



# О геномах

н.с. Лаборатории лесной геномики ИФБиТ  
СФУ

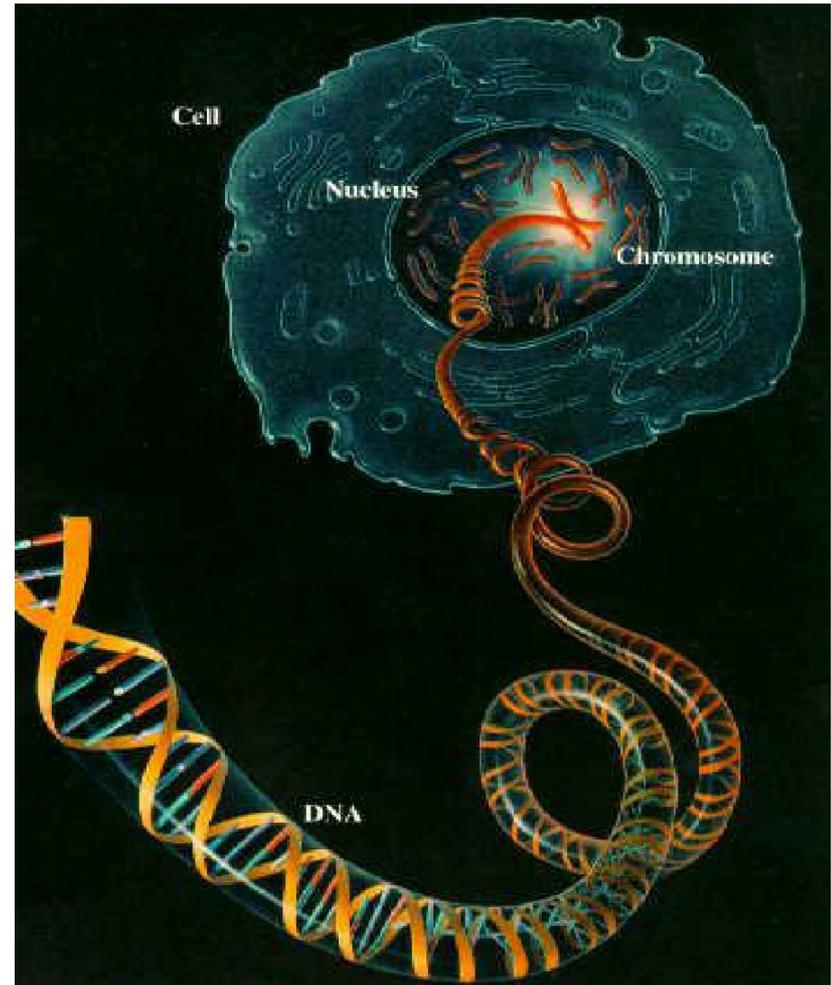
Инженер отдела генетики и селекции Филиала  
ФБУ «Рослесозащита» «Центра защиты леса  
Красноярского края»

**Дейч Ксения Олеговна**



# Геном

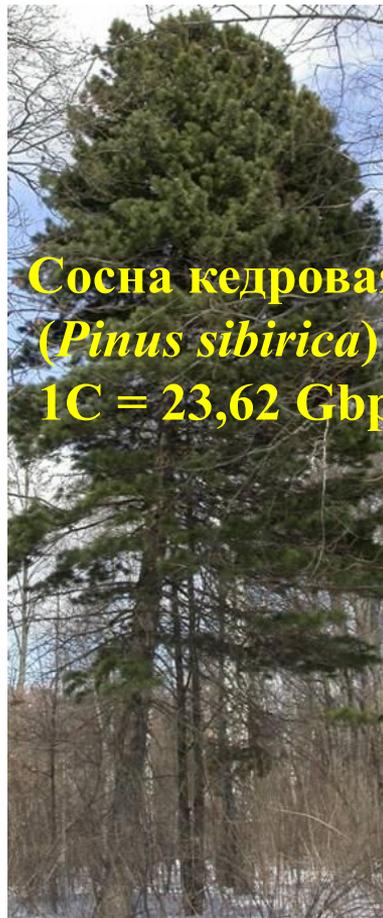
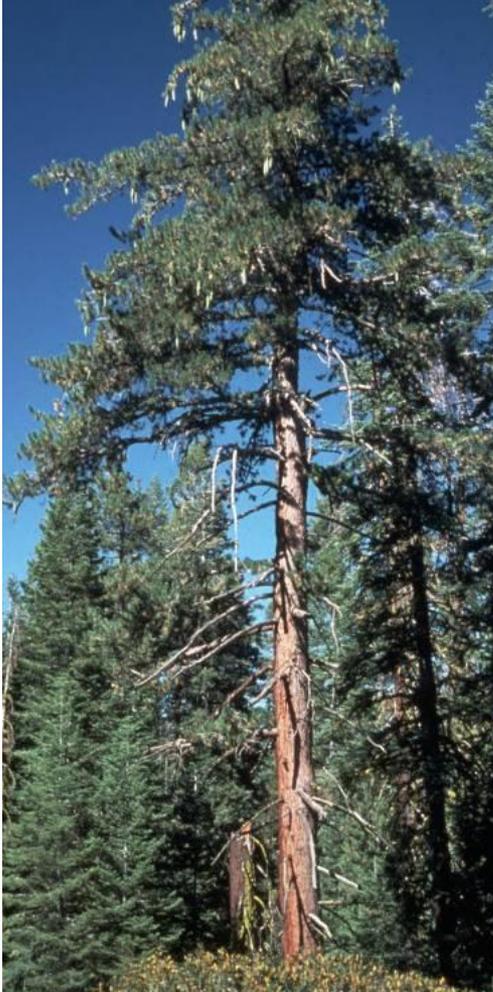
- Совокупность генов, содержащихся в одинарном наборе хромосом ( $1n$ ) данного организма.
- Измеряется в парах нуклеотидных оснований п.н.о. или п.о. (bp).



# Сравнительный размер генома

**Сосна сахарная**  
*(Pinus lambertiana):*

**1C = 28,90 Gbp**



**Сосна кедровая**  
*(Pinus sibirica):*  
**1C = 23,62 Gbp**



**Ель**  
**обыкновенная**  
*(Picea abies):*  
**1C = 19,57 Gbp**



**Лиственница**  
**сибирская**  
*(Larix sibirica):* **1C = 12,03 Gbp**

1 Gbp =  $10^9$  bp = млрд. (пар оснований)

•Геном человека  
*Homo sapiens:*  
**1C = 3,20 Gbp**



**Ядерные геномы цветковых растений в сравнении с другими организмами**

Представитель	Размер генома (п. н.)	Количество хромосом	Представитель	Размер генома (п. н.)	Количество хромосом
<b>Бактерии</b>					
<i>Escherichia coli</i>	$4.64 \times 10^6$	—	<i>Pyrus communis</i> (груша)	$4.96 \times 10^8$	2x = 34
<b>Грибы</b>			<i>Lycopersicon esculentum</i> (томат)	$4.99 \times 10^8$	2x = 24
<i>Saccharomyces cerevisiae</i>	$1.21 \times 10^7$	x = 17	<i>Cucurbita pepo</i> (тыква)	$5.02 \times 10^8$	2x = 40
<b>Животные</b>			<i>Brassica oleracea</i> (капуста)	$5.99 - 6.62 \times 10^8$	2x = 18
<i>Drosophila melanogaster</i>	$1.80 \times 10^8$	2x = 8	<i>Phaseolus vulgaris</i> (фасоль)	$6.37 \times 10^8$	2x = 22
<i>Homo sapiens</i>	$3.40 \times 10^9$	2x = 46	<i>Sorghum bicolor</i> (сорго)	$7.48 \times 10^8$	2x = 20
<b>Водоросли</b>			<i>Beta vulgaris</i> (свекла)	$7.58 \times 10^8$	2x = 18
<i>Cyanidioschyzon merolae</i>	$1.20 \times 10^7$	x = 15	<i>Spinacea oleracea</i> (шпинат)	$9.89 \times 10^8$	2x = 12
<i>Chlorella vulgaris</i>	$3.90 \times 10^7$	x = 16	<i>Trifolium repens</i> (клевер белый)	$9.99 \times 10^8$	4x = 32
<i>Chlamydomonas reinhardtii</i>	$1.00 \times 10^9$	x = 18	<i>Glycine max</i> (соя)	$1.12 \times 10^9$	4x = 40
<i>Coccinodiscus asteromphales</i>	$2.50 \times 10^{10}$	x = ?	<i>Brassica napus</i> (брюква)	$1.13 \times 10^9$	2x = 38
<b>Мхи</b>			<i>Petunia hybrida</i> (петуния)	$1.27 \times 10^9$	2x = 14
<i>Physcomitrella patens</i>	$6.00 \times 10^8$	x = ?	<i>Asparagus officinalis</i> (спаржа)	$1.31 \times 10^9$	2x = 20
<b>Голосеменные</b>			<i>Solanum tuberosum</i> (картофель)	$1.60 - 1.86 \times 10^9$	4x = 48
<i>Pinus resinosa</i>	$6.80 \times 10^{10}$	2x = 24	<i>Crepis capillaris</i> (скерда)	$1.87 \times 10^9$	2x = 6
<b>Цветковые растения</b>			<i>Nicotiana plumbaginifolia</i>	$2.29 \times 10^9$	2x = 24
<i>Arabidopsis thaliana</i> (арабидопсис)	$1.45 \times 10^8$	2x = 10	<i>Zea mays</i> (кукуруза)	$2.30 - 2.72 \times 10^9$	2x = 20
<i>Prunus persica</i> (персик)	$2.62 \times 10^8$	2x = 16	<i>Lactuca sativa</i> (латук)	$2.64 \times 10^9$	2x = 18
<i>Raphanus sativus</i> (редис)	$2.90 \times 10^8$	2x = 18	<i>Arachis hypogaea</i> (арахис)	$2.81 \times 10^9$	4x = 40
<i>Cucumis sativus</i> (огурец)	$3.67 \times 10^8$	2x = 14	<i>Helianthus annuus</i> (подсолнечник)	$2.87 - 3.19 \times 10^9$	2x = 34
<i>Citrus sinensis</i> (апельсин)	$3.67 \times 10^8$	2x = 18	<i>Pisum sativum</i> (горох посевной)	$3.95 \times 10^9$	2x = 14
<i>Carica papaya</i> (папайя)	$3.76 \times 10^8$	2x = 18	<i>Nicotiana tabacum</i> (табак)	$4.22 - 4.65 \times 10^9$	4x = 48
<i>Oryza sativa</i> (рис)	$4.15 \times 10^8$	2x = 24	<i>Hogdeum vulgare</i> (ячмень)	$4.87 \times 10^9$	2x = 14
<i>Ananas bracteatus</i> (ананас)	$4.44 \times 10^8$	2x = 50	<i>Triticum monococcum</i> (пшеница)	$5.75 \times 10^9$	2x = 14
<i>Cucumis melo</i> (дыня)	$4.54 \times 10^8$	2x = 24	<i>Vanilla planifolia</i> (ваниль)	$7.67 \times 10^9$	2x = 32
<i>Trifolium pratense</i> (клевер красный)	$4.68 \times 10^8$	2x = 14	<i>Secale cereale</i> (рожь)	$9.50 \times 10^9$	2x = 14
<i>Daucus carota</i> (морковь)	$4.73 \times 10^8$	2x = 18	<i>Avena sativa</i> (овес)	$1.13 \times 10^{10}$	6x = 42
<i>Vitis vinifera</i> (виноград)	$4.83 \times 10^8$	2x = 38	<i>Allium cepa</i> (лук репчатый)	$1.53 \times 10^{10}$	2x = 16
			<i>Triticum aestivum</i> (пшеница мягкая)	$1.60 \times 10^{10}$	6x = 42
			<i>Tulipa sp.</i> (тюльпан)	$2.47 - 3.07 \times 10^{10}$	2x = 24
			<i>Lilium longiflorum</i> (лялия)	$9.00 \times 10^{10}$	2x = 24
			<i>Fritillaria assyriaca</i> (рябчик)	$1.24 \times 10^{11}$	2x = 24

- Среди позвоночных одними из самых больших геномов обладают хвостатые амфибии и **двоякодышащие рыбы** – по **~120 Gb** ДНК.
- У наземных растений гигантские геномы обнаруживают представители **лилейных** (*Fritillaria assyriaca* – **~127 Gb**).
- Несоответствие между размером генома и фенотипической сложностью организма, который им обладает, получило название **«парадокса C»** (C-value paradox).

*Arabidopsis* = 125 Mbp

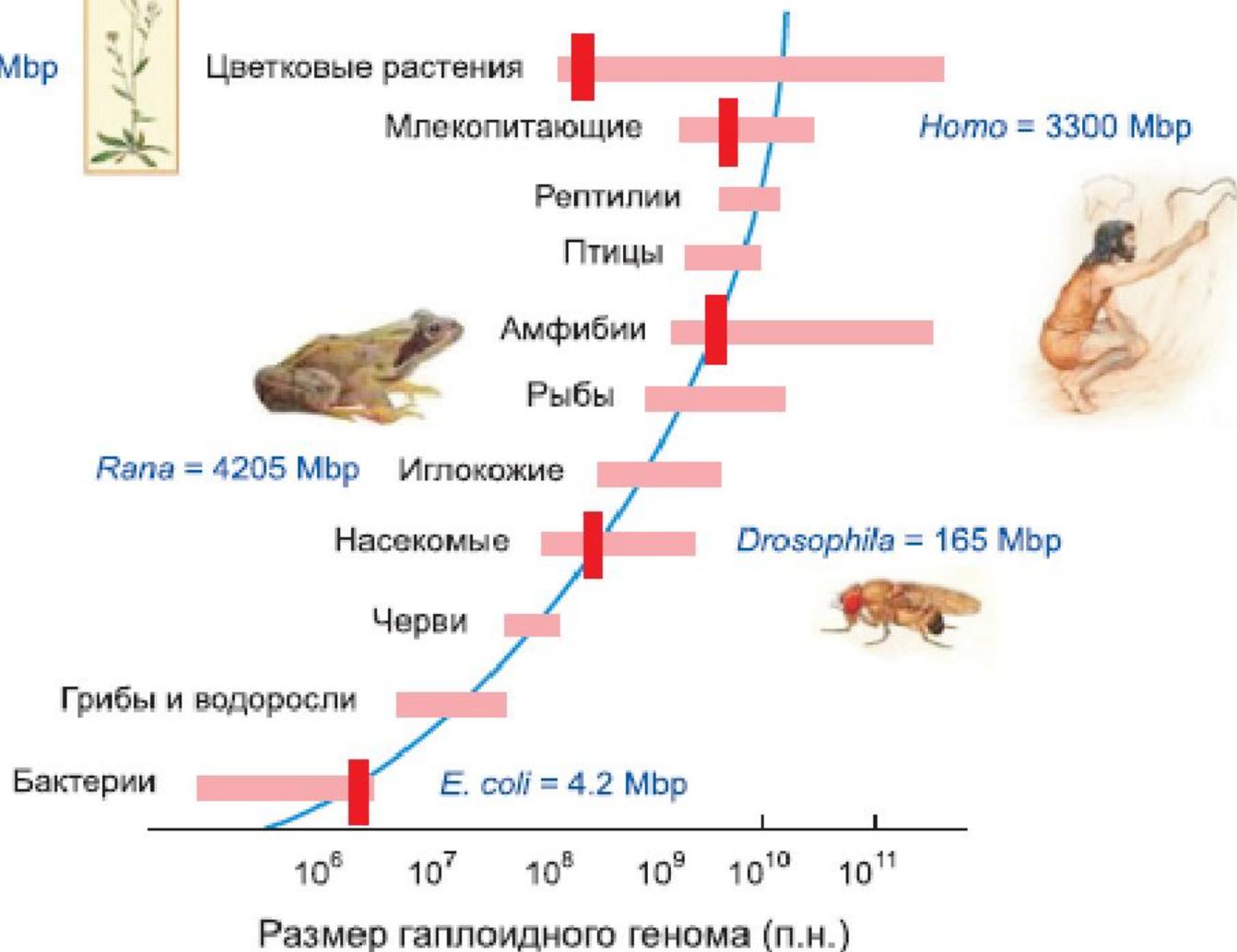
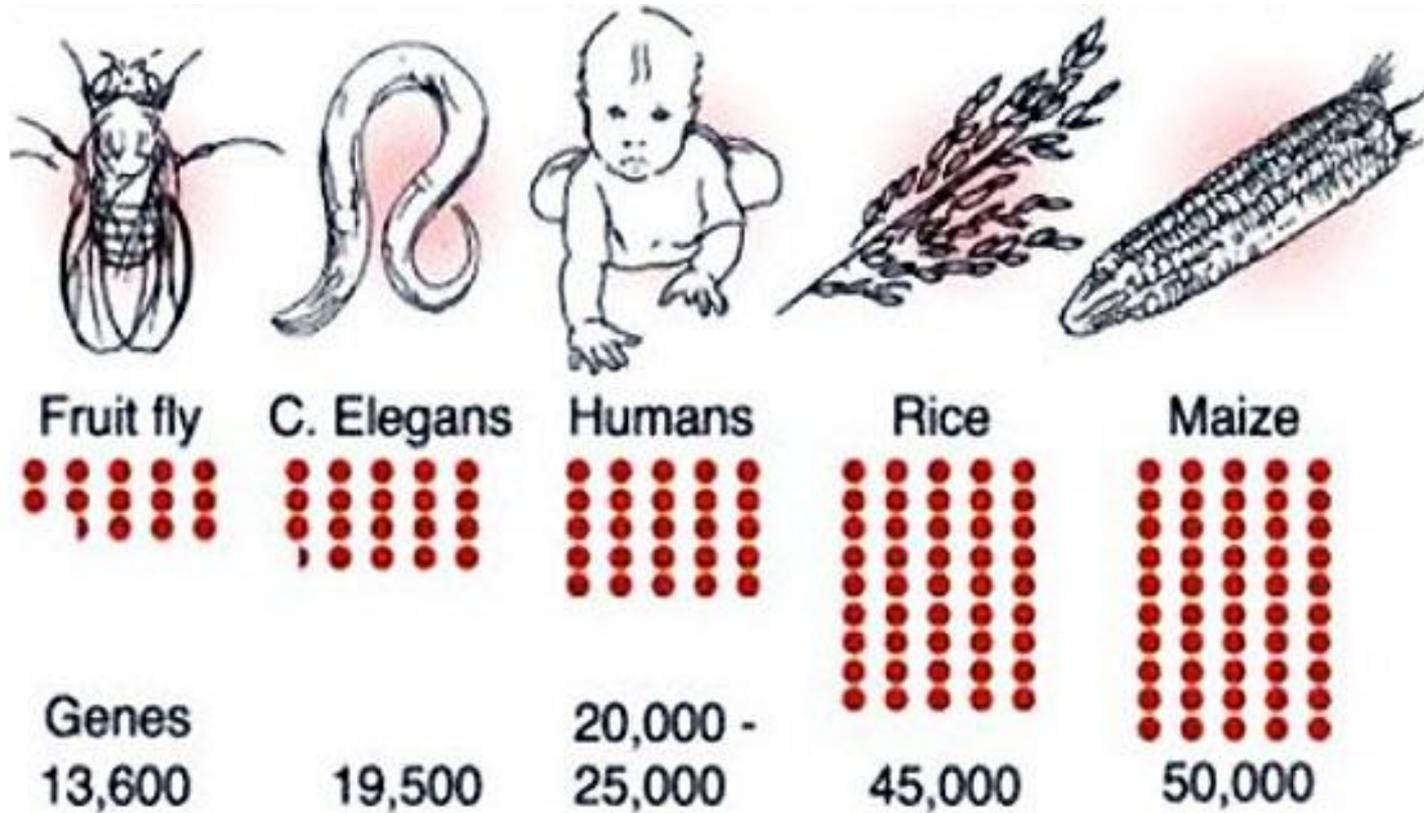


