

PGConf.Сибирь 2022

Геномный взгляд на эволюционный прогресс

ТРИФОНОВ Владимир Александрович

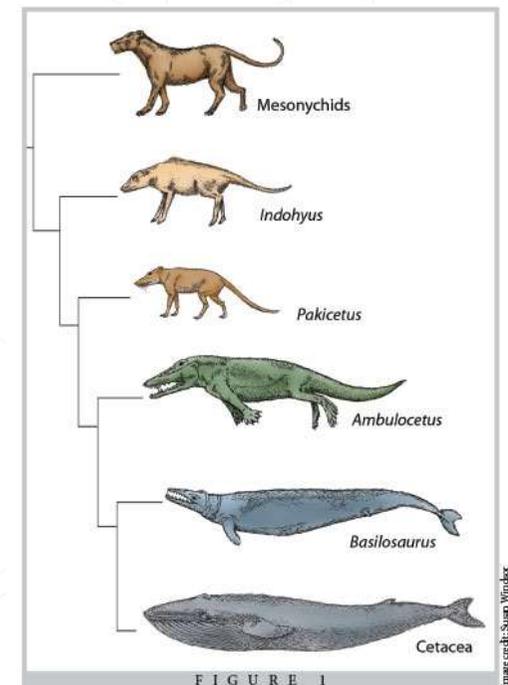
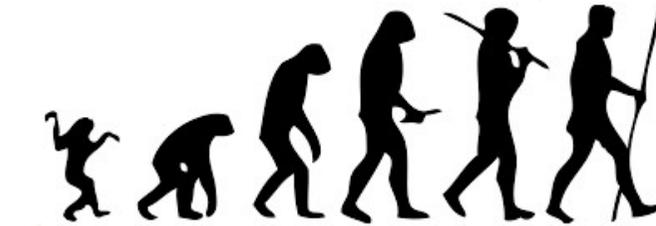
д.б.н., профессор РАН

Институт молекулярной и клеточной биологии СО РАН

НГУ

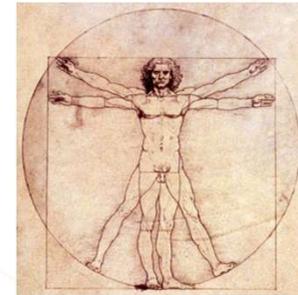
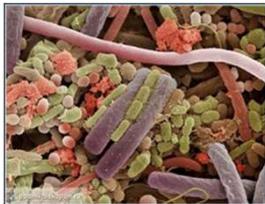
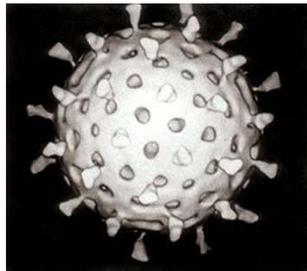
 PostgresPro

Почему происходит усложнение многих организмов в процессе эволюции и какие у этого усложнения генетические основы?



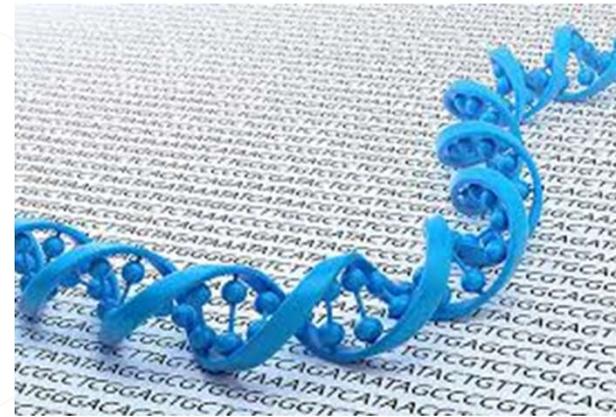
Что изучает геномика?

ГЕНОМЫ разных организмов, от вирусов до высших позвоночных

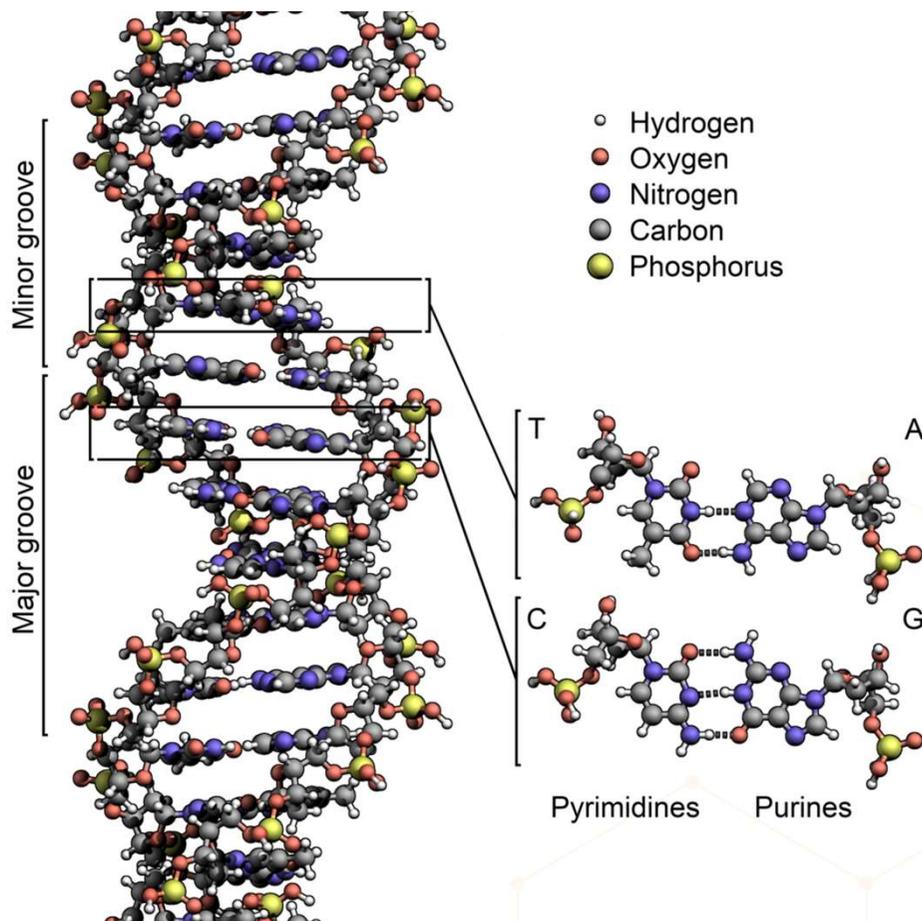


Что такое геном?

- Генетический материал – инструкция как построить клетки и организм и как ими управлять!
- Записан в виде последовательности нуклеотидов в ДНК

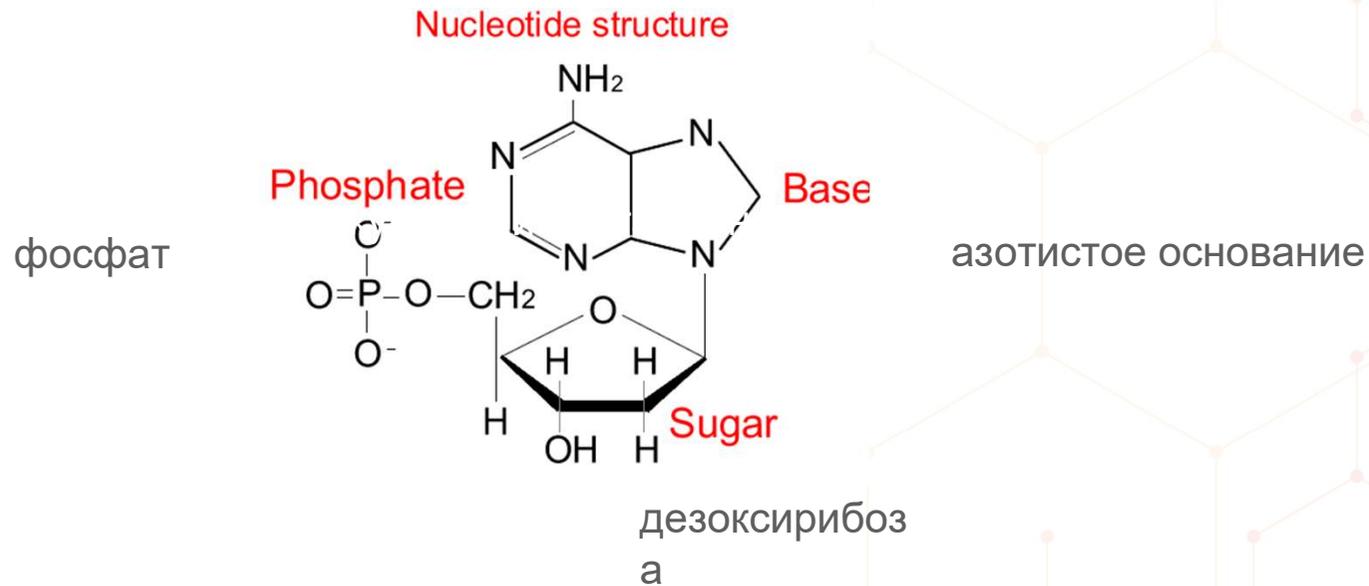


Что такое ДНК?



