

Геномика и протеомика

Лекция 1. Интегральные исследования геномов (бакалавры)

Составитель: проф. М.Р. Шарипова

Предмет

Геномика – наука, которая изучает молекулярную структуру и функцию полных геномов разных организмов

Цель

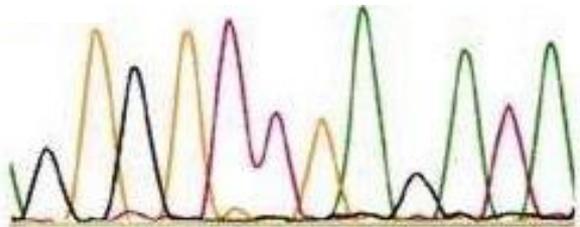
**Выяснение молекулярных механизмов
организации и функционирования
генома**

Задачи:

- Выяснение универсальных законов построения и организации геномов различного уровня сложности
- Характеристика полного транскриптома представителей различных биологических видов
- Характеристика полного протеома, определение границ протеома для биологического выживания вида



Фабрика автоматического секвенирования (расшифровки) ДНК



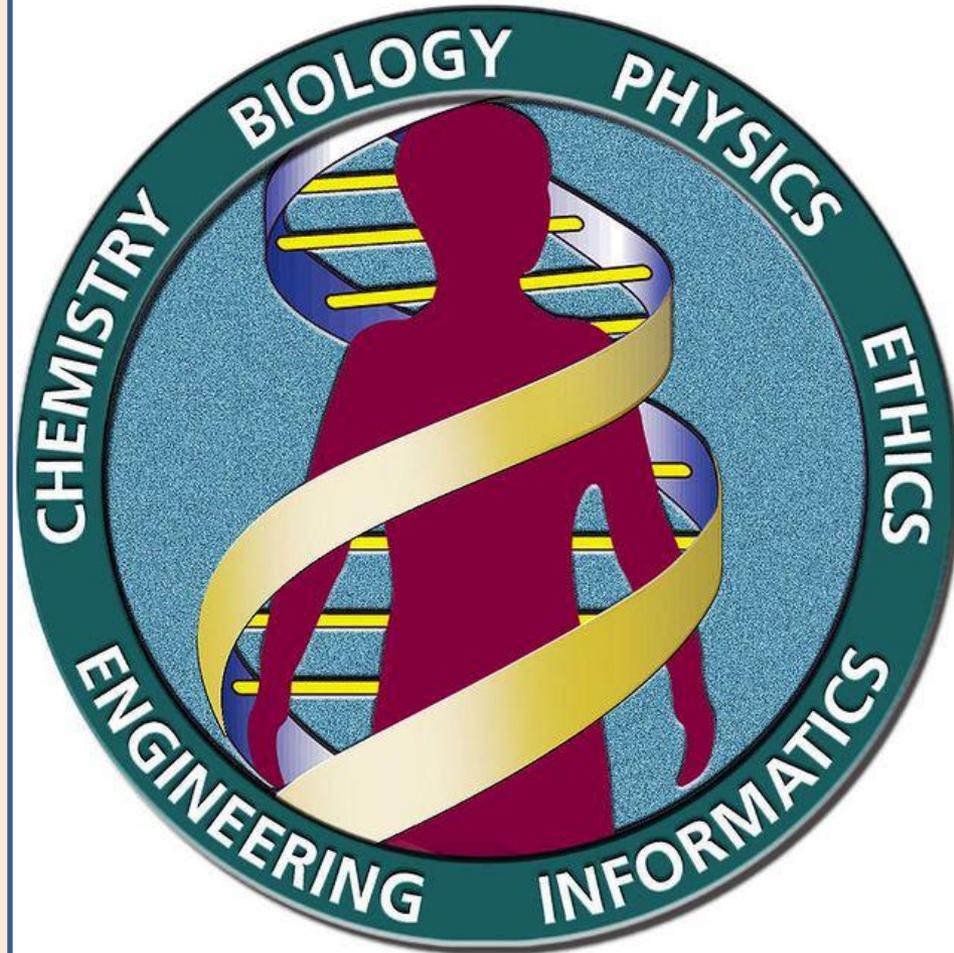
GGTTCATAAATTTGGA TCGTACTTATGGTTAATGC.

- 1977 – секвенирован геном бактериофага ф-Х174 (5 кб)
- 1981 – создан первый автоматический секвенатор
- 1983 – Кэри Мюллис открыл ПЦР для амплификации ДНК
- 1988 – создана мировая база данных NCBI
- 1990 – стартовал проект Геном человека, начались эксперименты по генной терапии
- 1995 – секвенирован первый геном бактерии *Haemophilus influenzae*
- 1996 - секвенирован геном одноклеточного эукариота - дрожжей *Saccharomyces cerevisiae*
- 1998 - секвенирован геном многоклеточного эукариотического организма – нематоды *C.elegans*
- 2001 – опубликован геном человека

Программа Геном человека

Геном человека

международная программа
целью которой является
определение нуклеотидной
последовательности –
секвенирование
ДНК человека,
идентификация генов и их
локализация в геноме -
картирование



Логотип программы

Программа Геном человека

**Международное научное сообщество
The Human Genome Project
HGP**

**Частная компания
CELERA**

**Объединила 4 000 ученых
Совет HUGO – 20 ученых
(во главе Джеймс Уотсон)**

Всего 400 ученых

Во главе:

Крейг Вентер

**поставил задачу быстрого и
дешевого секвенирования
человеческого генома (в
отличие от трехмиллиардного
международного проекта,
бюджет проекта ограничивался
300 млн**

США

ВЕЛИКОБРИТАНИЯ

КАНАДА

ФРАНЦИЯ

ГЕРМАНИЯ

ЯПОНИЯ

Условия взаимодействия HUGO

- **1)** Доверие, прозрачность, открытость во взаимодействии
- **2)** Вся информация выкладывается в базы данных с открытым доступом и доступна ученым любого уровня
- **3)** Вход в базы – бесплатный

Таким образом, были созданы все условия для научной интеграции ученых разных стран, это первый опыт масштабного взаимодействия

Подходы к секвенированию

HUGO (HGP)

ДНК



ХРОМОСОМЫ



КАРТИРОВАНИЕ ХРОМОСОМ



РЕСТРИКЦИОННОЕ
КЛОНИРОВАНИЕ



СЕКВЕНИРОВАНИЕ КЛОНОВ



ПОСТРОЕНИЕ КОНТИГА

CELERA

ДНК



РЕСТРИКЦИОННОЕ
КЛОНИРОВАНИЕ



СЕКВЕНИРОВАНИЕ КЛОНОВ

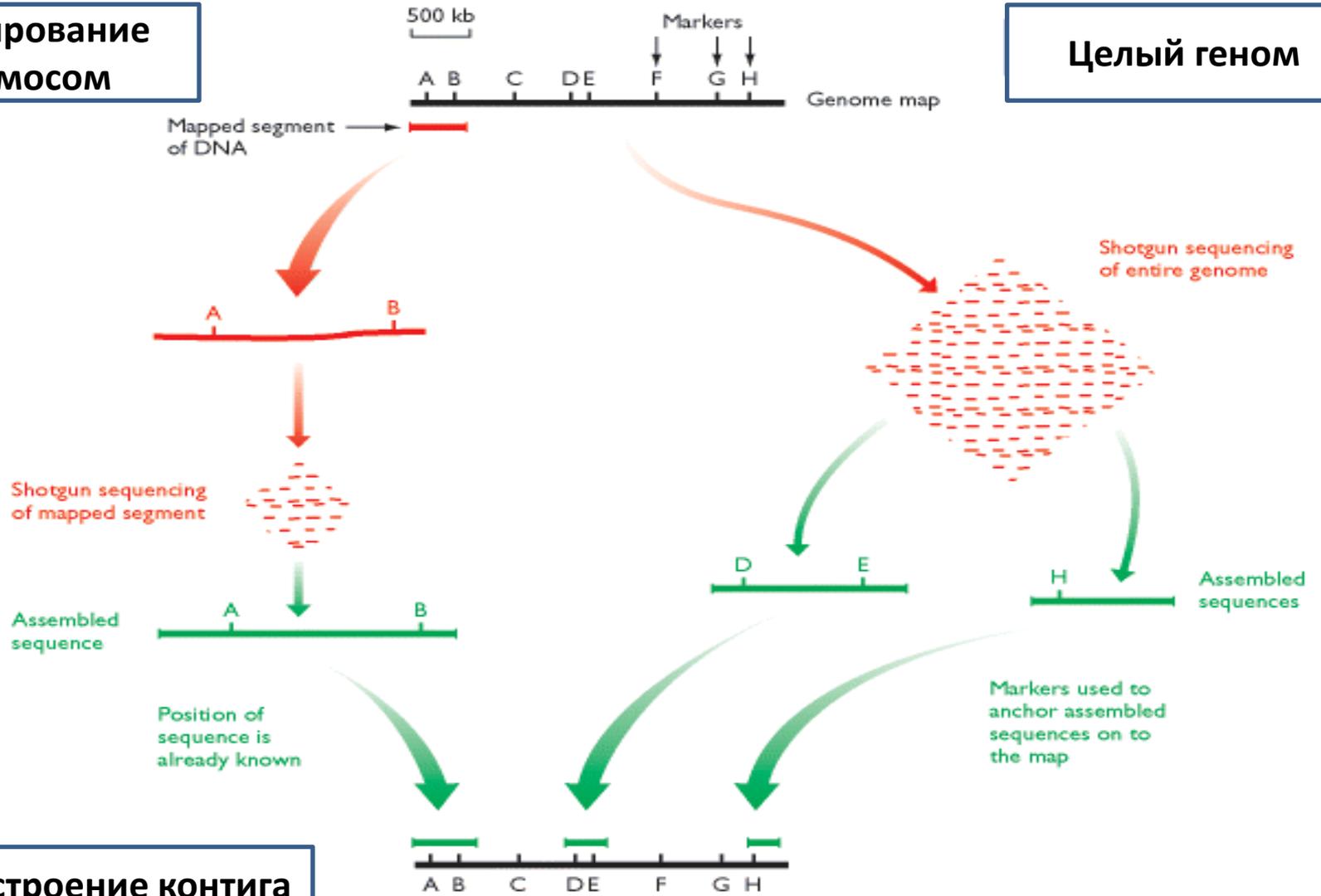


ПОСТРОЕНИЕ КОНТИГА

Shotgun-секвенирование

Клонирование
хромосом

Целый геном



Построение контига

Организация работы в фирме Селера

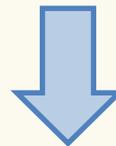
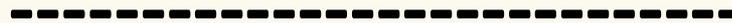
РОБОТЫ