Рис. 1. Размер генома у разных групп организмов (Mbp = млн п.н.).

# По количеству генов



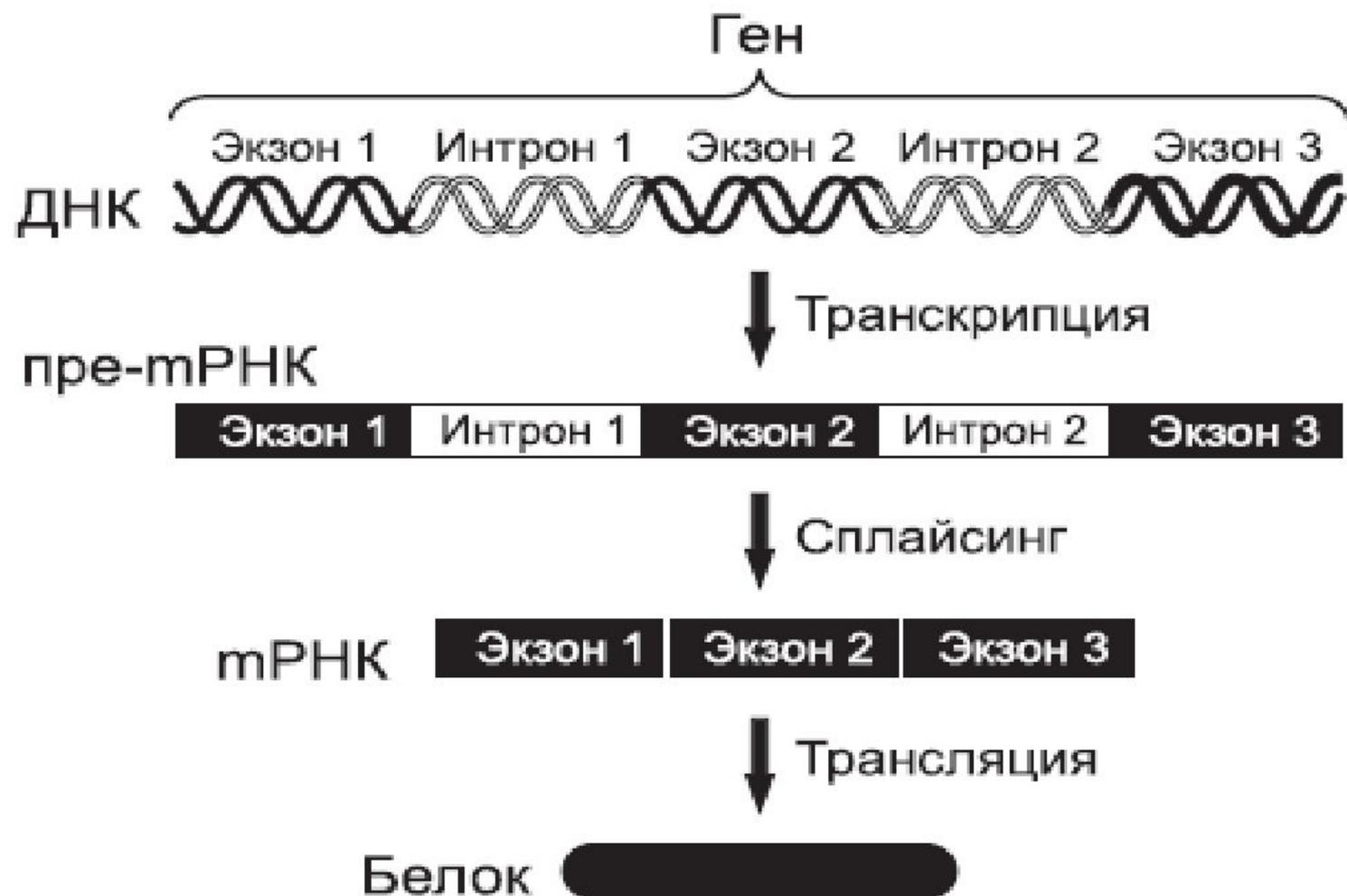
# Сложность строения генома

- Даже если геном человека по числу генов **практически равен геному одуванчика**, то по сложности и многообразию генных продуктов мы существенно отличаемся от одуванчика, что приводит к огромной разнице между этими формами жизни, видимой и невооружённым глазом».



# Альтернативный сплайсинг

- У человека средний ген, имеющий размер  $27 \times 10^3$  п.н., содержит 8 интронов со средней длиной 1500 п.н. Однако некоторые гены, например гены гистонов, гены отдельных трансмембранных рецепторов, интронов не содержат.
- У млекопитающих таких безынtronных генов 6 %, у плодовой мушки – 17 %, а у низших эукариот, таких как дрожжи, – 95 %.
- Именно альтернативным сплайсингом объясняется тот факт, что у человека синтезируется не менее 105 разных типов белков, кодируемых всего  $\sim 20 \times 10^3$  генами.



**Рис. 1. При созревании mРНК интроны вырезаются, а экзоны сшиваются.**

# Полиплоидия

- **Полиплоидия** – это кратное увеличение числа наборов хромосом в ядре.
- 1. **Автополиплоиды** образуются путем **кратного увеличения одного и того же** генома в результате спонтанного удвоения числа хромосом или формирования нередуцированных гамет.
- 2. **Аллополиплоиды** (амфиплоиды) образуются на **основе объединения двух** или нескольких целых геномов, принадлежащих разным видам и родам (гибридная полиплоидия).
- По разным оценкам от 30 до 70 % современных цветковых растений и большая часть мхов и папоротников являются полиплоидами (Meyers, Levin, 2006). К ним относятся и такие важные сельскохозяйственные культуры, как мягкая и твердая пшеница, овес, сахарный тростник и многие другие. (Источники данных по размеру геномов: М. Сингер, П. Берг. Гены и геномы, М.: Мир, 1998; KewDNAC-values database (<http://data.kew.org/cvalues>).

# > ПЛОИДНОСТЬ = > ПЛОД

- *Fragaria vesca* (земляника лесная)

– Диплоид ( $2n=14$ )



- *Fragaria moupinensis* (земляника луговая)

– Тетраплоид ( $2n=28$ )



- *Fragaria moschata* (земляника мускусная)

Гексаплоид ( $2n=42$ )



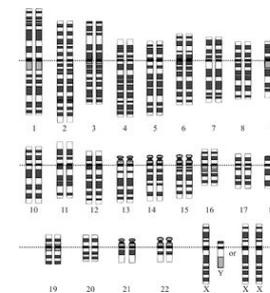
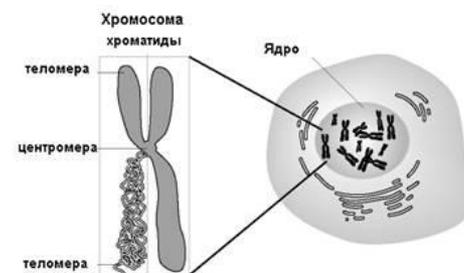
- *Fragaria x ananassa* (земляника ананасная)

- Октаплоид ( $2n=56$ )



# По числу хромосом в ядре

- Высокие числа хромосом ( $> 500$ ) характеризуют мхи и папоротники.
- Максимальное число хромосом ( $2n=1440$ ) зарегистрировано у папоротника *Ophioglossum reticulatum*, у цветковых растений – у пальмы *Voanioala gerardii* ( $2n=596$ ) из субтропических лесов Мадагаскара.
- У большинства высших растений числа хромосом варьируют в пределах от 10 до 50 ( $n=5-25$ ).
- Предполагают, что виды с числом хромосом  $n > 10$  являются полиплоидами.



# Размеры генов

- Средний размер гена в хромосоме приходится около **50 тысяч пар** нуклеотидов.
- Самые короткие гены содержат всего **два десятка** нуклеотидов, например, гены эндорфинов - белков, вызывающих ощущение удовольствия.
- Гены интерферонов - белков, защищающих человека от вирусных инфекций, имеют размер **около 700** нуклеотидов.
- Самый длинный ген, кодирующий один из белков мышц - миодистрофин, содержит **2,5 миллиона** нуклеотидов.

# Аналог языка

- «Если проводить аналогию с литературным текстом, то получается, что есть **четыре нуклеотида**, и это «**буквы**».
- Из них складываются **триплеты** — комбинации из трёх последовательно расположенных нуклеотидов, образующие кодоны, с помощью которых в информационных рибонуклеиновых кислотах кодирует последовательность расположения аминокислот в белках. Это «**слова**».
- Из кодонов складываются **гены** — законченные «**предложения**», а из совокупности генов формируется «**полный**» **текст**, который и называется **геном**.

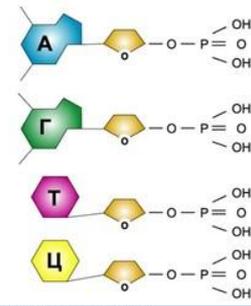
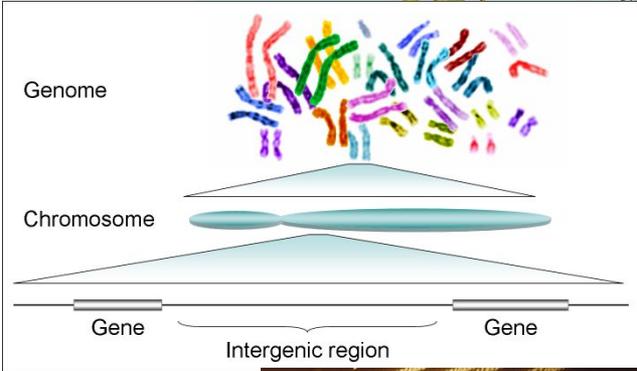
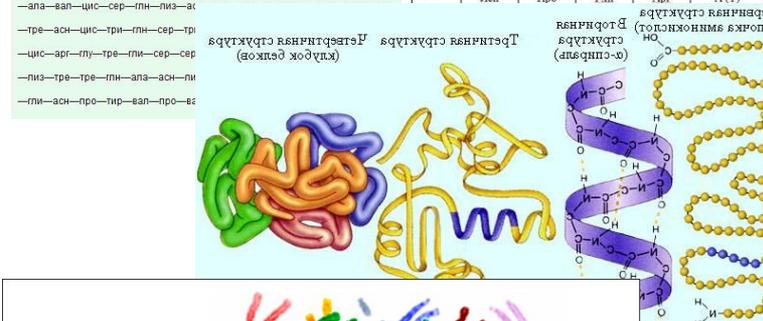


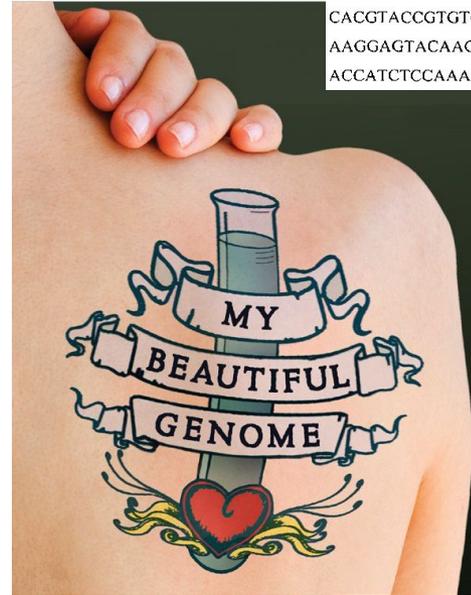
Таблица генетического кода

основание	Второе основание				Третье основание
	У(А)	Ц(Г)	А(Т)	Г(Ц)	
У(А)	Фен	Сер	Тир	Цис	У(А)
	Фен	Сер	Тир	Цис	Ц(Г)
	Лей	Сер	-	-	А(Т)
Ц(Г)	Лей	Про	Гис	Арг	У(А)
	Лей	Про	Гис	Арг	Ц(Г)
	Лей	Про	Гли	Арг	А(Т)

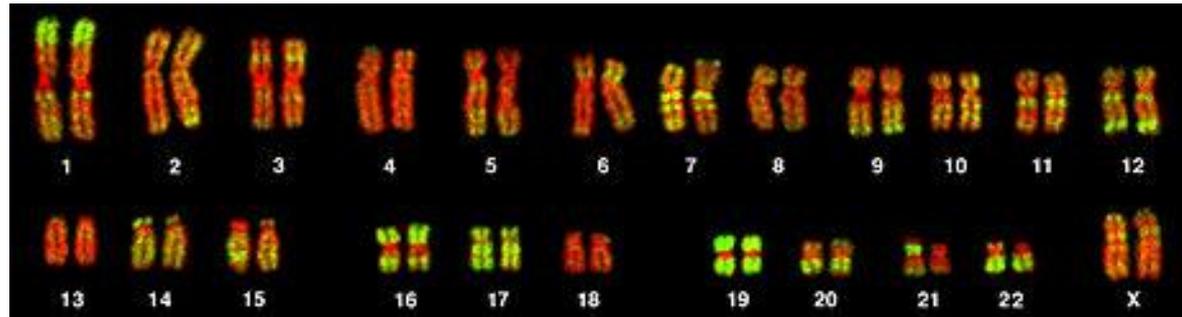
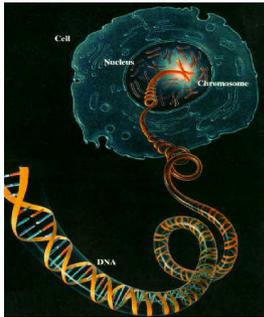