Нерегулярный биополимер, состоящий из остатков четырех нуклеотидов: А, Г, Т, Ц.

Молекулярный состав генома эукариот

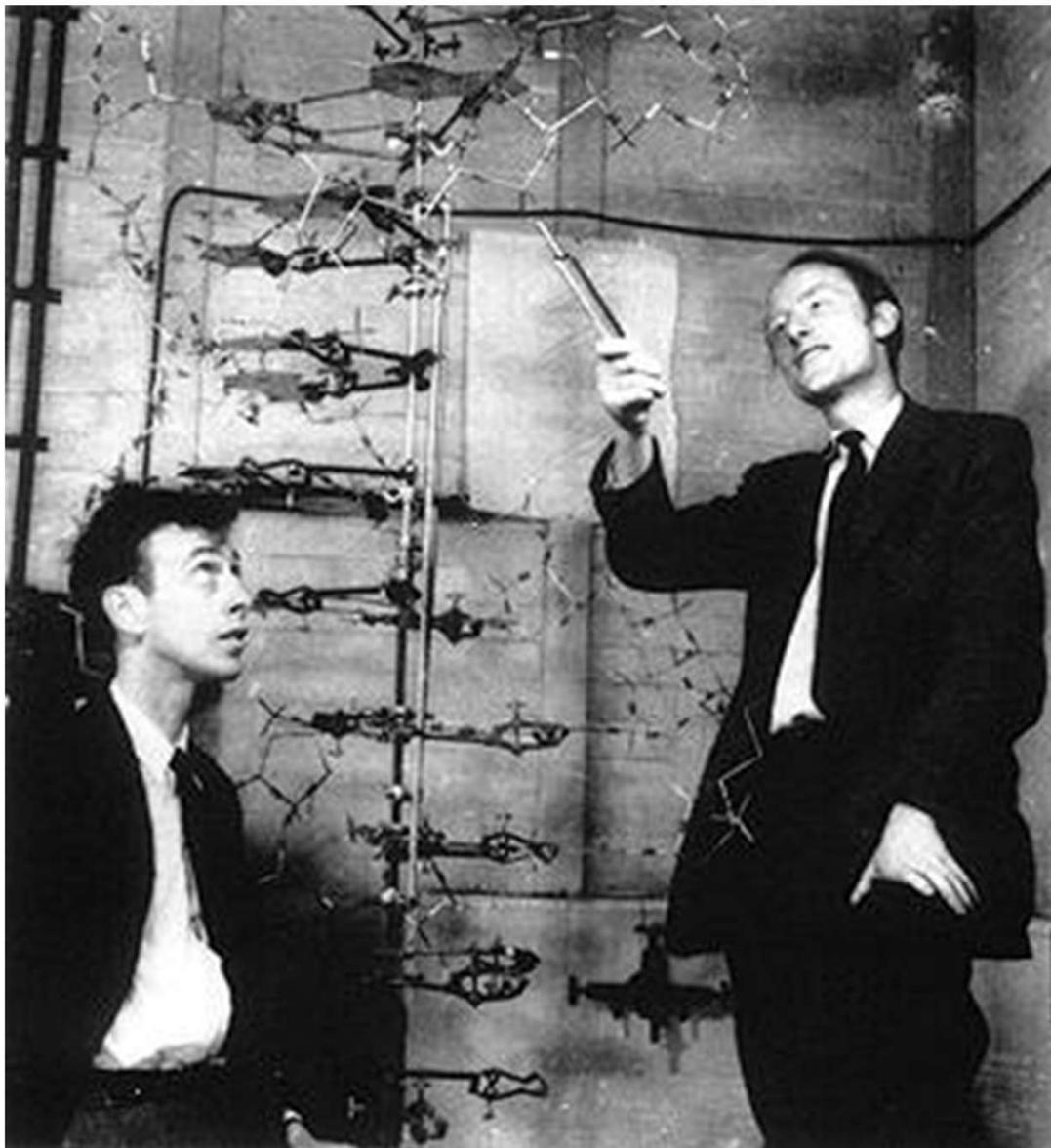


Нуклеотиды ДНК (А,Г,Т,Ц)

Генетическая информация

Информация о строении функциональных молекул (белков и РНК), закодированная с помощью последовательности нуклеотидов ДНК и РНК. Вся информация о структуре и деятельности клеток и организма в целом.





Открытие двойной
спирали ДНК

1953 г -
Джеймс Уотсон и
Фрэнсис Крик

Архей. Гипотеза мира РНК

функцию хранения генетической информации и катализ химических реакций могли выполнять ансамбли молекул РНК.

Даже сегодня РНК выполняют важнейшие функции в клетке:

- весь биосинтез белка (рРНК, тРНК, мРНК);
- в репликации РНК-праймер и РНК-содержащая теломераза эукариот;
- в процессе обратной транскрипции можно синтезировать ДНК из РНК;
- в ходе созревания РНК используются малые ядерные РНК;
- носитель энергии в клетках АТФ – рибонуклеотид!

Происхождение генетического кода



- А, G, T, C → глицин, аланин, валин, пролин, серин

- Как нуклеиновые кислоты (полимеры из четырех мономеров) стали кодировать белки (полимеры из 20 мономеров) является огромной проблемой, которая пока не разрешена
- **Стереохимическая теория:** некое сродство между последовательностью нуклеиновой кислоты и аминокислотой. Однако нет особых взаимодействий кодонов и кодируемых аминокислот! Протяженные участки РНК могут избирательно связывать определенные аминокислоты и последующая эволюция таких участков могла оставить только совсем минимальные их части (триплеты).
- **Теория коэволюции:** и кодонов и аминокислот изначально было гораздо меньше и шло вовлечение в код новых аминокислот из метаболических цепей, причем аминокислоты из одного ряда кодировались похожими кодонами.
- **Адаптивная теория:** эволюция генетического кода шла с отбором вариантов, более устойчивых к ошибкам.
- **Гипотеза коллективной эволюции:** ранние репликаторы активно обменивались инновациями возникающие ранние способы кодирования активно перемешивались, пока один из наиболее стабильных не вытеснил впоследствии все остальные варианты.

ДНК как хранилище генетической информации

- Свойства ДНК (двойная спираль, химическая устойчивость)
- Отсутствие у ДНК каталитических свойств (специализация по функции хранения и воспроизведения генетической информации)

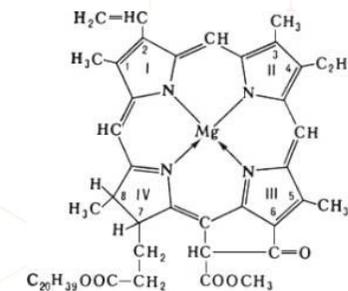


Клетка

- Изоляция репликаторов и стабилизация обмена информации
- **Теория Майкла Рассела**: неорганические сети микроячеек в гидротермальных источниках могли быть первыми компартментами до происхождения мембран.

Ранний докембрий. Архей

- Авто- и гетеротрофы возникают почти одновременно!
- Хлорофилл – сначала использовался как пигмент, защищающий от ультрафиолета, позже фотоавтотрофы используют его для зарядки АТФ
- Остатки хлорофилла – в породах 3.1 млрд лет
- Первый кислород не шел в атмосферу – тратился на окисление железа в первичном океане (образование залежей железных руд!)
- 2 млрд лет назад – большая часть железа уходит в ядро
- Кислород начинает насыщать атмосферу!