+

КОМПЬЮТЕРЫ

+

ИСПОЛНИТЕЛИ



ИНДУСТРИАЛЬНАЯ НАУКА

- **Цель программы «Геном человека» – секвенировать геном человека**

Работа над проектом «Геном человека» официально стартовала **1 октября 1990 г.** с момента открытия финансирования и была рассчитана **до 2004 г.**

Стоимость программы составила 3 млрд \$

Задачи программы Геном человека

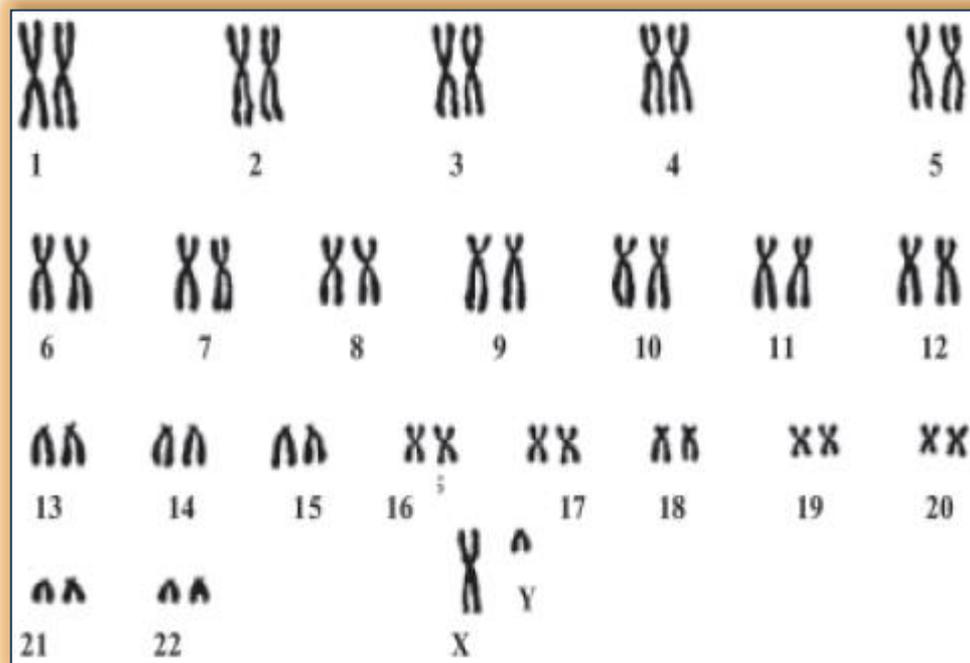
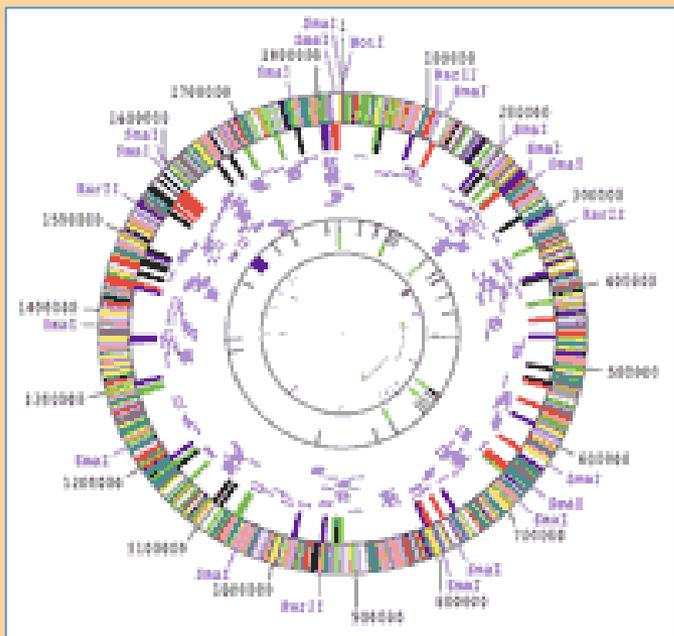
- 1) Картирование генома – составить детальную генетическую карту с разрешением 2 Мб
- 2) Составить физические карты каждой хромосомы, провести клонирование ДНК на фрагменты до 5 кб
- 3) Построить карту всего генома в виде клонов
- 4) Провести секвенирование клонов
- 5) Сборка последовательности ДНК хромосом и генома – построение контига
- 6) Нанести на полную нуклеотидную последовательность генома все гены человека

Подпрограммы

- Геномы бактерий
- Геном дрожжей
- Геном нематоды
- Геном дрозофилы
- Геном арабидопсиса
- Геном мыши
- Геном приматов

Цель:

1. Разработка новых методов (картирование, клонирование, хранение)
2. Возможность сравнения геномов
3. Изучение функции неизвестных генов на простых объектах

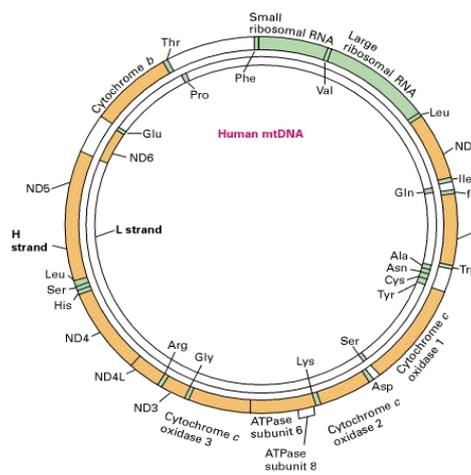


ДНК бактерий

хромосома +
плазмиды

ДНК человека

хромосомы +
митохондриальная ДНК



Первые пять лет главным в программе было разработка методов картирования геномов

Организм	Заболевание	Размер, Мб	Страна
<i>H. influenzae</i>	Менингит Пневмония	1.8	США
<i>Mycobacterium tuberculosis</i>	Туберкулез	4.4	США Франция
<i>Vibrio cholerae</i>	Холера	4.0	США
<i>Mycobacterium leprae</i>	Проказа	3.3	США
<i>Yersinia pestis</i>	Чума	4.7	США
<i>Salmonella typhi</i>	Брюшной тиф	4.5	США
<i>Shigella flexneri</i>	Дизентерия	4.6	США
<i>Plasmodium falciparum</i>	Малярия	23	США, Франция Япония
<i>Bacillus anthracis</i>	Сибирская язва	4.5	США

- В 2000 г. на ежегодной конференции в Ванкувере Крейг Вентер заявил, что секвенировано 90% генома
- 15 и 16 февраля 2001 г. журналы **Nature** и **Science** опубликовали итоги расшифровки генома человека

- Джеймс Уотсон:

«Все озабочены, чтобы расшифровать геном. Но это техническая задача. Рано или поздно, но это будет сделано. Ну вот после того как весь геном будет записан в суперкомпьютерах и все три миллиарда нуклеотидов будут известны – вот после этого я не знаю, сколько десятилетий нам понадобится, чтобы понять, что же мы расшифровали?»

Итоги секвенирования

1

**Гены: поиск и
подсчет**

2

**Технология
секвенирова
ния**

3

**Информацион
ные
технологии**

**Исследование этических, правовых и социальных
вопросов, возникших при расшифровке генома**

1. Поиск и подсчет генов:

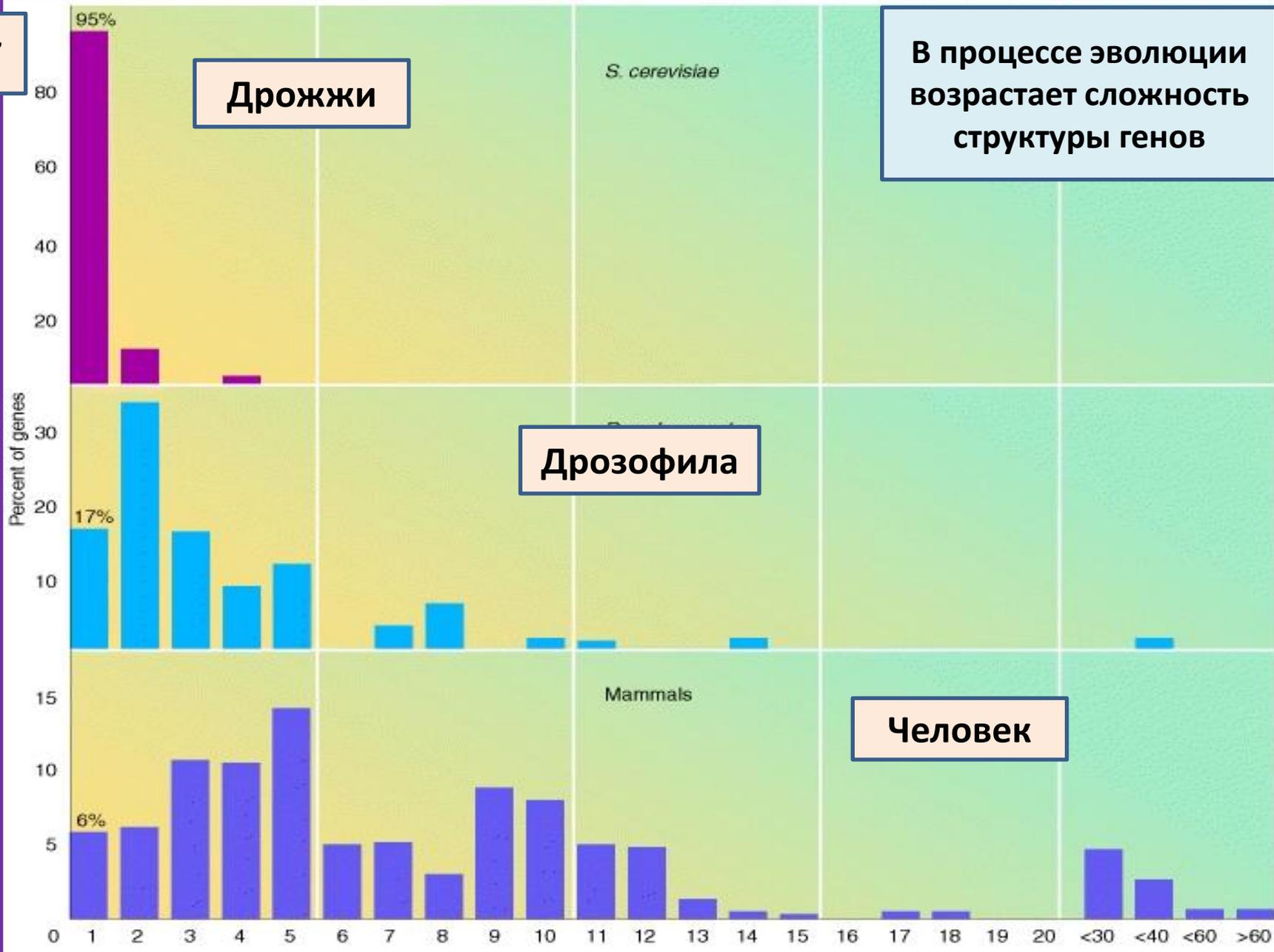
- 1) Всего в геноме человека около 25 000
- 2) Характерна высокая гомология с генами нижестоящих кланов, особенно по генам домашнего хозяйства
- 3) Большая часть генов в геноме с неизвестной функцией
- 4) Сложность строения генов и запутанность их экспрессии
- 5) В геноме человека доминирует некодирующая ДНК с избытком мобильных элементов, эволюция идет по пути разбавления геномов некодирующими последовательностями