- Роман Льва ТОЛСТОГО «Война и мир» = три миллиона ( $10^6$ ) букв. А у человека более трёх миллиардов ( $10^9$ ) «букв» — нуклеотидов в геноме. То есть условная **тысяча томов** романа.

```
ATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCT
TCGTTTCGCCCGGCCGCTATACTTGGTAGATCAGAACTCAGGAGTGTCTTTT
TTTAAATGCTAATTGGGAAAAAGACAGALCCAATCAAACCTGGTGTGAACCGTGT
ATGGTGACAAAGATAAACGGCGGCATTGTTTTGCTACCTGGAAGAATATTTCTGG
TTCCATTGAATAGTGAAACAAGGTTGTTGGCTGGATGATATCAAACCTGCTATGACA
GGACTGATTGTGTAGAAAAAAGACAGCCCTGAAGTATATTTCTGTTGCTGTGA
GGGCAATATGTGTAATGAAAAGTTTTCTTATTTTCCGGAGATGGAAGTCACACAG
CCCACCTCAAATCCAGTTACACCTAAGCCACCCACCGTGGTGGTGAACCTCACACAT
CCCCACCGTGCCAGCACCTGAACTCCTGGGGGAGCCGTACGTCTTCTCTTCCC
CCCCAAACCCAAGGACACCCTCATGATCTCCGGACCCCTGAGGTACATGCGGTG
GTGGTGAGCGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGAC
GCGGTGGAGGTGCATAATGCCAAGACAAAGCCCGGGAGGAGCAGTACAACAG
CACGTACCGTGTGGTCAAGCTCCTACCGTCTGCACCAGACTGGTGAATGGC
AAGGAGTACAAGTGAAGGTCTCCAACAAAGCCCTCCAGTCCCACGAGAAA
ACCATCTCAAAGCCAAAGGGCAGCCCGAGAACCACAGGTGTACACCTGCCC
```

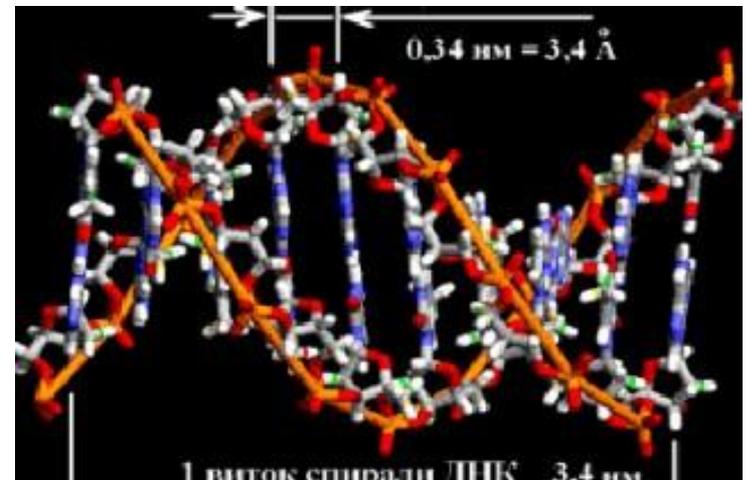


# Длина генома одной клетки человека (1n) около 1 м!



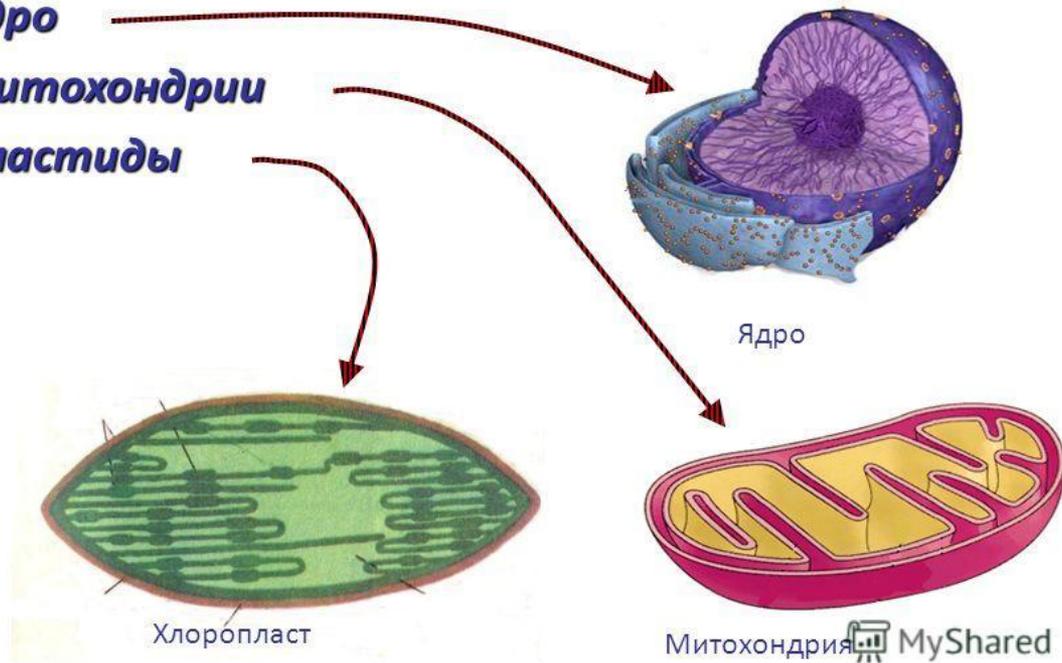
Термин **хромосома** (chromo – окрашенное, soma – тело) первым применил немецкий ученый В. Вальдейер (W. Waldeyer) в 1888 г., поскольку описанные структуры интенсивно окрашивались основными красителями.

Порода	Число хромосом 1n	Размер генома, Gbp	Длина генома, м
Лиственница ( <i>L. sibirica</i> )	12	12,3	4,2
Ель ( <i>P. abies</i> )	12	20	6,8
Сосна ( <i>P. sylvestris</i> )	12	22,9	7,8
Кедр ( <i>P. sibirica</i> )	12	24,1	8,2

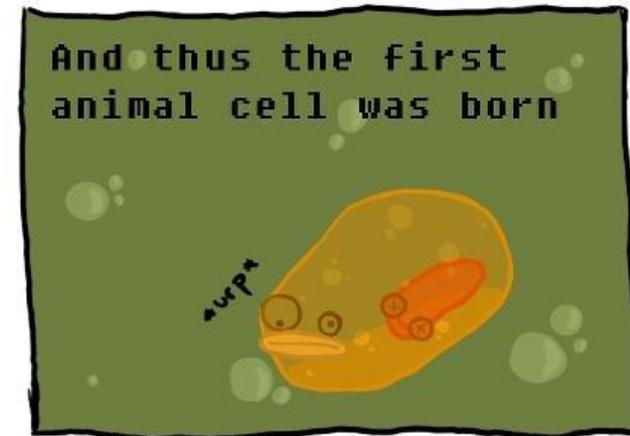
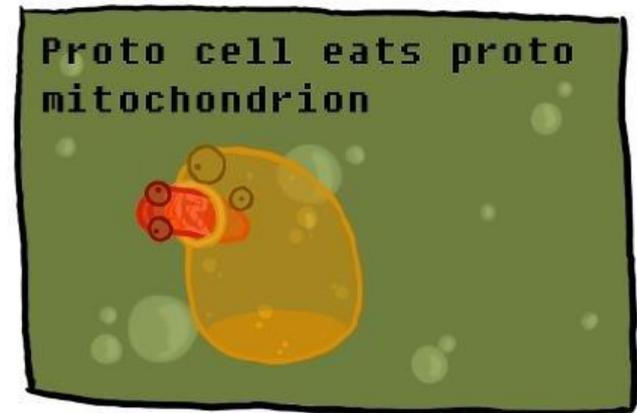
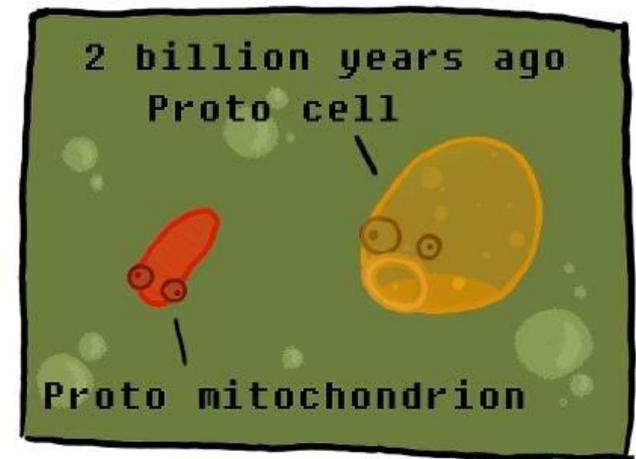


# Местонахождение ДНК в клетке

- Ядро
- Митохондрии
- Пластиды



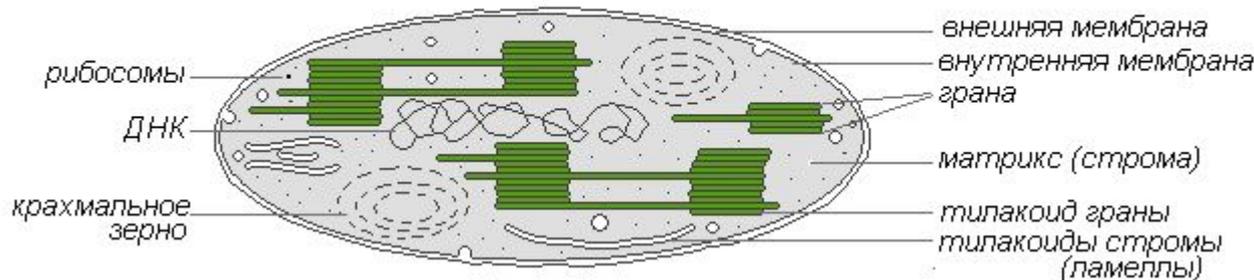
- Симбиогенное происхождение органелл сейчас общепризнанно. Предками хлоропластов, как и считал К.С. Мережковский, признаются цианобактерии, предками митохондрий, вероятнее всего, были бактерии, близкие к риккетсиям.



# Хлоропластная ДНК (хлДНК)

- представлена двуцепочечными кольцевыми молекулами (у высших растений варьирует от **120 до 200 т.п.н.** В часто в этих молекулах обнаруживаются повторы противоположной ориентации длиной 20–30 т.п.н., разделенные уникальными последовательностями.
- В молекулах хлДНК насчитывается **около 140 генов**, в число которых входят:
  1. гены, обеспечивающие синтез белка в органеллах (аппарат транскрипции и трансляции),
  2. гены белков. участвующих в процессе фотосинтеза.

*Строение хлоропласта*

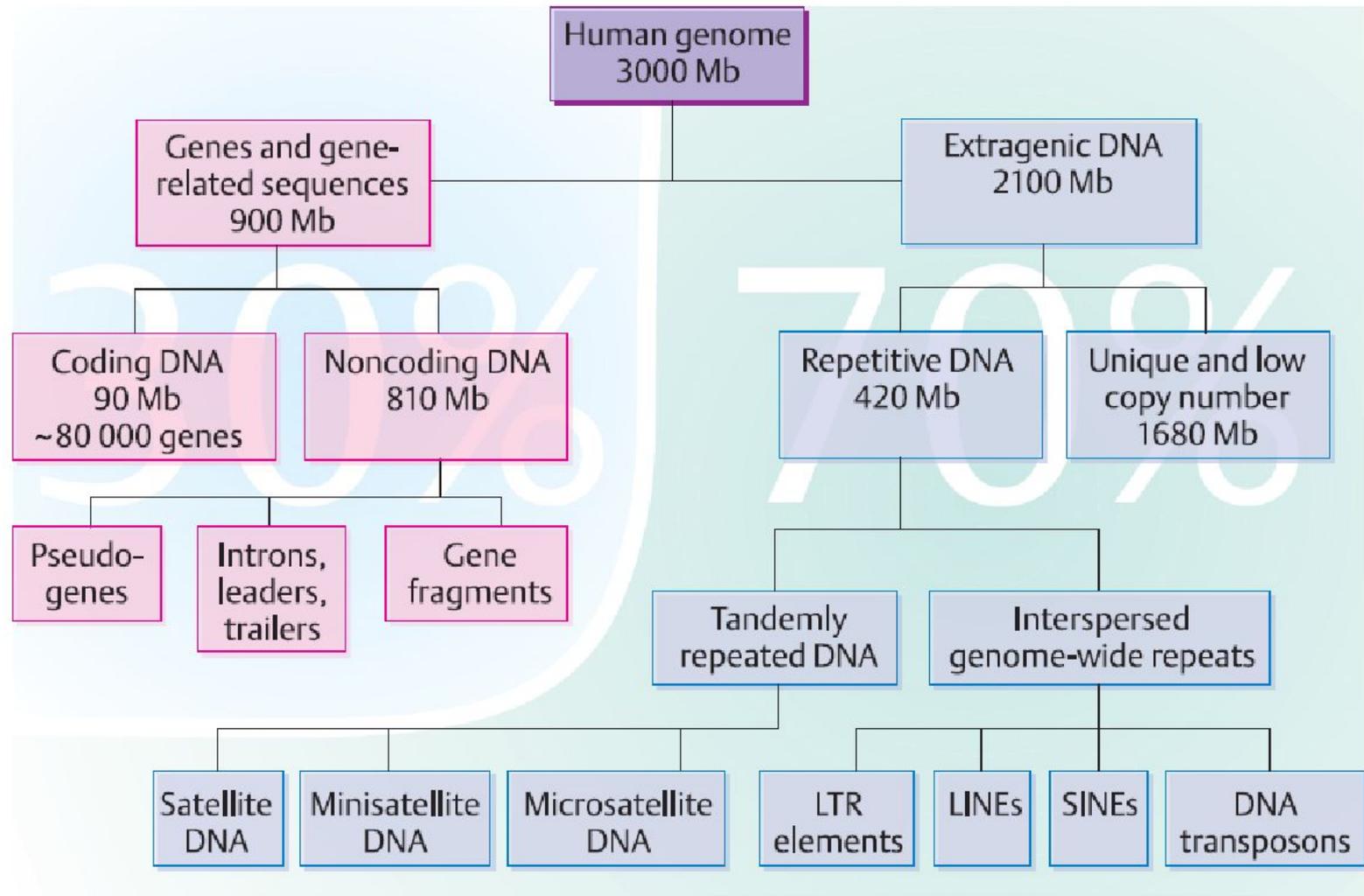


# Митохондриальная ДНК (мтДНК)

- **кольцевые**, лишь у немногих видов, в частности некоторых кишечнополостных животных, эти молекулы линейные.
- Для **животных**, обычно, размеры мтДНК около **16 т.п.н.**
- В мтДНК млекопитающих и других животных **37 генов**.
- Такой же набор генов присутствует в мтДНК
- **высших растений**, к нему добавляется **еще ген 5SРНК**.
- По размеру молекул **мтДНК растений** от **200 т.п.н.** у видов рода *Brassica* до **2500 т.п.н.** у арбуза. Увеличение размера молекул мтДНК происходит за счет некодирующих последовательностей, кроме них в мтДНК растений включены фрагменты хлоропластной ДНК.