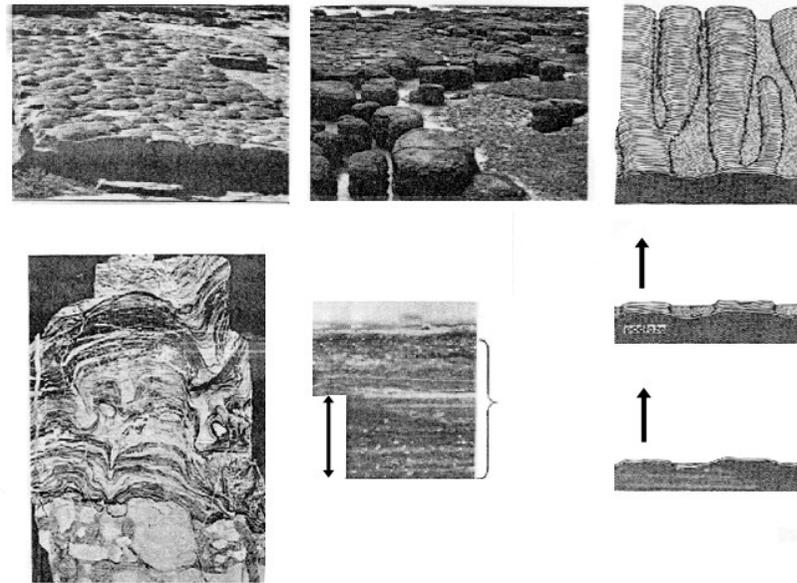
Ранний докембрий. Архей

Эволюция прокариот – развитие многообразия ферментов, катализирующих биохимических реакции. Регуляторные системы достигают некоторого разнообразия и больше не усложняются.

Ранний докембрий. Протерозой

содержание кислорода выше 1%

- 1.9 млрд лет назад – появление первых аэробов!
- Развивается озоновый слой – поглощение ультрафиолета – больше зон становится пригодными для жизни
- Основные организмы – прокариоты – бактерии и цианобактерии
- Древние симбиозы прокариот:
- маты и строматолиты



Палеопротерозойское обледенение



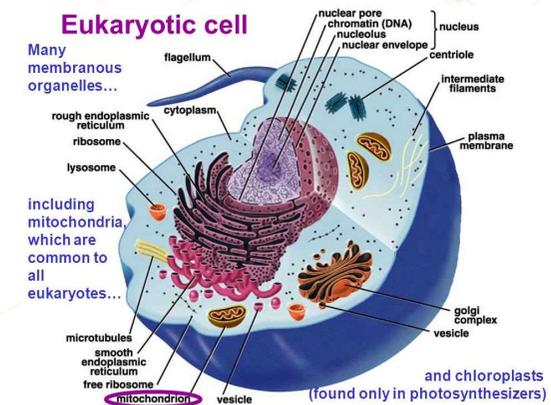
(850-635 млн лет назад)

- Сильнейшее похолодание, ледники дошли до экватора
- Возможная причина:
 - выветривание горных пород, удаление CO₂ из атмосферы
- $\text{CaSiO}_3 + \text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{Ca}^{2+} + \text{SiO}_2 + \text{HCO}_3^-$



Эукариотическая клетка

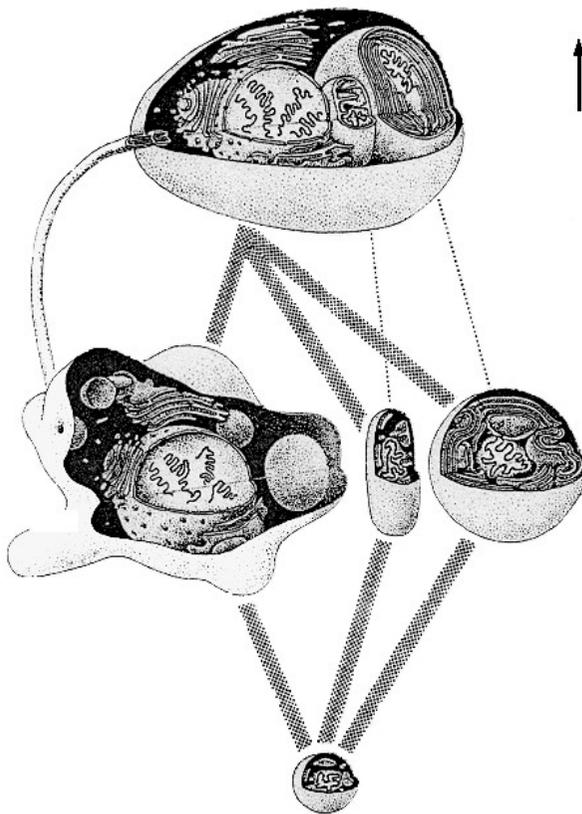
- Крупнее геном
- Больше объем клетки
- Эндомембранные структуры (компарментализация)
- Палеоэндосимбионты (митохондрии и хлоропласты)
- Актин-тубулиновый цитоскелет
- Клеточный транспорт
- Упорядоченное деление
- Пloidность
- Половое размножение
- Линейные хромосомы
- Интроны
- Разделение транскрипции и трансляции



Больше сложность! Сложная иерархия структур

Протерозой. Возникновение эукариот

PosgresPro

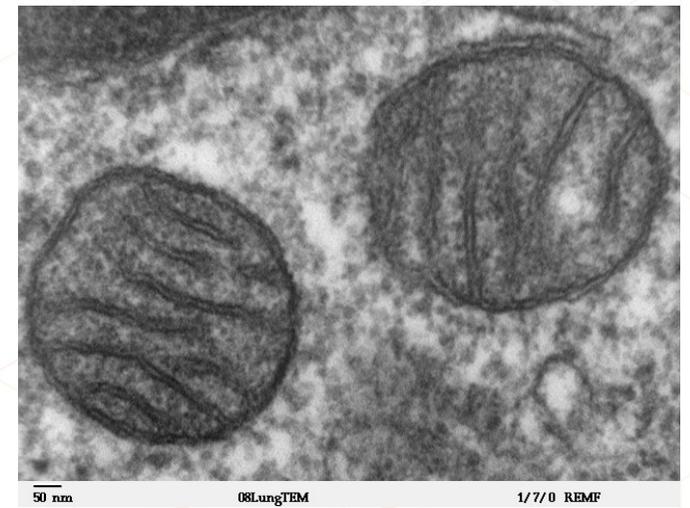


1,5-2 млрд. лет назад.
Симбиогенез. Упорядочение
деления. Увеличение
размера генома и клетки.



СИМБИОГЕНЕЗ

- Возникновение сложной эукариотической клетки было бы невозможно без существования энергетически выгодного метаболизма с участием митохондрий!
- Окислительное фосфорилирование – высокая эффективность, требует кислород!
- Геном митохондрий – от 10 тпн у животных до 1 млн пн у других организмов



Эукариоты

- Рост геномной сложности и скачек в регуляции экспрессии генов
- Отмена горизонтального переноса
- Появление полового размножения как способа сохранения генома в неблагоприятных условиях

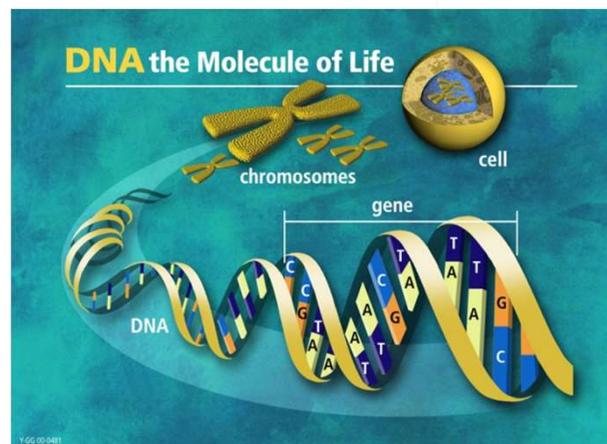
Поздний протерозой. Возникновение многоклеточности.



- Зачем многоклеточность?
- Одноклеточные – нужно перемещаться, т.к. быстро исчерпывается среда. Диффузия – потери синтезированной органики.
- Многоклеточные- специализация тканей, возможность запаса веществ, способность к активному передвижению.
- 600 млн лет назад – Эдиакарская фауна
- Особенности :
 - состояла из крупных или очень крупных беспозвоночных, размером до полутора метров
 - нарушенная билатеральная симметрия!

Сколько ДНК в одном геноме?

