

БИОИНЖЕНЕРИЯ

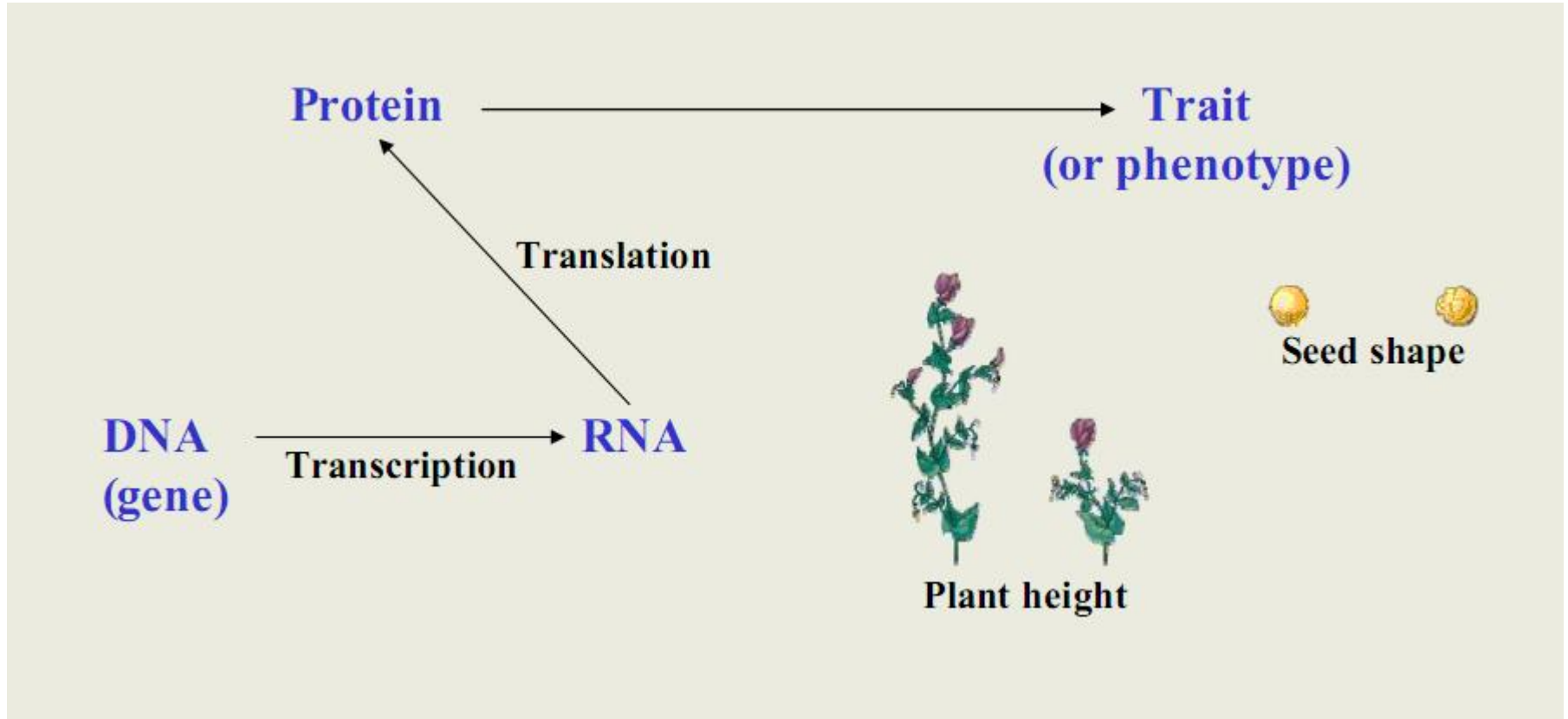
Сергей Лукьянов

Институт биоорганической химии имени академиков М.М. Шемакина и Ю.А. Овчинникова РАН

Лекция 1

**Центральная догма молекулярной биологии,
основные понятия - ген, геном, генотип, фенотип.
Основы генной инженерии: экзонуклеазы
рестрикции, вектора, молекулярное клонирование**

Центральная догма молекулярной биологии (Ф. Крик, 1958)



Формирование фенотипа происходит под действием генотипа.

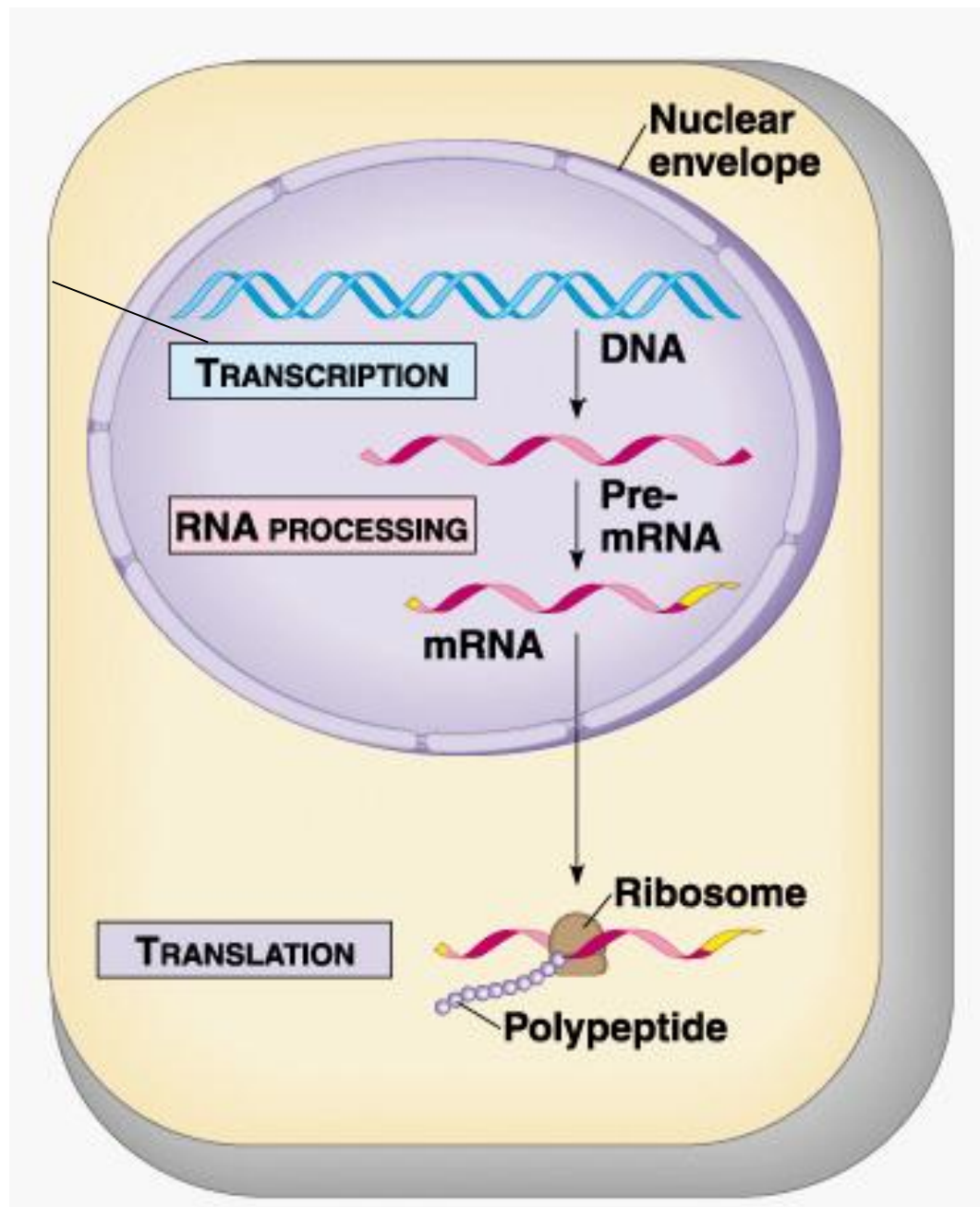
Клонированная самка муфлона, родившаяся у овцы, использованной в качестве приемной матери



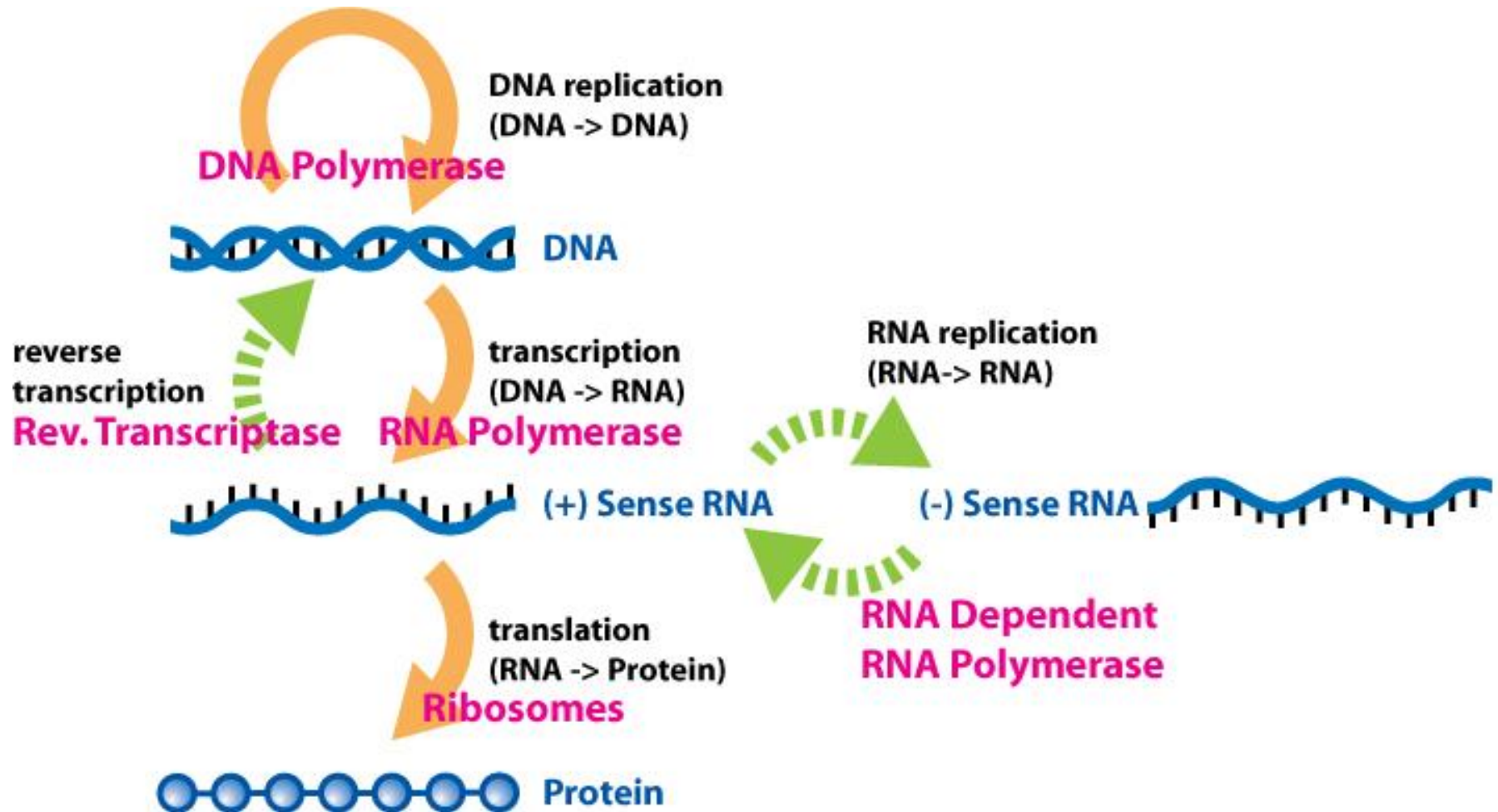
**Перенос ядра
соматической
клетки погибшего
муфлона в
энуклеированную
яйцеклетку овцы**

**Нормальные роды
после 155 дней
беременности**

Путь реализации генетической информации



Пересмотренная центральная догма молекулярной биологии



Определение терминов «Генотип» и «Фенотип»

- Генотип — совокупность генов данного организма
- Фенотип — совокупность характеристик, присущих особи (клетке, комплексу молекул, молекуле) на определённой стадии развития
- Наследуемые изменения фенотипа организма связаны с изменениями его генотипа
- Ген – фрагмент нуклеиновой кислоты (НК), в последовательности которой закодирована информация о последовательностях других НК или белков

Концепция гена

60-е годы

Ген – непрерывная кодирующая последовательность ДНК или РНК

Один ген кодирует один белок (РНК)

Последовательность гена колинеарна кодируемому белку

Регуляторная часть предшествует структурной

Ген имеет четкие границы

Ген имеет постоянную локализацию на хромосоме

Перемещения и изменения генов происходят только вследствие случайных мутаций

Современные представления

Ген содержит некодирующие участки (интроны)

Могут использоваться все три ОРС, альтернативный сплайсинг

Сплайсинг белков, редактирование РНК

Энхансеры находятся перед, внутри и за геном

Альтернативные сайты инициации и терминации транскрипции и трансляции

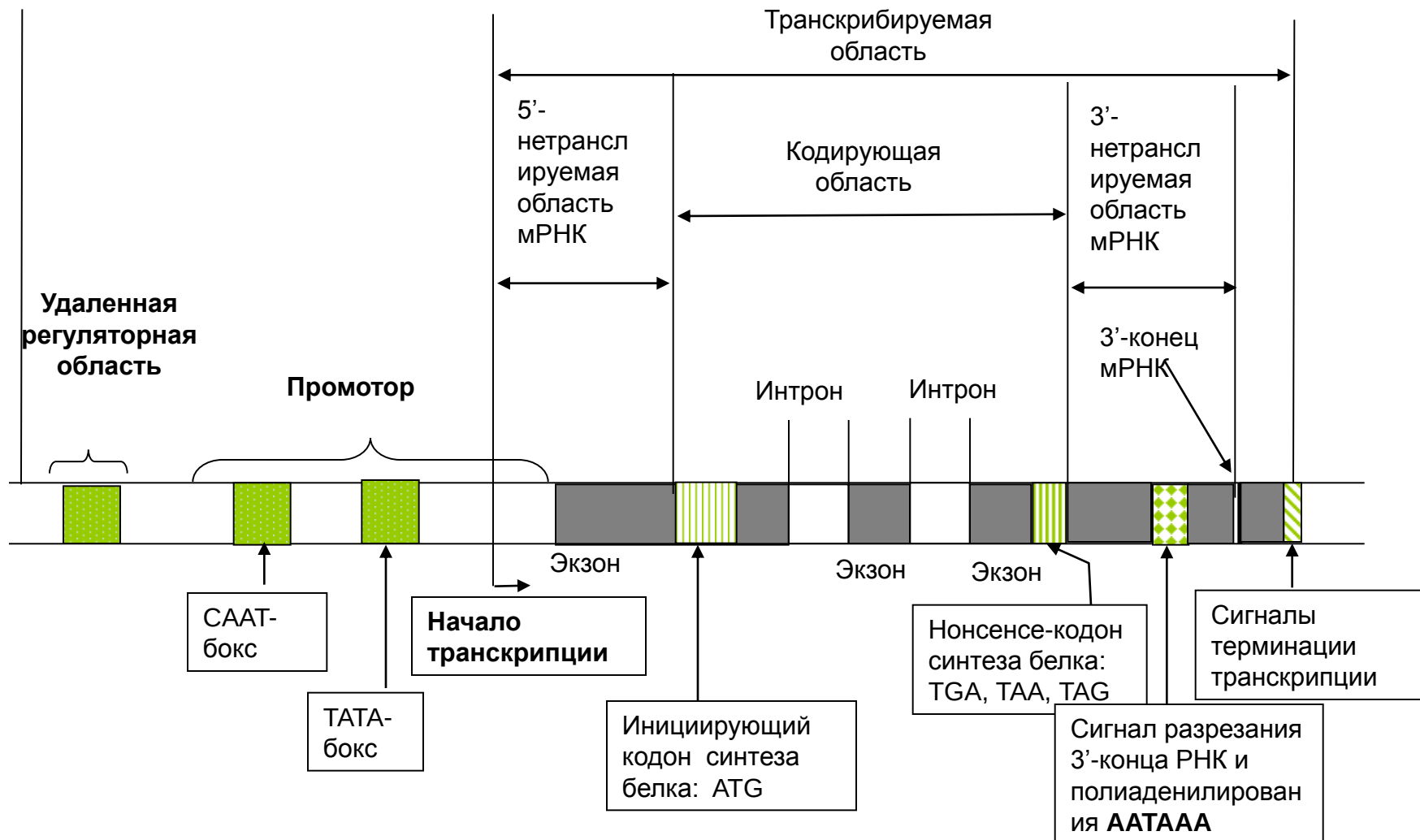
Мобильные генетические элементы

Запрограммированные перестройки генов Ig, соматический мутагенез, адаптивные мутации

