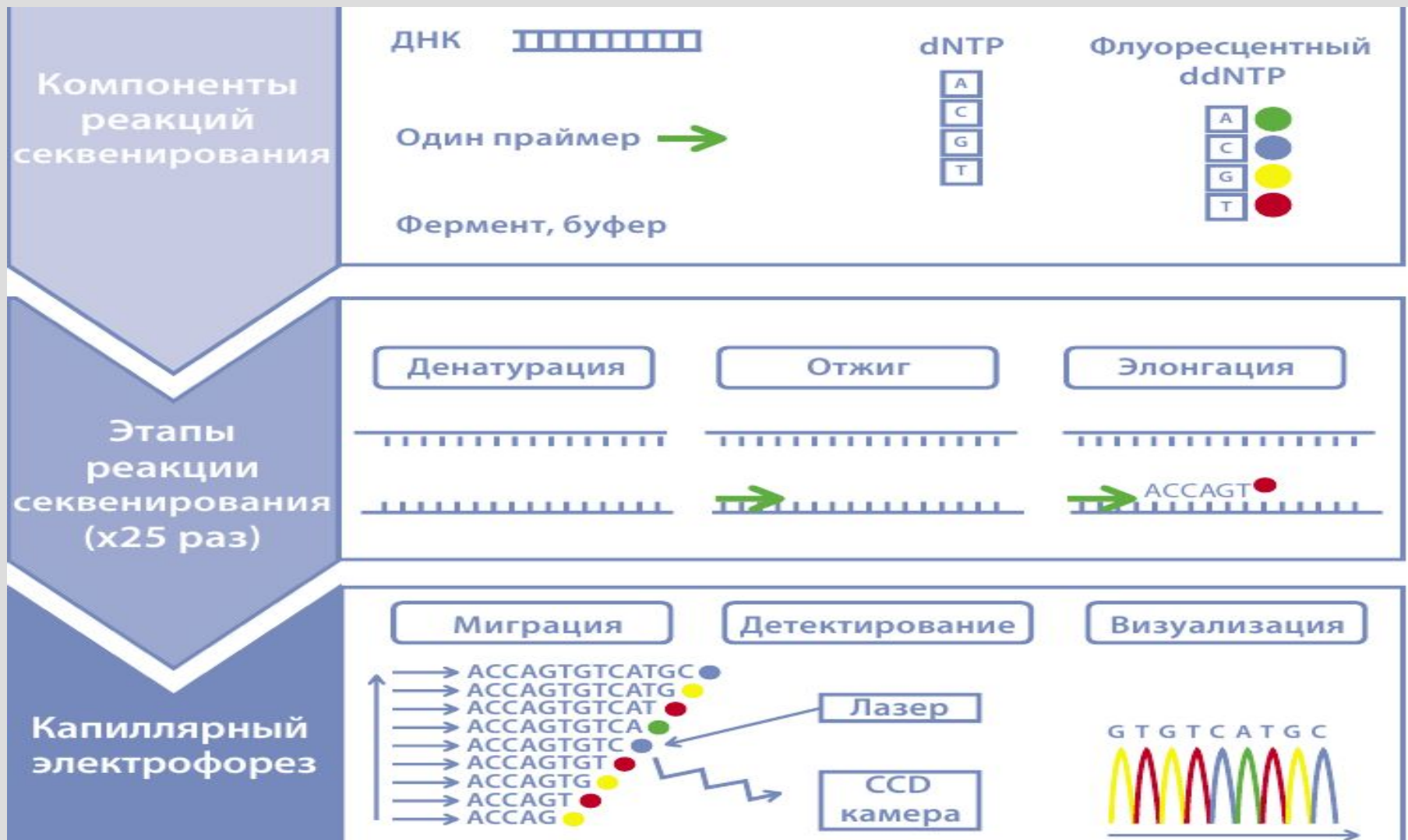
Deoxynucleotide (dNTP)