- Человек – в одной клетке (10 мкм) около 2 м ДНК, это 3 млрд. нуклеотидных остатков



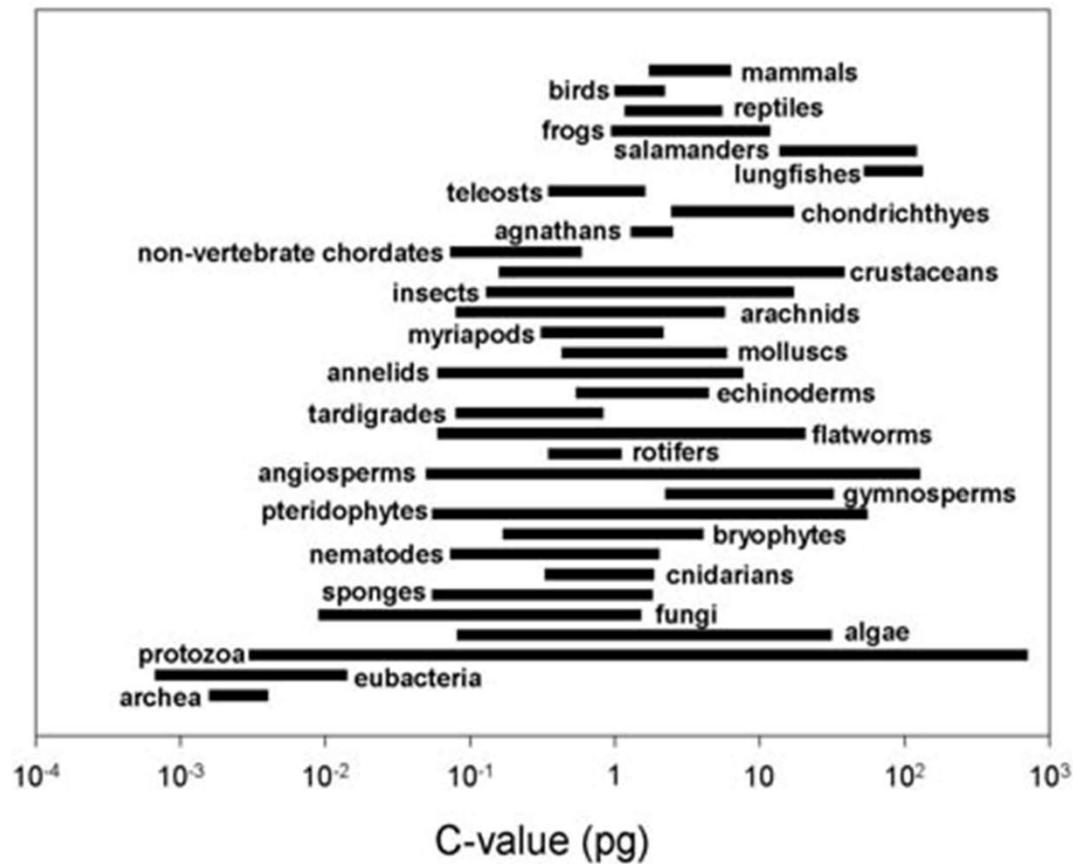
В разных геномах содержание ДНК
варьирует

У двоякодышащих рыб размер генома
43 млрд п.н.



Рогозуб (*Neoceratodus forsteri*)

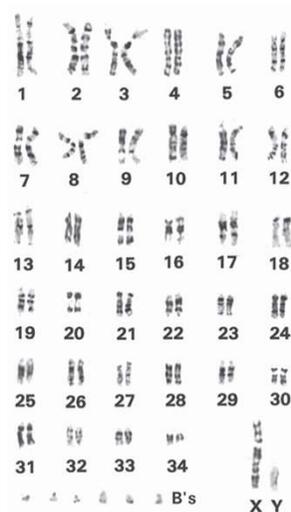
Размер генома и сложность организмов



Геном эукариот

Ядерный геном: линейные молекулы ДНК, каждая образует хромосому в комплексе с белками

Изменчивость: геном может быть фрагментирован по-разному у разных организмов и эта фрагментация обычно не сильно влияет на свойства генов



Сибирская косуля



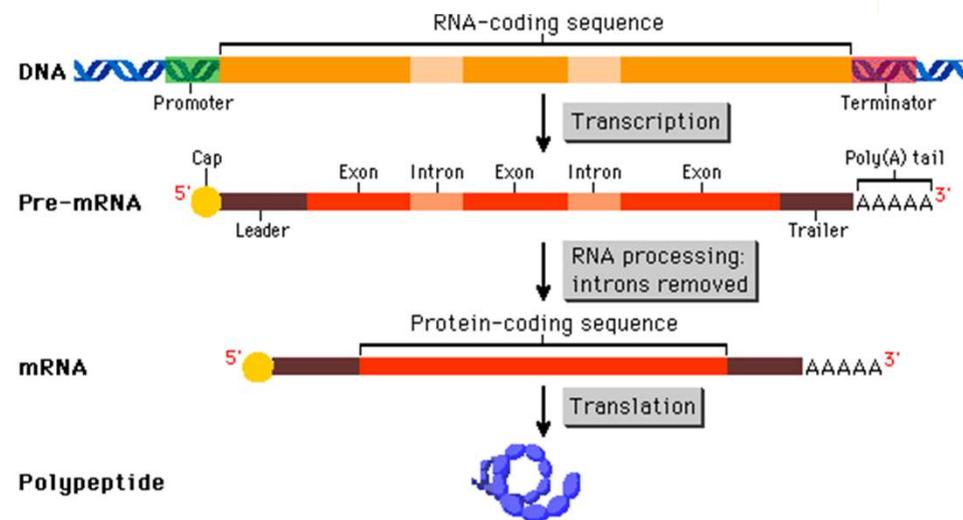
Индийский мунтжак

Термины геномики

- Последовательность по-английски
SEQUENCE
- Отсюда происходят новые слова и выражения в геномике:
- **Секвенировать** (определять последовательность нуклеотидов)
- **Секвенатор** (машина для автоматического секвенирования)

Кодирующая последовательность

Тот участок ДНК, который кодирует последовательность аминокислот в белках



Скорость накопления геномных данных

PosgresPro

- 1981 г – митДНК человека (16.6 тпн)
- 1986 г – геном хлоропластов табака и маршанции
- 1992 г – секвенирование хромосомы дрожжей (315 тпн)
- 1995 г – секвенирование генома гемофильной палочки *Haemophilus influenzae* (1.8 млн пн)
- 1996 г. – секвенирование первого эукариотического генома – пекарских дрожжей (*Saccharomyces cerevisiae*)
- 2001 г. – секвенирование и частичная сборка генома человека (полная сборка закончена только в 2007 г)
- 2012 г. – закончен проект The 1000 Genomes Project, включавший секвенирование геномов 1000 людей
- На сегодняшний день – отсеквенированы геномы нескольких тысяч вирусов, нескольких тысяч прокариот, около 200 эукариотических организмов



Развитие технологий высокопроизводительного секвенирования

Секвенаторы нового поколения – прочтения тысяч и миллионов последовательностей одновременно





Как разобраться с неупорядоченными
миллиардами последовательностей?

PosgresPro



Избыточность генома

- У эукариот (организмов с клеточным ядром) очень много повторенных последовательностей ДНК!
- Эти участки ничего не кодируют!

```
CATTACCAATTCTCCTGGGTTATACCATGTTTAAATCCACATTAACCTCTAGTCACCATATCATAATGCTTGCGTACATTAATTATTTAAGTACATA  
AGGCATGCTATGTTTAAATCCACATTAACCTCTAGTCACCATATCATAATGCTTGCGTACATTAATTATTTAAGTACATAAGGCATGCTATGTTTAA  
TCCACATTAACCTCTAGTCACCATATCATAATGCTTGCGTACATTAATTATTTAAGTACATAACGCATGCTATGTTTAAATCCACATTAACCTCTAG  
TCACCATATCATAATGCTTGCTCACSTTAAATTATATAAGCACATAAGGCATGCTATGTTTAAATCCACATTAACCTCTAGTCACCATATCATAATG  
TTTCATCTACCATCAAATGATATACACCATTTTCTCTATGTGCTCTAACATATCACTTCTGCAAATCATAATATGTAGTAAGAACCGAGCATTTTAA  
TCTGTCTAGAACATAAAGTTAATGAGATGAAAGACAATACCTGTAGAGTTTCATAACTGAATTATTACTGGCATCTGGTTCCTATTTTCAGGTCCAAA  
TTAACGATTATTTCCCATAACTAGGTCGCTACTGGCATCTGATTAATGTTAGAGGTA|
```