Структурные элементы гена

5'-конец
гена

3'-конец
гена



Коровый промотор:

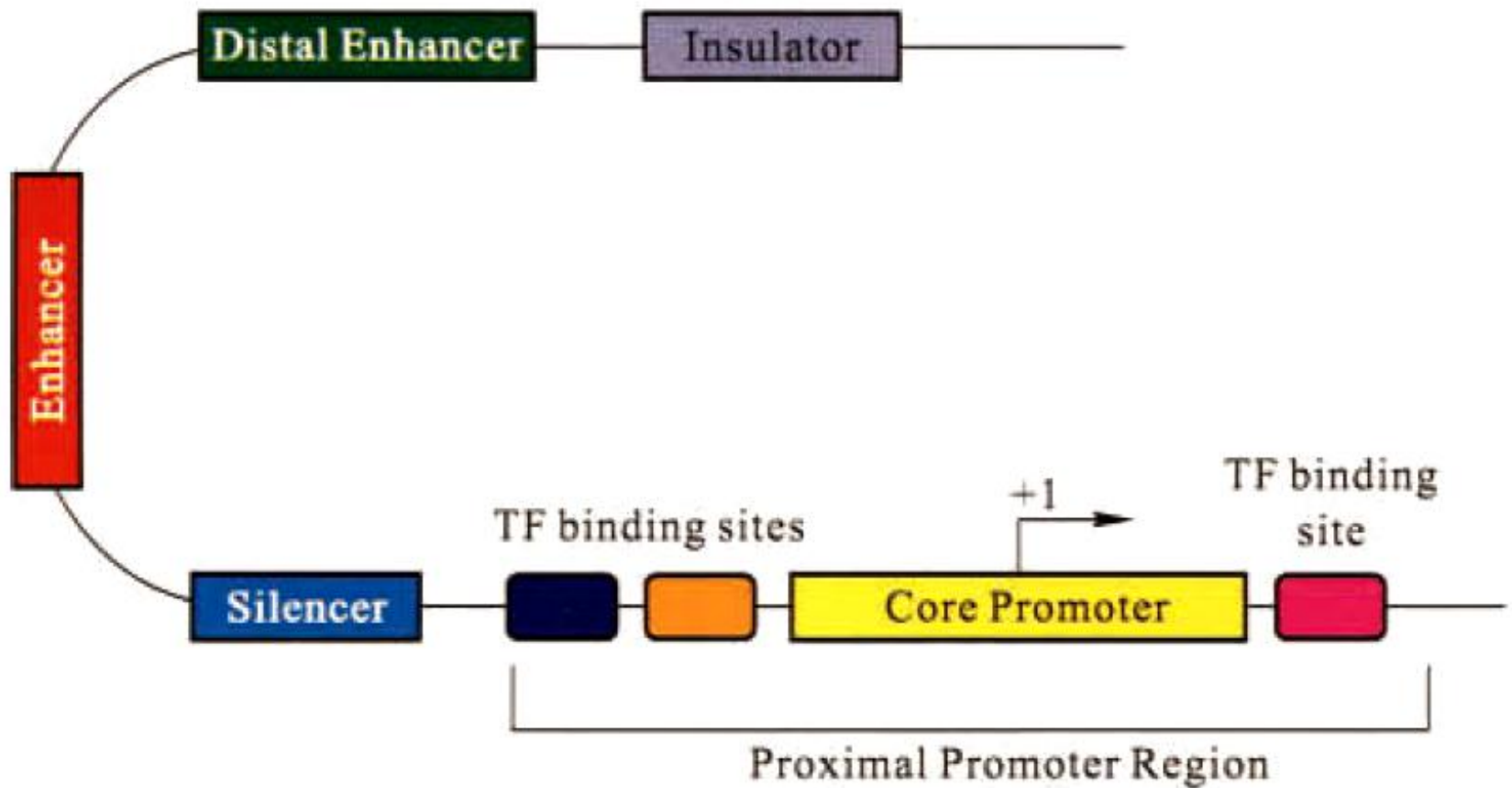
минимальная последовательность нуклеотидов, обеспечивающая правильную инициацию транскрипции в отсутствие других цис-действующих элементов

Энхансеры и сайленсеры обеспечивают (ткане)специфическую транскрипцию конкретных генов, стимулируя или подавляя их экспрессию, соответственно

Инсуляторы ограничивают действие энхансеров и сайленсеров на соседние гены

Проксимальная промоторная область включает коровый промотор и сайты связывания факторов транскрипции, влияющих на его активность

Модули, контролирующие транскрипцию, в эукариотических генах, кодирующих белки



TF - transcription factor

Определение термина «Геном»

Геном – совокупность всей ДНК гаплоидного набора хромосом, внехромосомных генетических элементов и органелл клетки зародышевой линии биологического вида

- Введен Г. Винклером в 1922 г.
- В отличие от термина «генотип» является биологической характеристикой вида в целом, а не отдельной особи
- Отличием генотипа от генома является включение в понятие "геном" некодирующих последовательностей, не входящих в понятие "генотип".
- Из-за большого числа аллельных вариантов генов и некодирующих последовательностей можно говорить лишь об усредненном геноме биологического вида (у человека ~1,500,000 SNP)
- Внехромосомные ДНК митохондрий и хлоропластов являются частью генома

Размеры геномов

Единицы измерения размера генома:

- kilobase (kb) - 10^3 пар оснований (т.п.н.)
- megabase (Mb) - 10^6 пар оснований (млн.п.н.)
- пикограммы (pg) \ 1 pg ДНК $\sim$ 1000 млн.п.н.

Размер генома

- Геномы вирусов – 10-1000 kb

Бактериофаг MS2 – один из самых маленьких вирусов, имеет всего 4 гена на единственной одноцепочечной РНК длиной 4000 нуклеотидов (4kb)

- Геномы бактерии – 1-10 Mb

Хромосома *Escherichia coli* – кольцевая ДНК длиной 4600 kb.

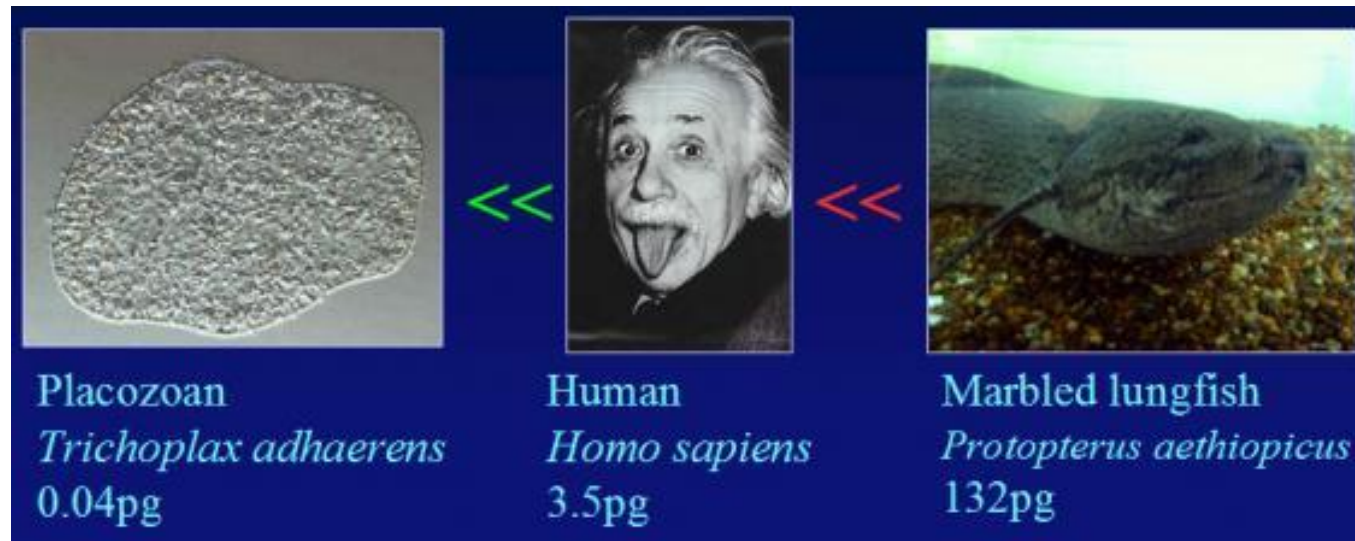
- Геномы эукариот – 10-100 000 Mb

Геном *Drosophila melanogaster* – 180 Mb

Различия в размерах геномов эукариот даже близких видов очень велики

Парадокс С (C-value paradox)

Размер эукариотического генома *не коррелирует* с биологической сложностью видов и их положением в эволюционной иерархии (С.А. Thomas, 1971 г.)



Вариации в размерах генома различных видов внутри разных систематических групп:

- Простейшие – 5800 раз
- Водоросли – 5000 раз
- Покрытосемянные – 1000 раз
- Рыбы – 350 раз
- Членистоногие – 250 раз

C-value = размер гаплоидного генома организма

Вариации в размерах генома различных видов внутри разных систематических групп

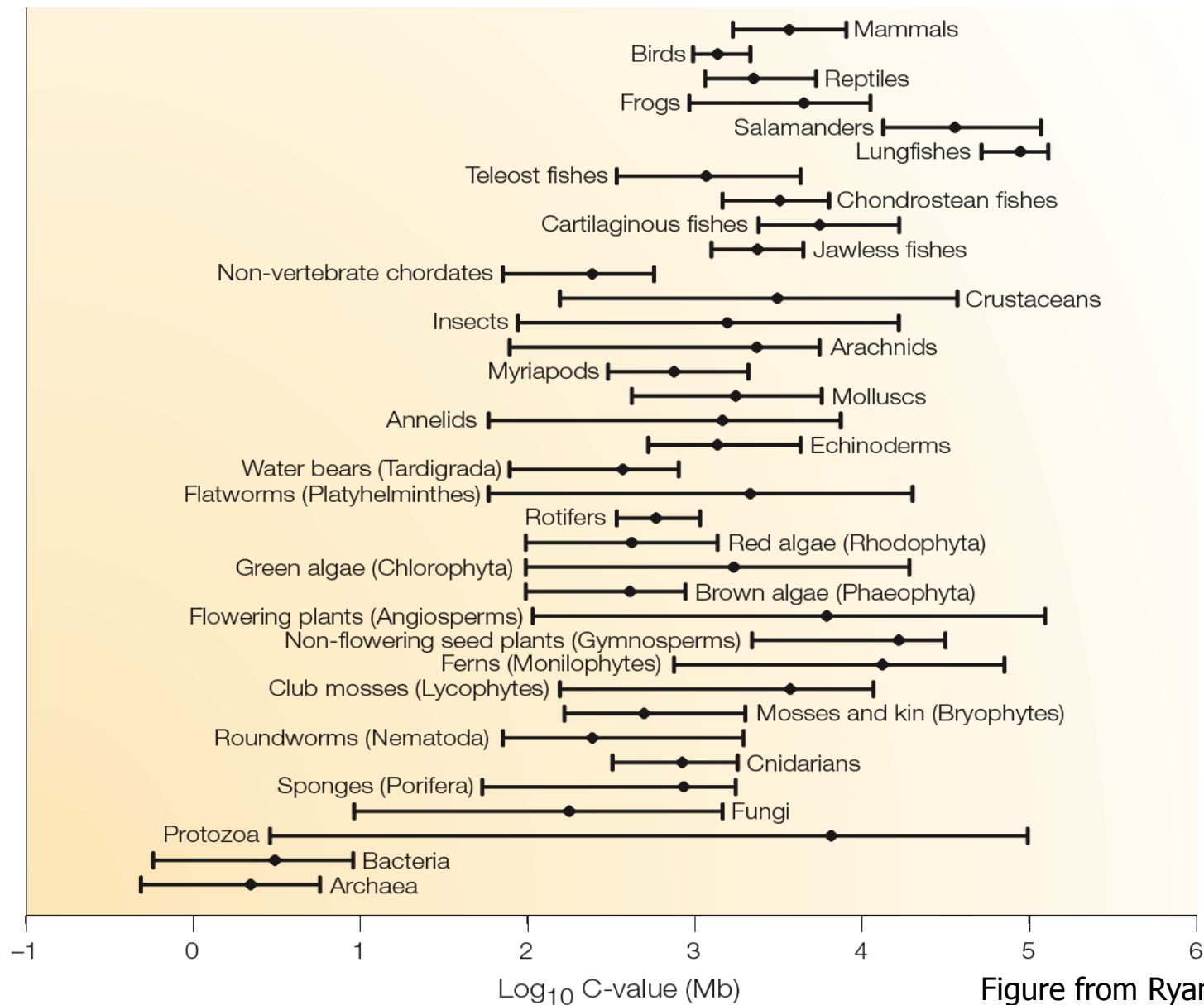
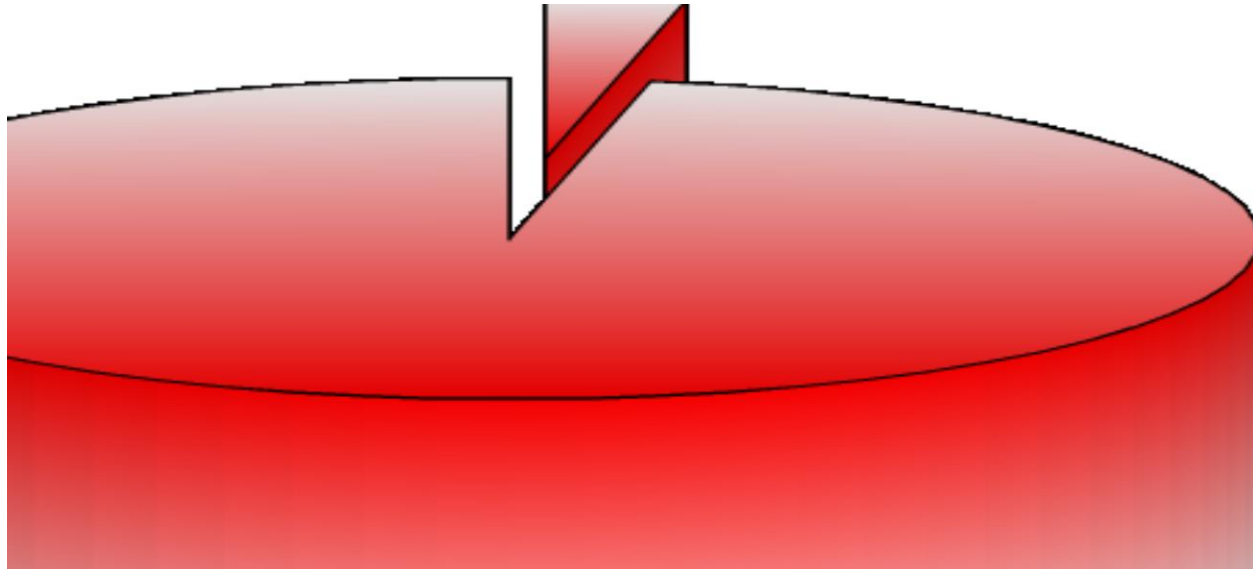


Figure from Ryan Gregory (2005)

Большие различия в размерах геномов определяются последовательностями, не кодирующими белки и нуклеиновые кислоты



В геноме человека доля кодирующих последовательностей $\sim 1.5-2\%$

Последовательности нуклеотидов генома человека



Размер эукариотического генома не коррелирует с количеством генов

Genome size comparison				
	Species	Chromosomes	Genes	Base pairs
	Human (<i>Homo sapiens</i>)	46 (23 pairs)	28-35,000	3.1 billion
	Mouse (<i>Mus musculus</i>)	40	22.5-30,000	2.7 billion
	Puffer fish (<i>Fugu rubripes</i>)	44	31,000	365 million
	Malaria mosquito (<i>Anopheles gambiae</i>)	6	14,000	289 million
	Fruit fly (<i>Drosophila melanogaster</i>)	8	14,000	137 million
	Roundworm (<i>C. elegans</i>)	12	19,000	97 million