$$[\text{ddNTP}]/[\text{dNTP}] = 1/20$$

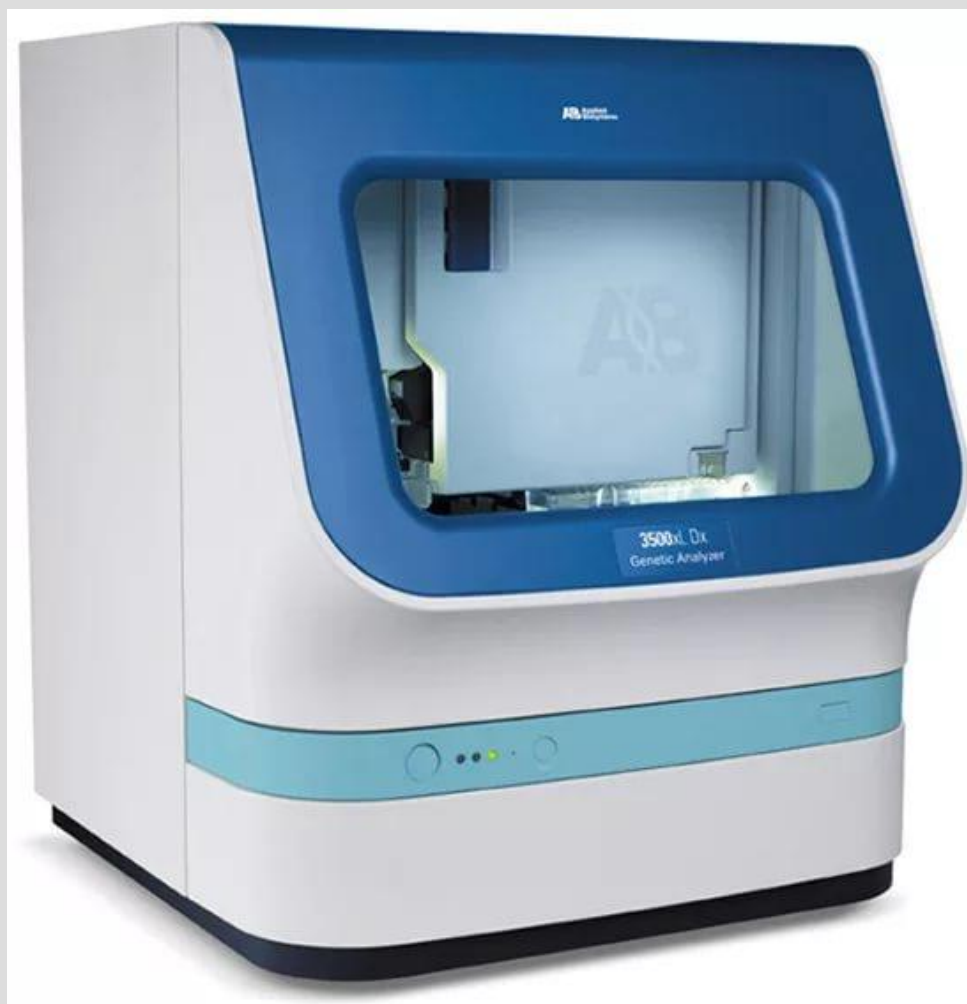
Классическая схема Сэнгера (1977 год)