ГЕНЫ,
%

Дрожжи

S. cerevisiae

В процессе эволюции
возрастает сложность
структуры генов



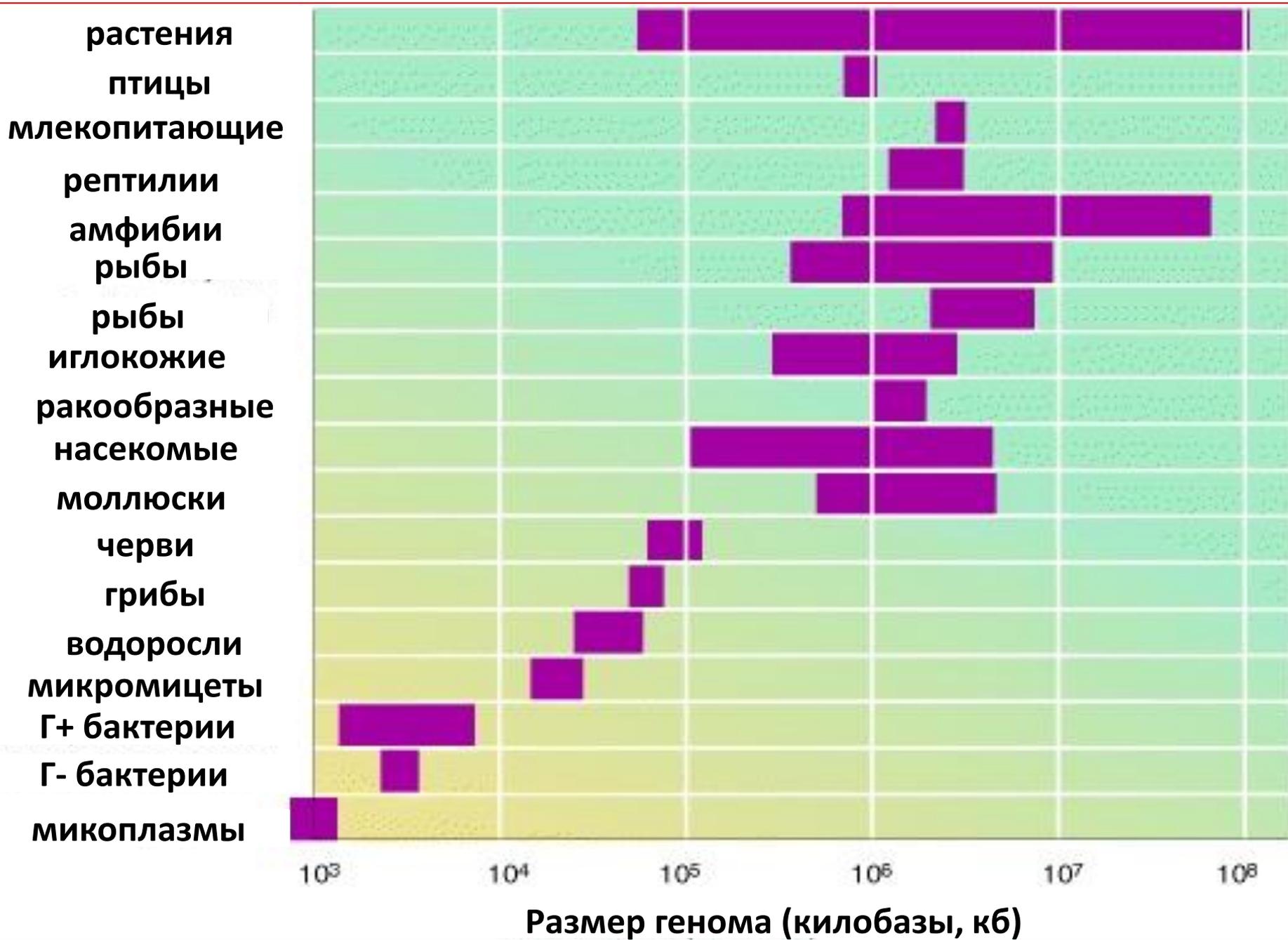
ЧИСЛО ЭКЗОНОВ

Секвенированные геномы: соотношение кодирующей и некодирующей ДНК

Геном	Размер, кб	Количество ORF	Год выполнения
<i>H. influenzae</i>	1 830	1 850	1995
<i>Mycoplasma genitalium</i>	580	468	1995
<i>E. coli</i>	4 639	4 289	1997
<i>B. subtilis</i>	4 214	4 099	1997
<i>Saccharomyces cerevisiae</i>	12 069	6 294	1997
Нематода	97 000	19 099	1998
Дрозофила	137 000	14 100	2000
Арабидопсис	115 428	25 498	2000
Человек	7 325 000	25 966	2001
Мышь	6 013 000	24 174	2002

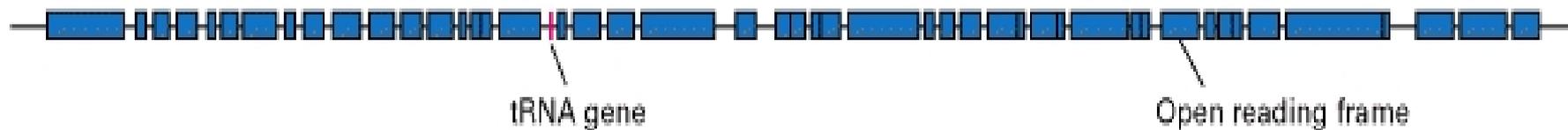
Секвенированные геномы: соотношение кодирующей и некодирующей ДНК

- 1) В процессе эволюции число генов растет не пропорционально количеству ДНК, а возрастает сложность их структуры
- 2) Эволюция эукариот идет по пути насыщения генома некодирующей ДНК – соотношение между размером генома и количеством генов растет от низших к высшим, у нематоды в 5 раз, а у человека в 300

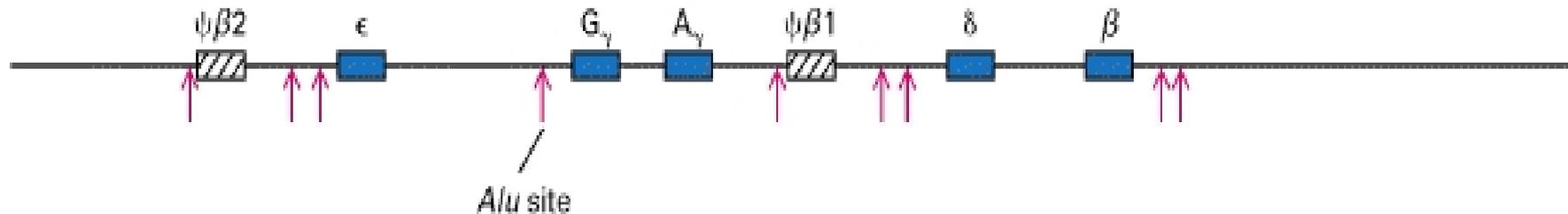


Сравнение 80-кб фрагментов: хромосомы III дрожжей *S. cerevisiae* (40 генов) и кластера бета-глобиновых генов человека (5 генов) на 11 хромосоме