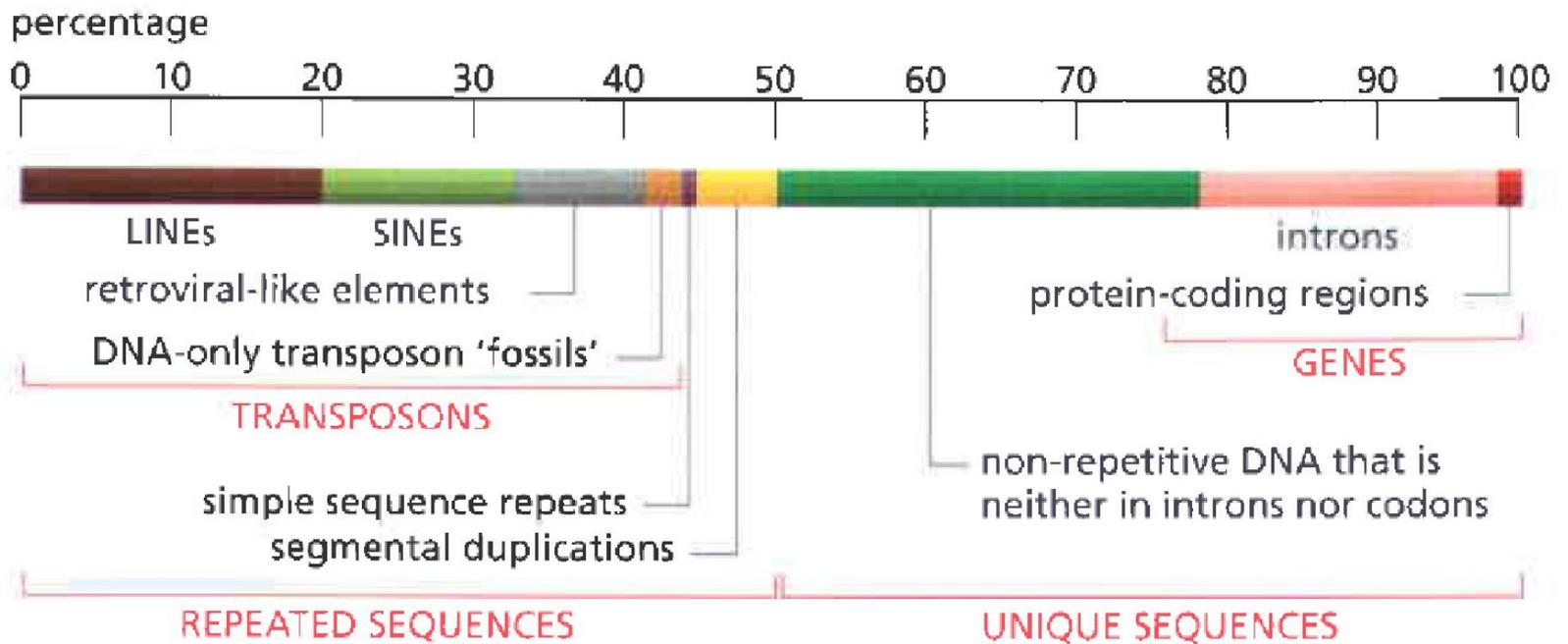


# Компоненты генома человека

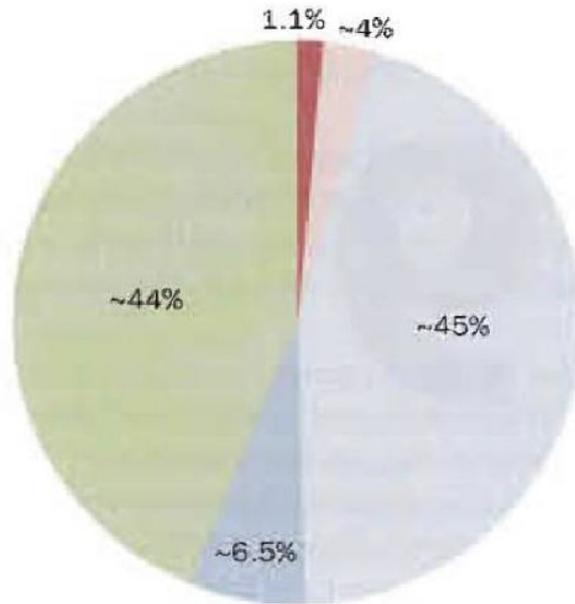


## A. The components of the nuclear human genome

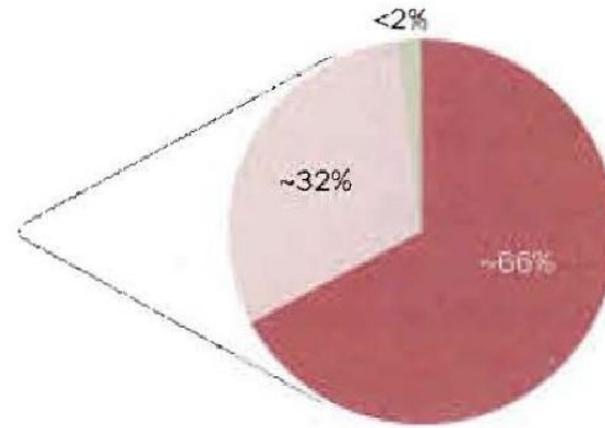
# Процент генов в геноме



# Сравнительный состав mt и n генома



nuclear genome



mitochondrial genome

highly conserved sequences

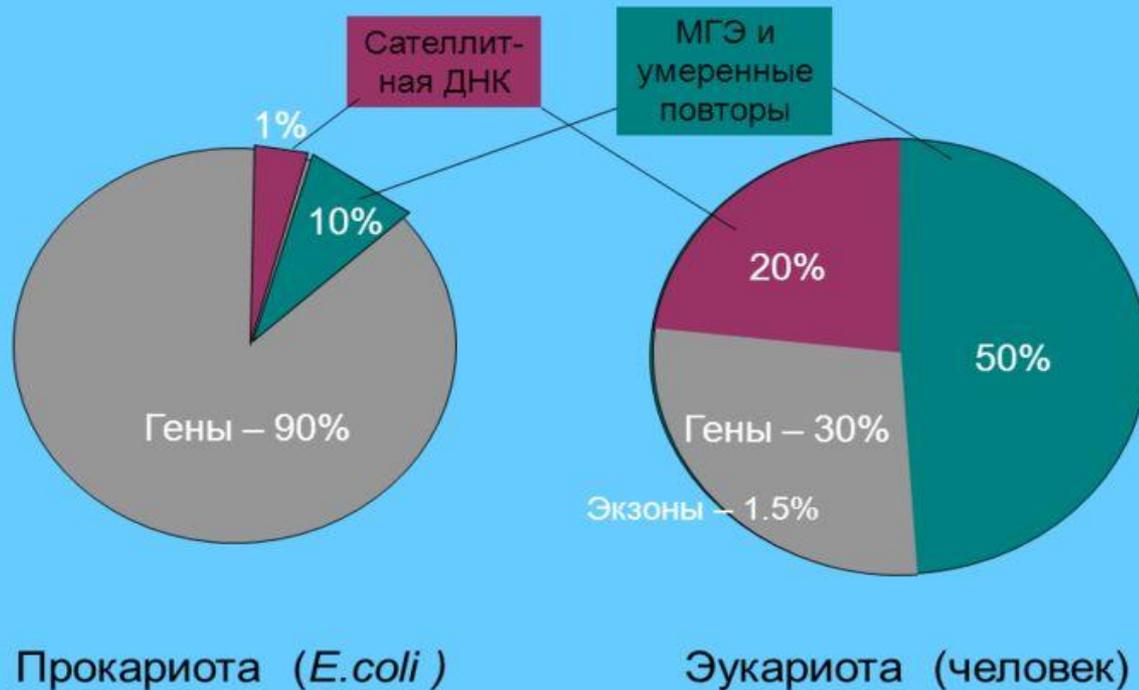
- protein-coding genes
- RNA genes, regulatory sequences

poorly conserved sequences

- transposon-based repeats
- heterochromatin
- other sequences

# Безъядерные и ядерные

## Геном прокариот и эукариот



MyShared

# Повторы в геноме

- У прокариот нет повторов в геноме, только плотно стоящие гены, друг за другом.
- Для человека до 50% всего генома занимают повторы перемежающиеся с функциональными генами.
- Для хвойных огромное количество повторов разного типа (до 80% всего генома).

# Основные этапы выделения

## ДНК



1. Разрушение клеточной стенки (лизис)
2. Осаждение грубых остатков (13000g)
3. Фенольно-хлороформная экстракция нуклеиновых кислот и осаждение белков
4. Осаждение ДНК этанолом (преципитация)



# Можно самим выделить свою ДНК «на кухне»

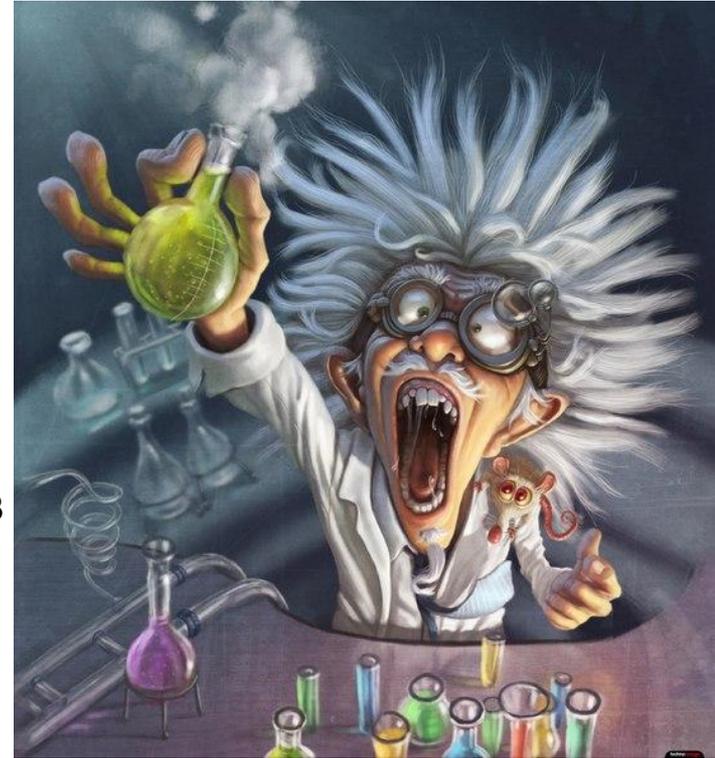
- 1. Смешать 500 мл воды и 1 ст. ложку соли. Размешать пока соль не растворится. Потом отлить 3 ст. ложки в др. стакан.
- 2. Прополоскать рот 1 мин. Сплюнуть обратно в стакан. Жидкость содержит клетки с внутренней стороны щек.
- 3. Аккуратно размешать жидкость с каплей мыла (старайтесь без пузырьков). Мыло разрушает мембраны клеток.
- 4. В отдельном стакане смешайте 100 мл изопропилового спирта и 3 капли красителя.
- 5. Наклонить соляной раствор и аккуратно влейте спирт, что бы образовался отдельный слой.
- 6. Подождите 2,5 мин. Вы увидите белый сгустки и нитевидные формирования. Это ваша ДНК
- [http://www.progene.ru/news/vydelenie\\_dnk\\_v\\_domashnikh\\_uslovijakh/2012-03-10-39](http://www.progene.ru/news/vydelenie_dnk_v_domashnikh_uslovijakh/2012-03-10-39)
- **Необходимое:** вода 0,5 л., соль поваренная 1 ст.л., стаканы и пробирки, моющее средство для посуды, спирт 100 мл, краситель пищевой, палочки.

# или выделить ДНК лука

1. Овощ почистить, порезать на куски и измельчить в блендере добавив 1 столовую ложку соли.
2. Кашицу процедить и добавить средство для мытья посуды. Перемешать и дать постоять 10 минут.
3. Ещё раз процедить и разлить по пробиркам.
4. Добавить пару капель раствора для хранения контактных линз.
5. Аккуратно, по краю пробирки влить спирт слоем в 2 см и оставить на 10-15 мин.
6. Всплывшие хлопья и есть очищенная ДНК.

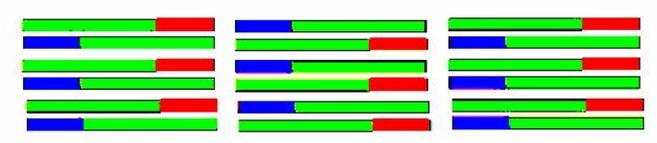
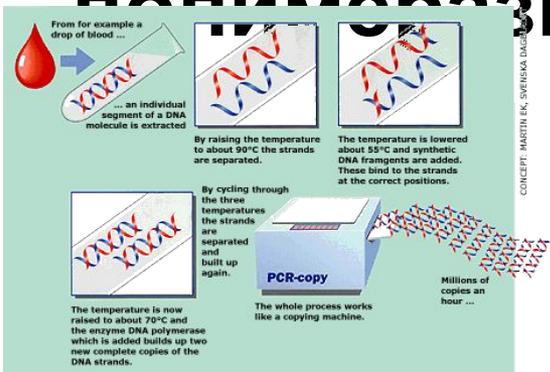
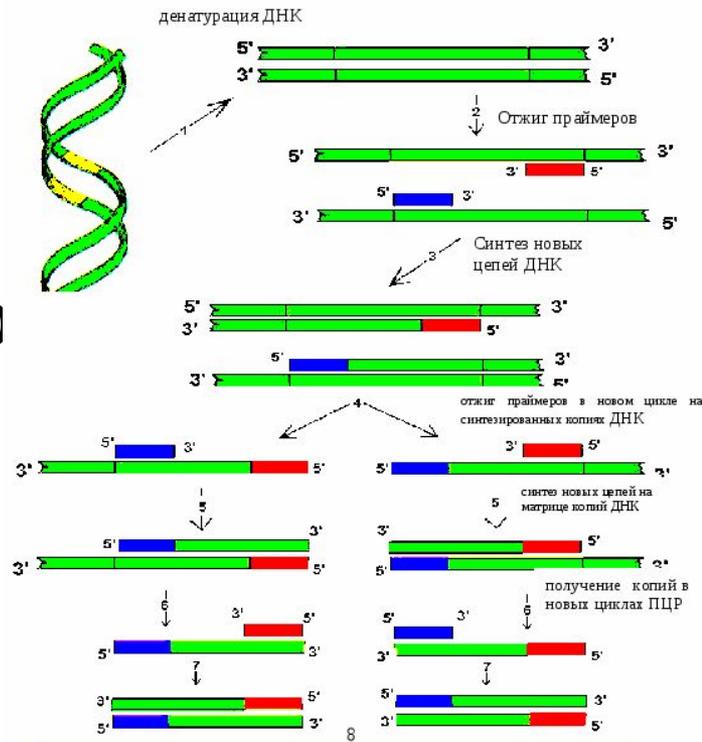
• <http://www.youtube.com/watch?v=ooZD3U62o50>

• **Необходимое:** вода 0,5л, соль поваренная 1 ст. л., овощ, моющее средство для посуды, жидкость для хранения контактных линз, спирт 100 мл., пробирки и стаканы, ситечко, блендер.