Причины избыточности эукариотического генома

1. Снятие ограничения на размер генома
2. Меньше размер популяции (хуже работает очищающий отбор)



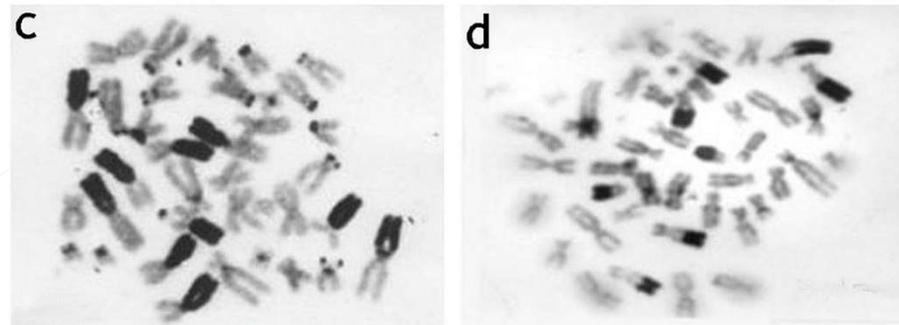
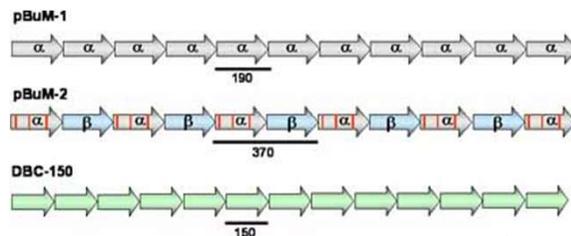
Выход: эпигенетический сайленсинг избыточной ДНК,
диминуция хроматина

Уровни избыточности генома

- Повторенные последовательности
- Амплификация генов, сегментные дупликации
- Возникновение добавочных хромосом
- Полногеномные дупликации

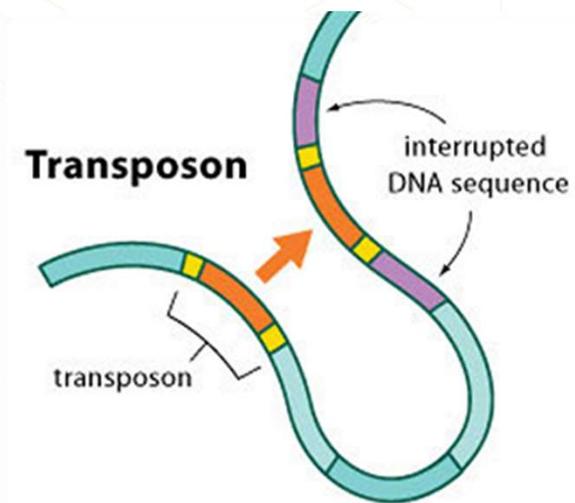
Сателлитная ДНК

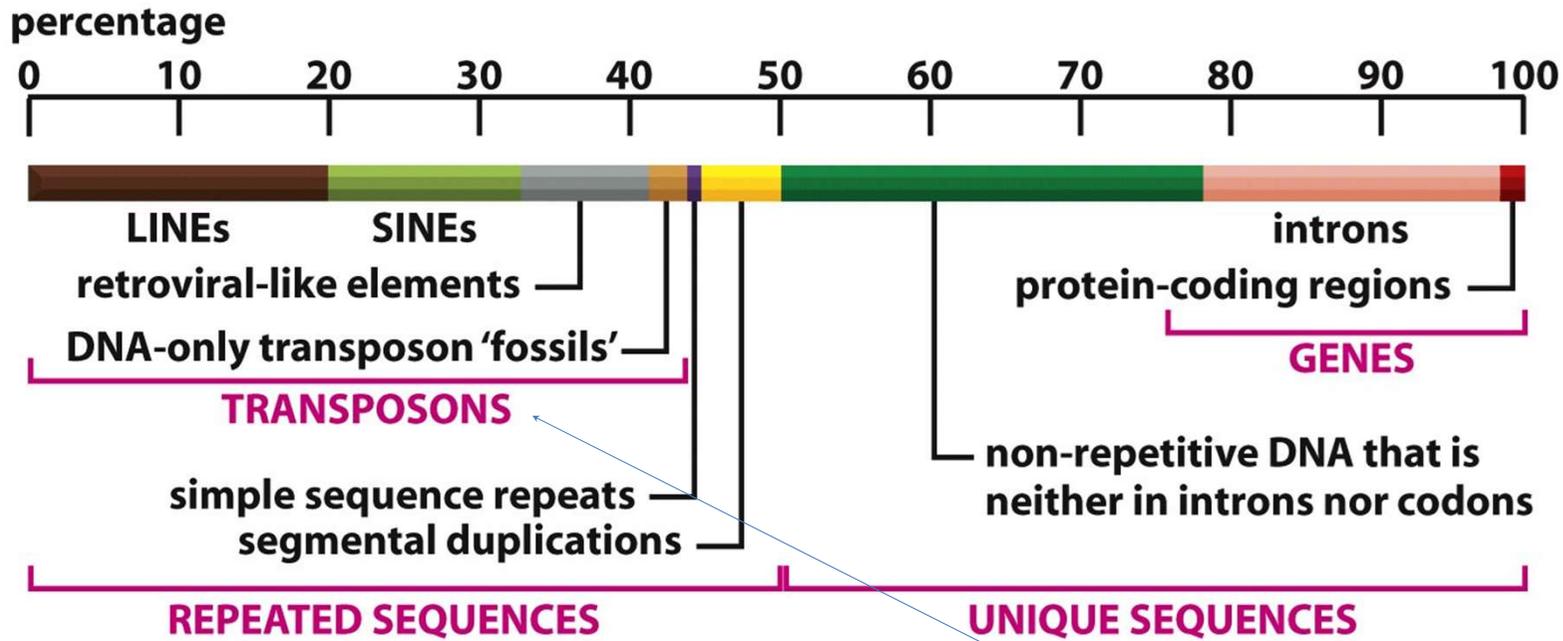
- Участки гетерохроматина (более плотный) и эухроматина (не плотные участки). В основном гены располагаются в эухроматине. Но встречаются и в гетерохроматиновых районах.
- Факультативный гетерохроматин - в зависимости от стадий клеточного цикла один и тот же участок хромосомы может быть в состоянии как гетеро-, так и эухроматина.
- Конститутивный гетерохроматин – всегда плотно упакован. В нем, как правило, генов нет.
- *Сателлитная ДНК часто располагается в центромерном районе.*
- В местах расположения сателлитной ДНК возможна максимальная компактизация. В конститутивном гетерохроматине все четыре уровня упаковки ДНК представлены даже в интерфазе.
- Сателлитная ДНК очень быстро эволюционирует!



Мобильные генетические элементы

Основная часть рассеянных повторов.
Способны менять свое местоположение в геноме
увеличивать количество копий





Геном человека - 45% составляют **мобильные элементы**

Важный фактор эволюции геномов

PosgresPro

80% мутаций у высших эукариот!

На уровне

Структуры генов (изменяются экзоны, интроны, промоторы)

Регуляторных элементов (новые энхансеры, инсуляторы)

Хромосомных перестроек (вызывают делеции, дупликации, инверсии)

Обратная транскрипция позволяет осуществлять горизонтальный перенос генов

Большинство встроек не имеют никакого эффекта!