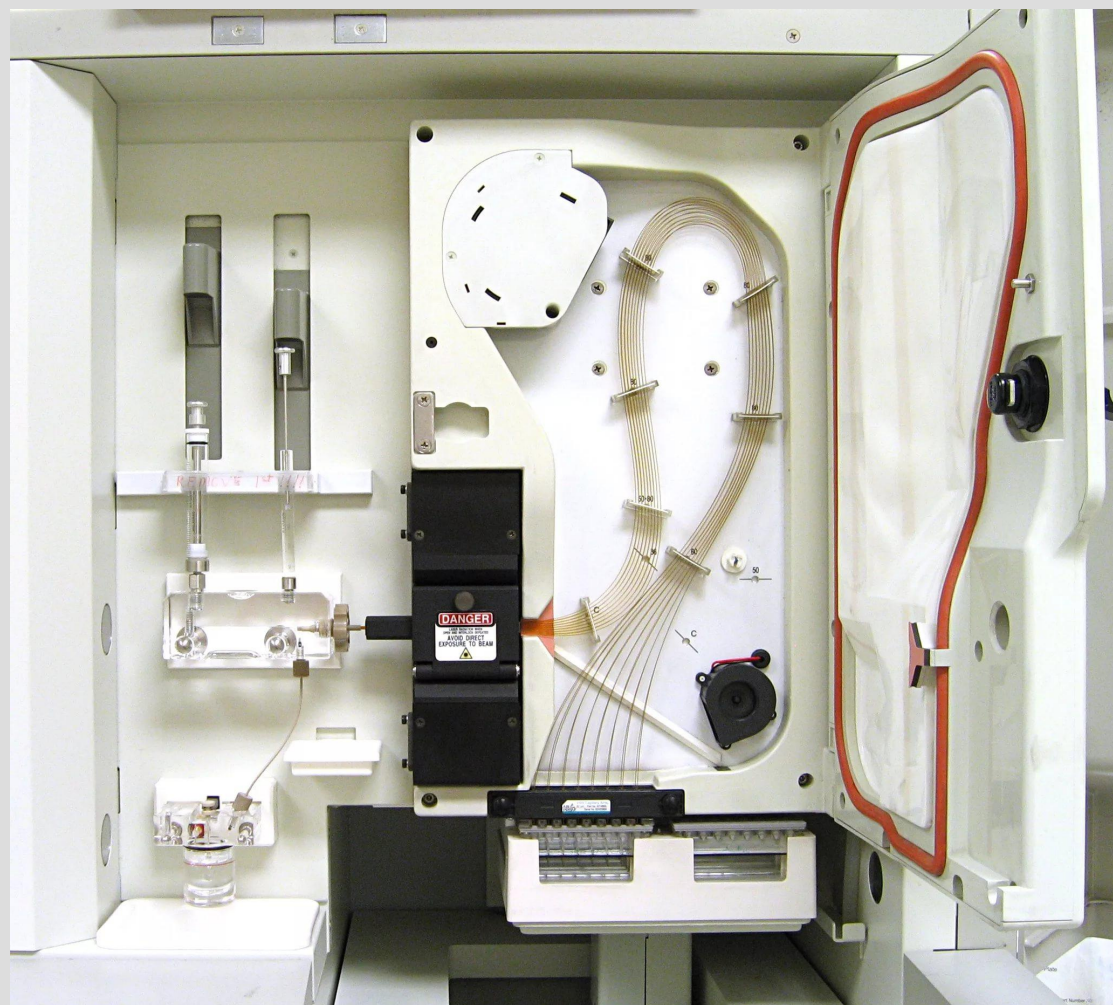
Автоматическое секвенирование по Сэнгеру (1990 год)