Работа с генами и геномами высших эукариот – их выделение и изучение функций – требует использования высокотехнологичных подходов из-за большой структурной сложности геномов (и самих генов)

Подходы к исследованию генов и геномов

In vivo

- Анализ закономерностей наследования признаков
- Создание трансгенных животных с направленно измененными генами
- Анализ локализации и взаимодействия белков в живых клетках

In vitro

- Клонирование фрагментов генома и определение их последовательностей
- Выделение продуктов экспрессии генома и их анализ
- Создание экспрессирующих конструкций
- Получение продуктов экспрессии и их анализ

In silico

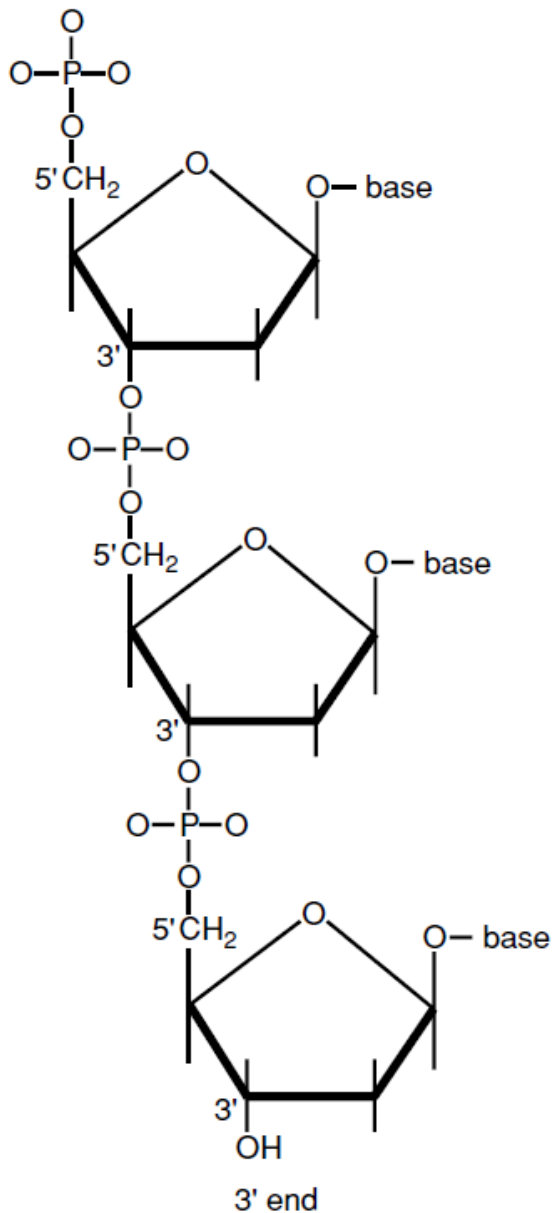
Компьютерный анализ для:

- Поиска генов
- Поиска регуляторных последовательностей
- Поиска повторяющихся элементов генома
- Сравнительный анализ геномов разных видов для поиска гомологий - гомологичные последовательности - аналогичные функции.
- Сравнительный анализ белков в поисках функций



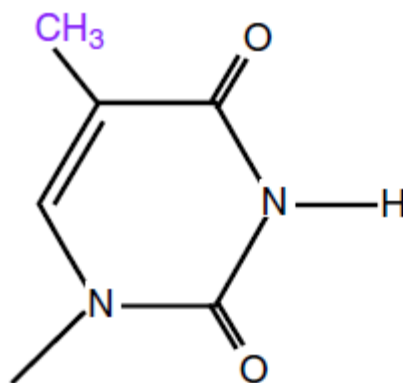
Генная инженерия - совокупность приёмов, методов и технологий для исследования структуры и функции генов и геномов

5' end

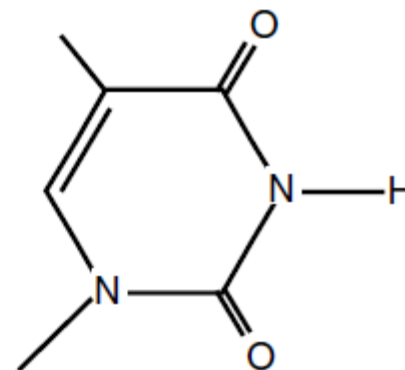


Сахаро-фосфатный остов НК

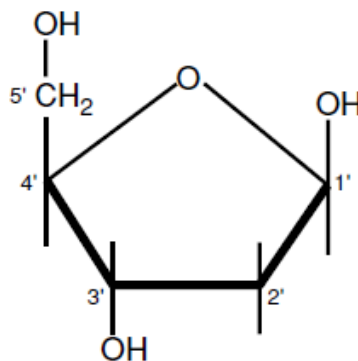
ДНК и РНК



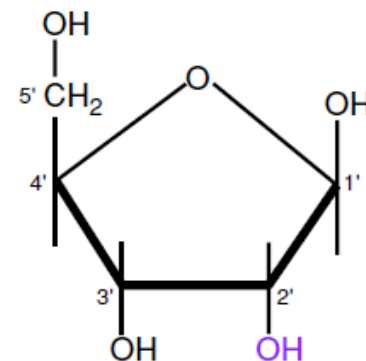
Thymine



Uracil

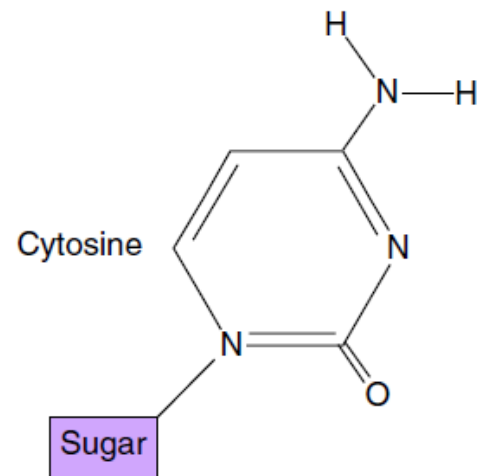
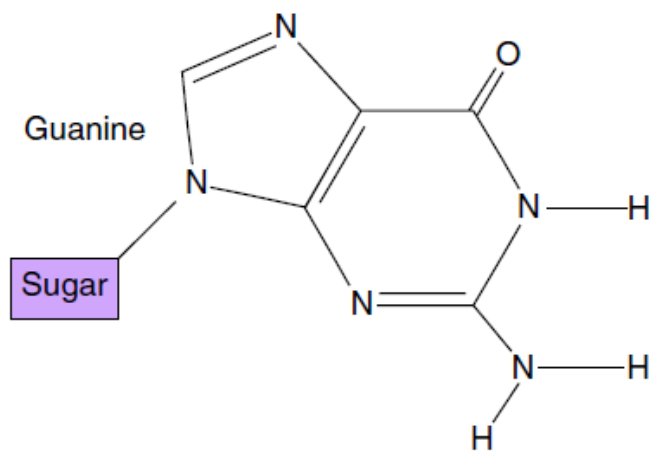
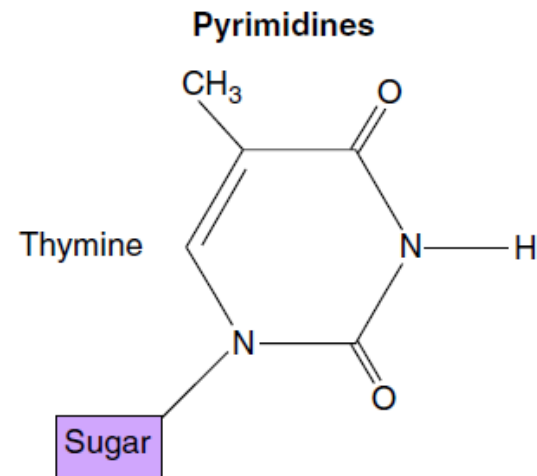
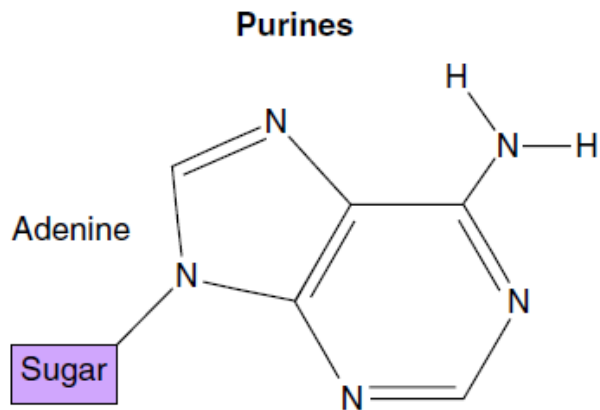


Рибоза

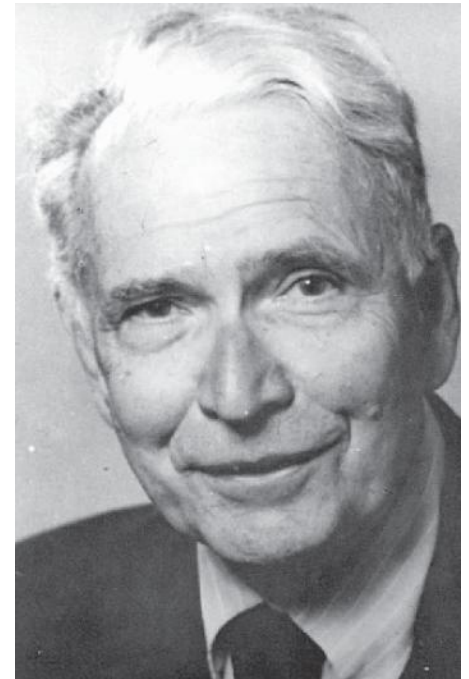
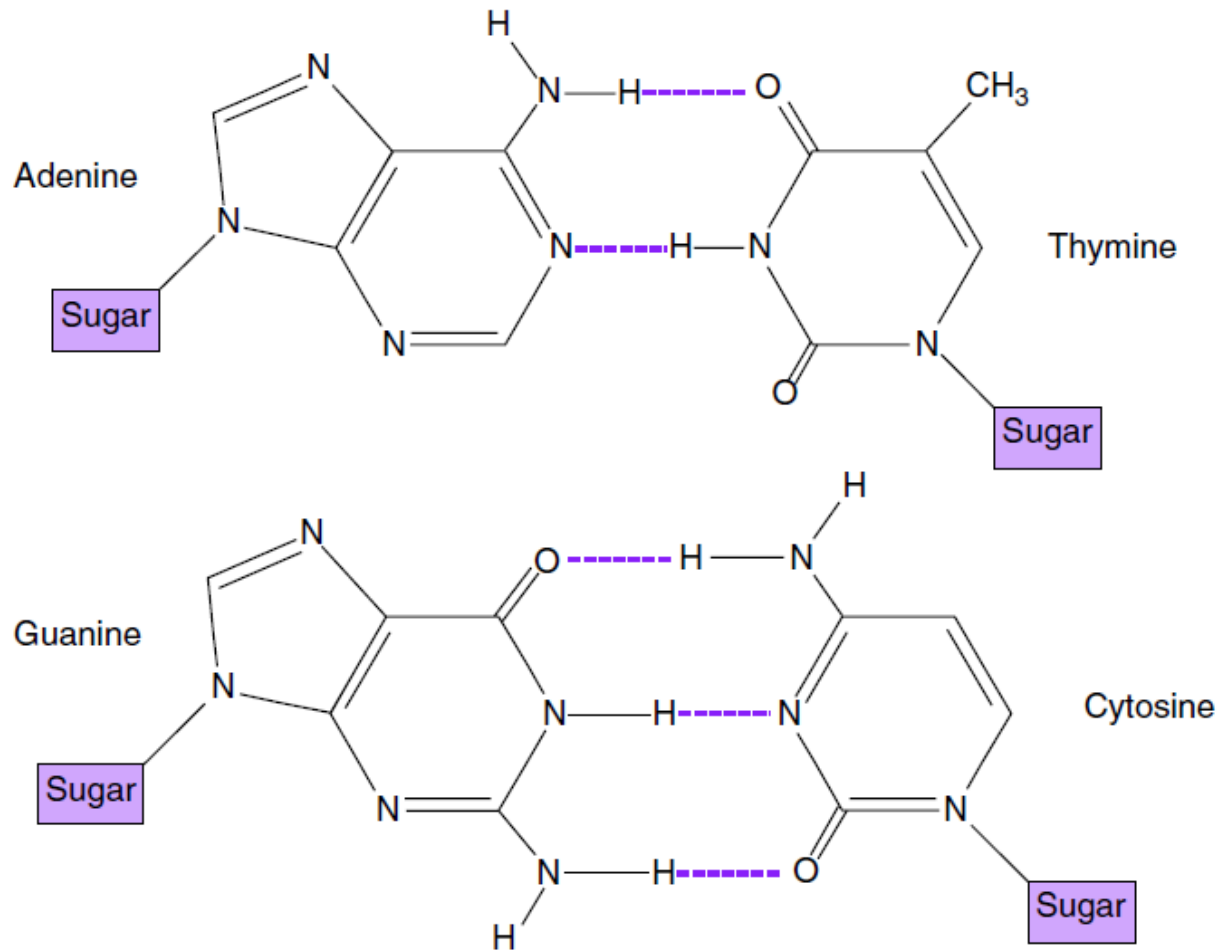