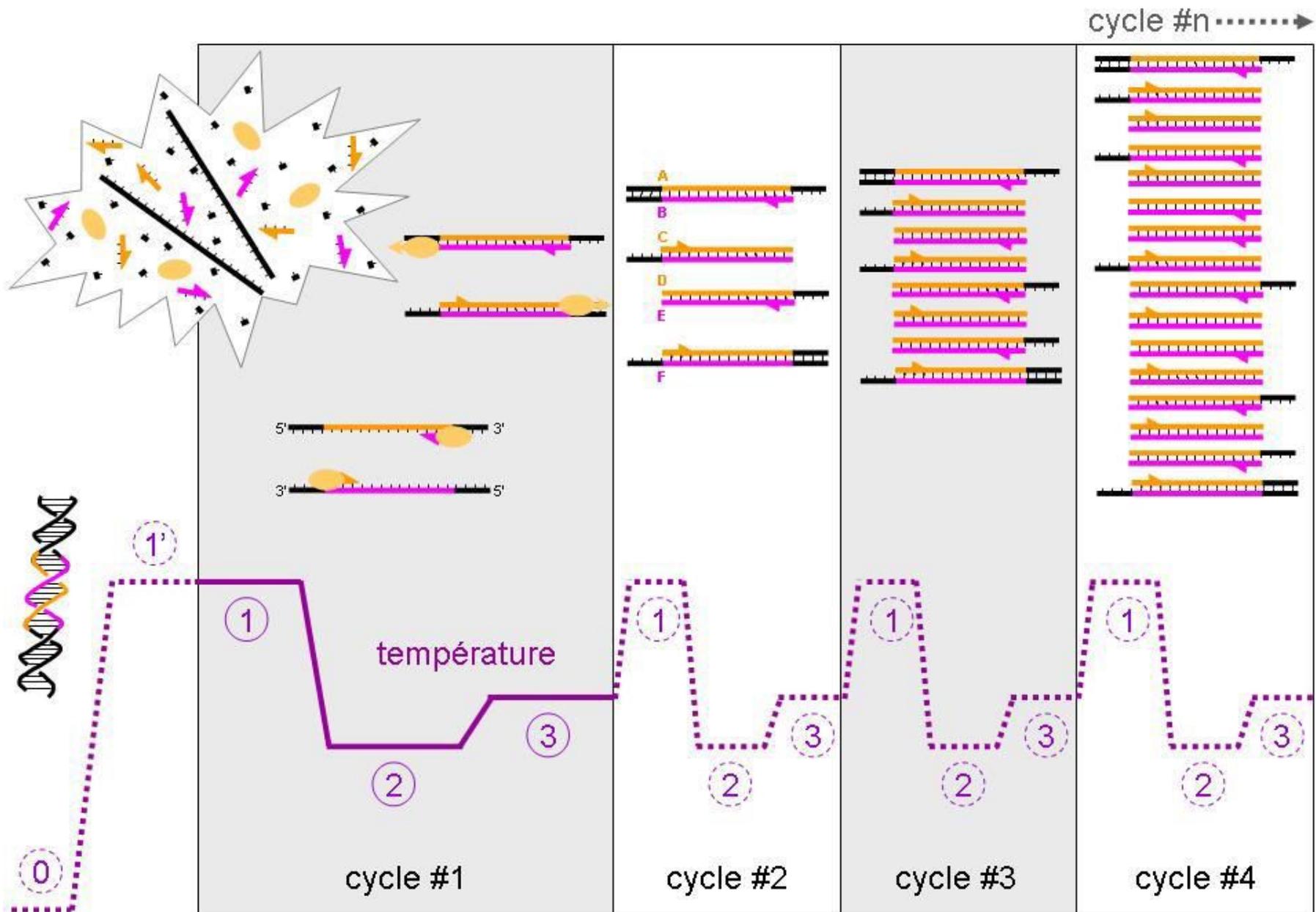
# Полимеразная цепная реакция

- Полимеразная цепная реакция (ПЦР) была открыта в 1983 году американским биохимиком К. Муллисом.
- Амплификация ДНК в ходе многократных удвоений исходной молекулы ДНК с помощью фермента **ДНК-полимеразы**.

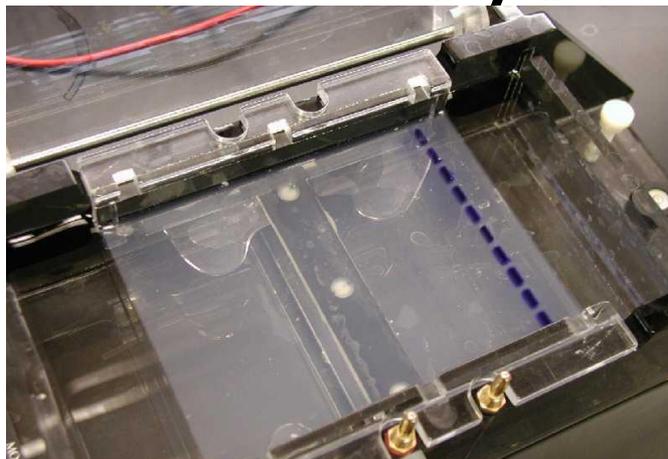


1 – денатурация двойных цепей ДНК; 2 – отжиг праймеров; 3 – синтез новых цепей на основе первоначальной матрицы ДНК; 4 – отжиг праймеров на синтезированных копиях в новом цикле; 5 – синтез новых цепей на матрице синтезированной копии ДНК; 6, 7 – в новом цикле отжиг праймеров и синтез копий на матрице полученных фрагментов ДНК; 8 – полученные копии.

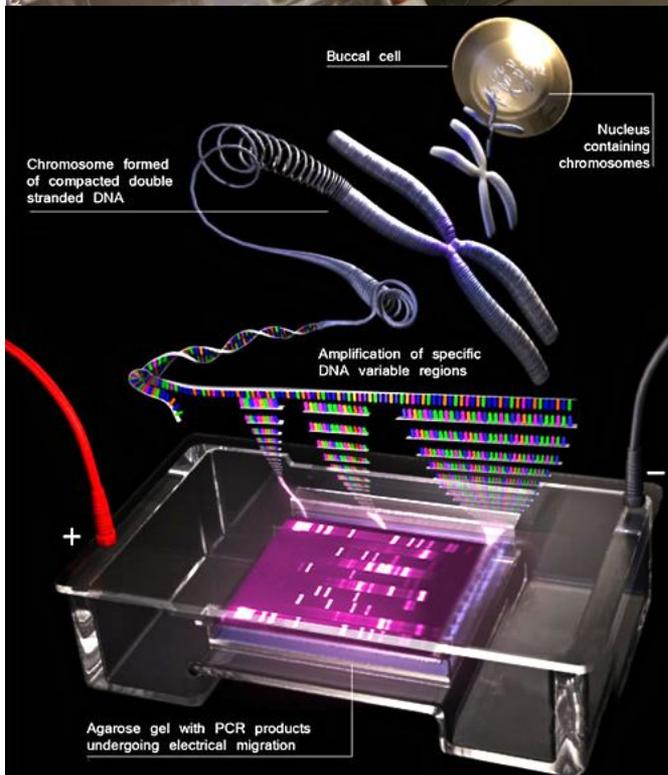
- ПЦР очень мощный **метод размножения фрагментов ДНК** вне организмов (*in vitro*). С момента его открытия метод используется сейчас практически во всех молекулярно-биологических лабораториях.
- **Первые опыты были нестабильные** и приходилось постоянно добавлять новую полимеразу после каждого цикла удвоения.
- Фермент, выделенный из термофильной бактерии ***Thermus aquaticus***, способной жить и размножаться в жестких условиях среды – **Taq-полимераза**, облегчил жизнь молекулярного биолога.



# Визуализация и контроль

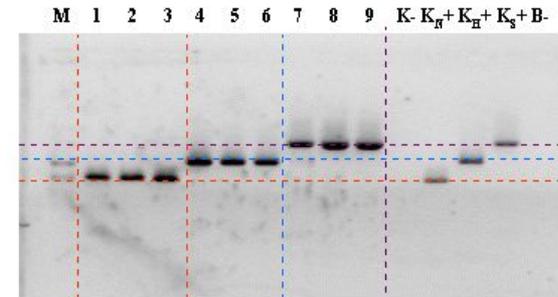


**Электрофорезом** называют движение заряженных частиц в растворе под действием электрического поля.

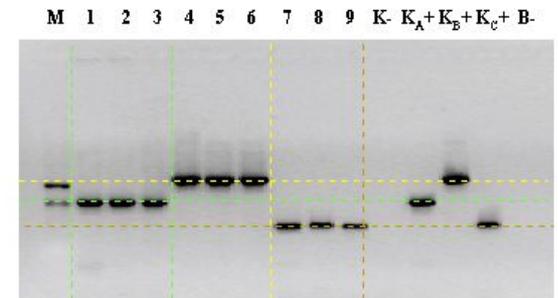


**Агароза** – это особая чистая фракция природного линейного поли-сахарида агара, который получают из морских красных водорослей (*Gracilaria*, *Gelidium*, *Ahnfeltia*).

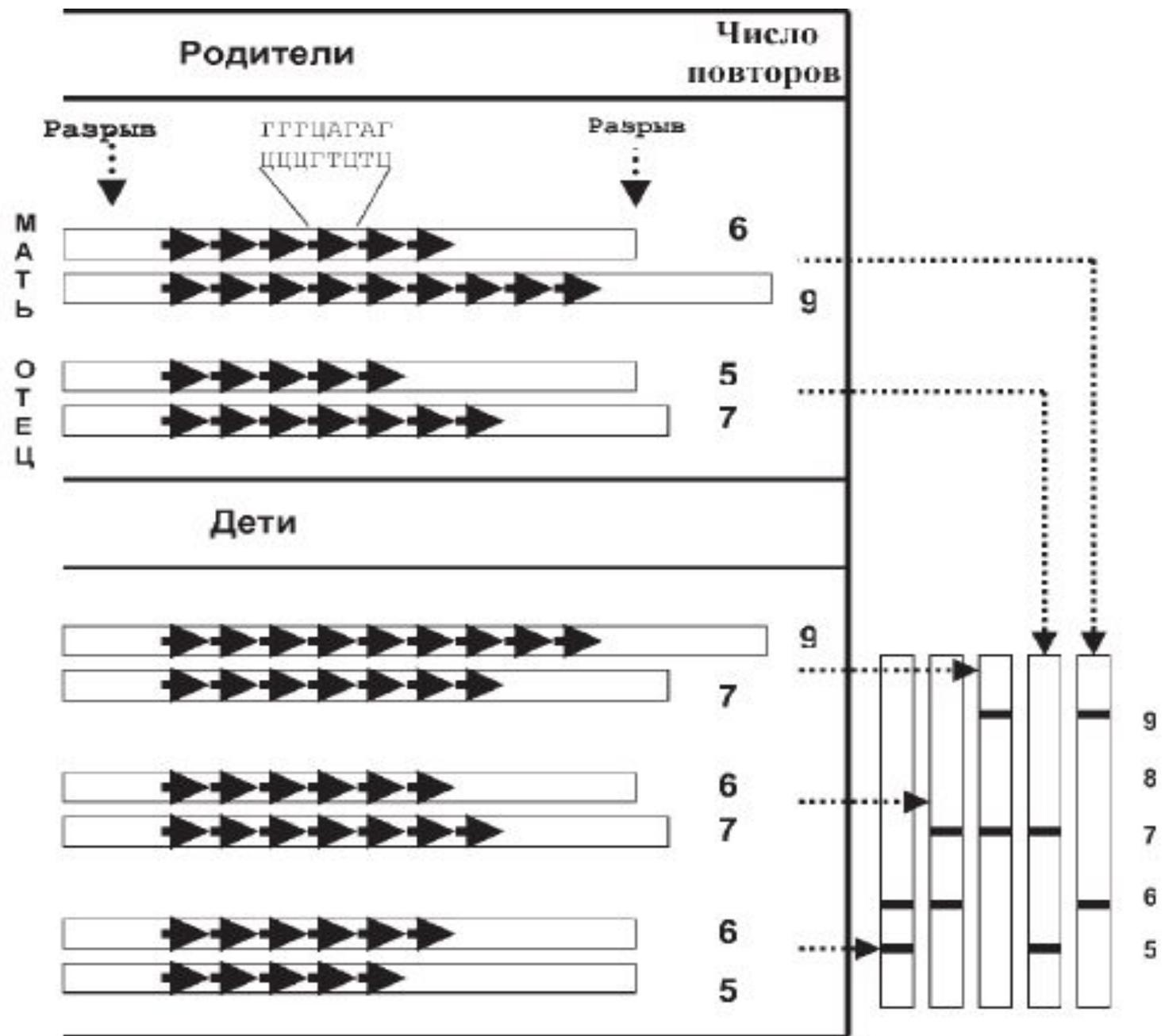
Рис. 1. Внешний вид результатов электрофореза продуктов ПЦР после амплификации с родоспецифичными (А) и серогрупповыми (Б) праймерами. Направление электрофореза - снизу вверх (на рисунке)



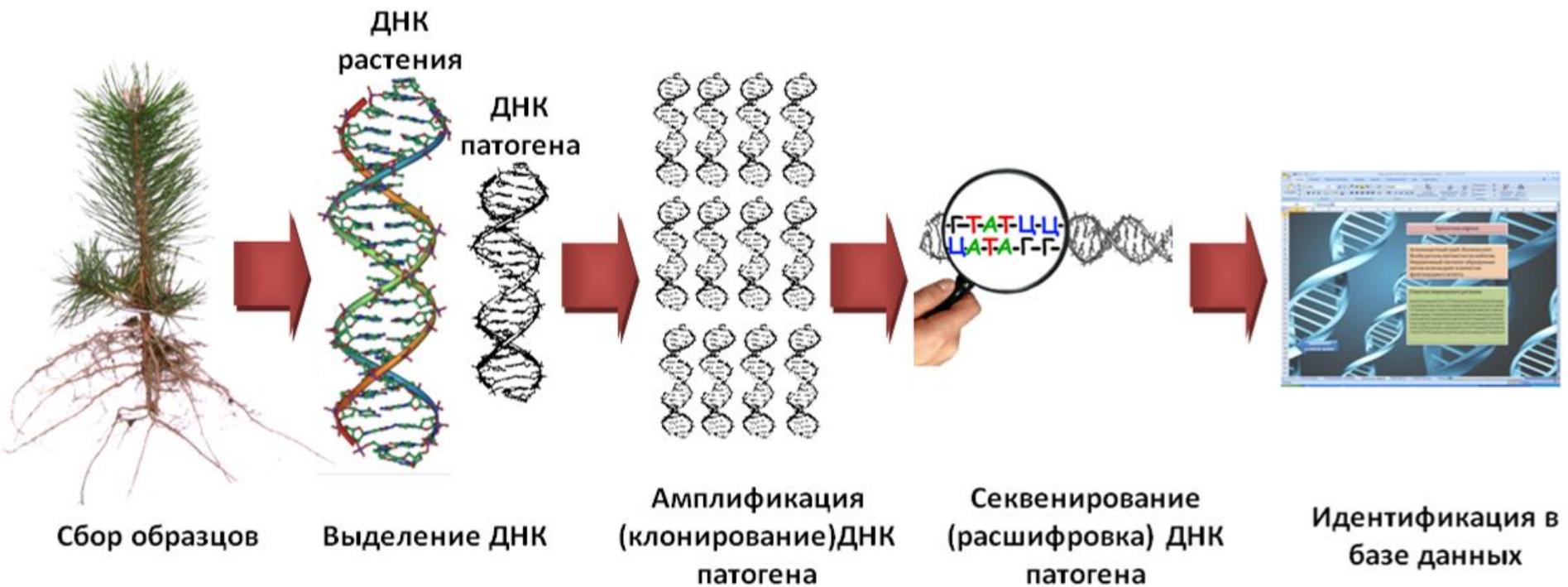
Дорожки 1-9 - образцы СМЖ от 9 больных ГВМ  
 К- - отрицательный контроль стадии 2 (ПЦР)  
 В- - отрицательный контроль стадии 1 (выделение ДНК)  
 К<sub>А</sub>+, К<sub>В</sub>+, К<sub>С</sub>+ = ДНК *N.meningitidis*, *H.influenzae*, *S.pneumoniae*, то есть положительный контроль стадии 2 (ПЦР)  
 М - маркер длины фрагментов ДНК



Дорожки 1-9 - образцы СМЖ от 9 больных менингококковым менингитом  
 К- - отрицательный контроль стадии 2 (ПЦР)  
 В- - отрицательный контроль стадии 1 (выделение ДНК)  
 К<sub>А</sub>+, К<sub>В</sub>+, К<sub>С</sub>+ = ДНК *N.meningitidis* серогрупп А, В, С, то есть положительный контроль стадии 2 (ПЦР)  
 М - маркер длины фрагментов ДНК



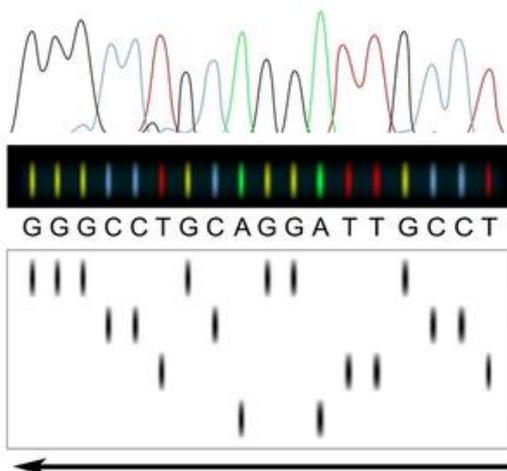
- **Секвенирование** биополимеров (белков и нуклеиновых кислот — ДНК и РНК) — определение их аминокислотной или нуклеотидной последовательности