Программируемая элиминация частей генома

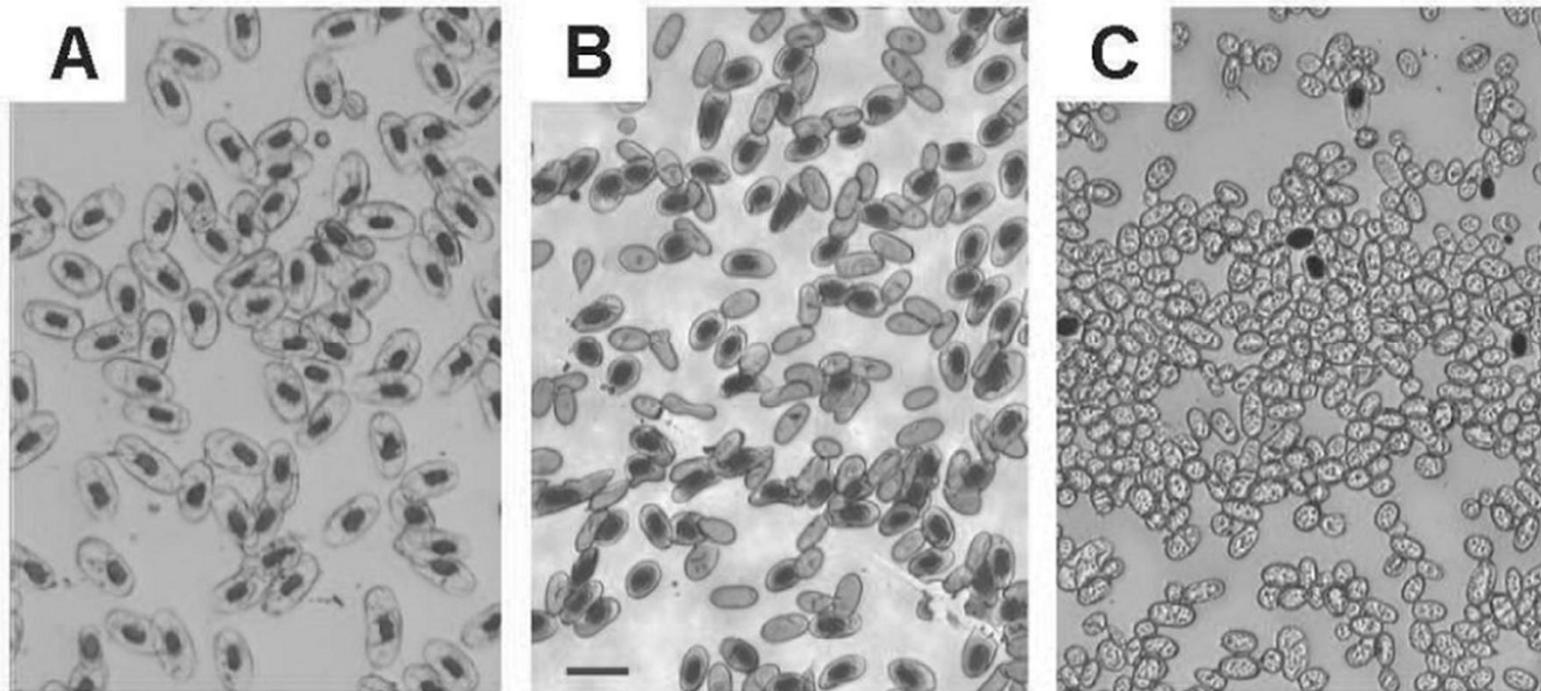
Обычно все клетки организма, произошедшие в результате митотических делений, содержат одинаковые геномы.

Некоторые ткани могут избавляться от генома полностью (эритроциты млекопитающих, клетки хрусталика), в некоторых тканях происходят специфические изменения (рекомбинация в клетках иммунной системы)

В некоторых случаях удаление части генома является обязательным пунктом программы развития организма.

Чаще всего это происходит при переходе от клеток зародышевого пути к соматическим клеткам.

Энуклеация эритроцитов у саламандр



В роде *Batrachoseps* у (A) *B. campi* (< 1 % ядер удаляется). У (B) *B. kawia* (~50 %).
У (C) *B. attenuatus* (~99 %)

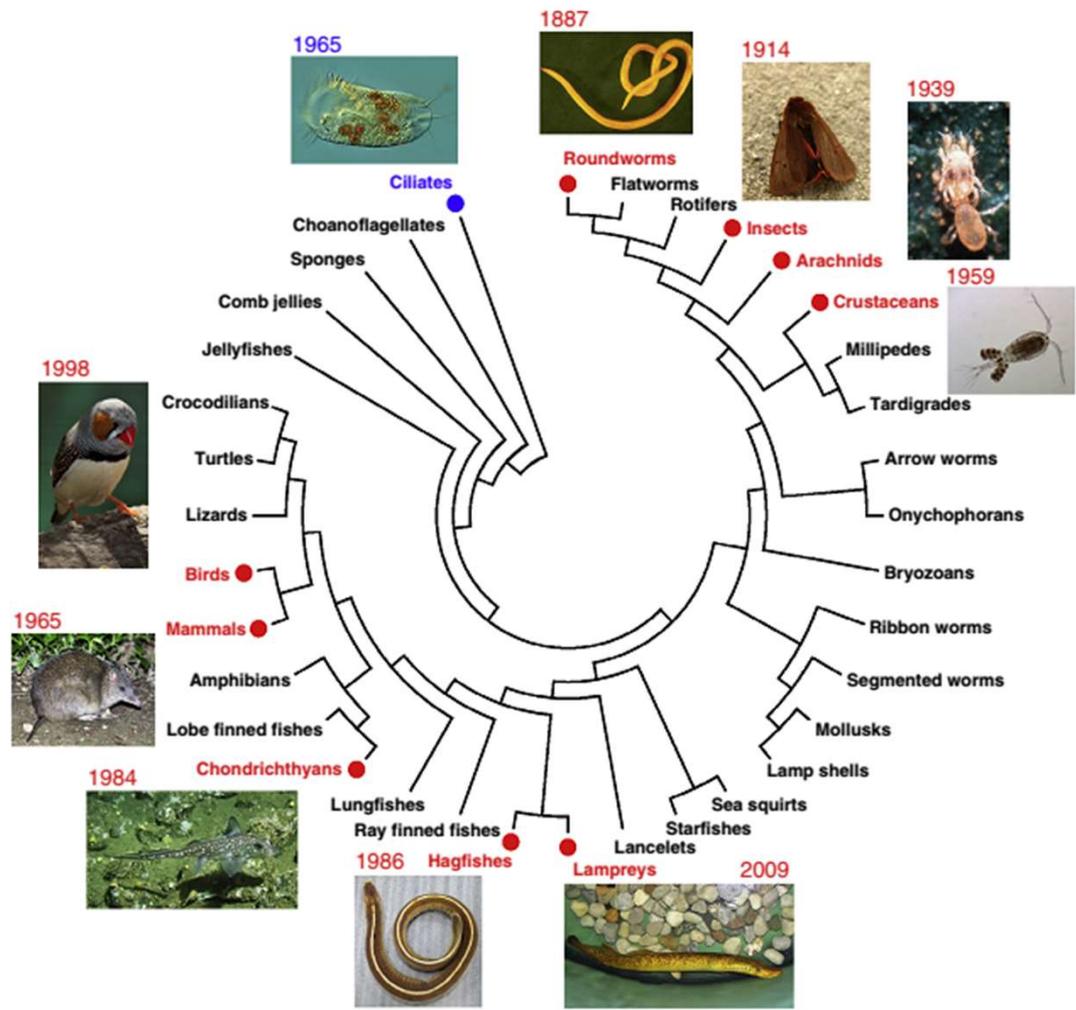
Zoology (Jena). 2008 ; 111(3): 218–230.

Genome size, cell size, and the evolution of enucleated erythrocytes in attenuate salamanders

Rachel L. Mueller^{a,b,*}, T. Ryan Gregory^c, Sean M. Gregory^c, Alice Hsieh^{a,b}, and Jeffrey L. Boore^{b,d,e}



Распространение программируемой элиминации



Статья 10 февраля 2021 г

PostgresPro

nature research

Search Q Login @

nature > milestone

MILESTONE | 10 FEBRUARY 2021

Genomic Sequencing

The past two decades have witnessed extraordinary technological and computational advances in nucleic acid sequencing. This Milestone timeline provides a perspective of major genomic sequencing-related developments in the 21st century — from the first human reference genome, through methodological breakthroughs, to the impact of sequencing on fields as diverse as microbiology, cancer and palaeogenetics. [show less](#)