Дезоксирибоза

Пуриновые и пиримидиновые основания

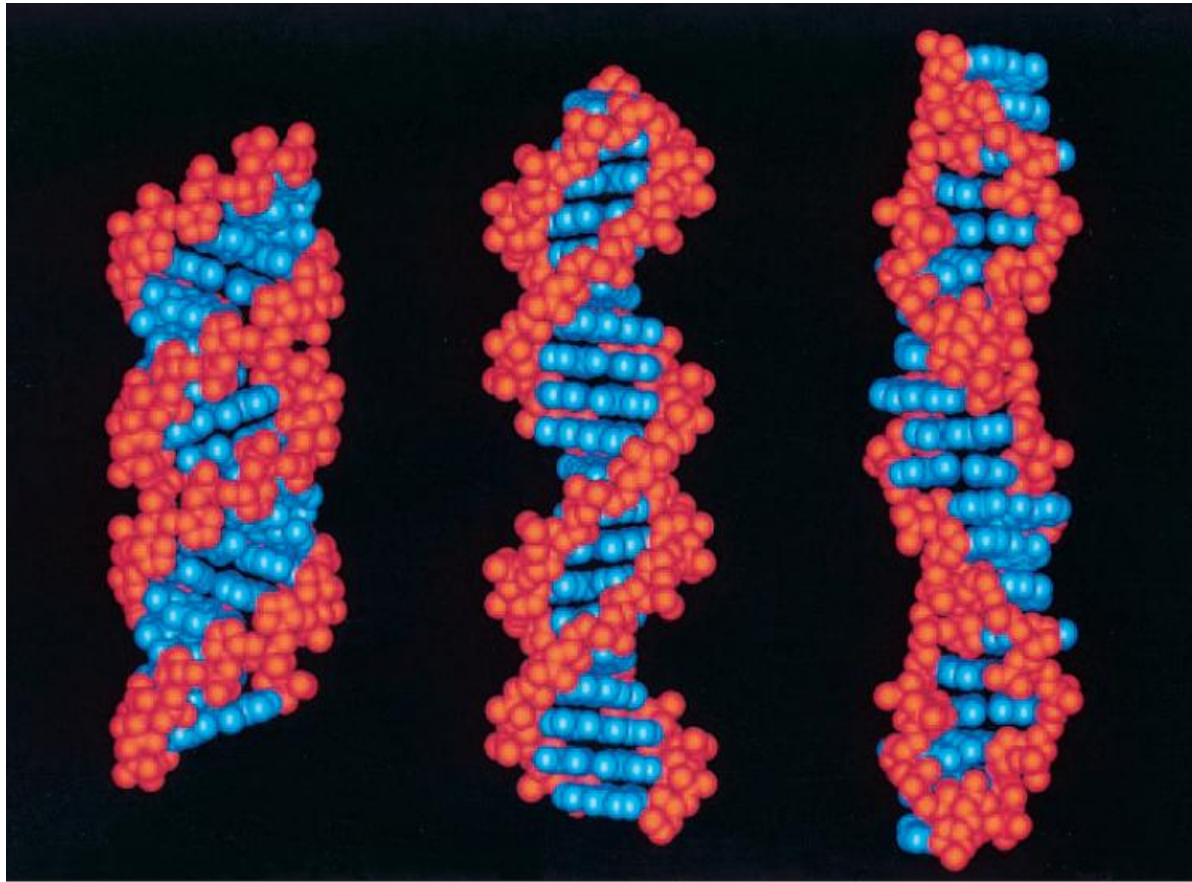


Комплементарные взаимодействия



Erwin Chargaff

Компьютерные модели А-, В- и Z-форм ДНК



A

B

Z

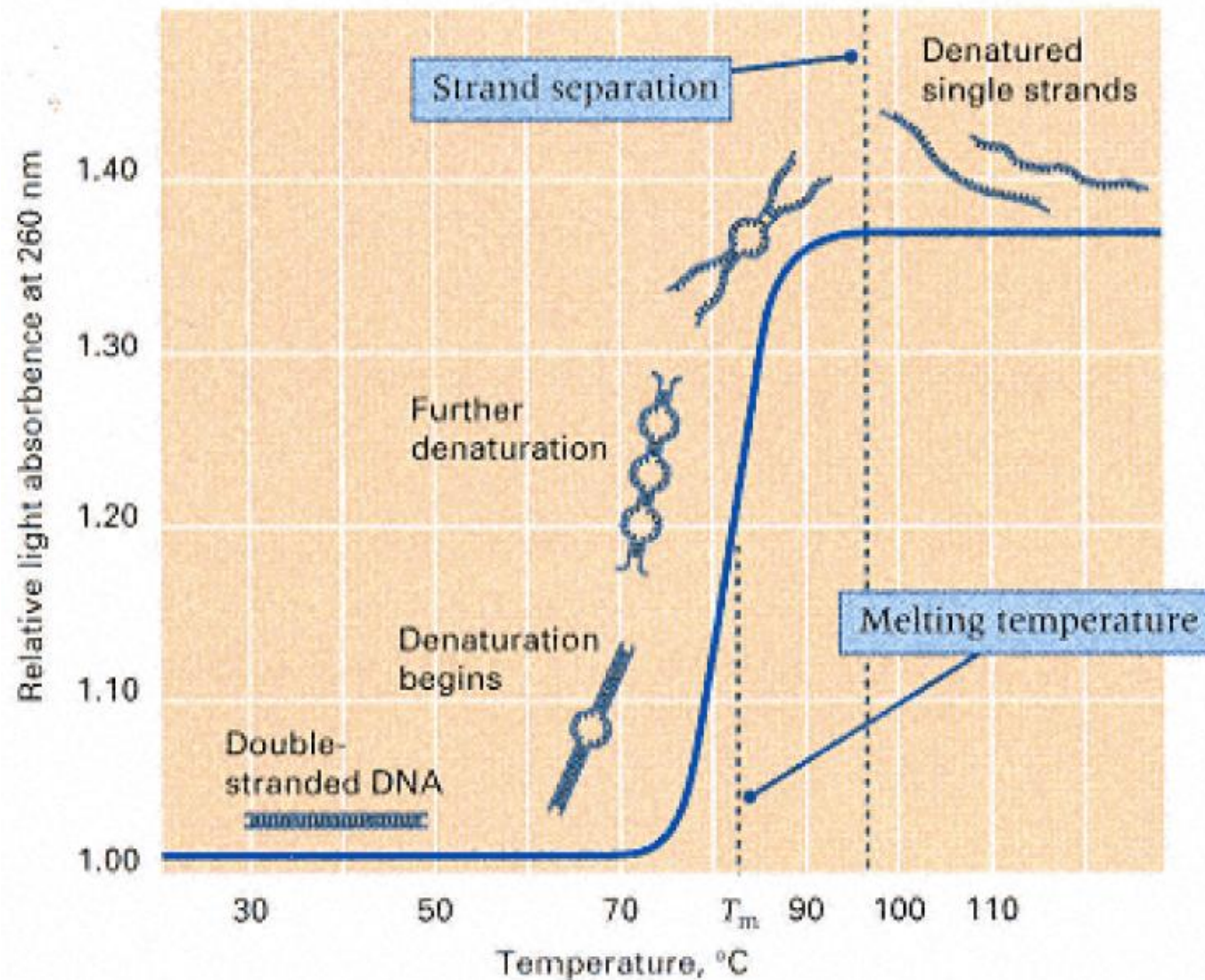
А-форма: ДНК-РНК-гбриды

В-форма:
обычная
конформация
ДНК в клетке

З-форма:
Левозакрученная
спираль
 $\text{poly}[\text{dG-dC}] \cdot$
 $\text{poly}[\text{dG-dC}]$

Отрезки ДНК
одинаковой
длины

Плавление (денатурация) ДНК



Комплементарные последовательности

5'–AGGCTG–3'
3'–TCCGAC–5'

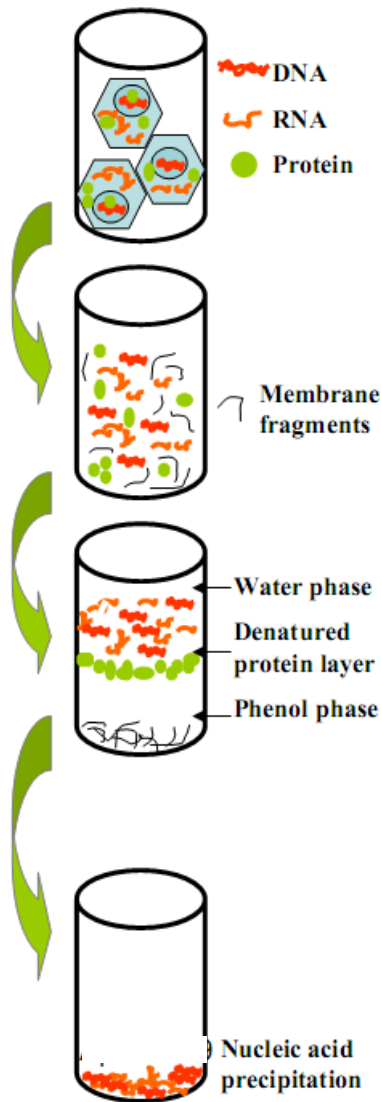
AGGCTG

**Последовательности ДНК
представляют в виде одной цепи, в
которой 5'-конец слева, а 3'-конец
справа**

CAGCCT

**Комплементарная инвертированная
последовательность**

Выделение нуклеиновых кислот



В зависимости от типа ткани используются различные подходы к выделению нуклеиновых кислот, однако все они имеют общие стадии

1 шаг – лизис клеток и клеточных мембран

Клетки разрушают механически или с помощью ферментативной обработки. Мембраны разрушают с помощью обработки детергентами (SDS, ..) или хаотропными агентами (гуанидин изотиоцианат)

2 шаг – удаление белков

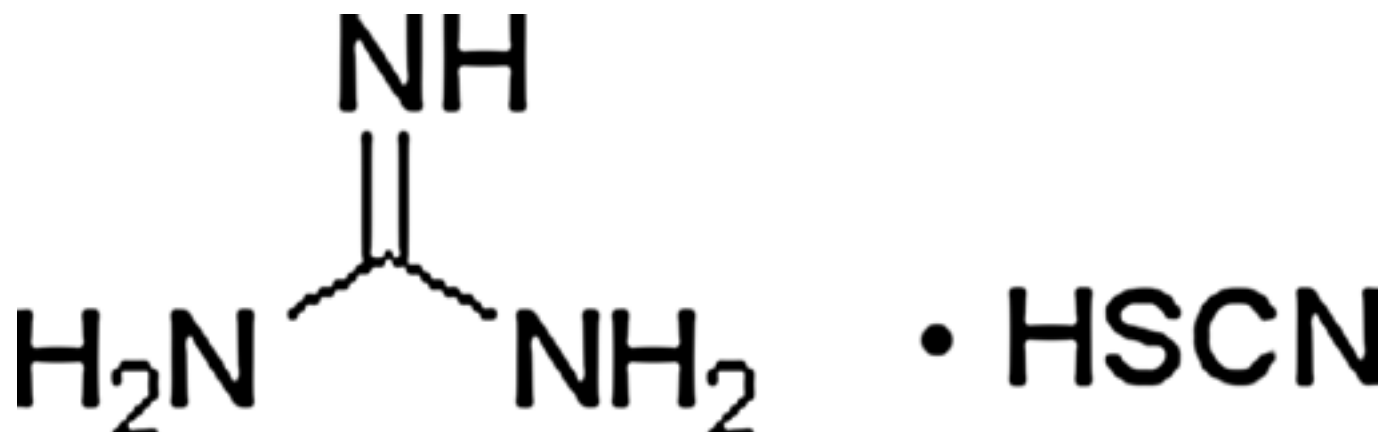
Используют экстракцию белков смесью фенола, хлороформа, изоамилового спирта для создания белковой фракции

Фенол должен быть уравновешен буфером

3 шаг - получение концентрированной фракции нуклеиновых кислот

Используют (1) осаждение ДНК этанолом (изопропанолом) в присутствии соли (ацетата натрия) с дальнейшим отделением осадка центрифугированием и растворением осадка или (2) адсорбцию ДНК на твердый носитель (стекло) с дальнейшей элюцией

Гуанидин изотиоцианат – ключевой компонент при выделении НК, особенно РНК



Другие названия:

Гуанидин тиоцианат

Гуанидин роданин

Фенольная экстракция ДНК

1:1 **фенол** : **хлороформ**

or

25:24:1 **фенол** : **хлороформ** : **изоамиловый спирт**

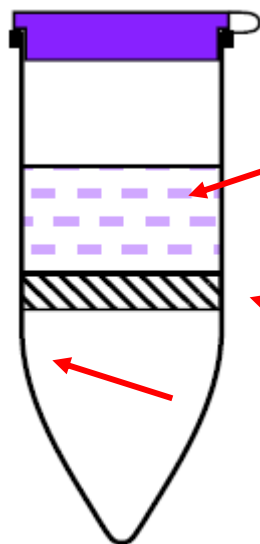
Фенол: денатурация белков, формирование интерфазы

Хлороформ: увеличение плотности органического слоя

Изоамиловый спирт: предотвращение вспенивания

Фенольная экстракция ДНК и РНК

Нейтральные и
слабощелочные
значения pH

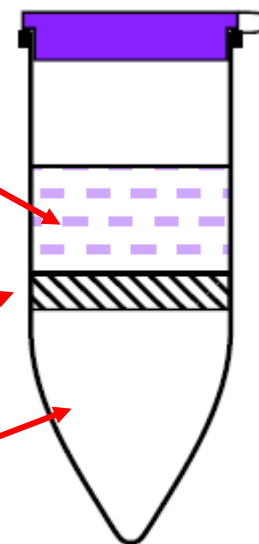


Водная фаза
(ДНК и РНК)

Денатурированные белки

Фенольная фаза

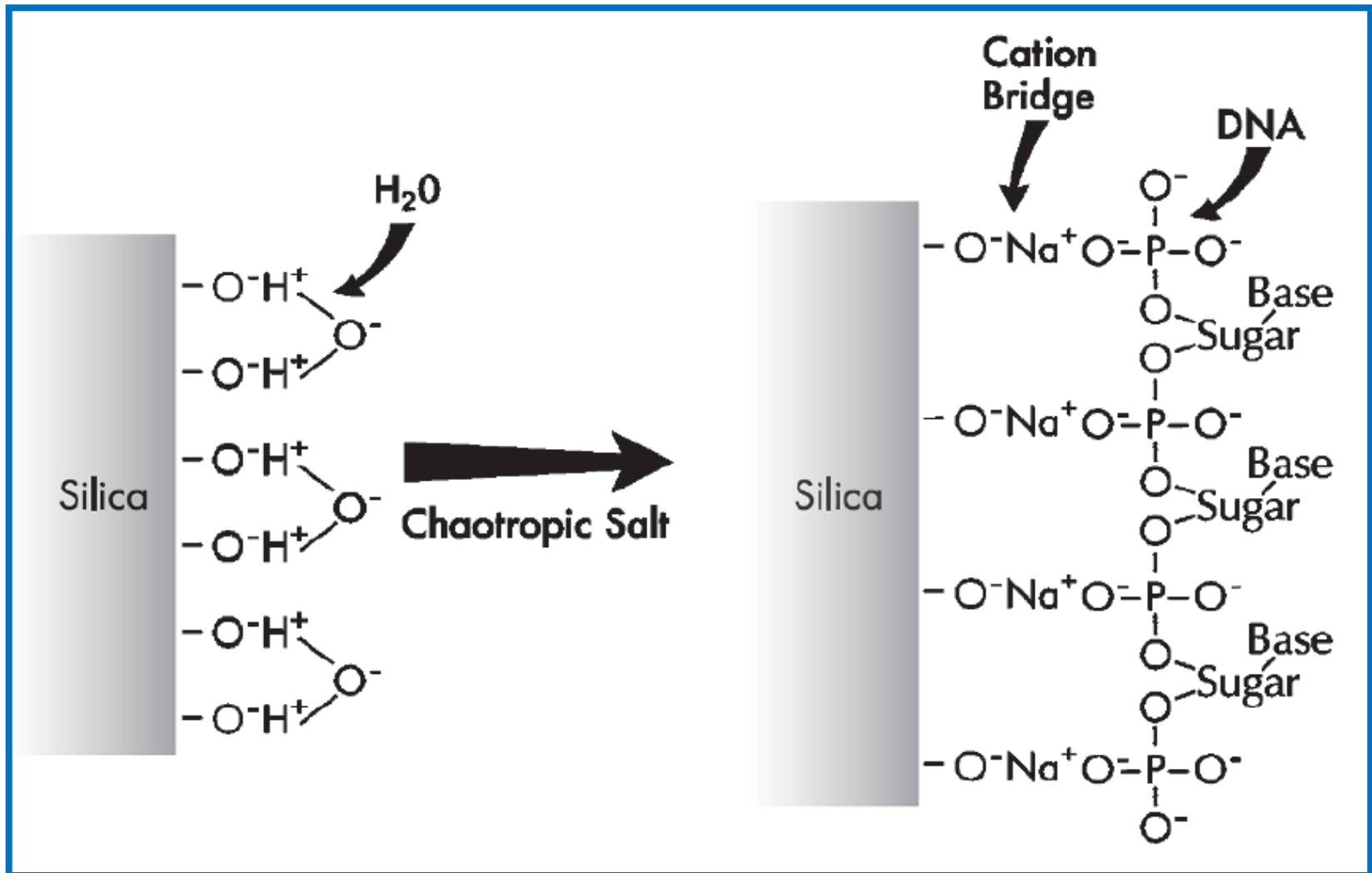
Кислые
значения pH



Водная фаза
(РНК)

Фенольная фаза
(ДНК)