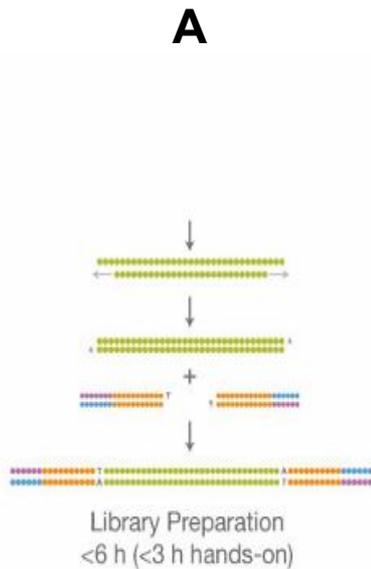
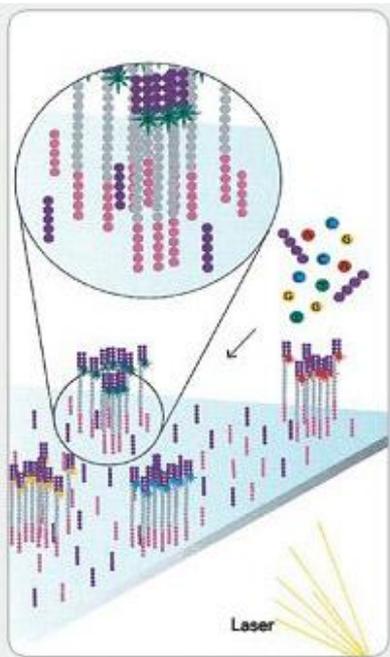
# Наиболее значимые даты

- **1944: идентификация ДНК как генетического материала для всех живых организмов**
- **1953: расшифровка генетического кода (Watson & Crick, *Nature* 171, 737: 1953).**
- **1977: первый полный сиквенс целого генома бактериофага phiX174; всего только 5386 нуклеотидов, в 60000 раз меньше генома человека (Sanger et al. *Nature* 265, 687: 1970).**
- **mid-1980s: бурное развитие автоматизации и компьютеризации секвенирования**
- **1990: начало проекта полного секвенирования генома человека**
- **1997: полный сиквенс генома дрожжей (~12 Mbp)**
- **1998: нематоды (~97 Mbp)**
- **2000: арабидопсиса (~125 (Mbp)**
- **2000: дрозофилы (~180 Mbp)**
- **2001: человека (~3,200 Mbp)**
- **2002: мыши (~3,500 Mbp) и риса (~420 Mbp)**
- **2006: тополя (~550 Mbp)**
- **2008: Новое поколение секвенирующих платформ - Next generation sequencing**
- **2013: неандертальца (~3,200 Mbp)**
- **2013: ели и 2014: сосны (~20,000 Mbp) 2015: кедр и лиственница?**

- Дидезоксинуклеотидный метод, или метод «обрыва цепи», был разработан Ф. Сенгером в 1977 году и в настоящее время широко используется для определения нуклеотидной последовательности **одного участка**

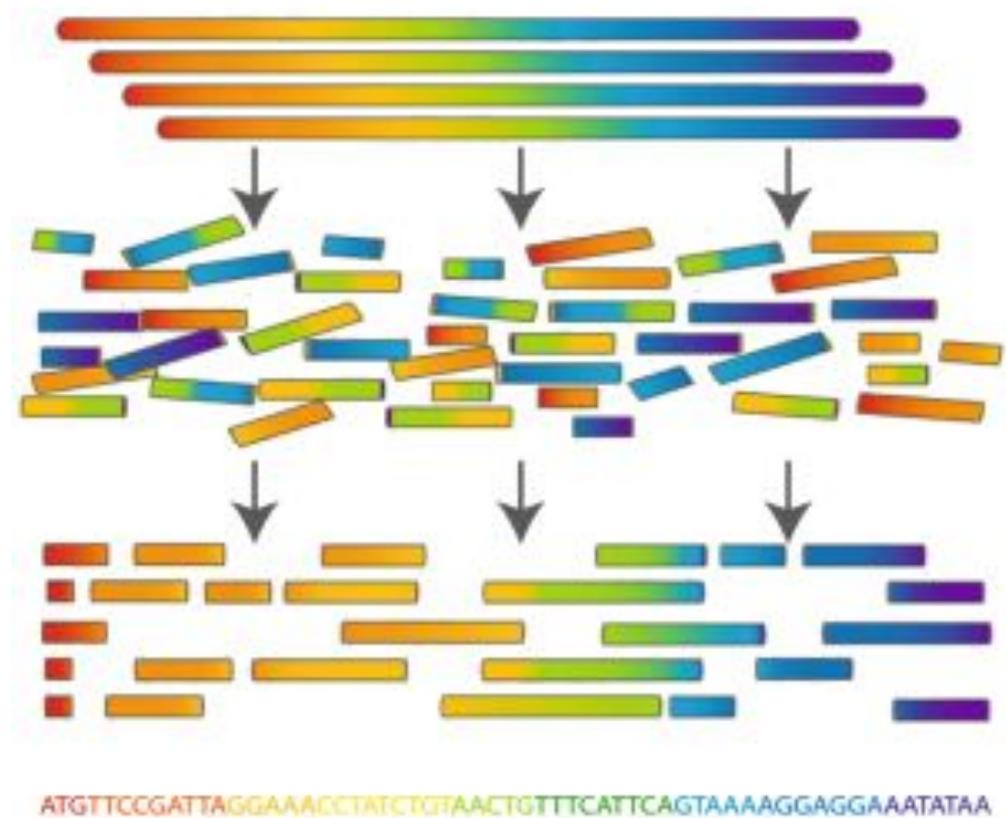


- Секвенирование методом синтеза новой цепи - отличительной особенностью методов секвенирования ДНК нового поколения является их направленность на получение нуклеотидных последовательностей **целых геномов** различных **организмов, включая человека.**

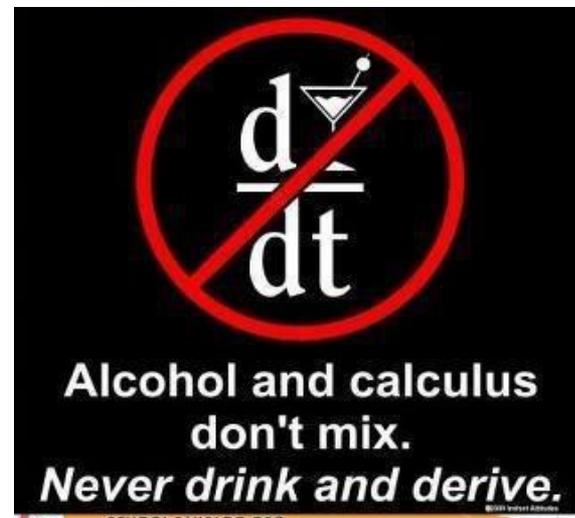
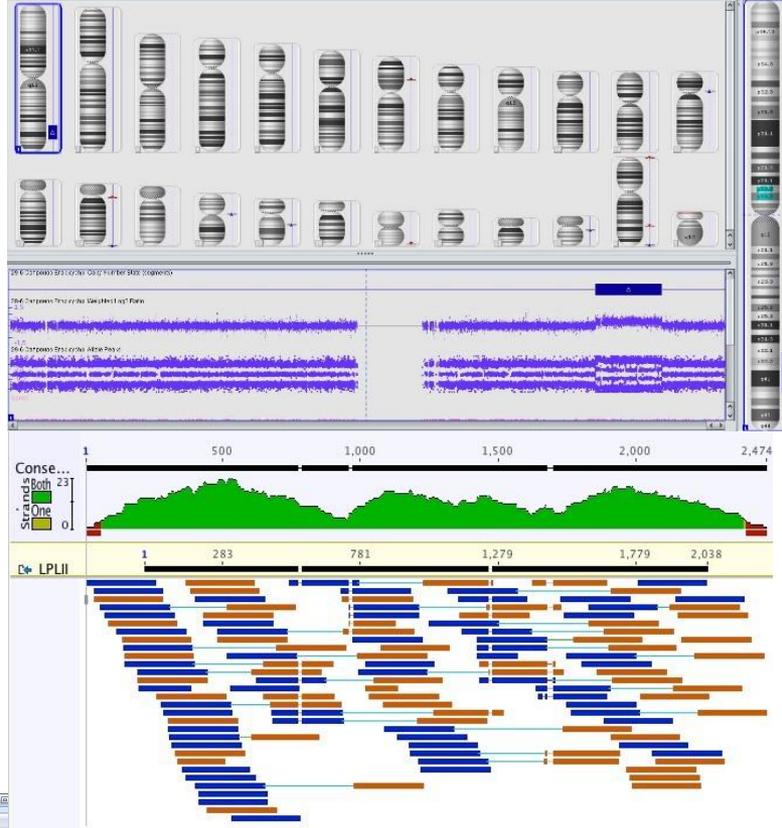
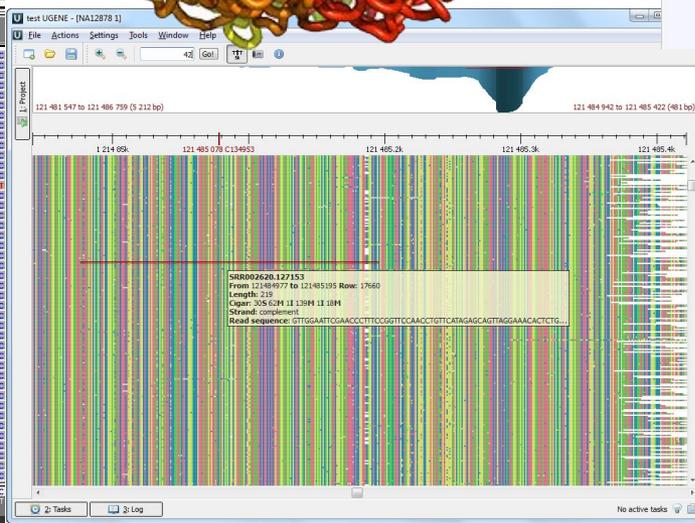
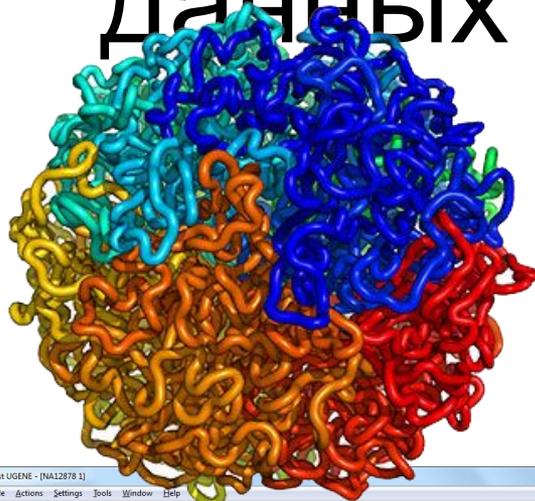
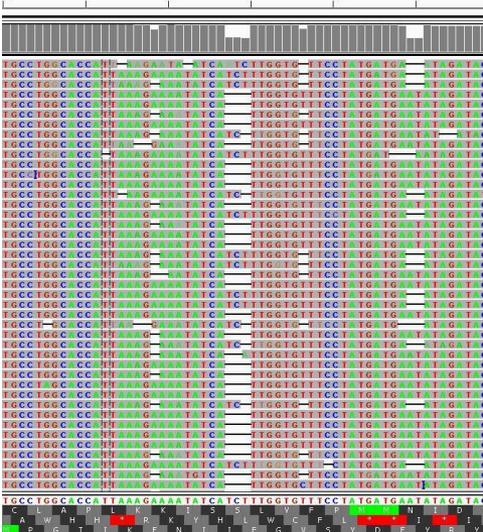
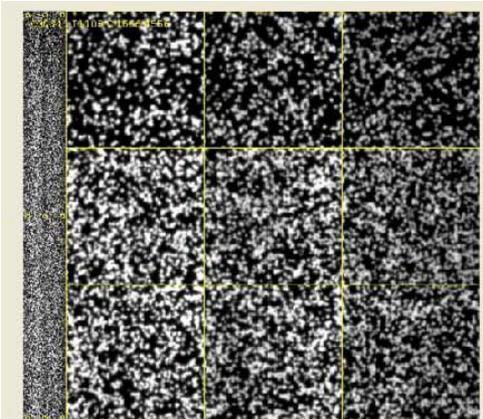