[Milestone home](#) [Editorial credits](#) [Multimedia](#) [Podcast](#) [Sponsor](#)

illumina®

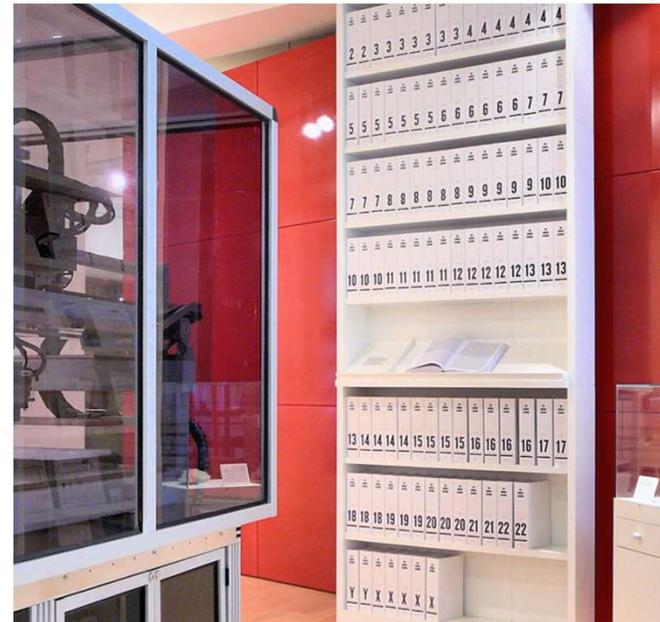
<https://www.nature.com/collections/dbieeeeeed> 42

Референсный (эталонный) геном

PosgresPro

Хорошо собранный геном представителя вида или таксона

Референсный геном человека



Распечатка первого
референсного генома
человека в виде книг в

Метагеномика

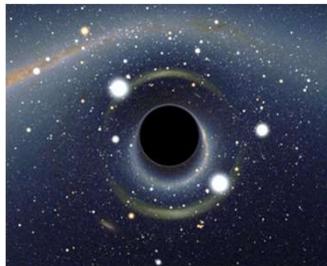


• выделяем ДНК из всего сообщества!

секвенируем

анализируем

• Многообразие вирусов!



Черные дыры!!! – много последовательностей имеет совершенно неясное происхождение



Environmental Genome Shotgun Sequencing of the Sargasso Sea
J. Craig Venter, *et al.*
Science 304, 66 (2004);
DOI: 10.1126/science.1093857

ARTICLE

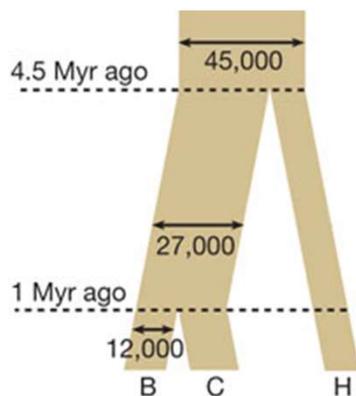
OPEN
doi:10.1038/nature12352

Insights into the phylogeny and coding potential of microbial dark matter

Christian Rinke¹, Patrick Schwientek¹, Alexander Sczyrba^{1,2}, Natalia N. Ivanova¹, Iain J. Anderson^{1,2}, Jan-Fang Cheng¹, Aaron Darling^{1,4}, Stephanie Malfatti¹, Brandon K. Swan², Esther A. Giles², Jeremy A. Dodsworth², Brian P. Hedlund², George Tsiamis², Stefan M. Siefert², Wen-Tso Liu^{1,5}, Jonathan A. Eisen³, Steven J. Hallam², Nikos C. Kyrpides¹, Ramunas Stepanauskas², Edward M. Rubin¹, Philip Hugenholtz¹ & Tanja Woyke¹

Геномы ближайших родственников человека

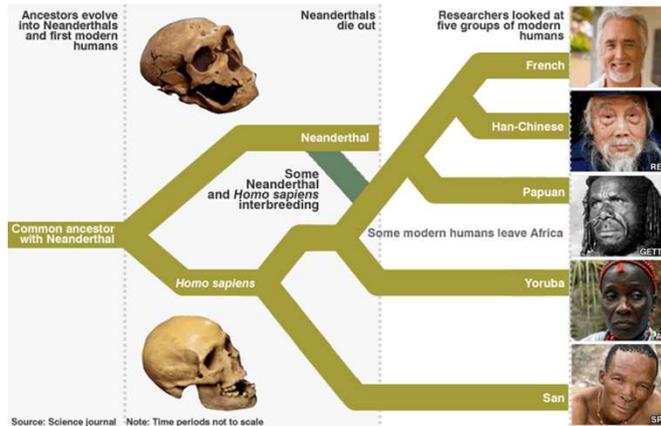
- Шимпанзе Mikkelsen et al 2005 Nature
- Бонобо Pruefer et al 2012 Nature



<http://www.sexandquantumphysics.com/2011/06/our-primate-cousins/>

Ч:Ш - 35 млн однонукл замен, 90 млн вставок и делеций
Большинство из этих отличий нейтрально!
Человек потерял около 80 генов в процессе эволюции – большинство из них – обонятельные рецепторы. Потеря одного из кератиновых генов (KRTNAP1) – изменение волосяного покрова.
Потеря гена MYH16 – уменьшение нижней челюсти, возможность для возрастания объема черепной коробки
HAR1A и HAR2 – связаны с развитием мозга и противопоставлением

Геном неандертальца



nature

Vol 444 | 16 November 2006 | doi:10.1038/nature05336

ARTICLES

Analysis of one million base pairs of Neanderthal DNA

Richard E. Green¹, Johannes Krause¹, Susan E. Ptak¹, Adrian W. Briggs¹, Michael T. Ronan², Jan F. Simons², Lei Du², Michael Egholm², Jonathan M. Rothberg², Maja Paunovic^{3,4} & Svante Pääbo¹



nature

Vol 449 | 18 October 2007 | doi:10.1038/nature06193

LETTERS

Neanderthals in central Asia and Siberia

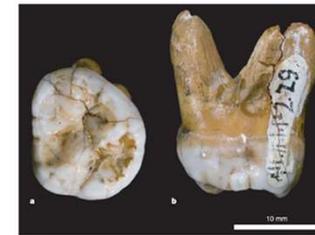
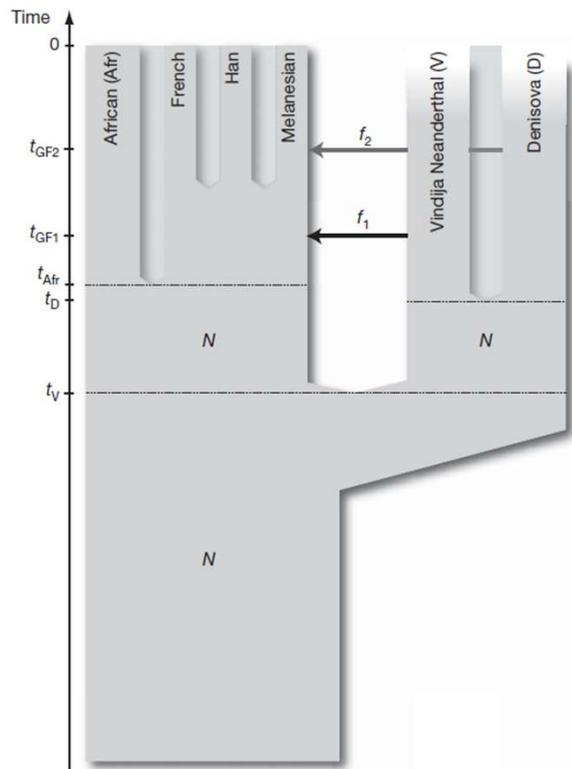
Johannes Krause¹, Ludovic Orlando², David Serre³, Bence Viola⁴, Kay Prüfer¹, Michael P. Richards¹, Jean-Jacques Hublin¹, Catherine Hänni², Anatoly P. Derevianko⁵ & Svante Pääbo¹



1-4 % генома европейцев

Геном денисовского человека

PosgresPro



doi:10.1038/nature08976

nature

LETTERS

The complete mitochondrial DNA genome of an unknown hominin from southern Siberia

Johannes Krause¹, Qiaomei Fu¹, Jeffrey M. Good², Bence Viola^{1,3}, Michael V. Shunkov⁴, Anatoli P. Derevianko⁴ & Svante Pääbo¹