S. cerevisiae (хромосома III)

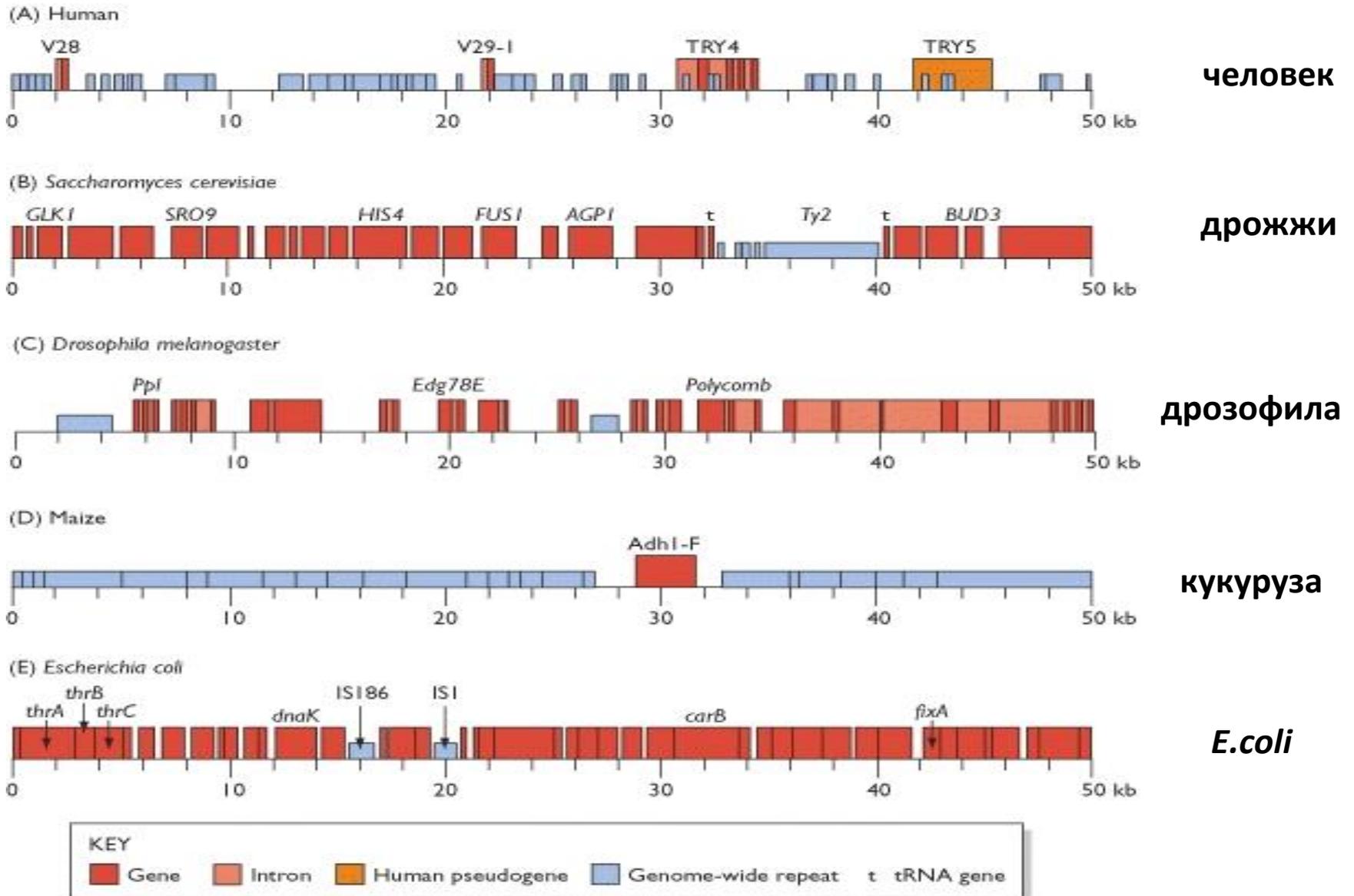


β -глобиновый генный кластер человека (хромосома 11)



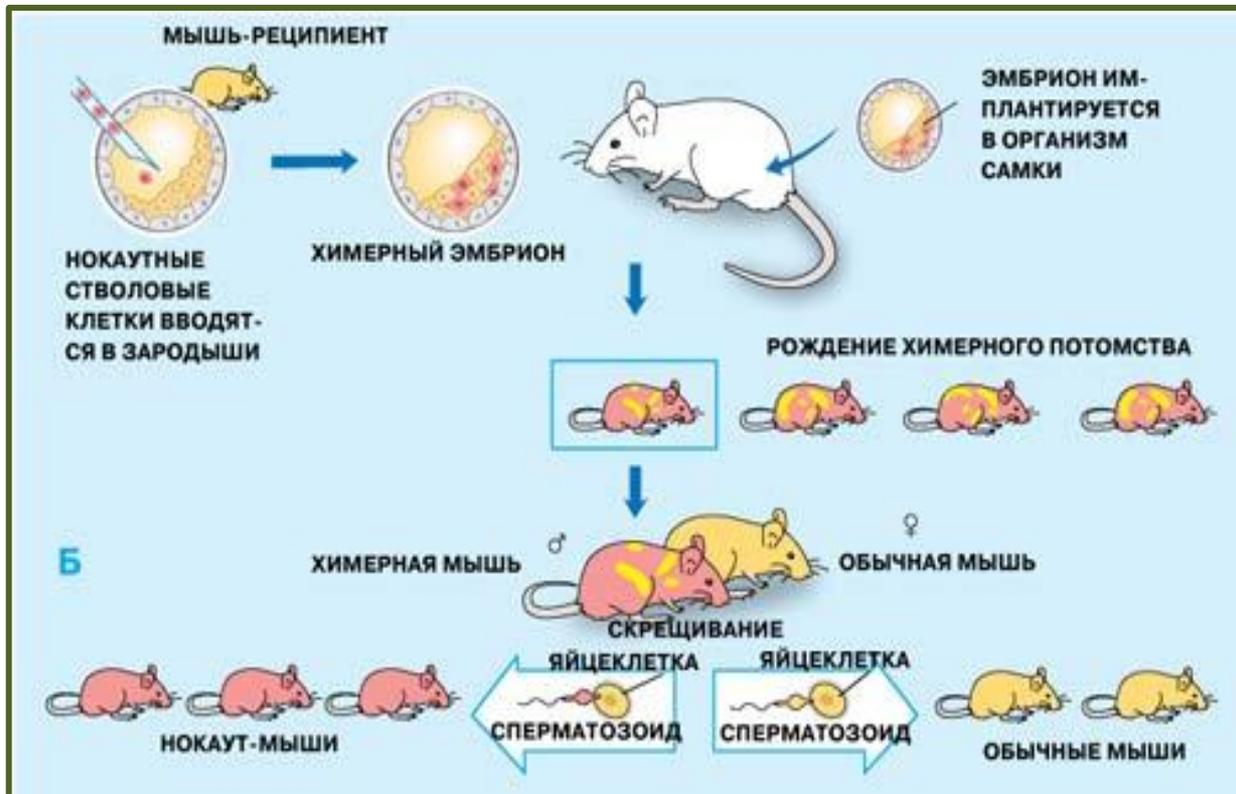
В процессе эволюции увеличивается некодирующая часть генома

Сравнение 50-кб фрагментов геномов человека, дрожжей, дрозофилы, кукурузы и *Escherichia coli*



КОМР: Knockout Mouse Project

- Нокаутируя последовательно гены в геноме мыши выясняют функции каждого из них



- Поскольку у человека и мыши многие гены сходны и выполняют одинаковые функции, нокаутные мыши служат для выяснения роли неизвестных генов

2. Развитие техники и технологии секвенирования:

- Визуализация хромосом, техника клонирования, создание искусственных хромосом, техника секвенирования и т.д.
- В итоге стоимость к завершению проекта сиквенса снизилась с 5\$ за 1 п.н. до 0.1\$
- Скорость секвенирования возросла с 10 000 п.н. до 1 000 000 п.н. в день
- Чтение 1-го млрд заняло 4 года, а 2-го – 4 месяца



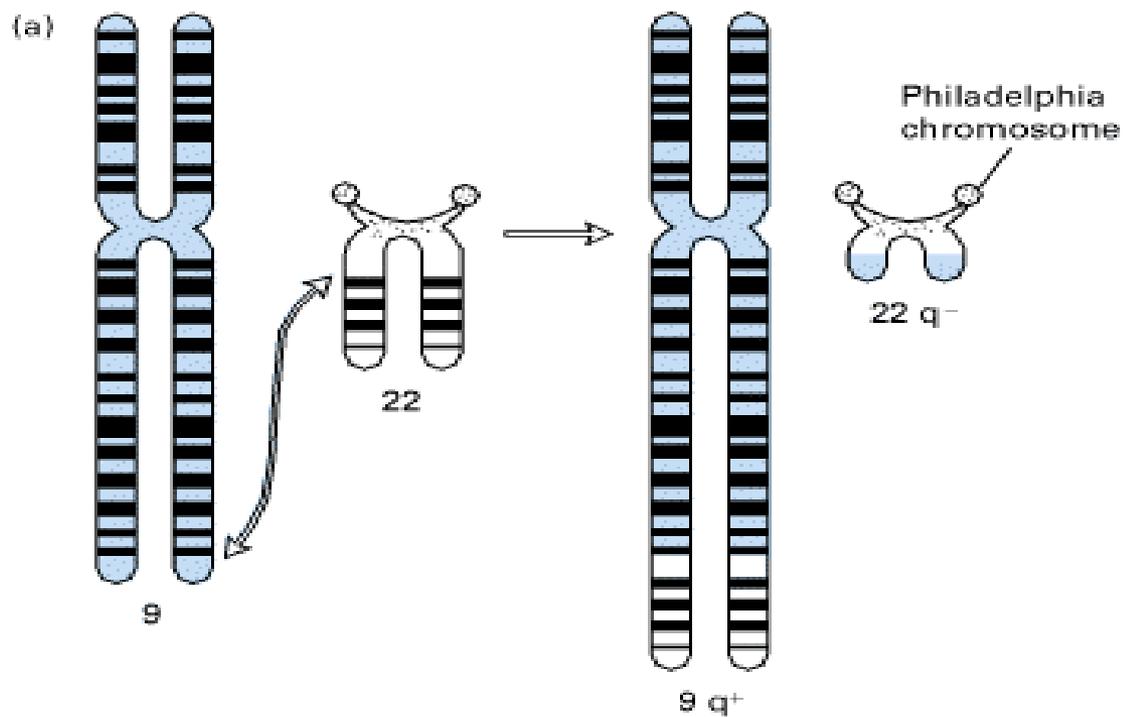
**FISH (fluorescence in situ hybridization) -
**флуоресцентная in situ
гибридизация -
окраска хромосом
человека****