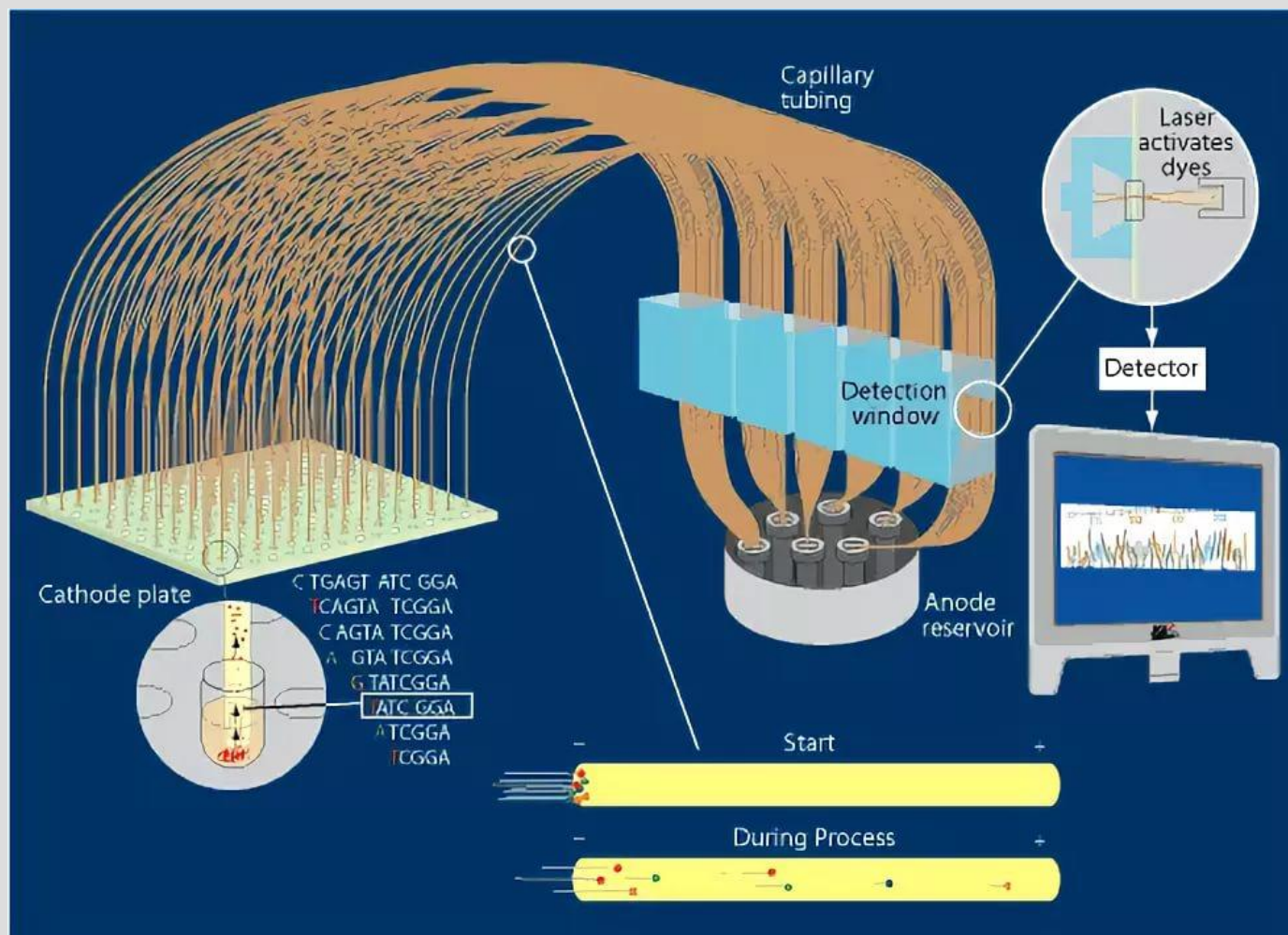
Автоматический секвенатор (генетический анализатор) (Applied Biosystems 3500 xL)