Взаимодействие ДНК с поверхностью стекла

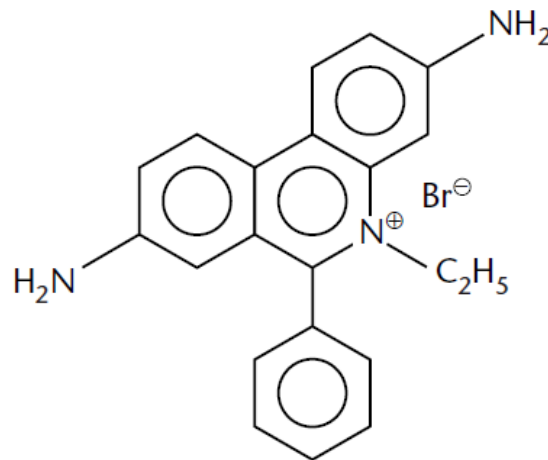


Несколько мифов, связанных с выделением нуклеиновых кислот

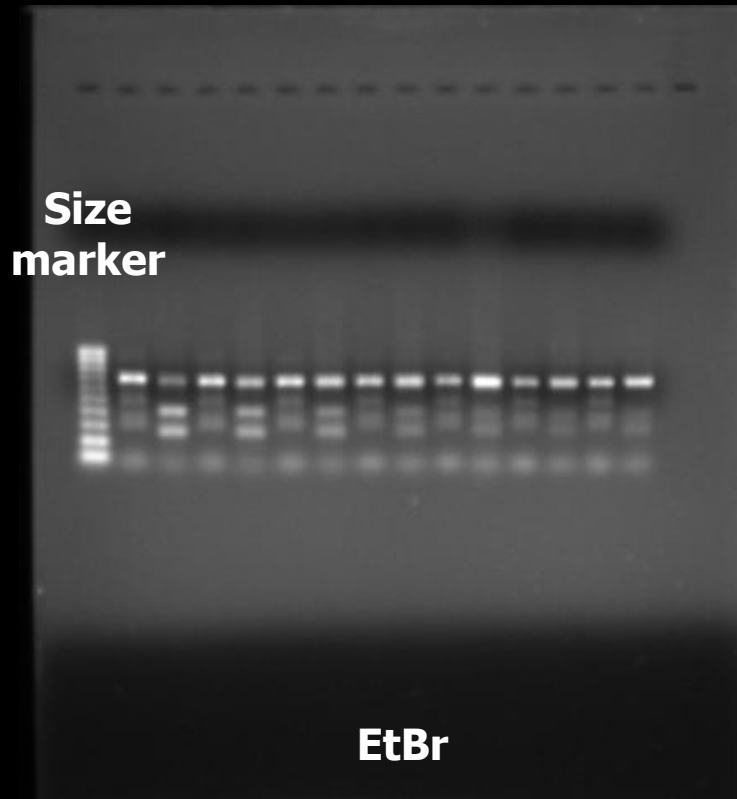
1. При осаждении спиртом нуклеиновые кислоты надо вымораживать – 2 часа на -20 С или 30 минут на – 70 С
2. РНК крайне нестабильна, поэтому хранить необходимо только на -70 С
3. Выделение РНК лучше проводить используя фенол с кислым рН (5-6), тогда как для выделения ДНК следует использовать фенол с рН 8.0

Гель-электрофорез

- Нуклеиновые кислоты отрицательно заряжены (PO_4^- группы)
- Нуклеиновые кислоты могут быть разделены по размеру с помощью гель-электрофореза
- В качестве основы геля используют агарозу или полиакриламид
- Бромистый этидий или SYBR green используют для окраски НК в геле



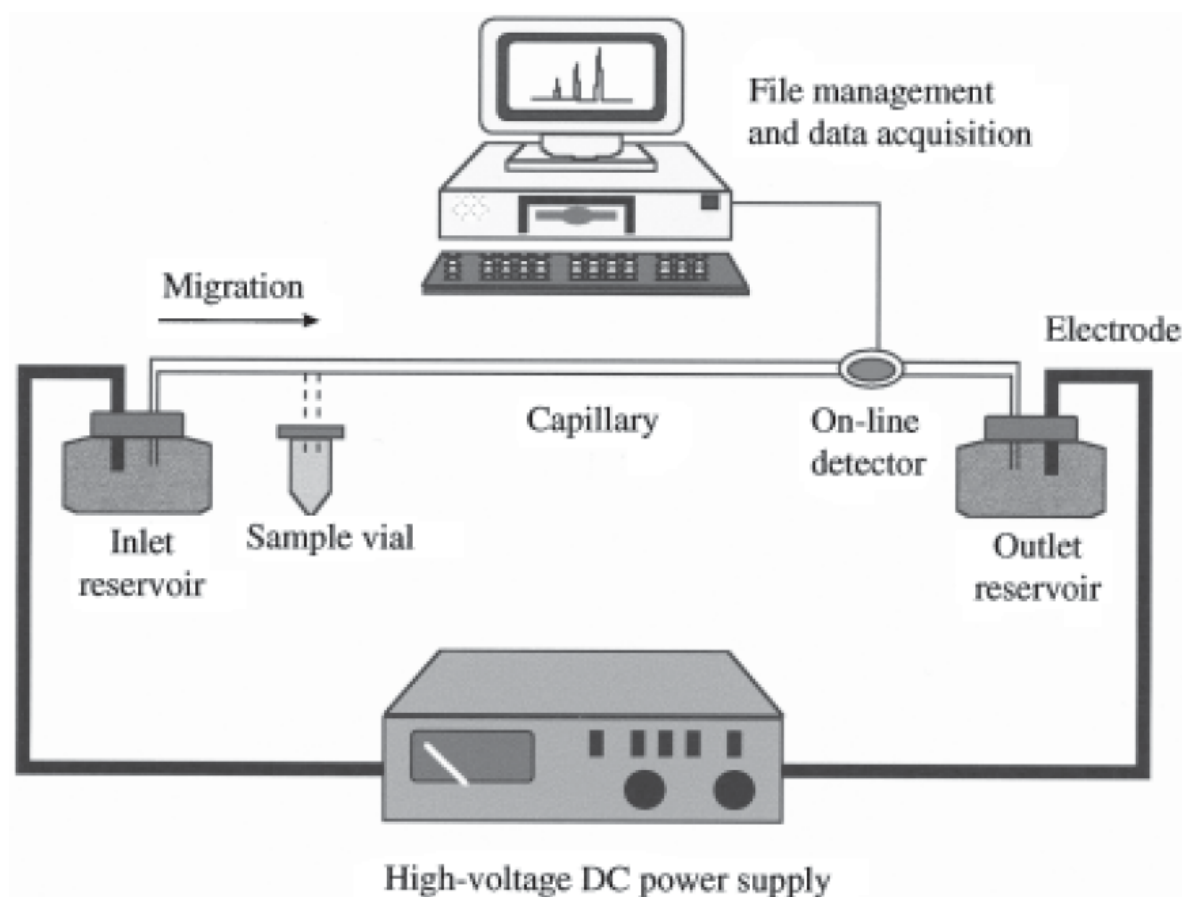
Визуализация ДНК фрагментов с помощью геле-электрофореза в агарозе



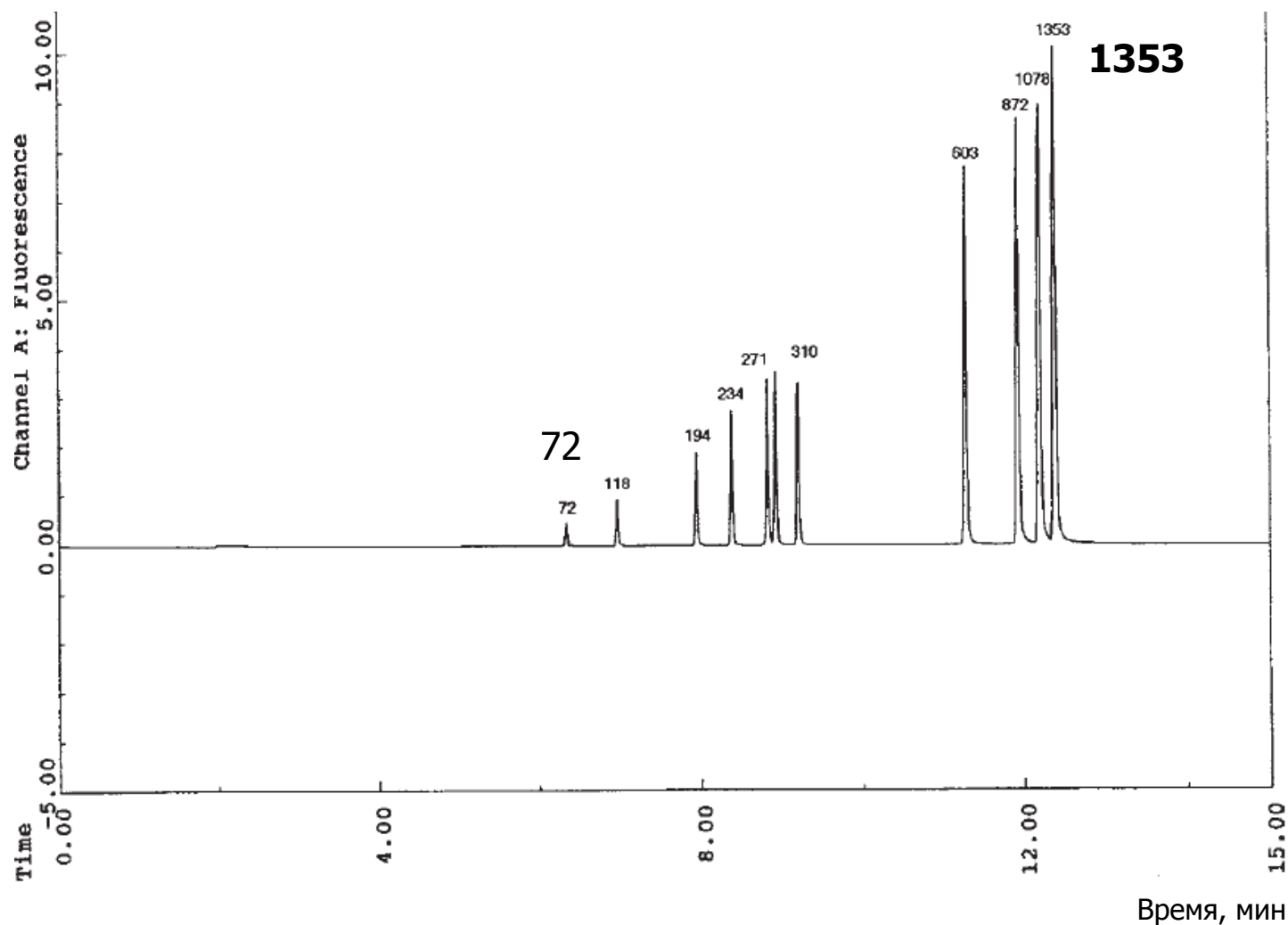
SybrGreen

Капиллярный электрофорез

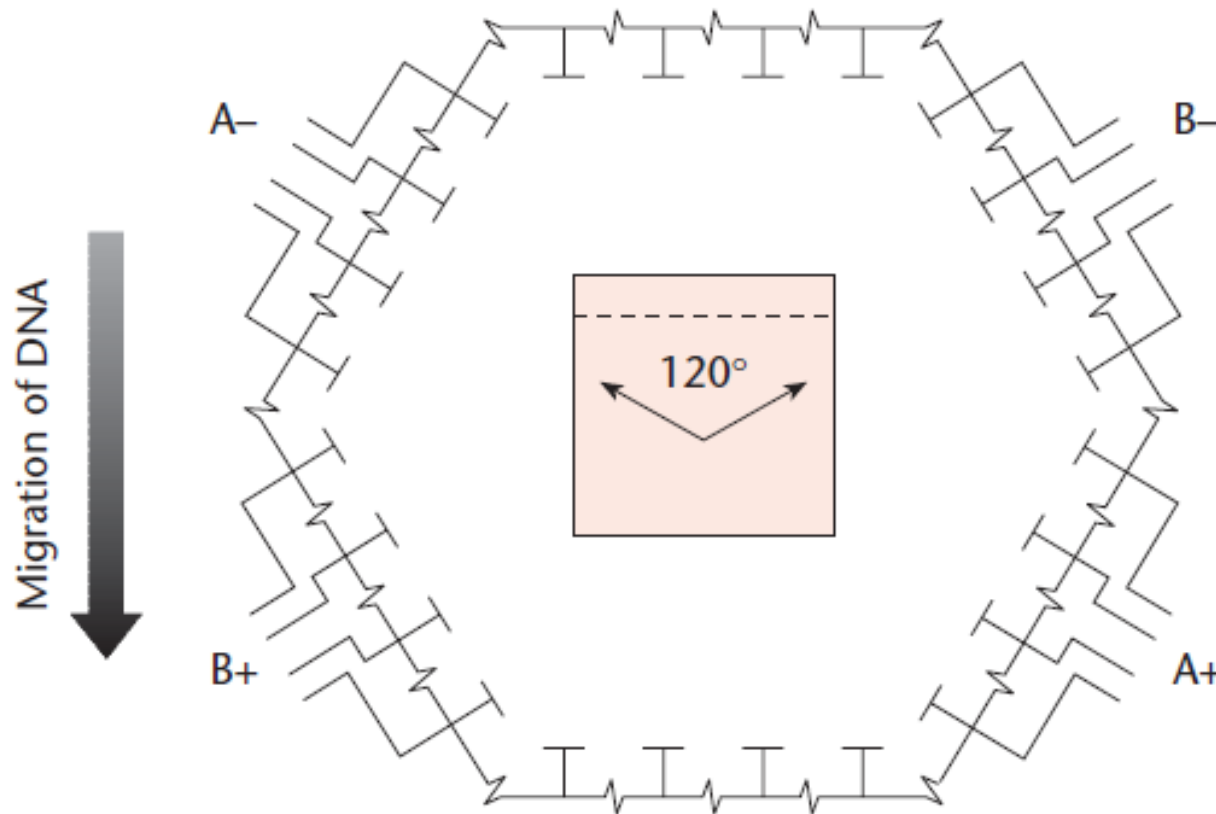
- Капилляр имеет диаметр 50-100 мкм, что обеспечивает быстрое теплорассеяние
- Флуоресцентный детектор обеспечивает высокую чувствительность
- Автоматический анализ результатов в реальном времени позволяет проводить фрагментный анализ ДНК



Разделение *Hae*III-фрагментов ДНК фага ØX174 с помощью капиллярного электрофореза

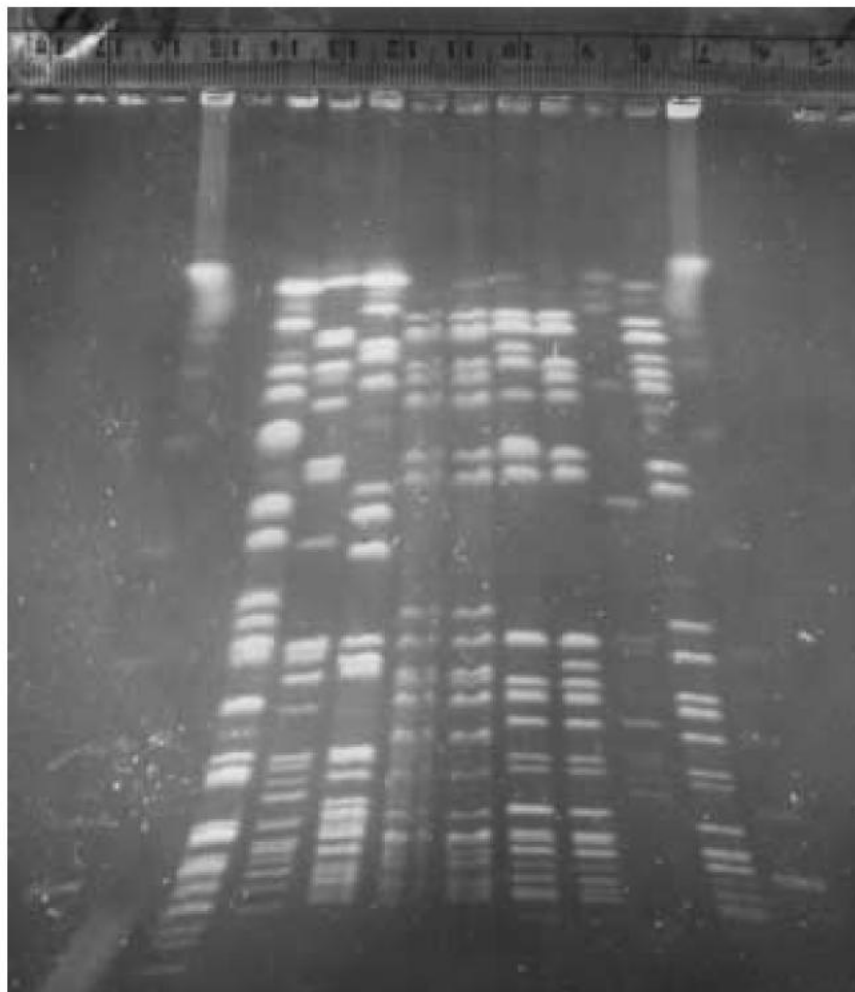


Электрофорез в импульсном электрическом поле

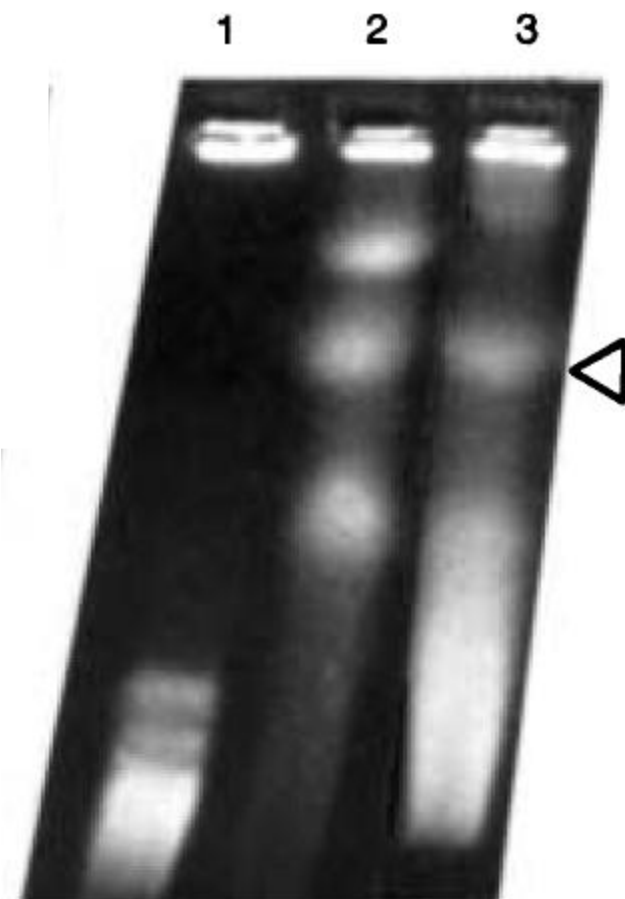


- Разделение ДНК до 10 м.п.н., рутинно – 200-300 т.п.н.
- Параметры полей программируются независимо, обеспечивая движение НК в одном направлении
- Недостаток – сложная картина