# Процесс обработки данных полногеномного секвенирования

- 1 полный запуск HiSeq 2000 = 300 полных геномов человека
- 1 дорожка = 50Gb



# Обработка а данных



Me



Myself

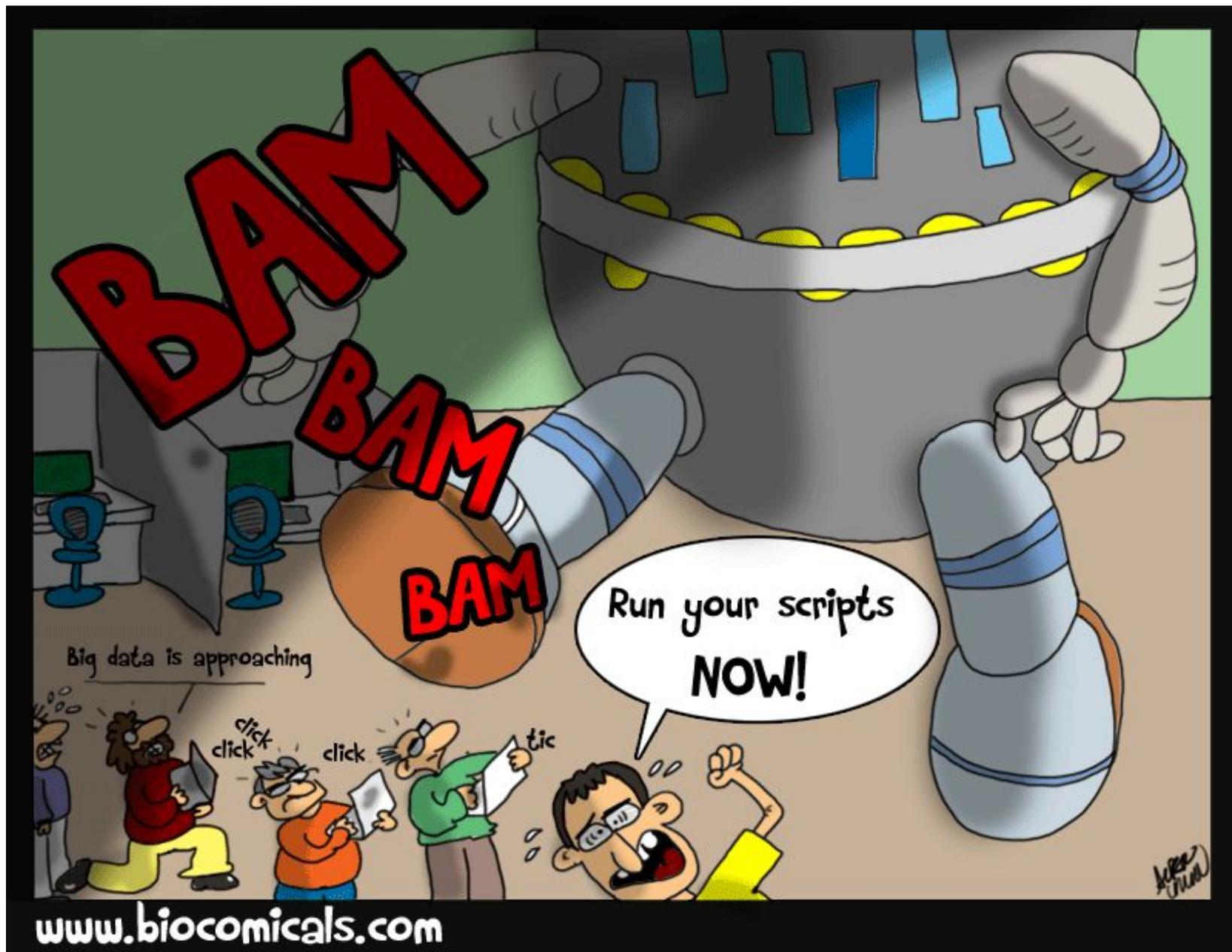


and  
Big DATA

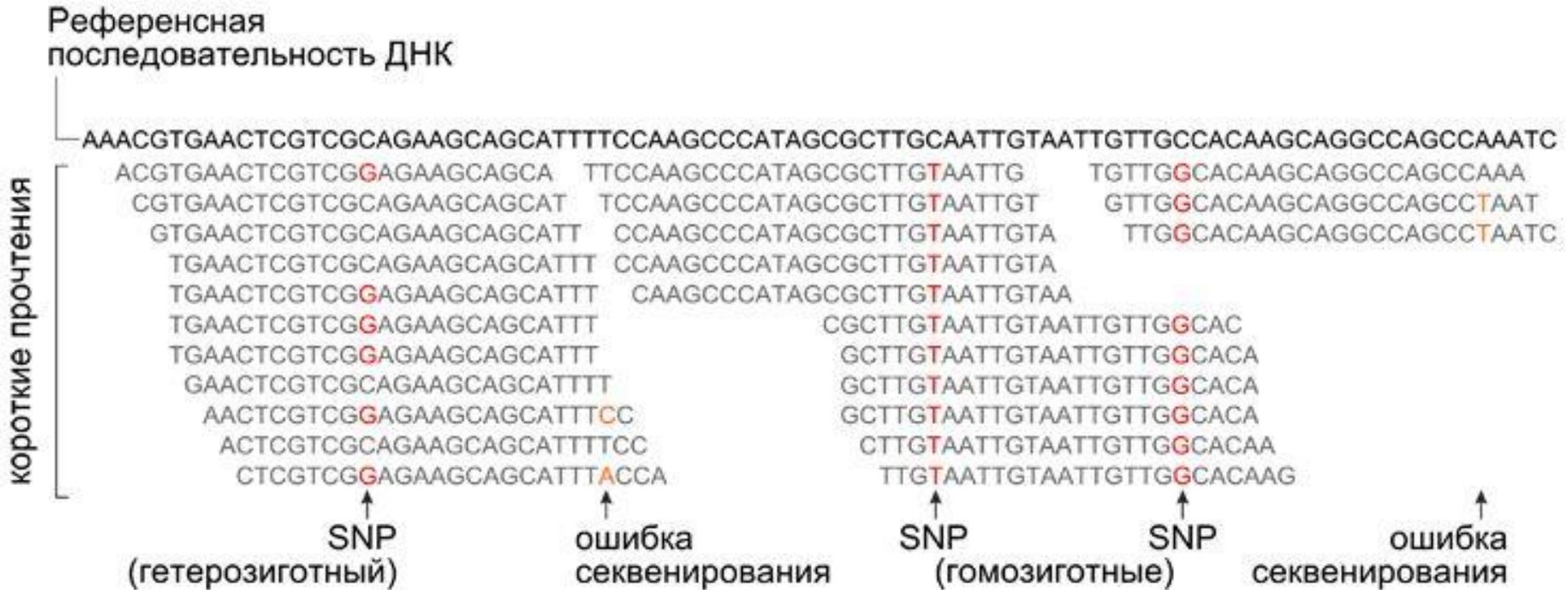


www.biocomicals.com

www.biocomicals.com



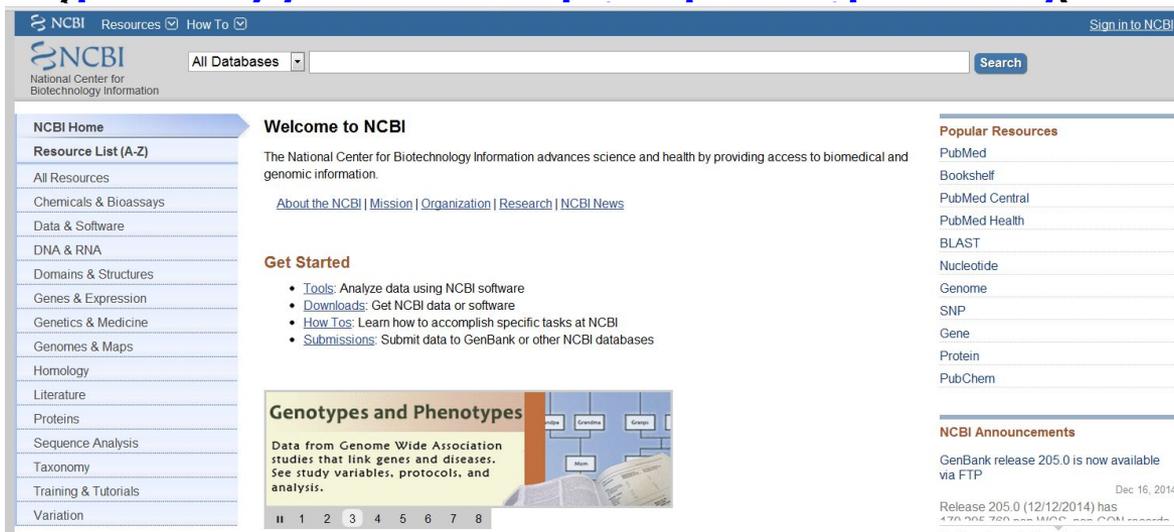
# Сравнение 11 объектов



Eukaryota	Cons eRF1	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 194	
	Homo sapiens	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 200	
	Xenopus laevis	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 200	
	Mesocricetus auratus	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 200	
	C. elegans	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 208	
	(a) Arabidopsis thaliana	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 198	
	(b) Arabidopsis thaliana	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 199	
	Saccharomyces cerevisiae	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 197	
	Schizosaccharomyces pombe	: GSAREVLQRFVTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 197	
	Podospora anserina	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 201	
	Plasmodium falciparum	: GNTREVLHKFTVDLPKKEGRGGQ	SALRFARLRMEKRHN	---	: 198	
	Archaea	Archaeoglobus fulgidus (arch)	: GKPIEVLWDWTSMPVPGKHRGGQ	SSVRFERLREIAIHE	---	: 191
		Methanococcus jannaschii (arch)	: GRNINILKLTSGVPGKFKAGGQ	SARRLERLIDLAAHE	---	: 200
		Methanobacterium thermoautotrophicum (arch)	: GKRIDLKLTSGVPGKHKAGGQ	SQRRFDRIDLAAHE	---	: 196
Pyrococcus horikoshii (arch)		: GKRIELDELTSMPVPGKTRAGGQ	SARRYERIREQETHE	---	: 197	
Cons RF1, 2		: VEIDINPADLRIDTYRASGAGGQ	HVNTTDSAVRITHLR	---	: 247	
Prokaryota	1 Bacillus subtilis (pro)	: VEVDIHEKDIRVDTFASSGPGGQ	SVNTTMSAVRLTHLP	---	: 248	
	1 Borrelia burgdorferi (pro)	: TEIDINEKDLRIDVYRSGAGGQ	HVNTTDSAVRITHLP	---	: 249	
	1 Coxiella burnetii (pro)	: DQIKINPAELRIDTFRSGAGGQ	HVNRTDSAIRITHLP	---	: 252	
	1 Escherichia coli (pro)	: ELPDINPADLRIDTFRSGAGGQ	HVNTTDSAIRITHLP	---	: 250	
	1 Haemophilus influenzae (pro)	: EMPEINPADLRIDTYRSGAGGQ	HVNTTDSAVRITHIP	---	: 249	
	1 Helicobacter pylori (pro)	: VEVSINPSDLKIEVFRAGGEGGQ	CVNTTDSAVRITHLP	---	: 248	
	1 Mycoplasma carpicolum (pro)	: VEIEIRSNDLRIDTYRASGAGGQ	HVNTTDSAVRITHLP	---	: 252	
	1 Mycoplasma genitalium (pro)	: VEITINPSDLRIDTYRASGAGGQ	HVNRTESAVRITHLP	---	: 251	
	1 Mycobacterium leprae (pro)	: GEVHIDESDLRIDVYRSGAGGQ	HVNRTESAVRITHLP	---	: 251	
	1 Mycoplasma pneumoniae (pro)	: VEVHINPADLRVDTYRASGAGGQ	GVNTTDSAVRITHLP	---	: 251	
	1 Mycobacterium tuberculosis (pro)	: GQVQIDESDLRIDVYRSGAGGQ	GVNTTDSAVRITHLP	---	: 251	
	2 Salmonella typhimurium (pro)	: ELPDINPADLRIDTFRSGAGGQ	HVNTTDSAIRITHLP	---	: 250	
	2 Synechocystis sp. (pro)	: VEVKIDPKDIEMSTARSGAGGQ	SALRFARLRMEKRHN	---	: 254	
	2 Borrelia burgdorferi (pro)	: IEITIKPEDIRIDTYRASGAGGQ	HVNKTSSAVRITHIE	---	: 247	
	2 Escherichia coli (pro)	: IDIEINPADLRIDVYRTSGAGGQ	HVNRTESAVRITHIP	---	: 267	
	2 Haemophilus influenzae (pro)	: IDIEINPADLRIDVYRSGAGGQ	HVNKTESAVRITHMP	---	: 267	
	2 Helicobacter pylori (pro)	: IDIEIDEKDVRYDYRSGAGGQ	HVNKTESAVRITHFP	---	: 266	
	2 Mycobacterium tuberculosis (pro)	: DHIDIPEGDVYRVDVYRSGGPGGQ	SVNTTDSAVRLTHIP	---	: 275	
	2 Salmonella typhimurium (pro)	: IDIDINPADLRIDVYRSGAGGQ	HVNRTESAVRITHIP	---	: 267	
	2 Streptomyces coelicolor (pro)	: DHIEIDSELRVDVYRSGGPGGQ	GVNTTDSAVRLTHIP	---	: 266	
	Mitoch.	mit1 Saccharomyces cerevisiae	: YERTFKPEIRVDIMRASGKGGQ	HVNTTDSAVRLTHIP	---	: 302
		mit1 C. elegans	: VSVVVPDSVKEAMRASGPGGQ	NVNRKSTAVRMTHE	---	: 274
mit1 Homo sapiens		: VDVKLDPKDLRIDTFRAGGQ	HVNKTDSAVRLVHIP	---	: 327	
mit1 Kluyveromyces lactis		: YERTFKPEIRVDIMRASGKGGQ	HVNTTDSAVRLTHIP	---	: 283	
mit1 Schizosaccharomyces pombe		: SSSLYDSSEVKIEVMRSRAGGQ	HVNRTESAVRLTHIP	---	: 285	

Рис. 1. Сравнение аминокислотных последовательностей в однобуквенном коде белков – факторов терминации трансляции I-го класса (RF1, RF2, eRF1) – у представителей трех царств живой природы: эукариот, археобактерий и прокариот. Фрагменты структуры заимствованы из банка SwissProt; справа указаны порядковые номера аминокислот, слева – видовые названия организмов и номера фактора митохондрий (mit), относящиеся к прокариотическому типу; черным цветом выделен мотив, общий для всех видов и функционально важный для человека

- первым шагом в исследовании нового гена или белка является сравнение со всеми уже известными последовательностями при помощи международного банка генетических данных NCBI



The screenshot shows the NCBI homepage with a navigation menu on the left, a central 'Welcome to NCBI' section, and a 'Popular Resources' list on the right. The navigation menu includes 'NCBI Home', 'Resource List (A-Z)', and various biological categories. The central section features a 'Welcome to NCBI' message, a search bar, and a 'Get Started' section with links to 'Tools', 'Downloads', 'How To's', and 'Submissions'. The 'Popular Resources' list includes PubMed, Bookshelf, PubMed Central, PubMed Health, BLAST, Nucleotide, Genome, SNP, Gene, Protein, and PubChem. The 'NCBI Announcements' section mentions the GenBank release 205.0.



# Пример полученных данных

- >gi|7524593:31594-31803 Pinus thunbergii chloroplast, complete genome  
ATGGAATGGATAGAACAGATTCTCCCGGAGAAT  
TCCCTCCGGGAAGGTCGAAATATTGAAAAGTTTT  
TTCTATTGGATTGGAAATGAACGAAAAGAATTC  
AATCCAATTCTTTTCCAAAAATGAAATCCTGTTG  
ACTTTTTCAACAATAGACTCAAGATCTGCATCTTT  
ATCGAACATATATTCCAGCTCCGATTTGATTAAG  
GAGTAG

# Неизвестный участок ДНК

- ATGGCACGTTTCGCTGAAAAAAAAAATCCTTTTGTGG  
СТААТСАТТСАТТGAGGAAAААТТСААААТСТААА  
СААААGGAAGAAAAGAGATAАТАGТGACTTG  
GTCCCGGGCАТСТGTCAТТGTACCTGCAATGATC  
GGTCATACAАТТGCTGTTCATAATGGGAGGGAA  
САТТТАСССАТТТATGTAAСAGATCGTATGGTAG  
ACCACAААТТGGGGGAАТТТGCACСТАСТСТАСТ  
TTTTCAAGGGACATGCGAGAAACGATAAGAAATC  
TCGTCGTТАА

# Пример первого участка

**complement(1..1062) gene=«psbA»**

**product="photosystem II protein D1**

- /translation="MTAIIERRESANLWGRFCDWITSTENR  
LYIGWFGVLMIPTLLTASVFIIAFIAAPPVDIDGIREPVSG  
SLLYGNNIISGAIPTSAAGLHFYPIWEAASVDEWLYNG  
GPYELIVLHFLLGVACYMGREWELSFRLGMRPWIAVA  
YSAPVAAASAVFLIYPIGQGSFSDGMPLGISGTFNFMIV  
FQAEHNILMHPFHMLGVAGVFGGSLFSAMHGSLVTSS  
LIRETTENQSANAGYKFGQEEETYNIVAAAHGYFGRLIFQ  
YASFNNSRSLHFFLAAWPVAGIWFTALGISTMAFNLN  
GFNFNQSVVDSQGRVINTWADIINRANLGMEVMHER  
NAHNFPLDLAAVESISIGG"

# Идентификация неизвестного организма

- <http://genome.sfu-kras.ru/deych> - где лежат неизвестные последовательности
- [http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) – где можно идентифицировать последовательность



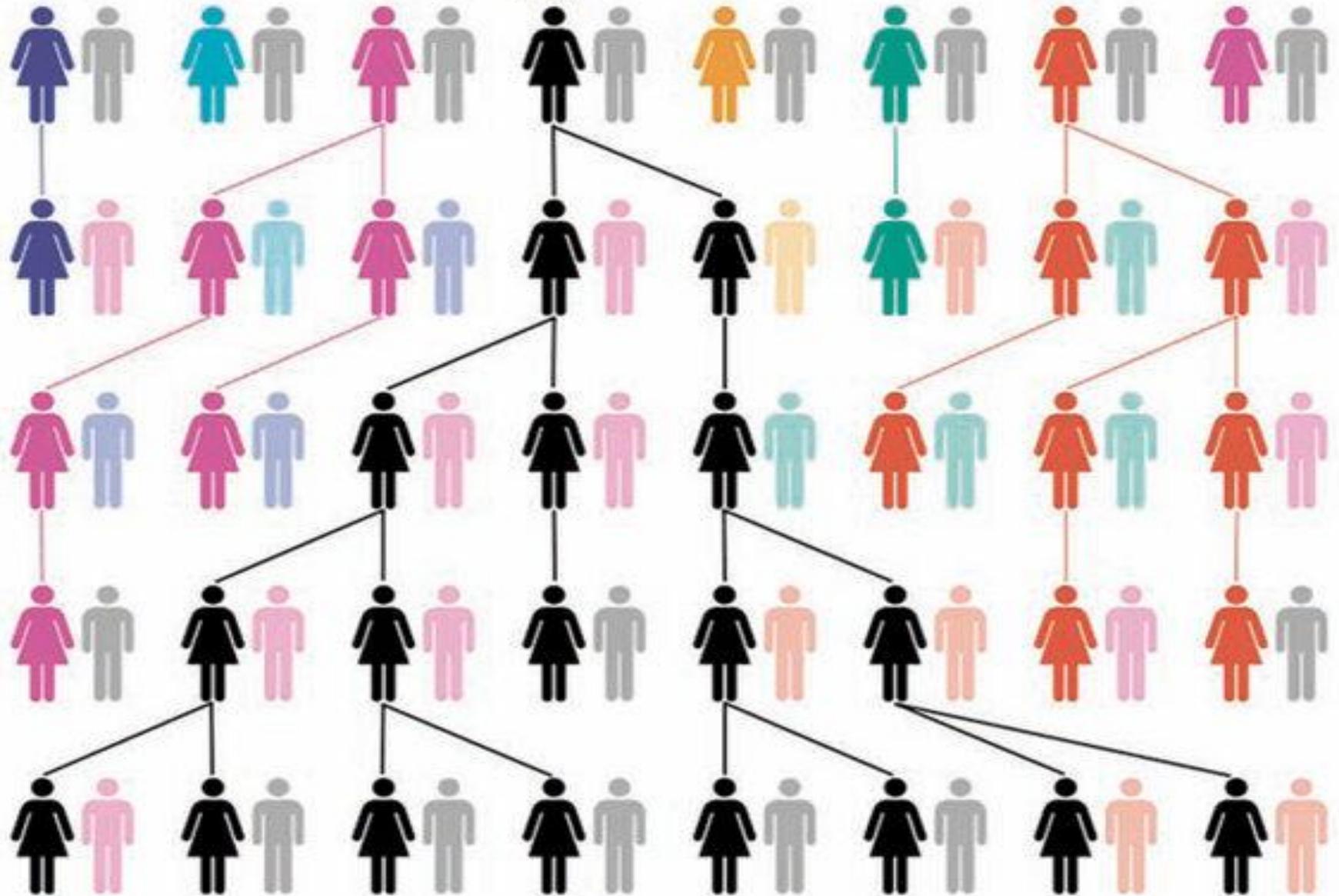
# Не метилирование МИТОХОНДРИАЛЬНОГО ГЕНОМА

- Высокая концентрация активных форм кислорода в митохондриях и слабая система репарации **увеличивают частоту мутаций мтДНК** по сравнению с ядерной на порядок.
- Радикалы кислорода служат причиной специфических замен Ц-Т (дезаминирование цитозина) и Г-Т (окислительное повреждение гуанина), вследствие чего, возможно, **мтДНК богаты АТ-парами**.
- Кроме того, все **мтДНК — не метилируются**, в отличие от ядерных и прокариотических ДНК. Известно, что метилирование (временная химическая модификация нуклеотидной последовательности без нарушения кодирующей функции ДНК) — один из механизмов программируемой инактивации и сохранения генов [Гвоздев В.А. // Сорос. образоват. журн. 1999. №10. С.11—17.].