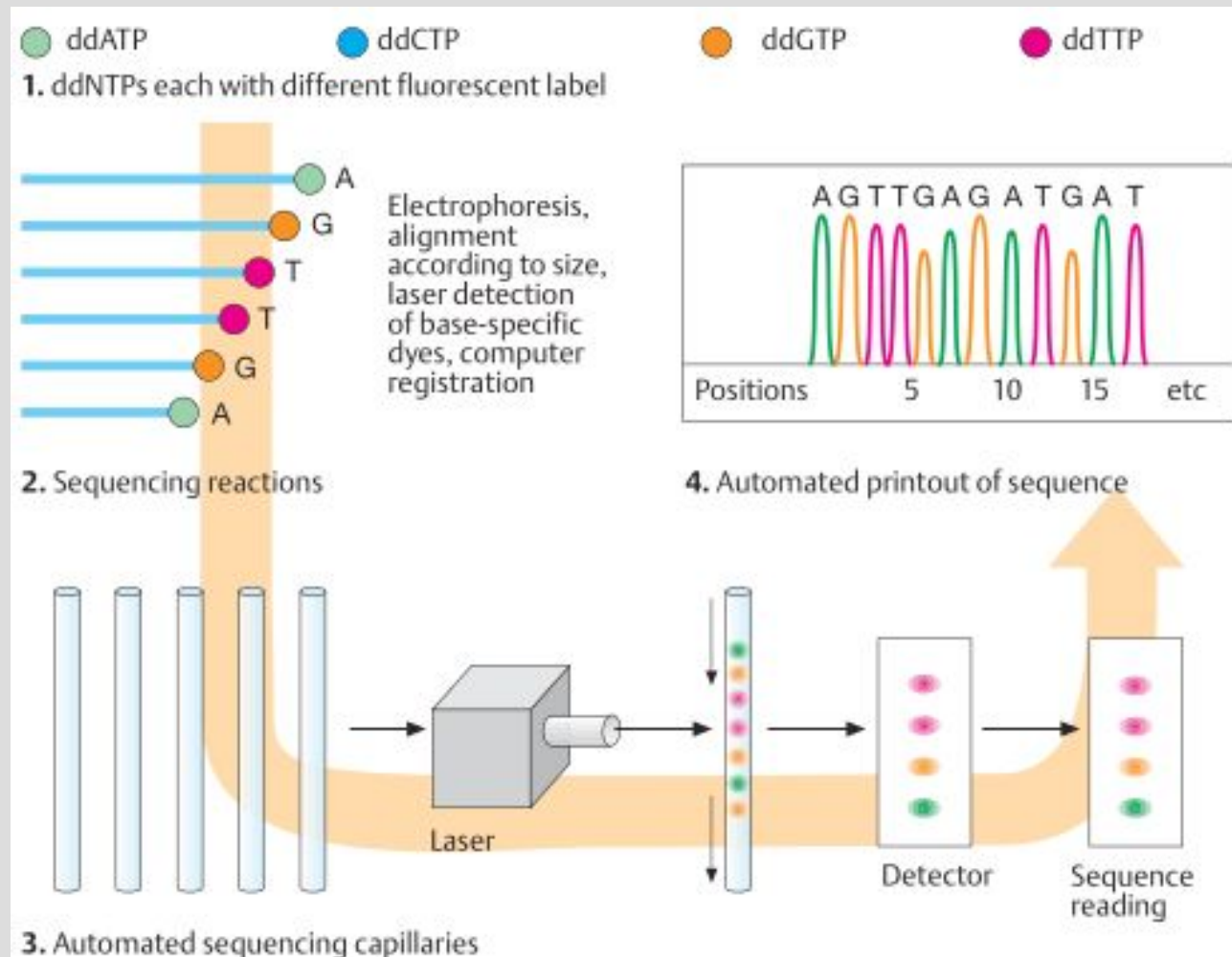
Генетический анализатор с капиллярным электрофорезом изнутри