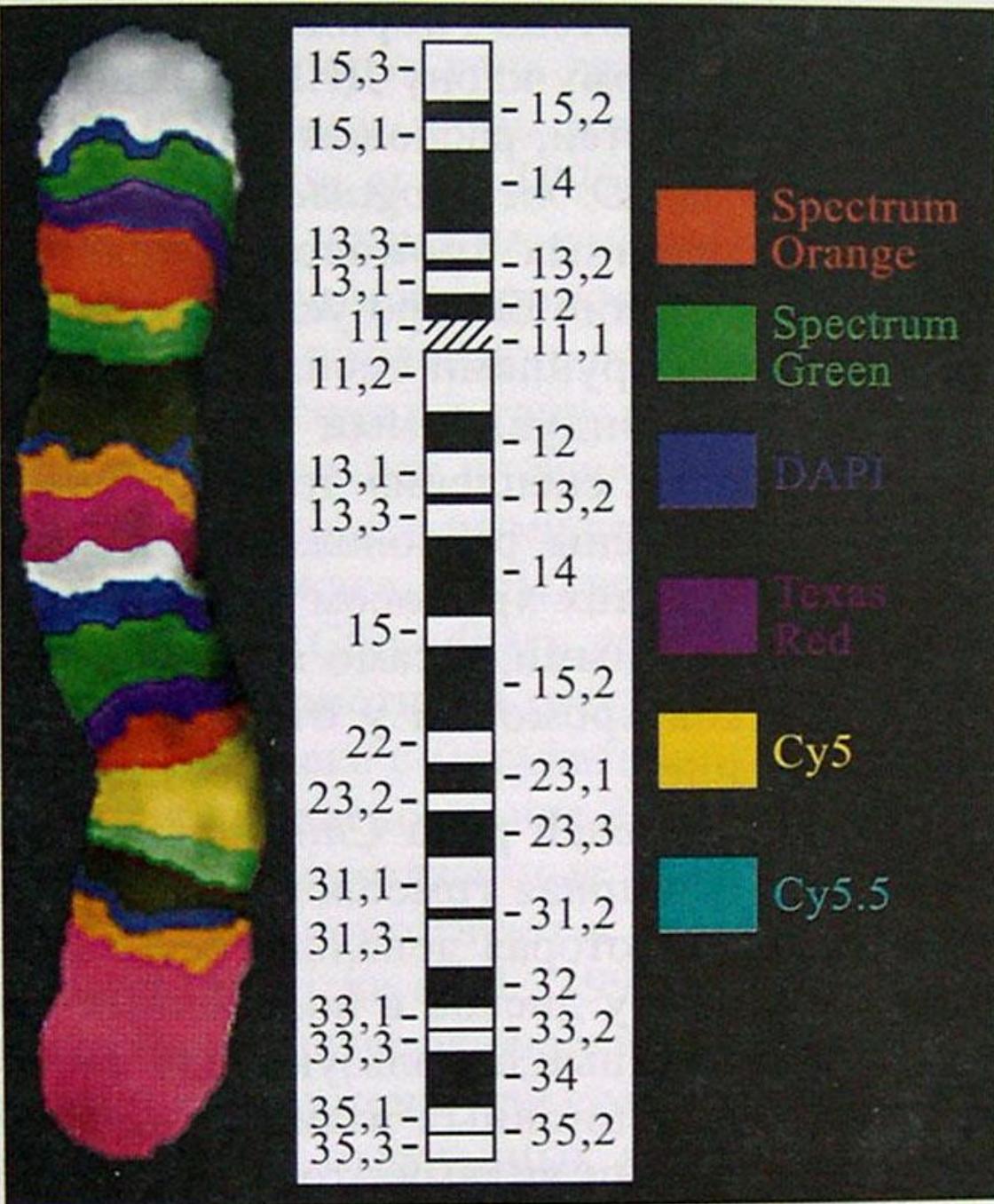


А) - Каждая хромосома
окрашена в свой
индивидуальный цвет

Б) – наличие зеленого и
красного цветов в 8-й и 11-й
хромосомах свидетельствует
о транслокации между ними

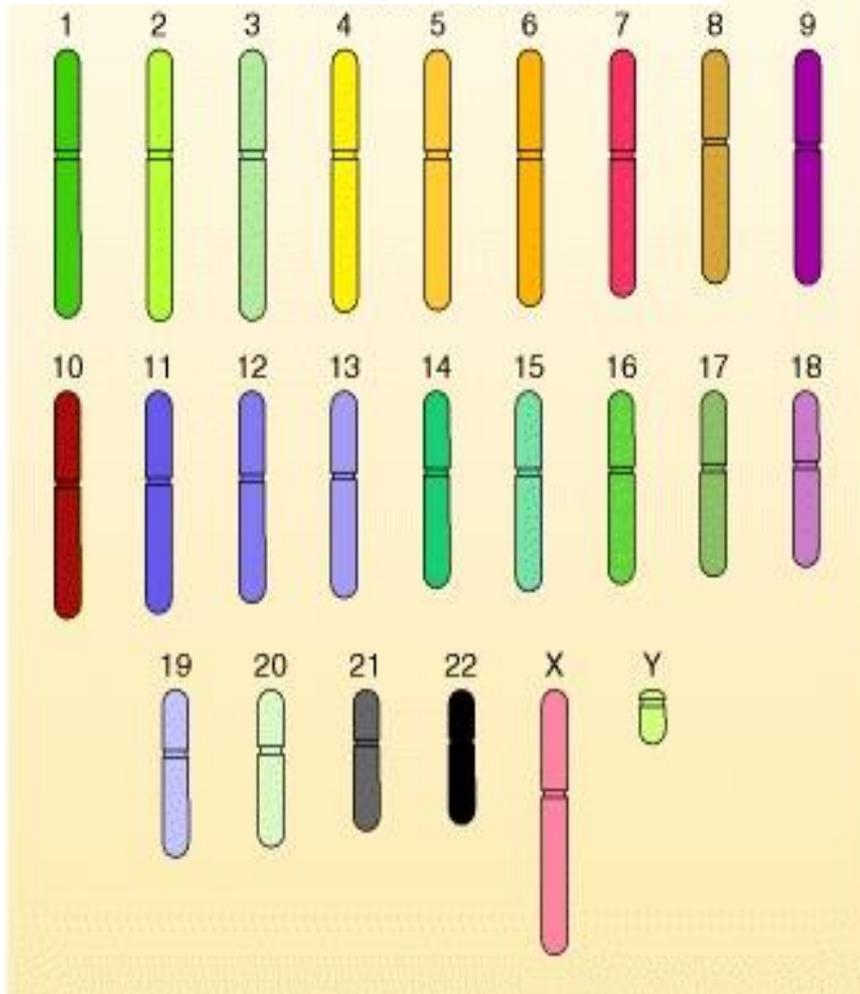
Транслокация между 9 и 22 хромосомой



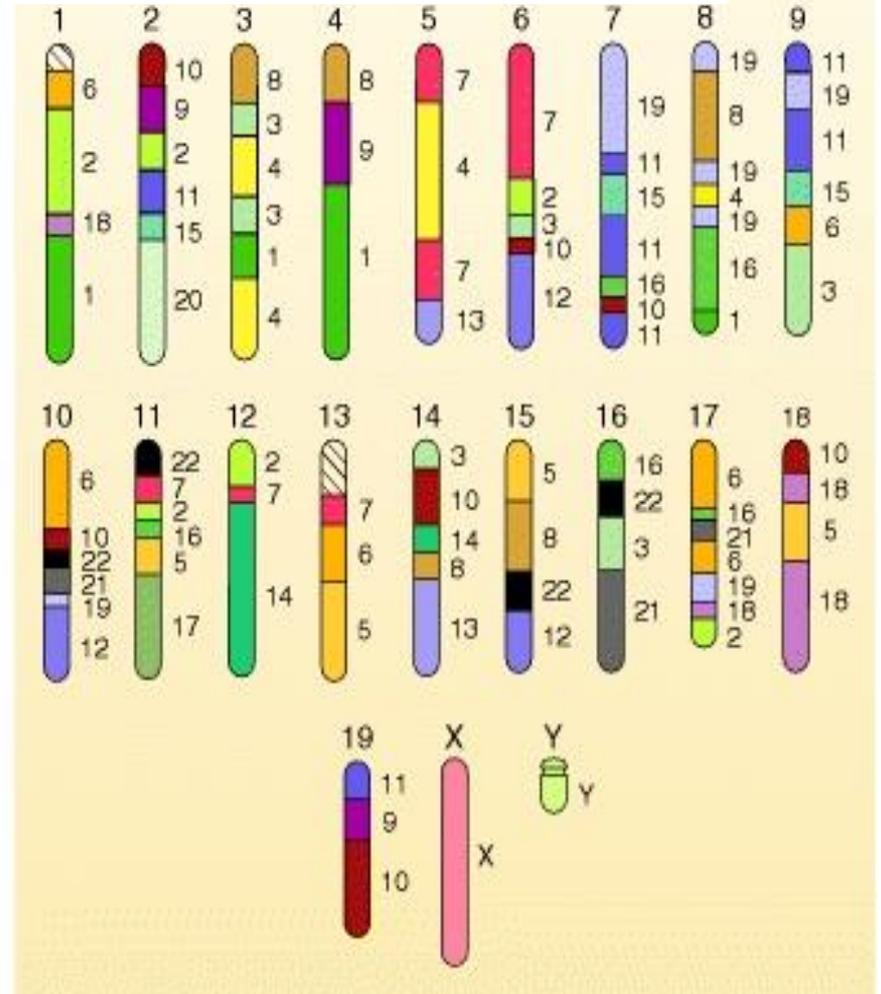


**Многоцветное
окрашивание
пятой
хромосомы
человека**

Человеческие хромосомы



Мышиные хромосомы



Синтенность человеческих и мышиных хромосом. В процессе эволюционной дивергенции происходит перегруппировка гомологичных блоков генов в различных комбинациях. Эти данные важны для изучения эволюции геномов

Принципы секвенирования нового поколения

- Не используется метод терминации цепи Сенгера, для создания ампликонов используется ПЦР
- Отпадает необходимость в клонировании ДНК
- В основе поэтапная реакция определения последовательности единичных молекул ДНК/РНК - пирофосфорилирование (Roche), лигазная реакция (SOLiD) , полимеразная реакция (Illumina)
- **Нанотехнологический** подход на базе миниатюризации и автоматизации

Информационные технологии и базы данных для хранения информации:

- Базы данных (НСВІ)
- Новые алгоритмы чтения и расшифровки
- Качественный уровень компьютеров

NCBI - Национальный центр биотехнологической информации

NCBI:

- **GenBank** - база данных о ДНК, РНК и белках
- **PubMed** - база данных для научной литературы
- **TaxBrowser** - база данных о таксономической информации
- **Taxonomy** - обеспечивает поиск данных о конкретном биологическом виде
- **BLAST** - стандартные программы биоинформатики
- Базы данных доступны через поисковую систему **Entrez**

Задачи NCBI:

- Хранение и анализ данных
- Компьютерная обработка данных
- Программное обеспечение для работы с последовательностями
- Координирование мировой информации, связь с другими базами данных

NCBI (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION)

PubMed

OMIM

Статьи

Книги

NCBI

GenBank

Нуклеотиды

Домены

Белки

Taxonomy

Геномы

Заключение по результатам программы Геном человека

- Автоматизация секвенирования и информационные технологии привели к созданию наукоемкой биотехнологической индустрии
- Был достигнут качественно новый уровень в развитии молекулярной генетики – появляется **НОВАЯ НАУКА - ГЕНОМИКА**

Геномика - принципы новой научной идеологии

1. Геномика является обратной генетикой:

сначала секвенируют фрагменты генома и потом выясняют, что они кодируют, в отличие от генетики, где наоборот, исследования направлены от признака к идентификации гена

2. Переход от изучения индивидуальных генов к исследованию полномасштабных геномов

3. Выяснение закономерностей экспрессии геномной последовательности

Принципы новой научной методологии

- **Тотальная автоматизация, роботизация и китизация экспериментальной работы**
- **Создание информационных баз данных, которые обеспечивают сравнительный анализ структур на основе гомологий**
- **Исследования генома путем амплификации и манипуляции его фрагментов**

Новые подходы позволили:

- **Перейти к интегральному исследованию геномов разных организмов**
- **Проводить сравнительный анализ геномов различного уровня сложности**
- **Показать, что геном функционирует не как линейный набор генов, а как сложная многоуровневая система на основе обратных связей**

Концептуальные положения новой науки

Геномика подразделяется на два основных направления – структурная и функциональная

- **Структурная** геномика характеризует физическую природу целых геномов
- **Функциональная** геномика характеризует *транскриптомы*, полный набор транскриптов, продуцируемых организмом и *протеомы*, полный набор кодируемых белков

Структурная Геномика

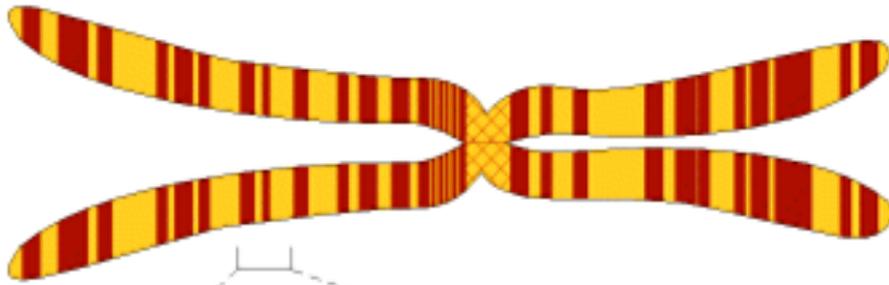


Секвенирование генома



Выяснение структурной организации
генома

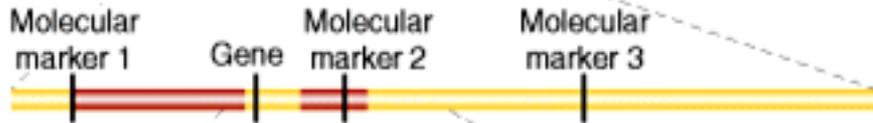
(a)



ASSIGNMENT OF GENE(S) TO A CHROMOSOME

- Linkage
- In situ hybridization
- Cell hybrids
- PFGE
- Rearrangement break points

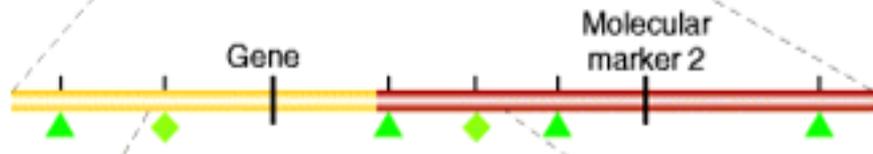
Этапы геномного анализа – картирование и секвенирование



CHROMOSOME MAPPING

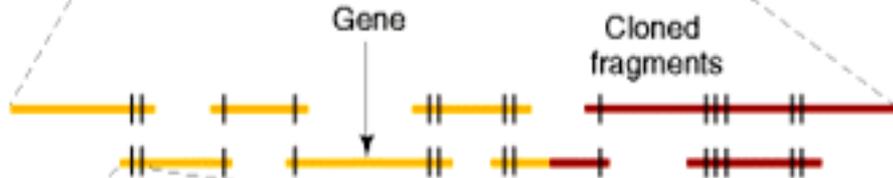
- Recombination
- Radiation hybrids

Структурная геномика: хромосомное и физическое картирование



RESTRICTION MAPPING

- Long cutters (sites ◆, ▲)
- PFGE



PHYSICAL MAPPING

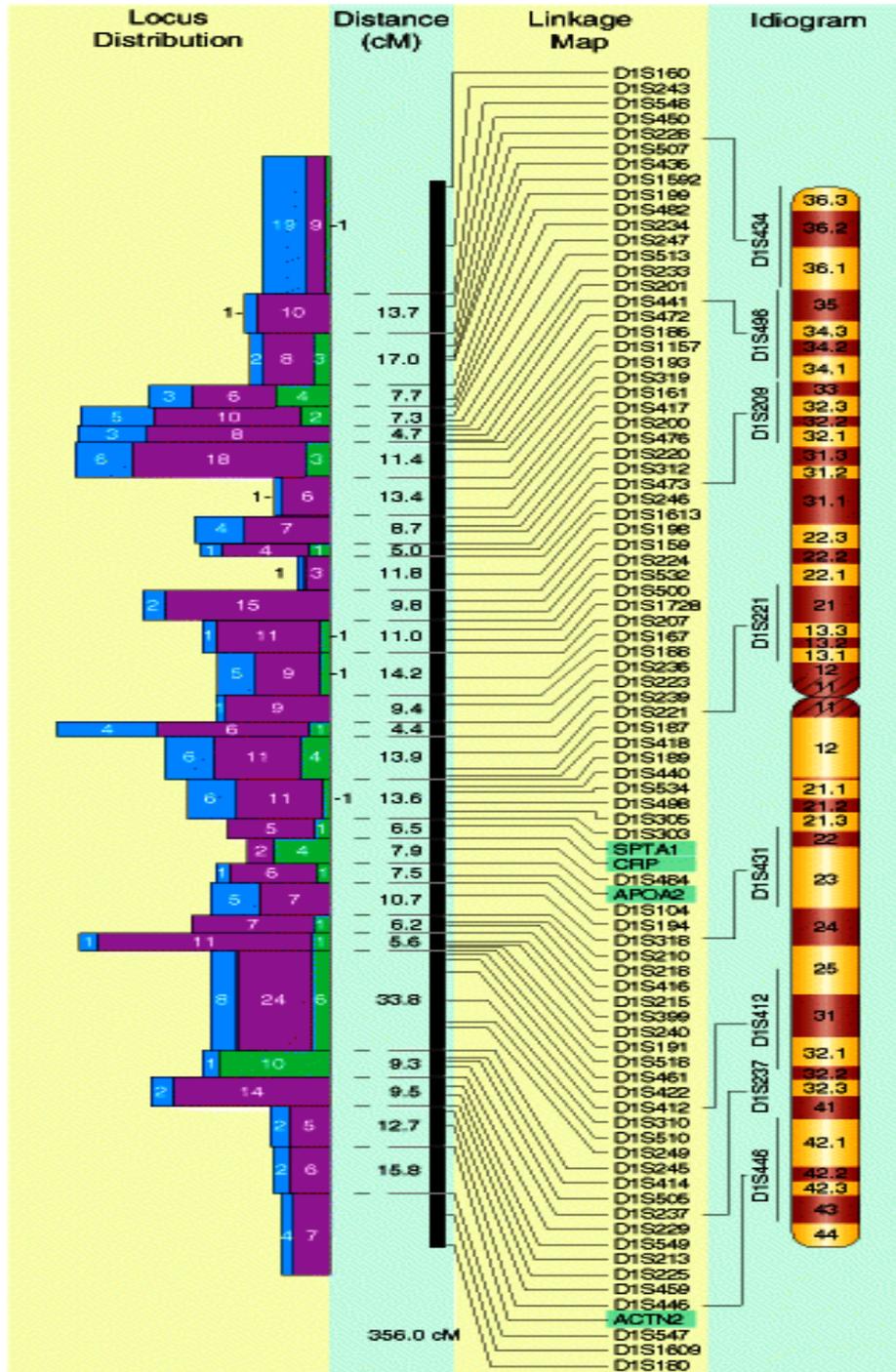
- YACs
- Cosmids
- Alignment tags (I)

TTAGCTTAACGTA CTGGTACCGTACCGTGGCTTAT

DNA Sequencing

Маркеры

1. Гены
2. Однонуклеотидные замены
3. Сайта рестрикции с мутацией
4. Полиморфные короткие повторы



Маркерные гены и молекулярные маркеры 1-й хромосомы человека

Key

- Short sequence length polymorphisms
 - Other DNA polymorphisms
 - Genes
 - Genes included on the linkage map
- } DNA markers