# «Митохондриальная Ева»

- Митохондрии человека содержат кольцевую молекулу ДНК (мтДНК), состоящую из **16 500 пар** нуклеотидов
- мтДНК передается **только по материнской** линии и не участвует в рекомбинации
- В клетке присутствует 100-1000 копий мтДНК (по сравнению с единичными копиями ядерных генов). Кроме того, кольцевая форма мтДНК делает ее более устойчивой к разрушению, которое часто начинается со свободных концов молекулы.

митохондриальная  
Ева



# Порядок наследования мтДНК

Мужчины носят в себе материнскую мтДНК, но только женщины передают ее своим детям

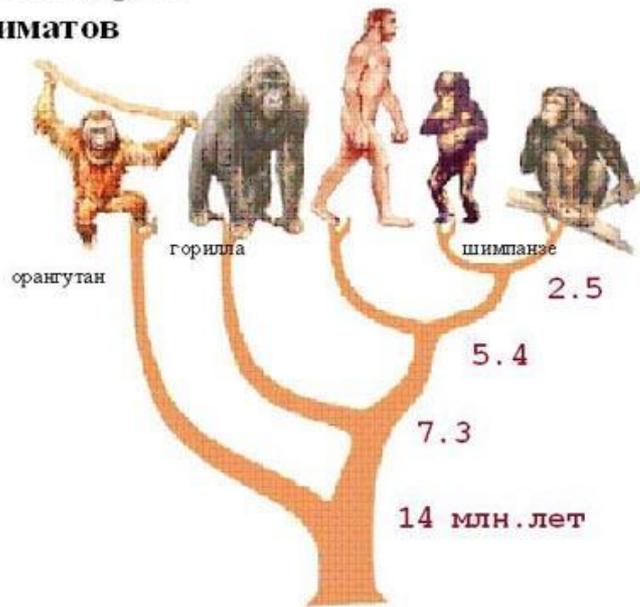
**красный** = носитель



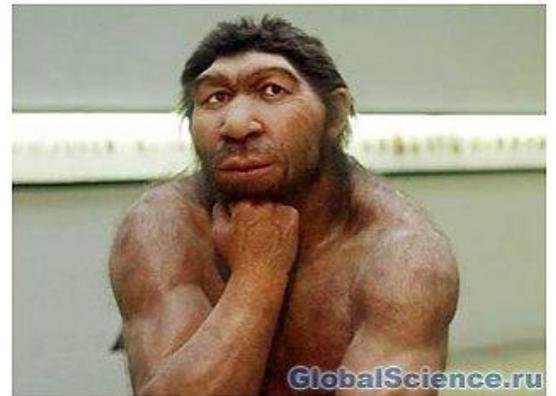
- Речь идет лишь о сохранении до настоящего времени одной из нескольких линий мтДНК
- Изучение разнообразия ДНК других генов, из разных хромосом, показало, что численность популяции в период видообразования составляла около 10 000 человек.

# Все люди на 99% по составу ДНК одинаковые

Филогенетическое древо  
высших приматов



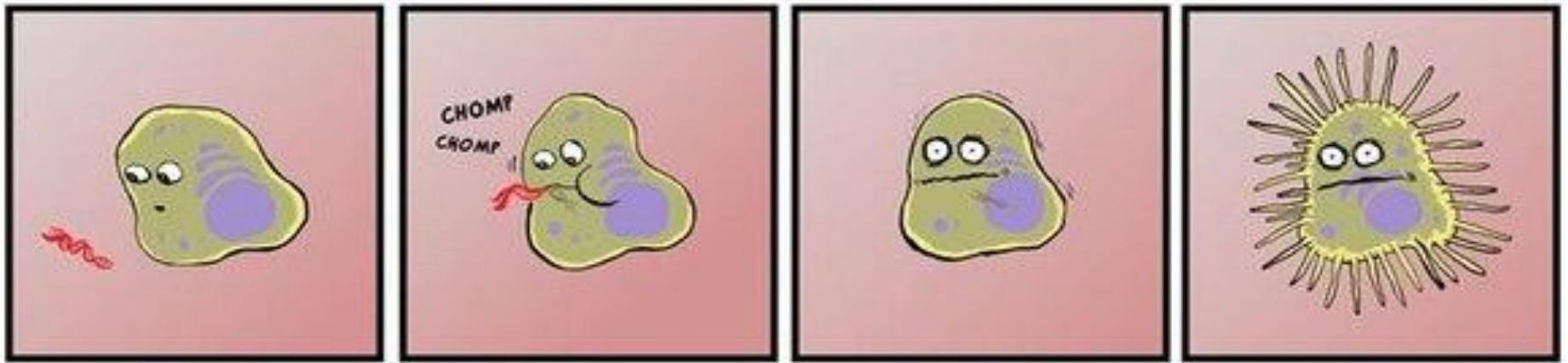
Генетические различия на уровне ДНК  
между людьми: 1 нуклеотид из 1000  
между человеком и шимпанзе : 1 нукл. из 100



- В каждом европейце 1-4% генома неандертальца при совпадении геномов на 99,5% (человек-неандерталец).

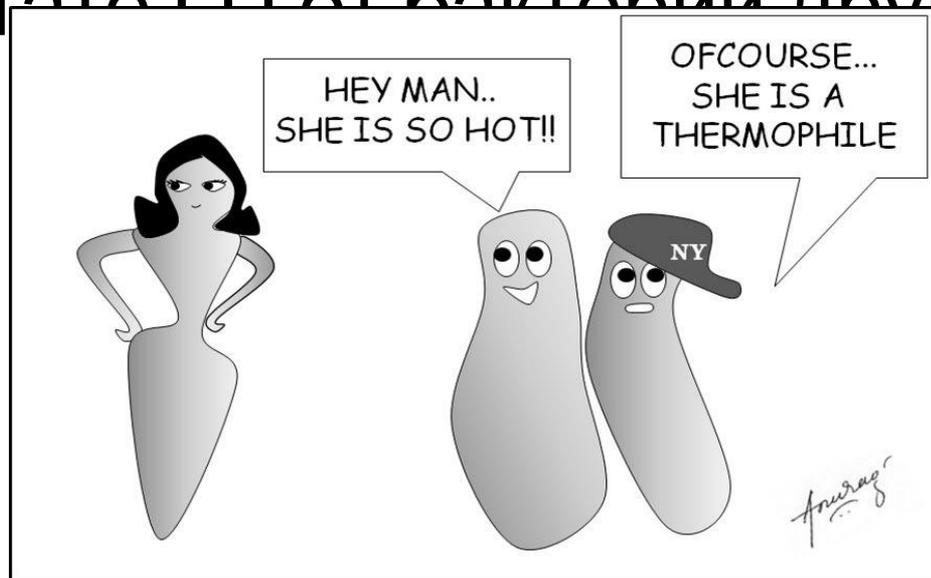
# Скорость эволюции

- При этом человеческий ребёнок приобретает примерно **10 новых вредных мутаций** на весь геном, которых не было у его родителей. И так каждое поколение.
- Скорость мутации (эволюции)  $10^{-8}$  на нуклеотид за поколение (геном человека 7 млрд. нуклеотидов  $2n \times 10^{-8} = 70$  мутаций в геноме у 1 потомка)
- (лекции А. Кондрашова)



- **Горизонтальный перенос** — обмен генетическим материалом — может происходить между таксономически далекими существами.
- Например, *Metanosarcina* — типичная архея, но треть ее генов имеют бактериальное происхождение, и эти гены обслуживают практически весь ее метаболизм, в то время как механизмы транскрипции, трансляции, репликация, устройство мембраны у метаносарцины характерны для архей.

- У термофильной зубактерии *Aquifex aeolicus* – 1512 генов, из них **16%** получены путем горизонтального переноса от архей
- У *E.coli* – из 4289 генов **755** получены в результате ГП от бактерий других видов.



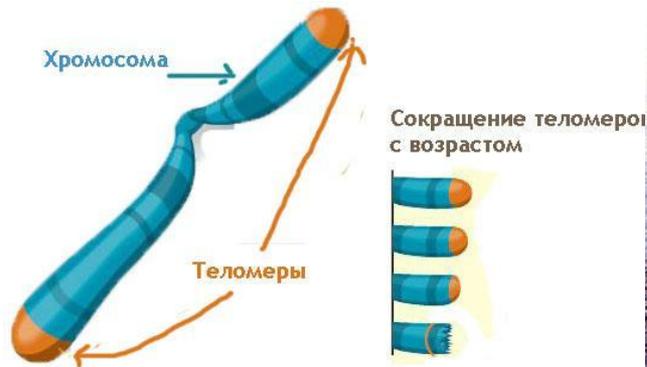
# Изменение генома под действием средового отбора

- Наиболее известным примером изменения генома, индуцированного культурой, является распространение аллеля, обеспечивающего **утилизацию лактозы**, которая по мере развития домашнего молочного скотоводства стала частью рациона человека.
- Причем исследования показали, что молочное животноводство развивалось прежде геномных изменений и создавало селекционное давление в пользу аллеля, обеспечивающего утилизацию лактозы (Burger *et al.*, 2007).



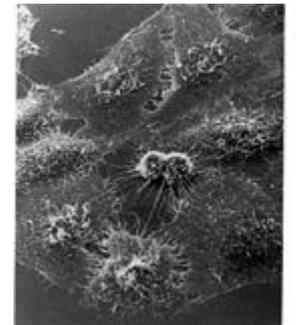
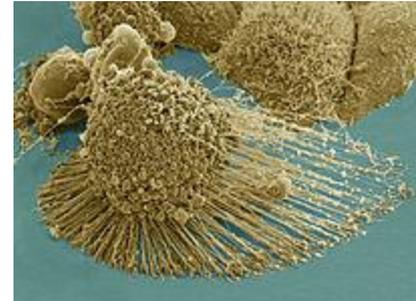
# Старение ДНК

- При каждом делении клеток происходит укорочение длинных повторов на теломерах, соответственно каждая следующая репликация (удвоение) «сморщивает» хромосому и в определенный момент она не может поделиться и умирает.
- Благодаря кольцевому строению ДНК клетки прокариот не стареют в череде поколений.
- У ученых, академиков и в целом умных людей теломеры больше (длиннее).



# Какая женщина дала начало бессмертной линии клеток?

- В 1951 году американке **Генриетте Лакс** был поставлен диагноз рак шейки матки, в этом же году она скончалась. Врачи в госпитале взяли у неё образцы здоровой и поражённой ткани, которые попали к исследователю Джорджу Гею, который обнаружил что, клетки опухоли Лакс были способны выживать и делиться бесчисленное число раз. Они положили начало клеточной линии **HeLa**, которая была использована для создания вакцины против полиомиелита, в исследованиях рака, ВИЧ, воздействия радиации и множестве других испытаний учёными по всему миру.
- Подсчитано, что с тех пор выращено **более 20 тонн клеток HeLa**, что превышает вес Генриетты при жизни в 400 раз.



- Клетки **HeLa** называют «бессмертными», они способны делиться бесконечное число раз, в отличие от обычных клеток, имеющих предел деления Хейфлика (всего около 50 делений, по 3-6 н.о. за деление и теряя 150-300 н.о. с концов хромосом).
- Это происходит потому, что как и при многих типах раковых опухолей, клетки HeLa производят фермент **теломеразу**, которая наращивает теломеры на концах ДНК хромосом.
- Клетки HeLa были с самого начала заражены вирусом папилломы, что часто случается с клетками рака, от которого умерла Генриетта.
- Клетки HeLa обладают **аномальным кариотипом**, различные сублинии HeLa имеют 49 — 78 хромосом, в отличие от нормального кариотипа человека, содержащего 46 хромосом.

# Отнеситесь к этому с юмором!

- клетки HeLa были описаны Леем ван Валеном как пример создания в современности нового биологического вида, *Helacyton gartleri*, названного в честь Стенли М. Гартлера, исследовавшего эти клетки.
- Аргументы за выделение в отдельный вид таковы:
  1. несоответствие числа хромосом у HeLa и людей;
  2. экологическая ниша клеток HeLa;
  3. способность клеток HeLa сохраняться и размножаться за пределами возможного для культур обычных человеческих клеток.
- Это определение вида **не получило широкого распространения** в научном сообществе.

# Образовательные курсы и лекции

- <https://www.coursera.org/>
- <https://stepic.org/>
- <http://postnauka.ru/>
- <http://vk.com/molbio>
- <http://icg.nsc.ru/lectures/>
- И др.

19 <b>K</b> Potassium 39.0983	63 <b>E</b> Europium 151.964	63 <b>E</b> Europium 151.964	15 <b>P</b> Phosphorus 30.973762	
20 <b>Ca</b> Calcium 40.078	116 <b>L</b> Livermorium [293]	25 <b>M</b> Manganese 54.938045		
79 <b>A</b> Gold 196.966569	60 <b>Nd</b> Neodymium 144.242			
105 <b>D</b> Dubnium [268]	8 <b>O</b> Oxygen 15.9994			
21 <b>Sc</b> Scandium 44.955912	53 <b>I</b> Iodine 126.90447	63 <b>E</b> Europium 151.964	7 <b>N</b> Nitrogen 14.0067	58 <b>Ce</b> Cerium 140.116

# Сайт НОЦ

- <http://genome.sfu-kras.ru/>

The image shows the header of the website for the Genomics Center at SFU. At the top, there is a navigation bar with links: "СФУ | Поступление | Наука | Обучение | Мой университет | Спорт | Сайты СФУ" and a search bar labeled "Поиск". Below this is a blue banner with the text "Центр геномных исследований" and "Научно-образовательный центр СФУ". To the left of the text are images of a DNA double helix and molecular models, along with two blocks of DNA sequence code. To the right is a logo featuring a stylized orange and blue 'X' shape. Below the banner, there is a grey bar with a language selector showing a UK flag and the word "English". On the far left, a vertical menu lists "Новости", "Главная", and "Лаборатория лесной".

СФУ | Поступление | Наука | Обучение | Мой университет | Спорт | Сайты СФУ

Поиск

Центр геномных исследований  
Научно-образовательный центр СФУ

English

Новости  
Главная  
Лаборатория лесной

# Спасибо за внимание!