Разделение крупных фрагментов ДНК электрофорезом в импульсном электрическом поле

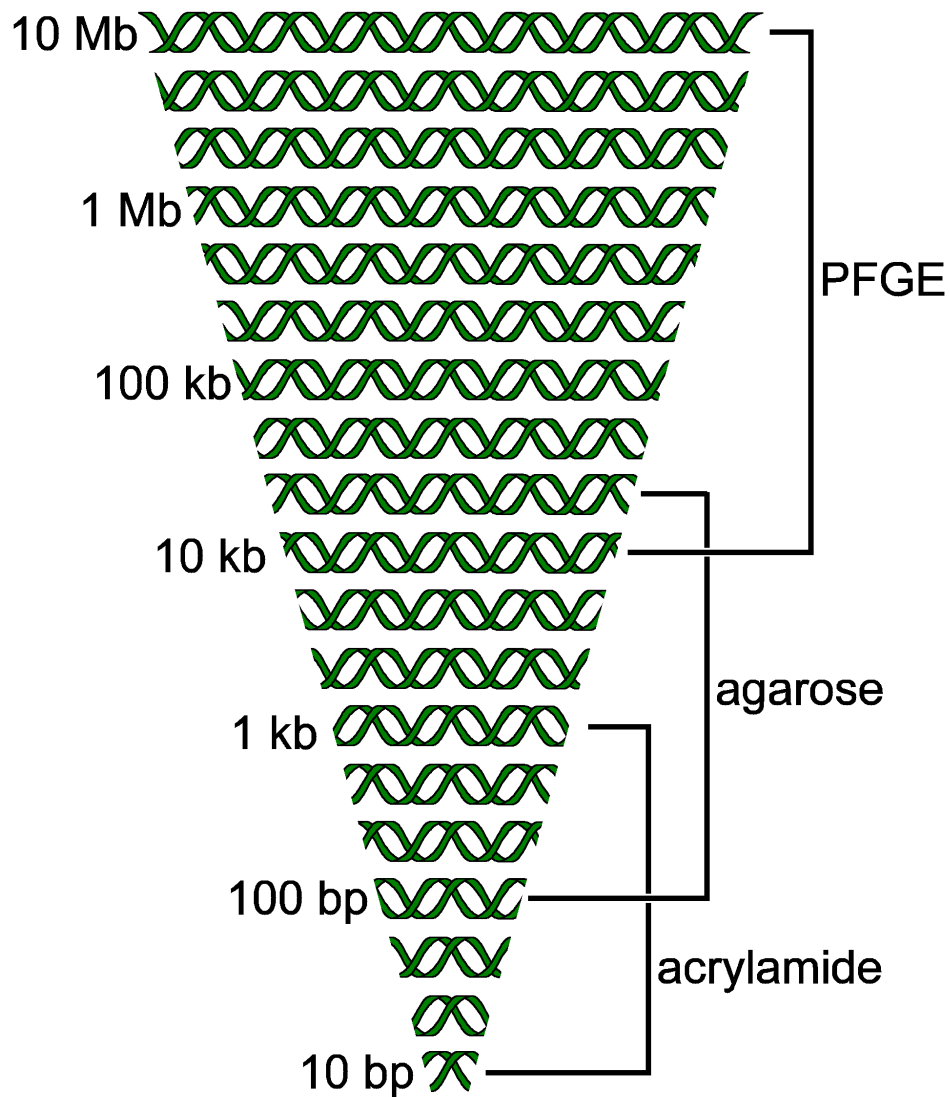


Рестрикция хромосом разных штаммов *E.coli*



Хромосомы дрожжей: 2 – *S. pombe*, 3 – *S. cerevisiae* самая большая хромосома ~1,5 м.п.н.

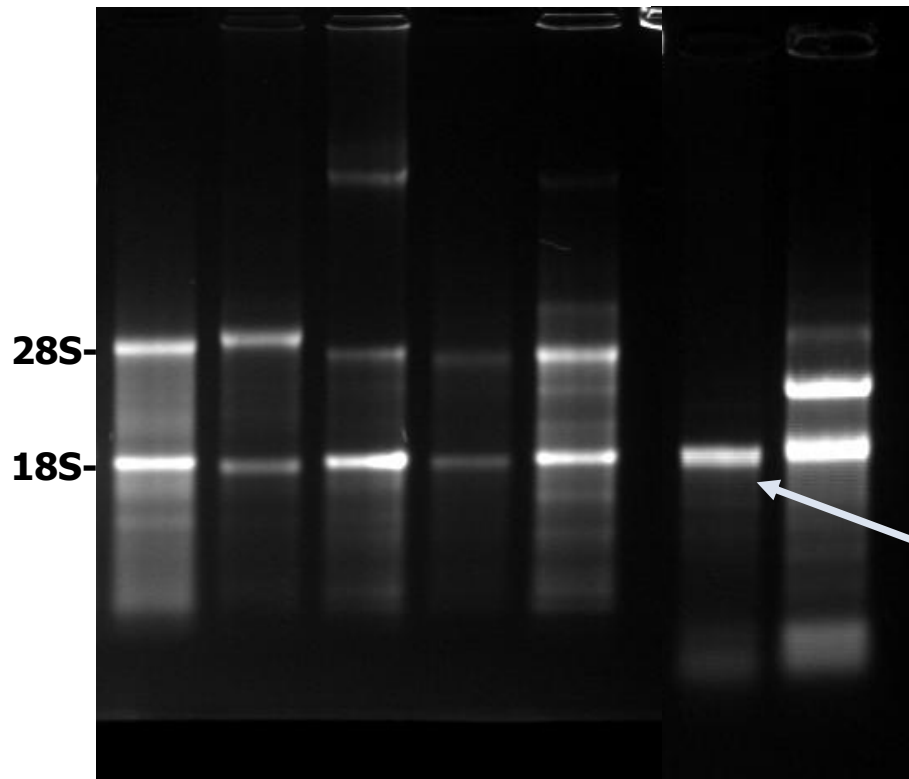
Гель-электрофорез



Разделение средних
фрагментов ДНК
электрофорез в агарозе

Разделение небольших
фрагментов ДНК
электрофорез в ПААГ

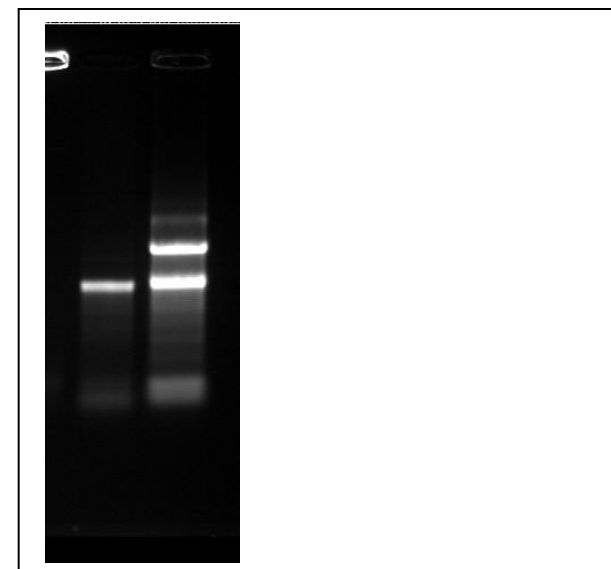
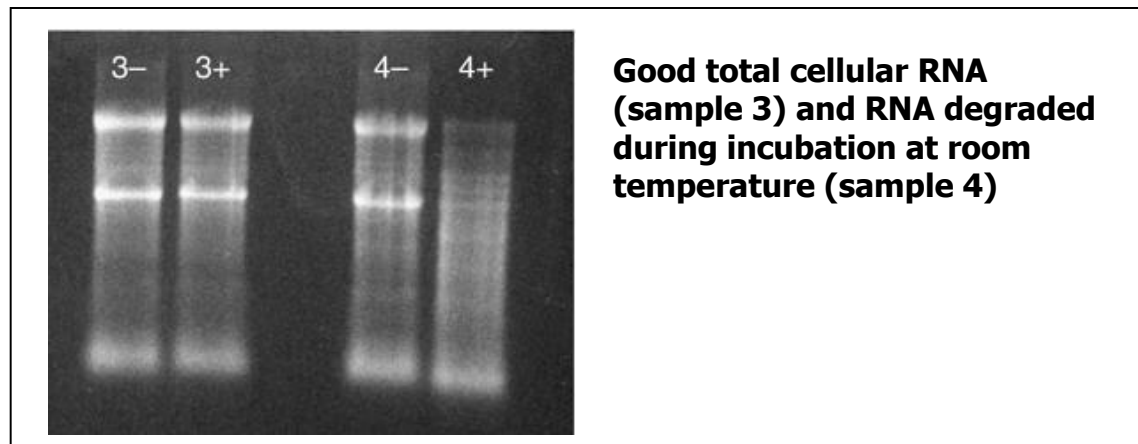
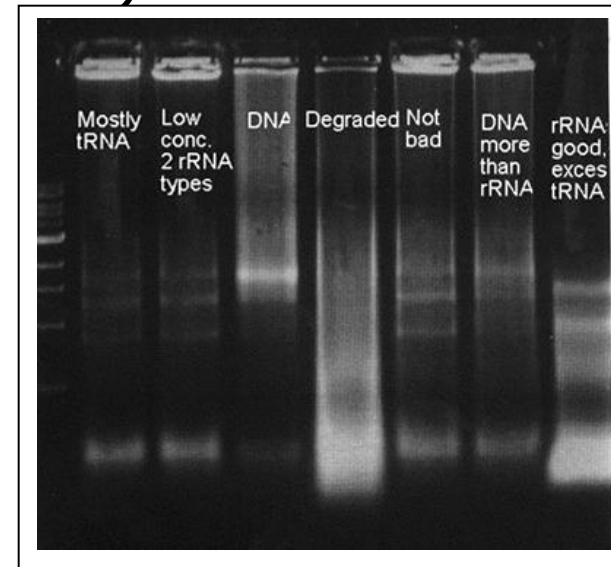
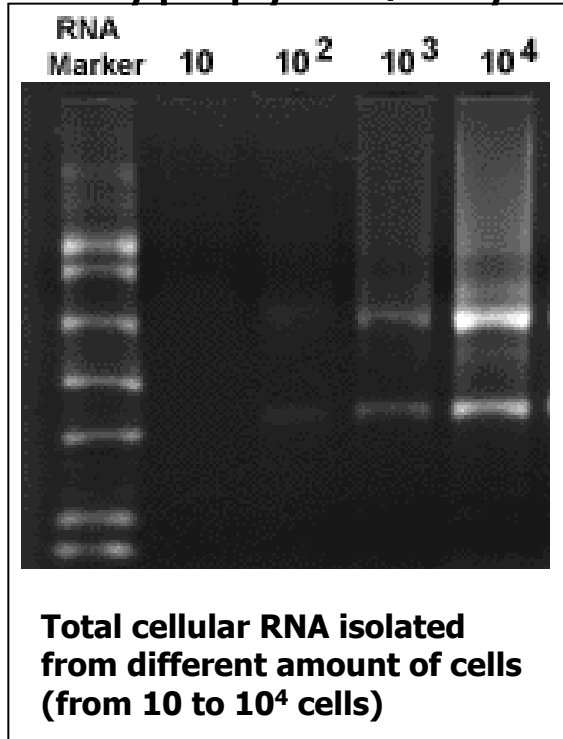
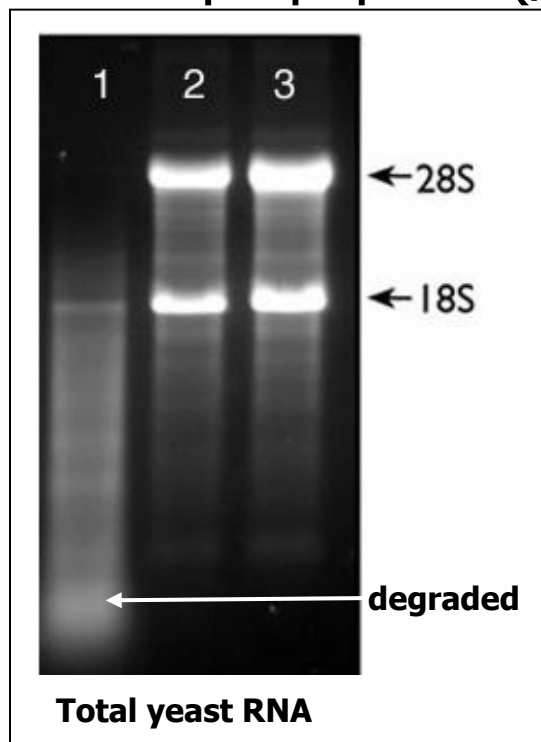
Визуализация РНК на геле-электрофорезе



Тотальная РНК из тканей млекопитающих выглядит как шмер, содержащий две интенсивные полосы 28S и 18S рибосомальной РНК. Соотношение интенсивностей полос 28S к 18S РНК должно быть не менее 1:1. Изменение соотношения в пользу 18S РНК свидетельствует о частичной деградации образца.

РНК из других организмов может иметь иные характеристики: например у насекомых может быть видна только одна полоса рибосомальной РНК на уровне 18S (Ishikawa, 1977) или шмер может быть длиной до 2-3 kb.