Капиллярный электрофорез



Лазерная детекция флюоресценции

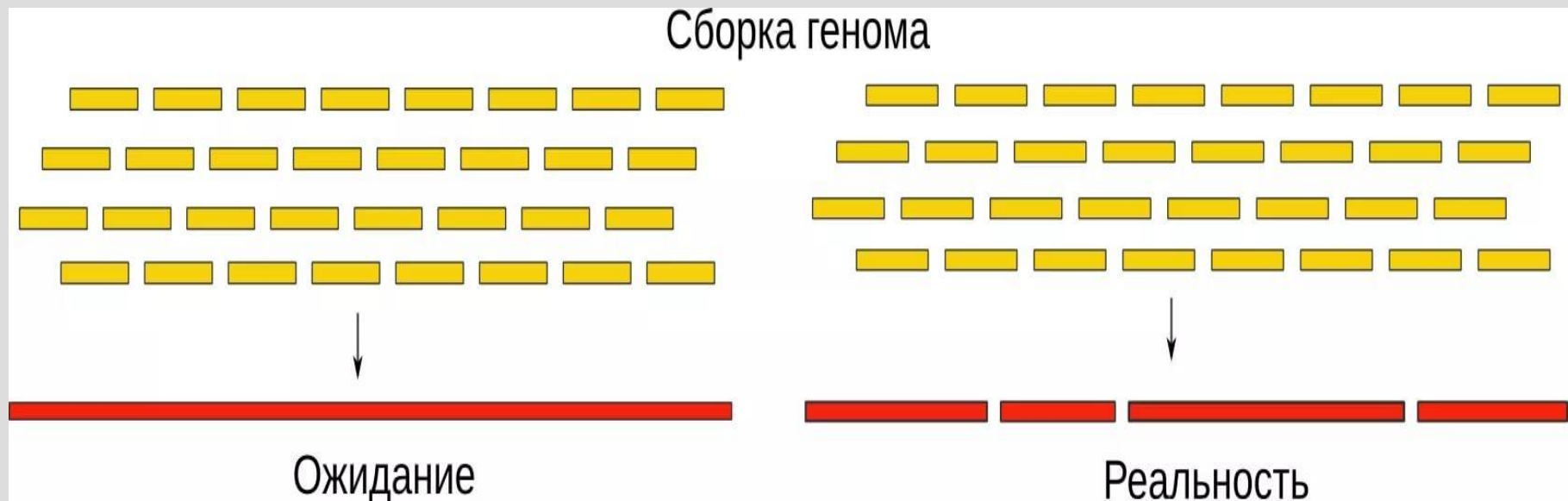


Параметры метода (платформы) секвенирования

- Длина прочтения (длина рида, read length)
- Производительность (выход, output)
- Уровень ошибки
- Время одного запуска
- Количество образцов на один запуск
- Стоимость секвенирования

Секвенирование по Сэнгеру: Длина прочтения (длина рида)

- Результат секвенирования – набор последовательностей (прочтений) сходной длины.
- Длина рида для метода Сэнгера составляет до 1000 п.о.
- Фрагмент ДНК “читается” не целиком, если он слишком большой (при форезе сложнее различить 50 и 51 п.о., чем 500 и 501 п.о.)



Секвенирование по Сэнгеру: Выход секвенирования (производительность)

- Выход равен количеству прочитанных нуклеотидов за один запуск:
 - $O = l * n$
 - l – средняя длина рида
 - n – количество ридов за запуск
- Метод Сэнгера: 70-96 кбаз (кб, kb) (ср.: геном человека ~ 3 Gb)

Секвенирование по Сэнгеру: Количество образцов на один запуск

- Количество образцов определяется исходя из выхода секвенирования и размера секвенируемых последовательностей.
- Для метода Сэнгера количество образцов не может быть больше количества капилляров в автоматическом анализаторе (т.е. ≤ 96).

Секвенирование по Сэнгеру: Уровень ошибки секвенирования

- **Уровень ошибки** – доля нуклеотидов, прочитанных ошибочно (%).
- Источники ошибок в методе Сэнгера:
 - 1). Качество ридов начинает падать ближе к концу (после 700 нуклеотида), т.к. форез не позволяет различать длинные фрагменты, отличающиеся на 1 нуклеотид.
 - 2). Плохо секвенируются гомополимеры (..TTTTTTT..), т.к. их сигналы могут сливаться.
- Уровень ошибки в методе Сэнгера: **0,1-1% (самый точный метод!)**

Секвенирование по Сэнгеру: Время одного запуска

- **2-3 часа:**
 - – 2 часа на терминирующую ПЦР
 - – 40 минут на электрофорез
- Метод позволяет запускать генетический анализатор бесперебойно в течение 48 часов.

Секвенирование по Сэнгеру: Стоимость (в пересчёте на производительность)

- Самый дорогой метод: стоимость выше ~ в 10 тыс. раз (!) по сравнению с другими методами (в пересчёте на производительность).
- Основная статья расходов – реагенты для терминирующей ПЦР (**флюоресцентно меченные ddNTP**)

Характеристика секвенирования по Сэнгеру

- **Преимущества:**
 - Низкая частота ошибок (0,1-1%)
 - Большая длина прочтения (300-1000 п.о.)
- **Недостатки:**
 - Низкая производительность
 - Высокая стоимость в пересчёте на производительность
 - Высокая трудоёмкость

Оптимальное применение секвенирования по Сэнгеру

- 1). Секвенирование небольших геномов (напр., вирусы) или целевое секвенирование
- 2). Корректировка ошибок в данных высокопроизводительного секвенирования (NGS)