Транскриптомика



мРНК



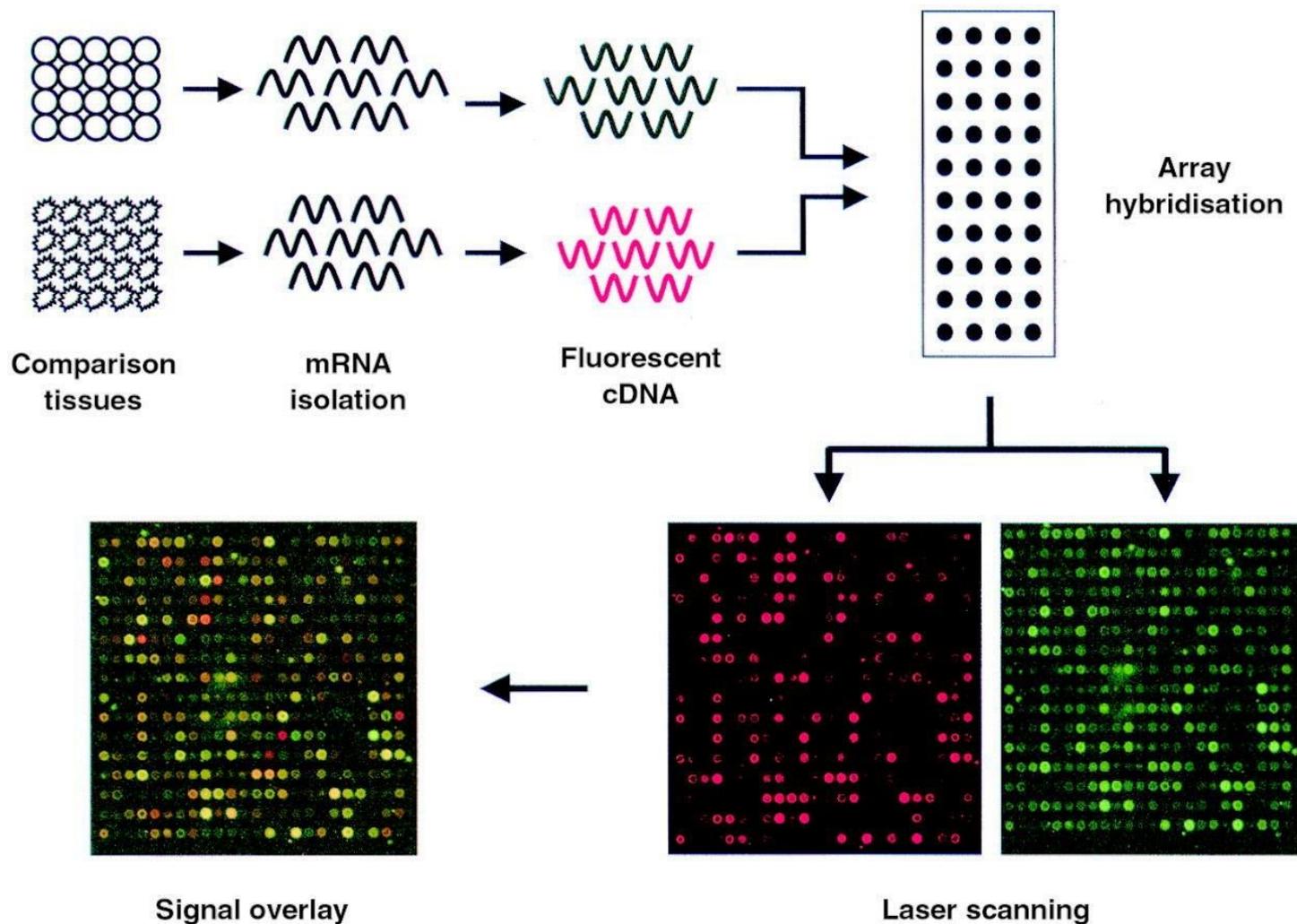
Библиотека кДНК



Перечень всех кДНК
(базы данных)

ДНК-микрочипы для анализа экспрессии генов

DNA microarray - технология



Протеомика



Идентификация всех белков

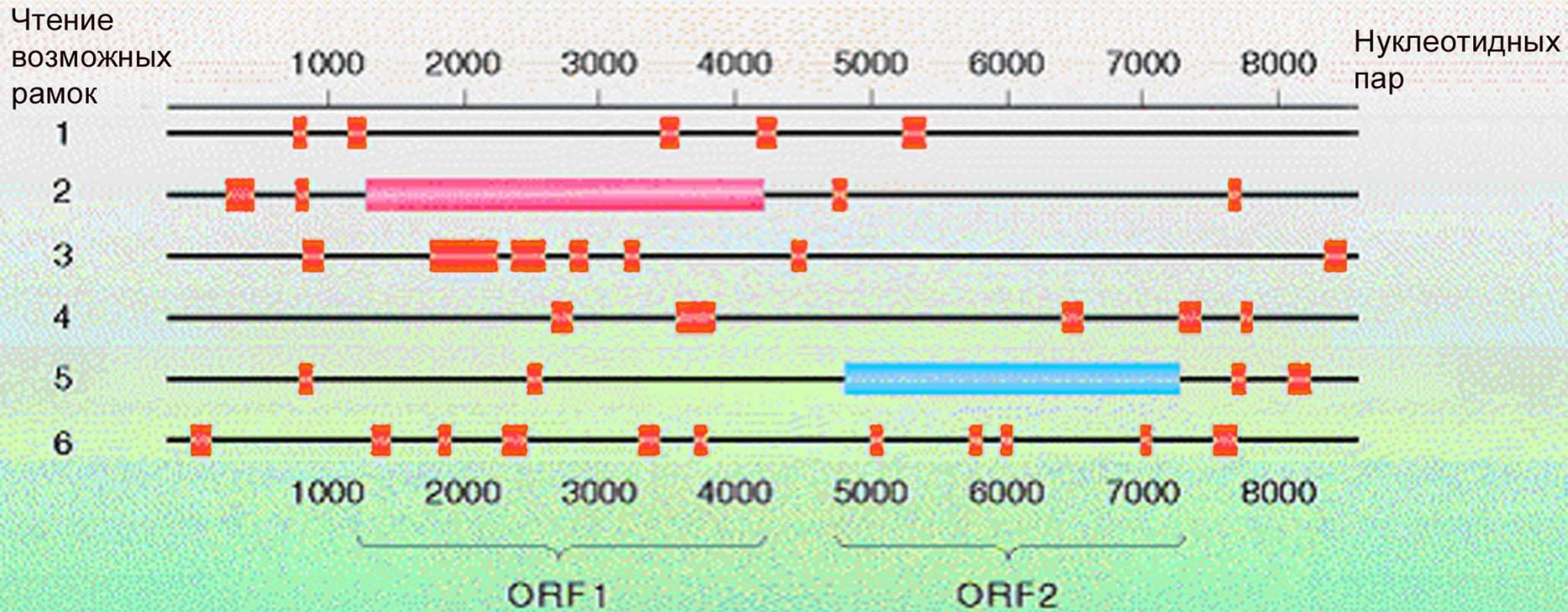


2D-электрофорез, Масс-спектрометрия



Перечень всех белков
(базы данных)

Поиск открытых рамок считывания



Практическая Геномика

(идентификация всех генов):

- **Генодиагностика**
- **Генотерапия**
- **Генопаспортизация**
- **Фармакогеномика**
- **Превентивная геномика**
- **Создание новых вакцин**

- **Генотерапия** – комплекс генно-инженерных методов, направленных на внесение изменений в ДНК соматических клеток в целях лечения заболевания
- **Фармакогеномика** – отрасль фармакологии, которая исследует влияние генетической вариации каждого человека в его ответе на лекарственное средство
- **Гено- или ДНК-диагностика** – комплекс методов, позволяющих обнаружить ДНК патогенов или изменения в ДНК, вызывающие заболевание

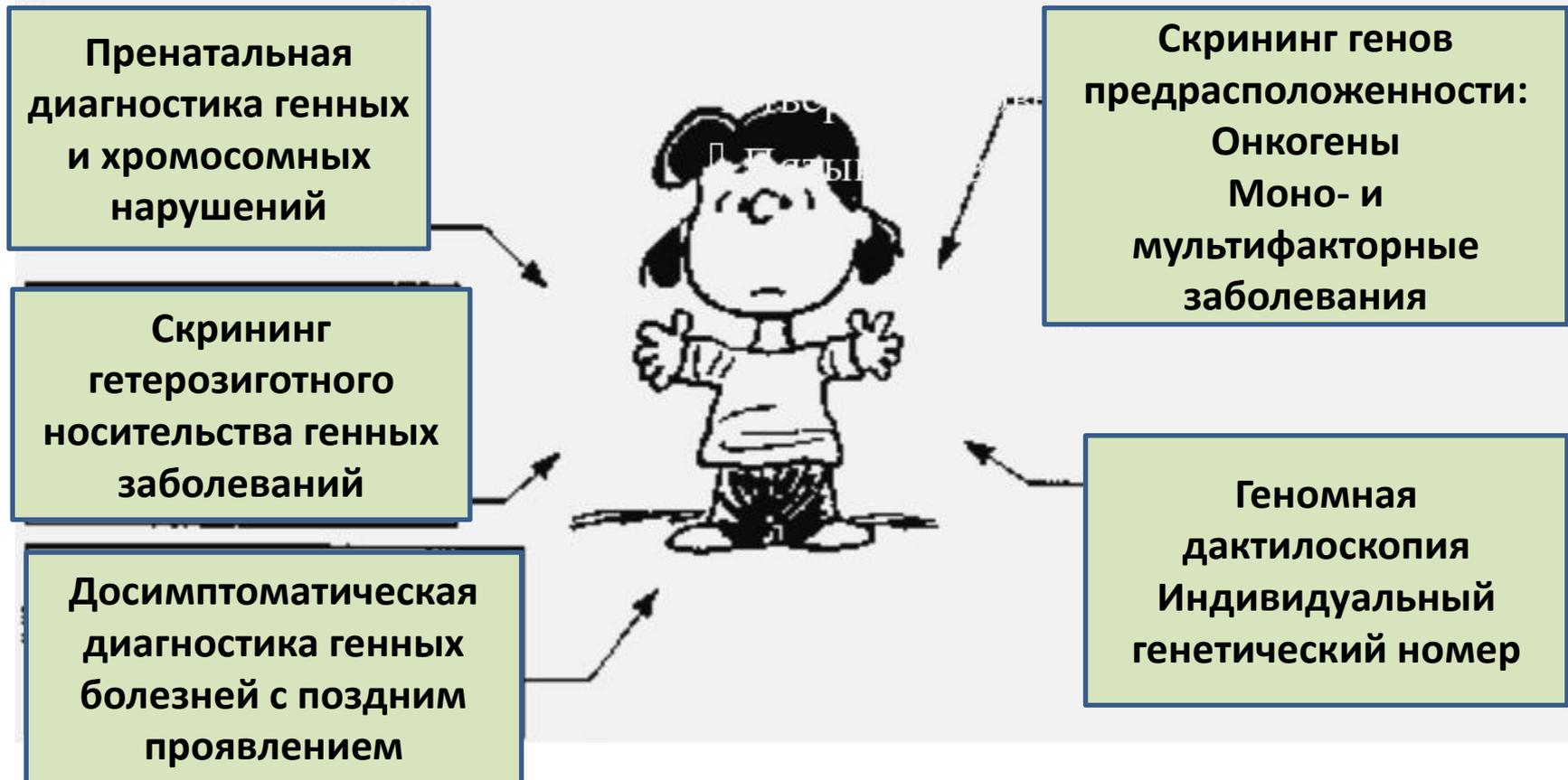
ГЕНОДИАГНОСТИКА

- **Моногенные заболевания – поврежден один ген**
- **Полигенные заболевания – повреждены несколько генов**
- **Заболевания при нарушении систем регуляции**
- **Мультифакториальные – экспрессия генов зависит от условий среды**