Визуализация РНК на формальдегидном гель-электрофорезе (денатурирующие условия)



Молекулярное клонирование

Получение препарата («клона») идентичных молекул ДНК

Впервые было осуществлено в 1970 году

Ключевые составляющие технологии

- *выделение эндонуклеаз рестрикции*
- *открытие плазмидной ДНК бактерий - кольцевых молекул, способных к саморепликации в бактериальных клетках*

Эндонуклеазы рестрикции

Эндонуклеазы рестрикции – ферменты, защищающие бактерий от чужеродной ДНК

Разрезают фосфодиэфирную связь между соседними нуклеотидами в ДНК

Опознают и связывают специфическую последовательность ДНК, называемую сайтом рестрикции или сайтом узнавания (recognition sequence or restriction site)

Рестриктные фрагменты ДНК имеют фосфатную группу на 5'-конце и гидроксильную группу на 3'-конце

Эндонуклеазы рестрикции по-разному чувствительны к метилированию:

Sau 3A (и GATC и G^{Me}ATC)

Dpn I (только G^{Me}ATC)

Mbo I (только GATC)

В генной инженерии используют эндонуклеазы рестрикции II типа

- В большинстве узнают симметричные сайты (обладающие центральной осью и считывающиеся одинаково в обе стороны от оси симметрии)
- Разрезают двухцепочную ДНК в определённой точке внутри сайта рестрикции
- Гомодимеры, не имеют собственной метилтрансферазной активности
- Mg^{2+} - зависимые



<i>EcoRI</i>	G [^] A <u>A</u> TTC	(N6m <u>A</u>)
--------------	-------------------------------	-----------------

<i>HhaI</i>	G <u>C</u> G [^] C	(5m <u>C</u>)
-------------	-----------------------------	----------------

<i>PvuII</i>	CAG [^] <u>C</u> TTG	(N4m <u>C</u>)
--------------	-------------------------------	-----------------

Подчеркнуты метилируемые нуклеотиды

Субстратная специфичность рестриктаз II типа

Палиндромные сайты

мелкощепящие: *Bgl*I (GGCC), крупнощепящие: *Eco*RI (GAATTC), *Not*I (GCGGCCGC)

Частично вырожденные сайты

Hinc II (GT**YR**AC, **Y** = p**Y**rimidine, **R** = pu**R**ine),

Разорванные сайты

*Bgl*I (GCC**N**₅GGC, **N** = a**N**y)

Квазисимметричные сайты

*Btr*I (**C**AC[^]G**T****C**, класс IIQ)

Двойные сайты:

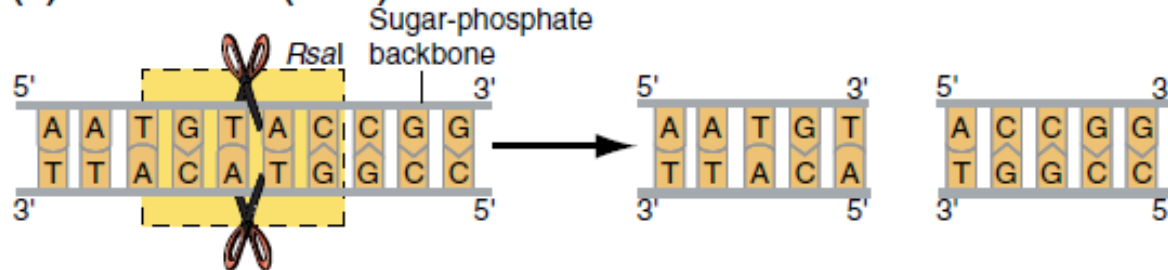
*Sfi*I (GGCC**N**₅ GGCC)

В неоптимальных условиях может происходить изменение субстратной специфичности:

*Eco*RI - GAATTC, *Eco*RI* - AATT (pH , NaCl, органические растворители)

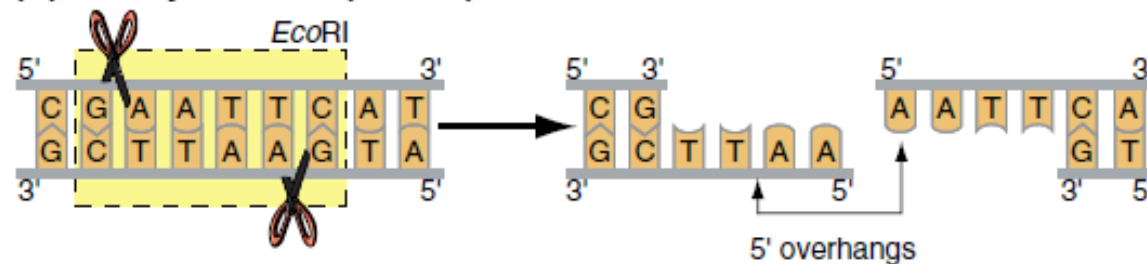
Формы разрывов ДНК, образующихся под действием рестриктаз II типа

(a) Blunt ends (*RsaI*)



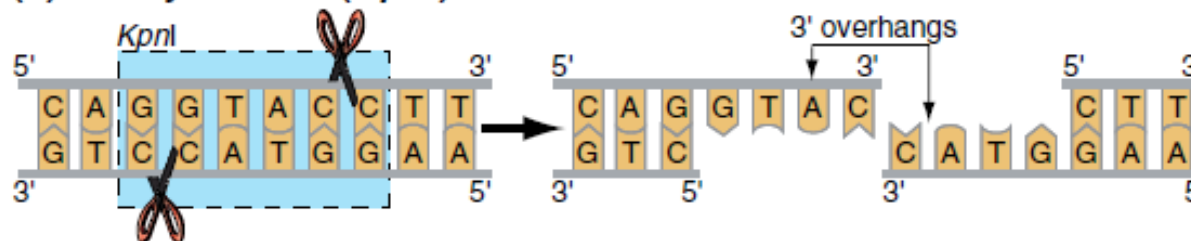
«Тупые» концы

(b) Sticky 5' ends (*EcoRI*)



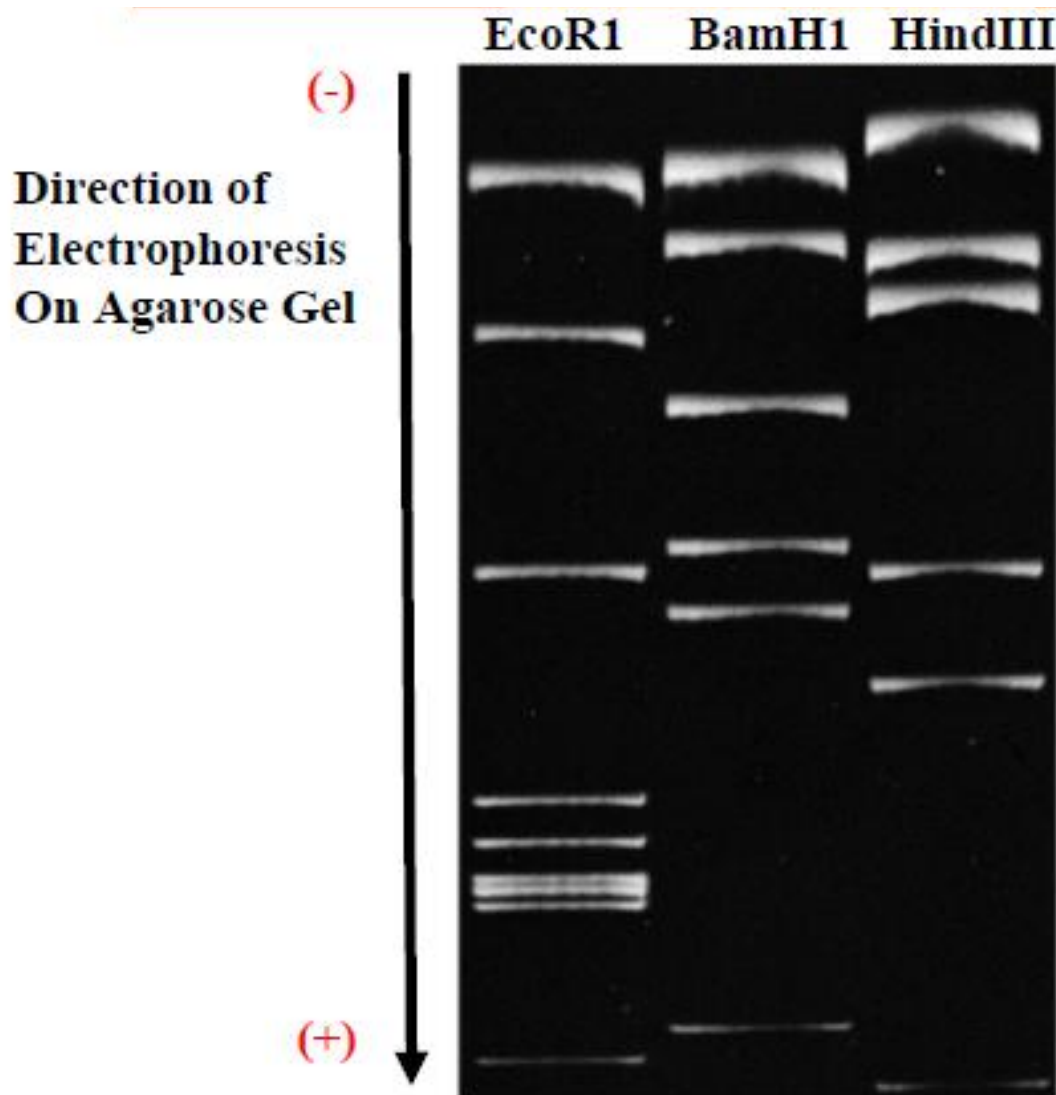
5'-выступающие
«липкие» концы

(c) Sticky 3' ends (*KpnI*)



3'-выступающие
«липкие» концы

Фрагменты ДНК, полученные с помощью разных эндонуклеаз рестрикции



Рестриктаза *Fok* I (тип IIS)

5' –**GGATG**NNNNNNNNNNNNNNNNNN
3' –**CCTAC**NNNNNNNNNNNNNNNNNN

FokI recognition and cleavage site

Catalytic domain

DNA-binding domain

Сперва мономерная субъединица связывается с ДНК, потом с этим комплексом связывается разрезающая ДНК субъединица. Рестриктаза быстрее разрезает ДНК, содержащую два сайта связывания, чем ДНК, содержащую один такой сайт

Образующиеся последовательности липких концов уникальны

FokI crystal structure.



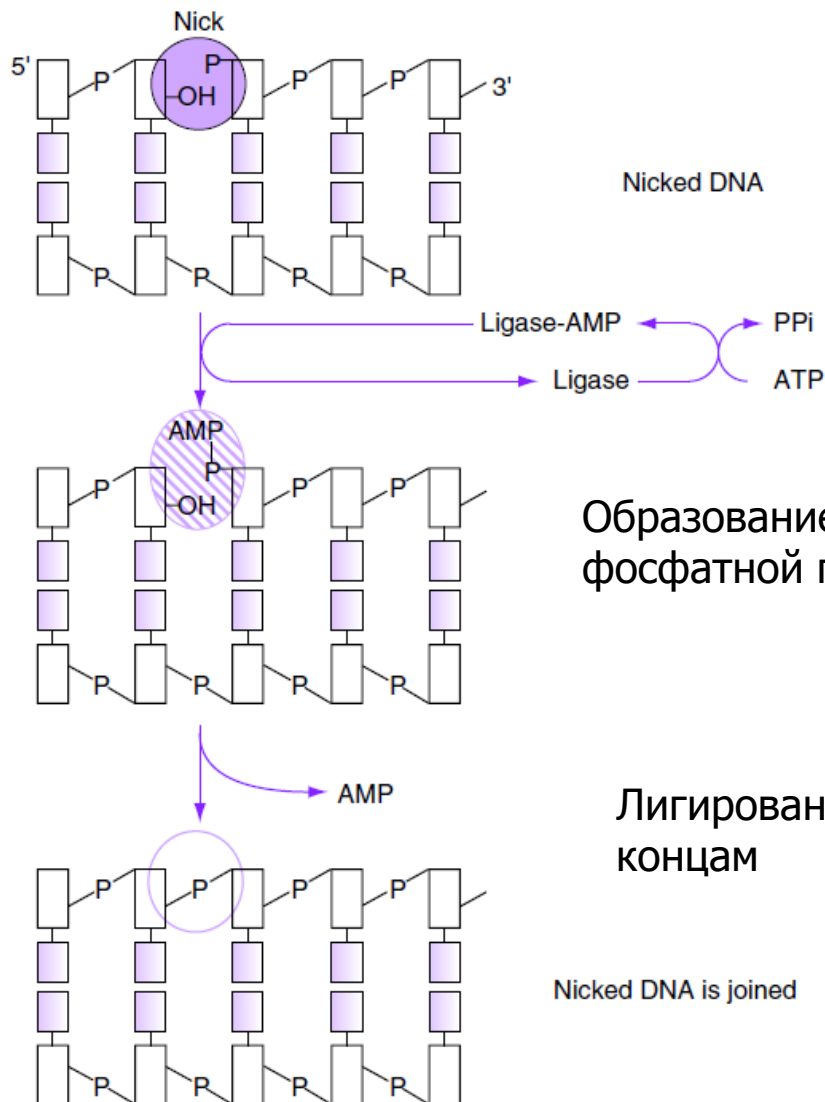
ДНК лигазы

- Формируют фосфодиэфирную связь между двумя молекулами ДНК
- ДНК субстрат должен иметь фосфатную группу на 5'-конце

В генной инженерии чаще всего используют ДНК лигазу фага Т4

- Субстрат – только двухцепочечная ДНК
- Может «сшивать» как молекулы с комплементарными липкими концами, так и молекулы с тупыми концами

Механизм действия Т4 ДНК лигазы



Образование фосфодиэфирной связи между 3'-ОН- и 5'-фосфатной группами одноцепочечных ДНК

Лигирование фрагментов ДНК по «тупым» концам

Nicked DNA is joined



Вектор – рекомбинантная самореплицирующаяся молекула ДНК, в которую встраивается фрагмент чужеродной (клонлируемой) ДНК

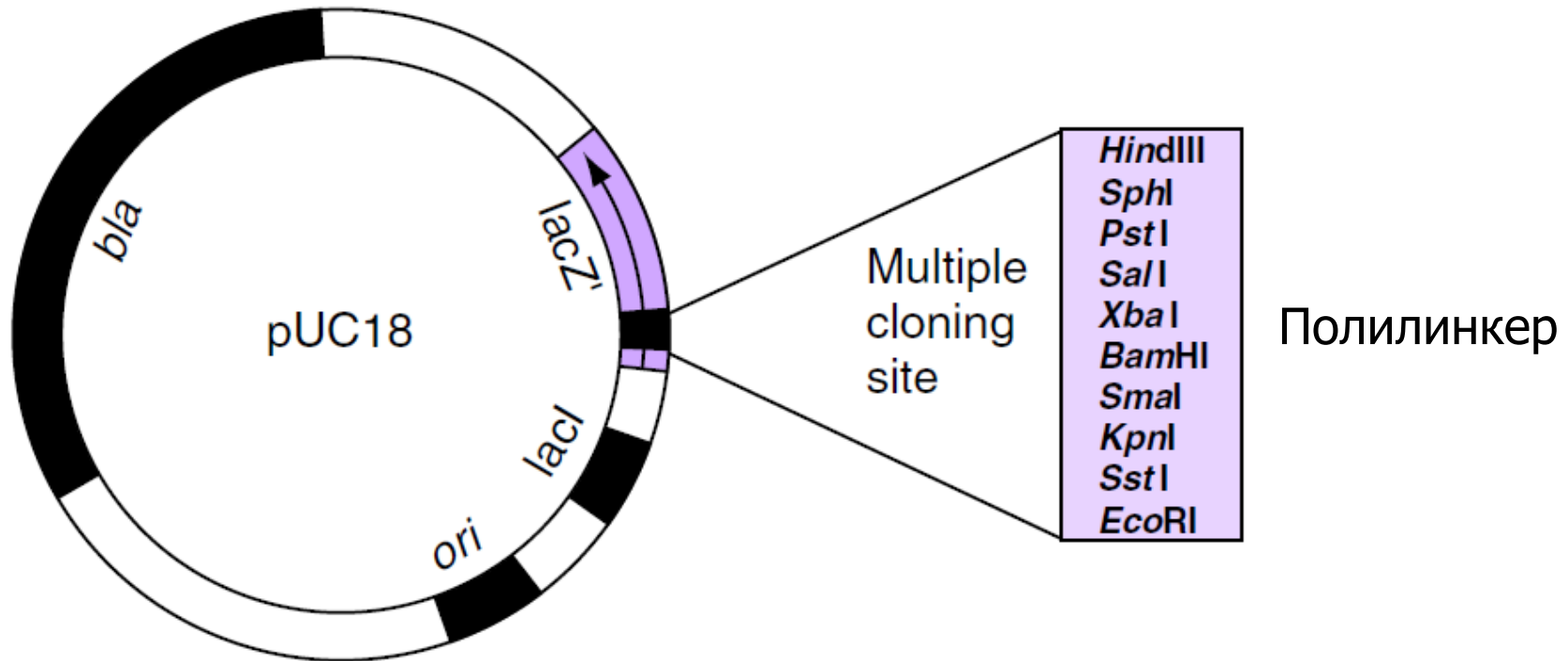
- Способность к длительному существованию в клетках-хозяевах (репликация автономная или в составе хромосом)
- Наличие биохимических или генетических маркеров, которые позволяют обнаруживать его присутствие в клетках
- Должны допускать встраивание чужеродной ДНК без нарушения своей функциональной целостности

Плазмиды – внехромосомные кольцевые молекулы ДНК



- Плазмиды обладают автономной репликацией
- В зависимости от ориджина репликации в клетке может содержаться различное число копий плазмиды
 - Низкокопийные (1-2 копии на клетку)
 - Высококопийные (10-100 копий на клетку)
- Плазмиды обладают небольшим размером
 - Минимальный размер определяется элементами внутриклеточной автономии
 - Существуют ограничения на размер вставки чужеродной ДНК
- Совместимость разных плазмид в клетке
 - Группы несовместимости
 - Случайная сегрегация в дочерние клетки
- Легко выделять
- Легко манипулировать
- Плазмиды часто содержат гены-репортеры (например, гены устойчивости к антибиотику)