ARTICLE

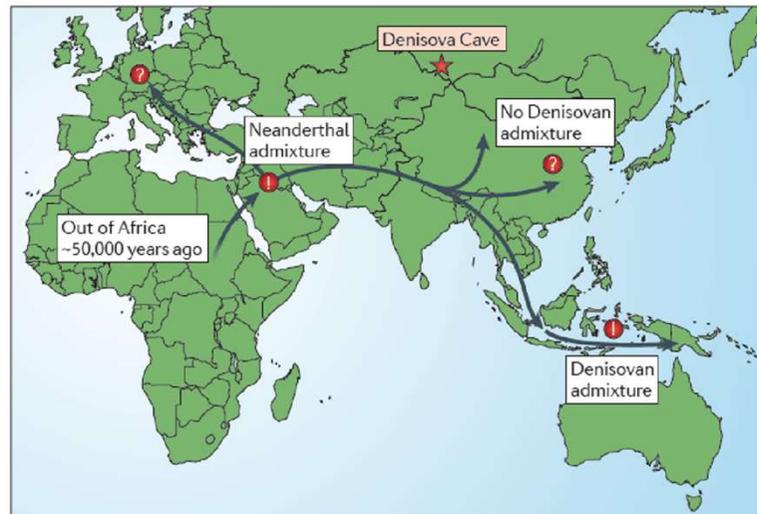
doi:10.1038/nature09710

Genetic history of an archaic hominin group from Denisova Cave in Siberia

David Reich^{1,2*}, Richard E. Green^{3,4*}, Martin Kircher^{3*}, Johannes Krause^{5,6*}, Nick Patterson^{2*}, Eric Y. Durand^{6*}, Bence Viola^{3,7*}, Adrian W. Briggs^{1,3}, Udo Stenzel³, Philip L. F. Johnson⁸, Tomislav Maricic³, Jeffrey M. Good⁹, Tomas Marques-Bonet^{10,11}, Can Alkan¹⁰, Qiaomei Fu^{3,12}, Swapan Mallick^{1,2}, Heng Li², Matthias Meyer³, Evan E. Eichler¹⁰, Mark Stoneking³, Michael Richard^{7,13}, Sahra Talamo⁷, Michael V. Shunkov⁴, Anatoli P. Derevianko⁴, Jean-Jacques Hublin⁷, Janet Kelso³, Montgomery Slatkin⁶ & Svante Pääbo³

Происхождение современного человека

- Ген FOXP2, связанный с речью, присутствует у неандертальца.
400 тыс лет назад – разделение линий неандертальца и кроманьонца.



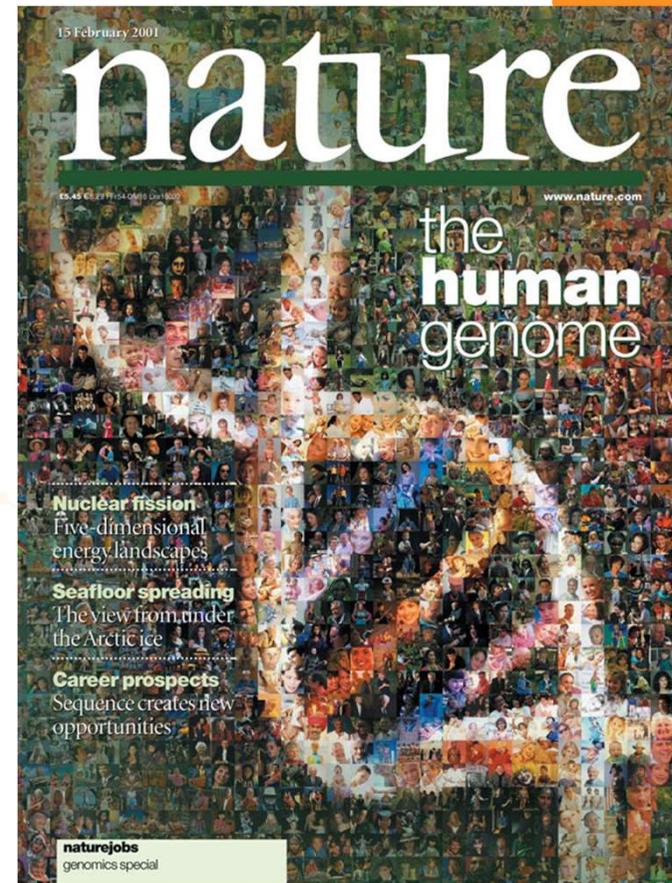
APPLICATIONS OF NEXT-GENERATION SEQUENCING

Learning about human population history from ancient and modern genomes

Mark Stoneking* and Johannes Krause†

Геном человека

- Белок-кодирующие последовательности занимают только 1.5% генома
- Роль остальных компонентов генома сейчас активно изучает программа ENCODE
- 90% варибельности генома сосредоточена в некодирующей области, т.е. проявление большинства фенотипических особенностей, в том числе и болезней, связано с изменениями регуляции работы генов.
- Геном человека является референсным для большого числа исследований как в генетике человека так и в сравнительной геномике

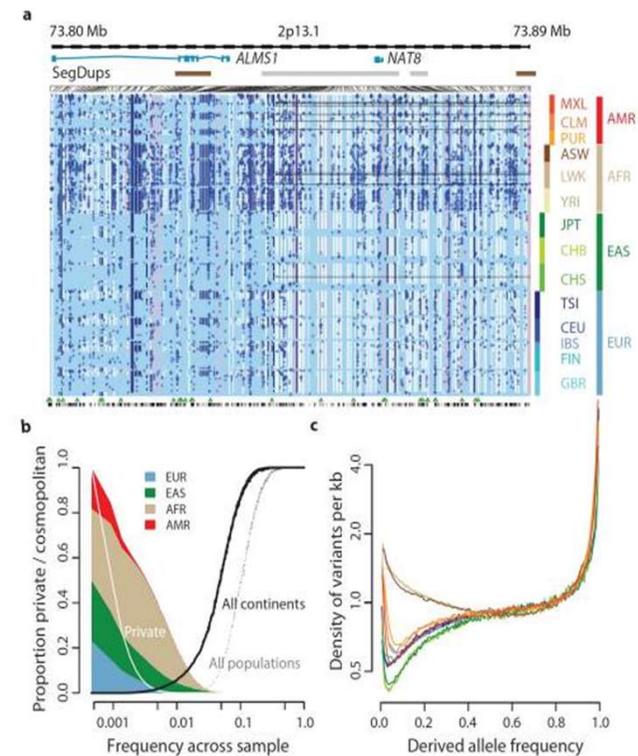


Проект 1000 геномов «The 1000 Genomes Project»

PosgresPro

Одного генома мало, чтобы делать выводы о всей популяции.

- Описано:
- 15 млн однонуклеотидных полиморфизмов.
- 1 млн коротких делеций/дупликаций
- 20 тыс крупных структурных вариаций
- Каждый человек несет 50-100 мутаций связанных с описанными наследственными болезнями.
- Оценена скорость накопления замен в геноме человека - одна замена на 100 млн пн в поколение.

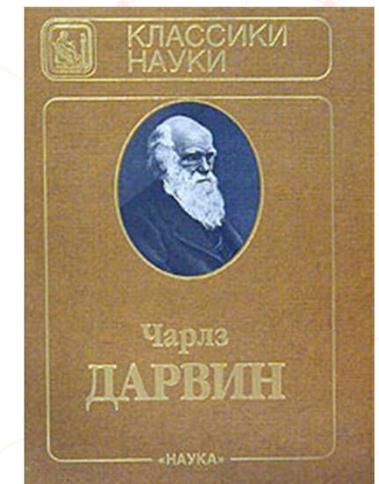
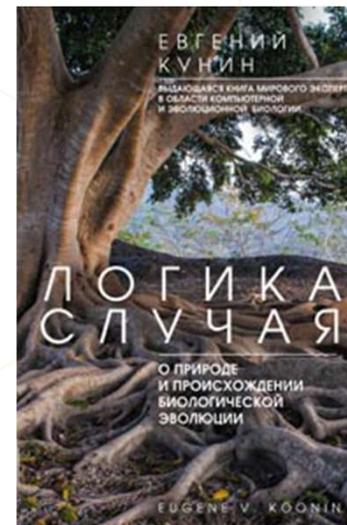
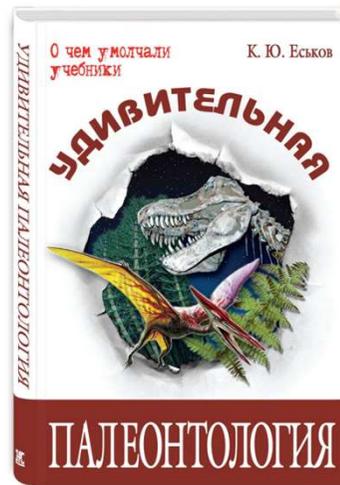


Перспективы

- Каждый новый геном приносит удивительные открытия, проясняя особенности происхождения, развития и функционирования живых организмов, объясняет причины как возрастания сложности, так и эволюционного регресса

Рекомендуемая литература

PosgresPro



PGConf.Сибирь 2022

postgrespro.ru