- **Создание новых вакцин** - ДНК-вакцины, генно-инженерная конструкция, которая после введения в клетки обеспечивает синтез белков патогенов или опухолевых антигенов и вызывает иммунную реакцию
- **Превентивная геномика** – ДНК-тестирование с целью персонализированной профилактики заболеваний и коррекции образа жизни, чтобы избежать опасных заболеваний
- **Генопаспортизация** – создание генетических паспортов

Проекты создания «генетического паспорта» гражданина

*Исследование этических, правовых и социальных вопросов, возникших при расшифровке генома



ГЕНЕТИЧЕСКАЯ КАРТА РЕПРОДУКТИВНОГО ЗДОРОВЬЯ

(2)

Кариотип

(1)

Медико-генетическое
консультирование
супружеской пары

(3)

Диагностика
гетерозиготного
носительства :

- Муковисцидоз;
- Миодистрофия Дюшенна;
- Гемофилия А;
- Фенилкетонурия;
- Аденогенитальный синдром;
- Спинальная мышечная атрофия.

(4)

СВЕДЕНИЯ О СУПРУГЕ:

1. Кариотип (2);
2. Тесты на гетерозиготное носительство мутаций наиболее частых моногенных болезней (3);

КОНСУЛЬТАЦИИ ГЕНЕТИКА и АКУШЕРА;
ИНФОРМАЦИЯ ДЛЯ ВРАЧА и БЕРЕМЕННОЙ,
ВЫРАБОТКА ТАКТИКИ ВЕДЕНИЯ БЕРЕМЕННОСТИ
ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

(5) Тестирование наследственной “предрасположенности”

Тромбофилия: *FV, MTHFR, PAI-1, PLAT, GPIIIa, Pr, Fb* (7)

Гестозы: *GSTPi; PAI-1, TNF- α ; eNOS; ACE; PON; GP-IIIa, HLA-G GSTV-1, mEPHX* (10)

Привычное невынашивание :
GSTM1; GSTT1; GSTPi. ; DRB1; DQA1, DQB1, MTHFR (7)

Диабет 1 типа -*HLA DR и DQ (DR3 и DR4) Mic-A; VDR3 CTLA4*(6)

Диабет 2 типа -*DQB1, ACE, TNF α PRARA, PRARD, TCF7L2*

Эндометриоз: *GSTT1; GSTM1; CYP19 miEPOX, NAT-2 TNF α ; IL4R; CYP1A1* (8)

Остеопороз - *VDR 3; COL1A1; CALCR; ER-1* (4)

Бронхиальная астма *GSTT1 GSTM1 TNFA, IL4, IL4R, Nos1* (6)

Нерасхождение хромосом в мейозе и дефекты зародка нервной трубки *MTHFR; MTTR.* (2)

ВАРИАНТ «ГЕНЕТИЧЕСКОЙ КАРТЫ ЗДОРОВЬЯ РЕБЕНКА»

Год рождения -
Национальность -

Кариотип

Диагностика гетерозиготного носительства :

- Муковисцидоз;
- Миодистрофия Дюшенна;
- Гемофилия А;
- Фенилкетонурия;
- Адрено-генитальный синдром;
- Спинальная мышечная атрофия.

ИНФОРМАЦИЯ ДЛЯ ВРАЧА РОДИТЕЛЕЙ, ВЫРАБОТКА ТАКТИКИ КОРРЕКЦИИ ОБРАЗА ЖИЗНИ, ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

Тестирование наследственной «предрасположенности»

Бронхиальная астма: GSTM1, GSTT1, GSTP1, CC16, IL4, IL4R, NOS1, TNF-alpha

Остеопороз - VDR 3; COL1A1

Диабет I -DQA1, DQB1, MIC-A, CTLA4

Артериальная гипертензия – ACE, AGT, AGTR1, AGTR2, BkR, REN, ADRB2, ADRB1, MTHFR, NOS3, MTRR, ApoE, ApoCIII, PRAR-γ

Наследственная тромбофилия - MTHFR, FV, PAI-1, FGB, GPIIIa/b, FII

Метаболический синдром - ApoE ApoCIII, AGT, ACE, AGTR1, AGTR2, BKR, REN, ADRB1, ADRB2, DQB1, TNFA, DRD-2A, SR, IGF1, PRAR-δ, PRAR-α, PRAR-γ, UCP2, UCP3

Лейкозы - транслокации; CYP1A1, GSTM1, CYP2C9, TPMT

Трансплантология - CYP2D6, CYP2C9, CYP2C19, TPMT/HLA

Анализ генов, влияющих на формирование зависимости к алкоголю и наркотикам - DRD-2A, SR

Предрасположенности к определенным видам спорта, фитнеса - ACE, AGT, AGTR1, BkR, REN, AGTR2, MTHFR, ADRB2, ADRB1, ApoE, NOS3, GPIIIa/b, VDR, AR, AMPD1, PGC1A, CnB, ACTN3, DRD-2A, SR, IGF1, PRAR-δ, PRAR-α, PRAR-γ, UCP2, UCP3

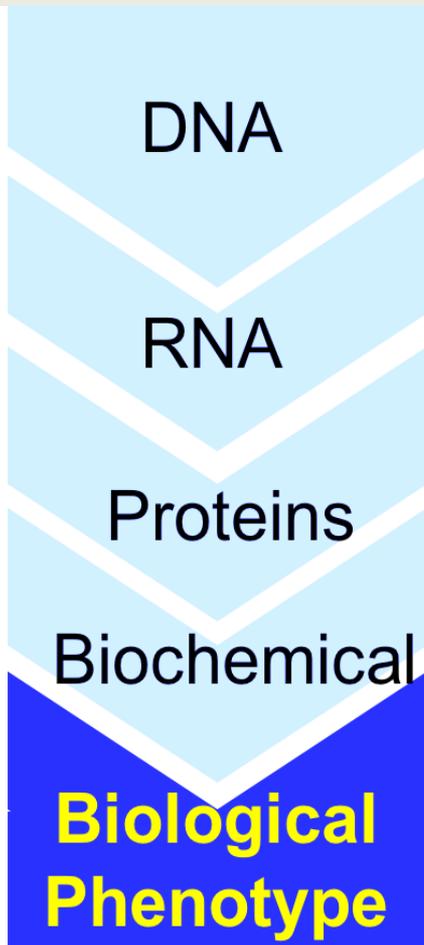
Устойчивость к ВИЧ инфекции - 32delCCR5/+

Международная программа «Протеом человека» (Сидней, 2010), The Human Proteome Project, HPP



Программа Метаболом человека

Human Metabolome Project (HMP)



Genomics



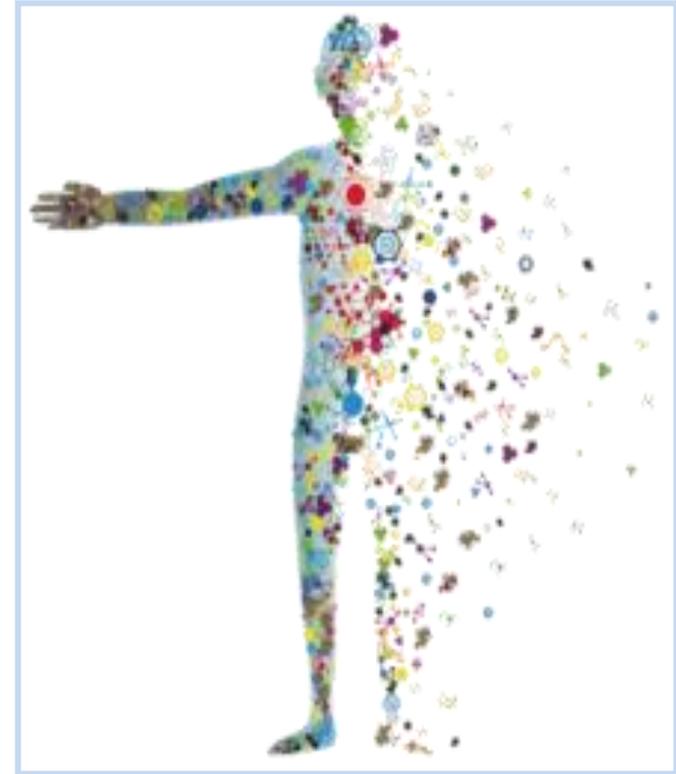
Transcriptomics



Proteomics



Metabolomics



Центральная догма биологии
показывает поток информации от ДНК к
фенотипу: от геномики к метаболомике

Применение метаболомики в медицине

- **Основная идея** метаболомики заключается в обнаружении специфических биомаркеров в биологическом образце для диагностики ряда заболеваний
- **Методы:** масс-спектрометрия, ядерный магнитный резонанс, хроматография
- Исследования направлены на описание качественных характеристик метаболома биологического образца пациента с патологией