Плазмидный вектор pUC18



bla – Ген β-лактамазы, селектируемый маркер (устойчивость ампициллину)

ori – Точка начала репликации (origin)

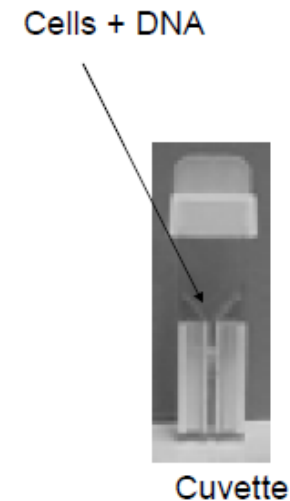
lacZ' – N-концевая часть гена β-галактозидазы (кодирует 146 из 1021 АК-остатков), селектируемый маркер (хромогенный субстрат X-Gal)

lacI – Ген Lac-репрессора (Индуктор – IPTG)

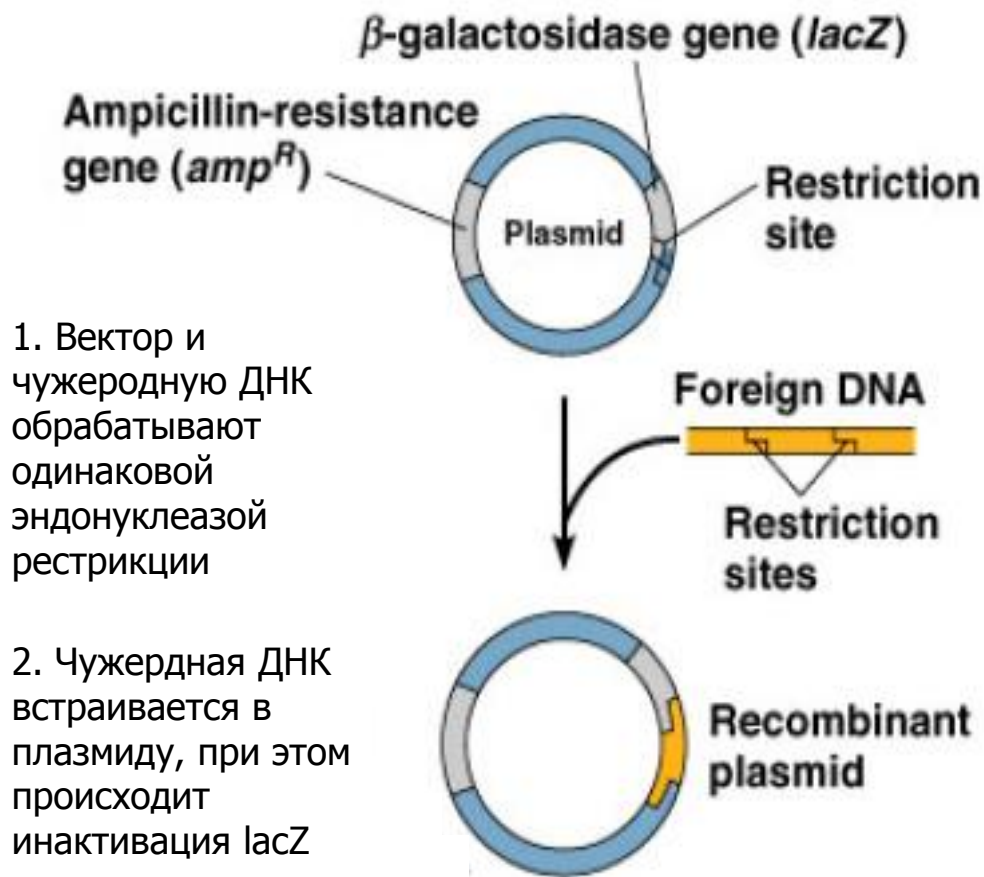
Трансформация

Перенос рекомбинантной ДНК в клетки бактерий (*E. coli*)

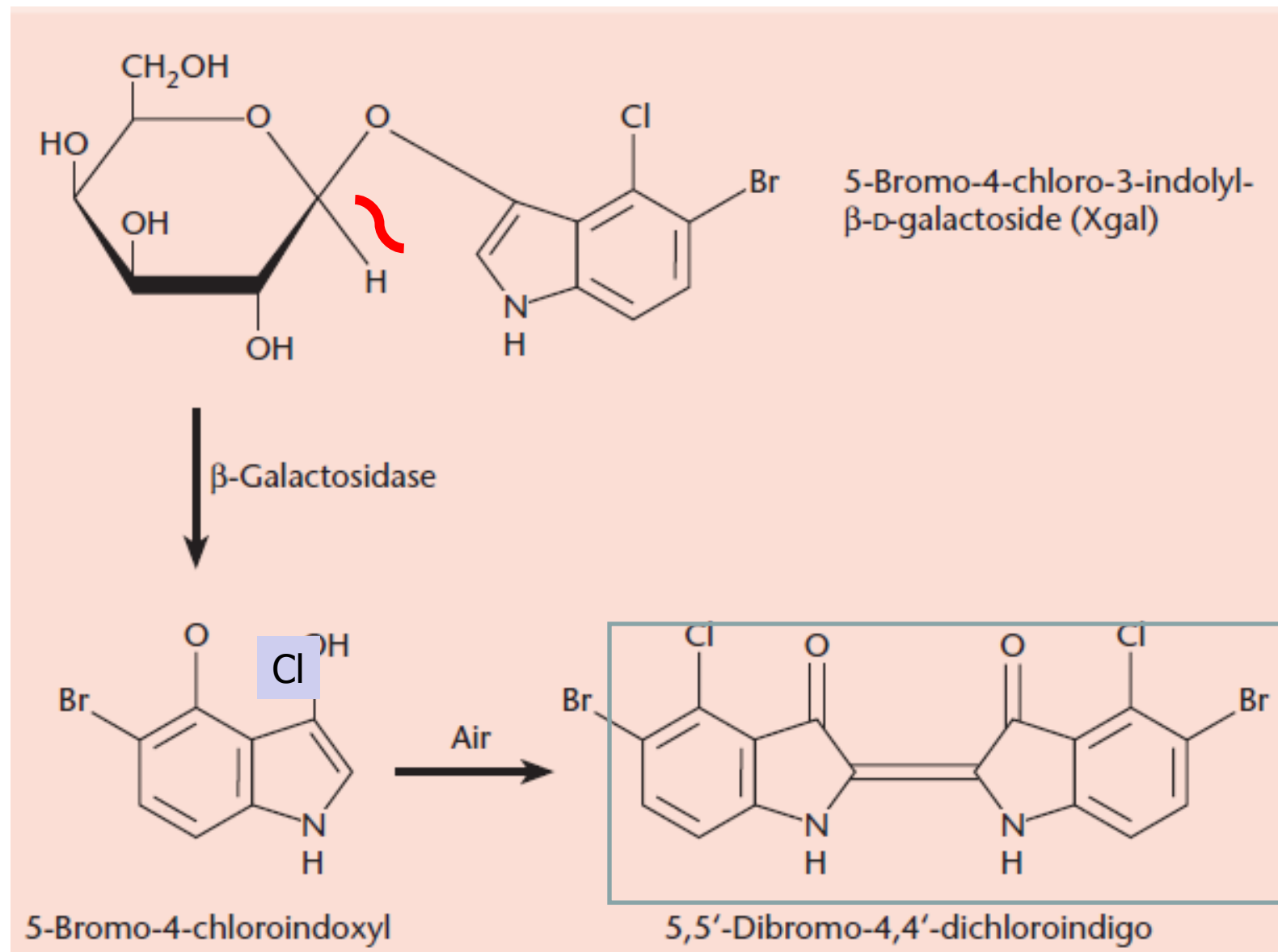
- Химическая трансформация
 - Инкубация *E. coli* в CaCl_2
 - Тепловой шок
 - 10^5 - 10^9 трансформированных клеток на мг ДНК
- Электропорация
 - 10^8 - 10^{10} трансформированных клеток на мг ДНК
 - Так же используется для переноса ДНК в клетки животных, дрожжей и растений



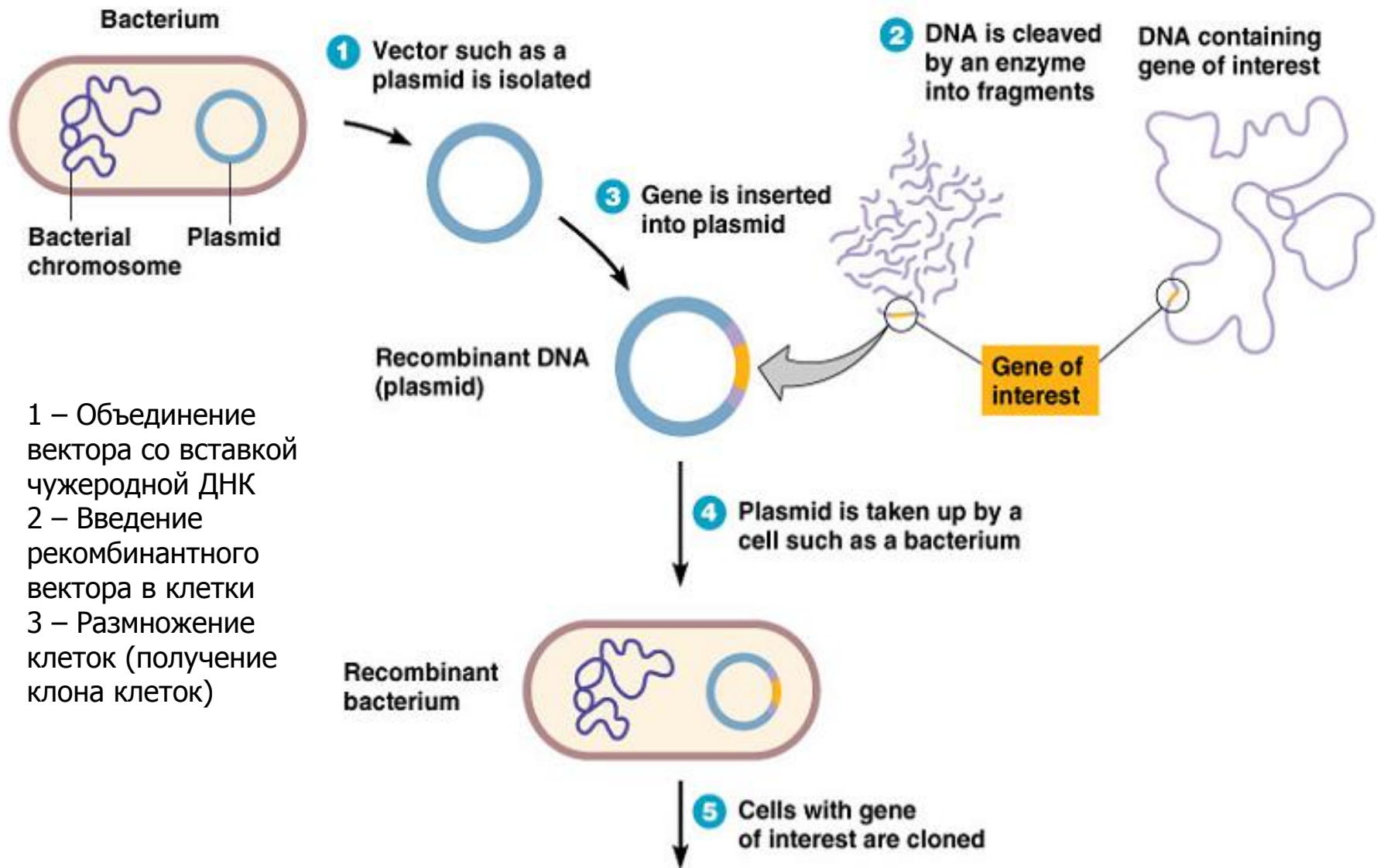
Бело-голубая селекция



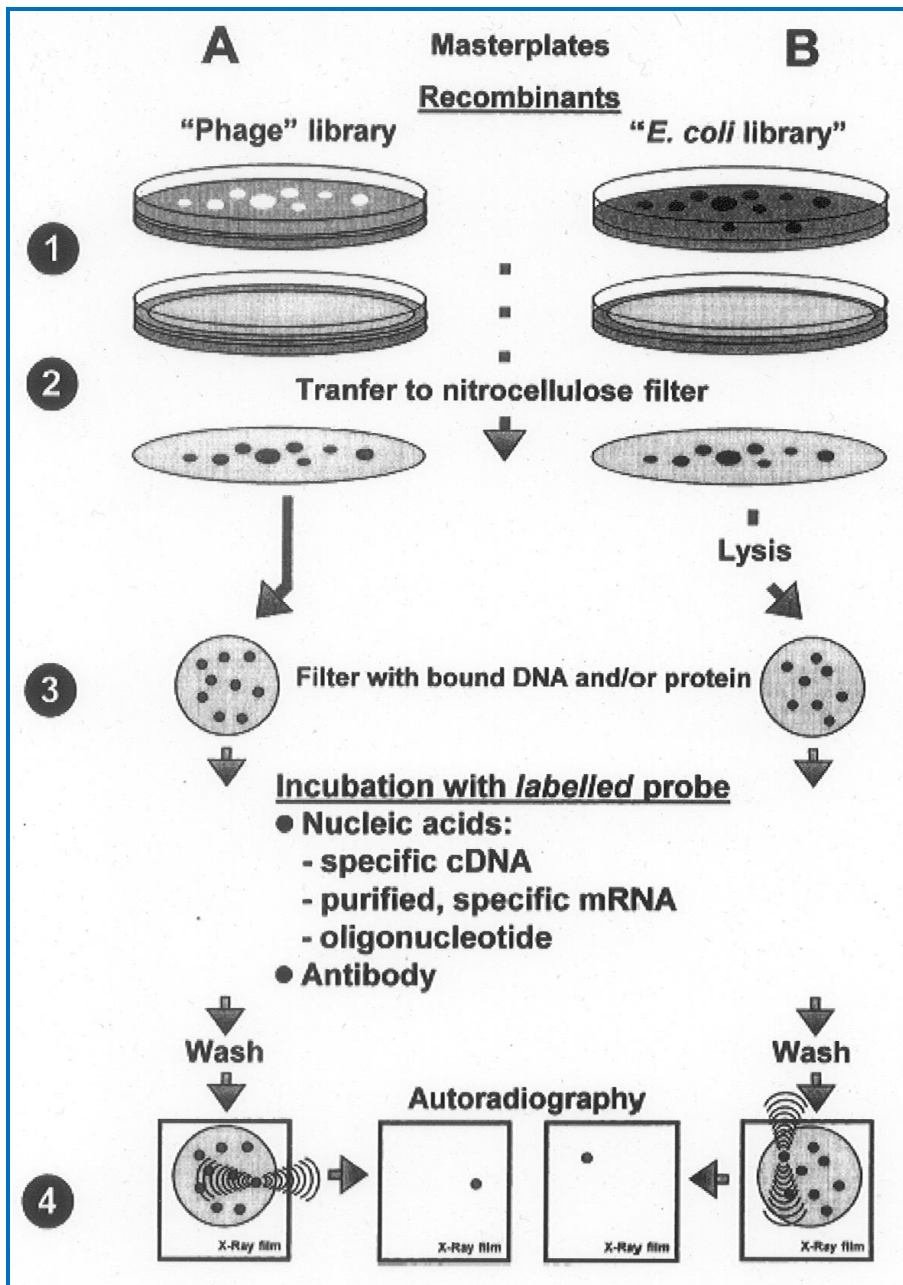
Расщепление Xgal β-галактозидазой



Молекулярное клонирование



- 1 – Объединение вектора со вставкой чужеродной ДНК
- 2 – Введение рекомбинантного вектора в клетки
- 3 – Размножение клеток (получение клона клеток)



Отбор нужных клонов из фаговых и бактериальных библиотек

Часть каждой бактериальной колонии или фаговой бляшки переносят на мембрану, ДНК денатурируют и гибридизуют с меченым олигонуклеотидным зондом

Отбор с помощью ПЦР