



# Долгожительство

- Ранее считалось, что рекорд долгожительства принадлежит **сосне остистой** (*Pinus longaeva*), произрастающей в лесах Северной Америки. Дереву, названному в честь ветхозаветного патриарха Мафусаила, **почти 5000 лет**. Это значит, что проросло оно примерно 30 веков до нашей эры. А вот в Австралии ныне живёт и здравствует **эвкалипт**, который, конечно, не относится к хвойным, но зато он отпраздновал своё **13000-летие!** Его называли деревом ледникового периода, в мире сохранилось всего пять



# Сравнительный размер генома

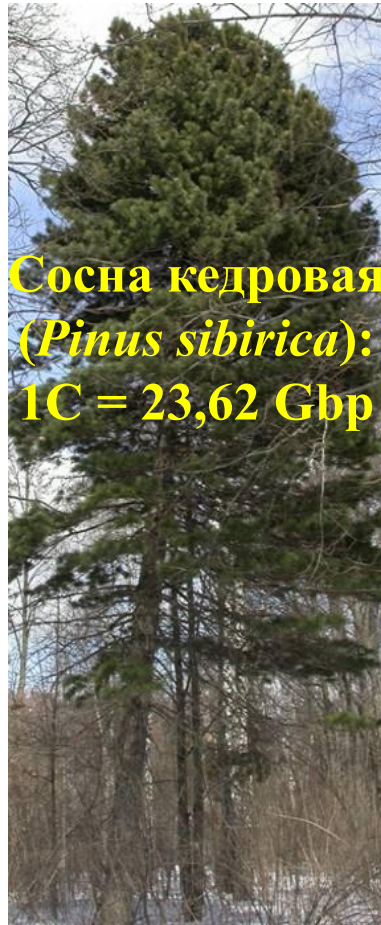
**Сосна сахарная**  
*(Pinus lambertiana):*

**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**



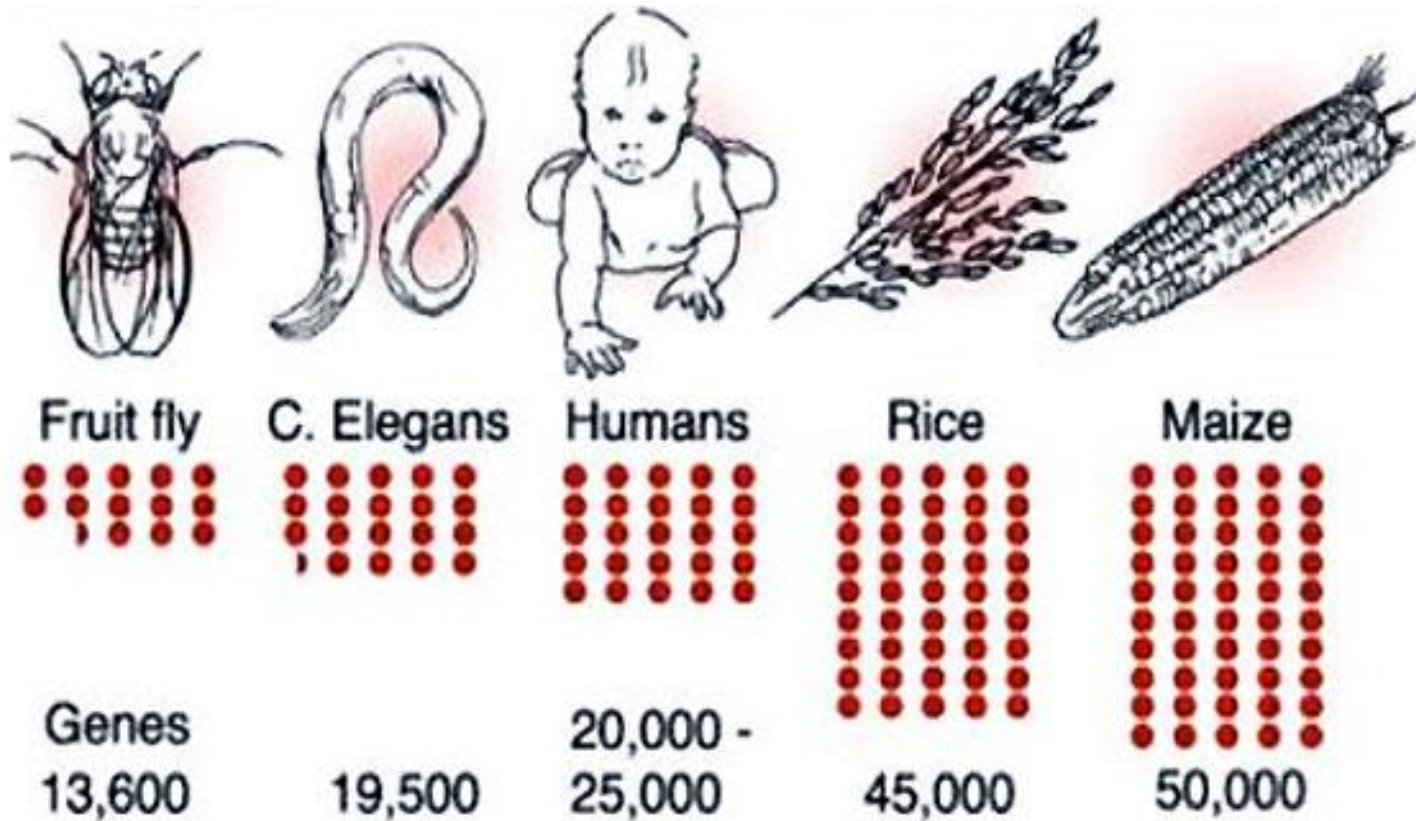
• Геном человека, *Homo sapiens*:  
**1C = 3,20 Gbp**



## 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>Coccolodiscus 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)	$1.24 \times 10^{11}$	$2x = 24$

# По количеству генов



# Сложность строения генома

- Чем сложнее экзонно-интронная структура генов, тем сложнее организм.
- Если у него только один кодирующий элемент (экзон) в гене, то организм относительно прост (бактерии), если несколько — это повышает вариативность образующихся белков, из которых, как из кирпичиков, построено всё живое.
- Поэтому 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».

# Размеры генов

- Средний размер гена в хромосоме приходится около 50 тысяч пар нуклеотидов.
- Самые короткие гены содержат всего два десятка букв-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 букв.

- У примитивных организмов, таких как бактерии, гены занимают около **80-90% всей ДНК**.
- У человека на гены приходится, по-видимому, **не более 5%** нуклеотидных последовательностей. Остальную ДНК раньше называли избыточной, но со временем стало ясно, что она выполняет важные функции, в том числе содержит информацию о том, как, в каком порядке должны включаться гены.
- В начале и в конце гена находятся регуляторные последовательности, которые определяют в каких тканях, на каких этапах развития и при каких внешних или внутренних (например, гормональных) сигналах будет работать данный ген.



# 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 в молекуле нуклеиновой кислоты, 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**.

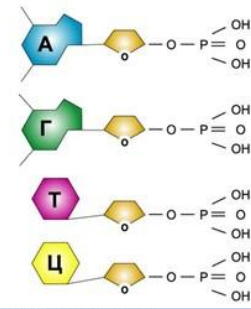
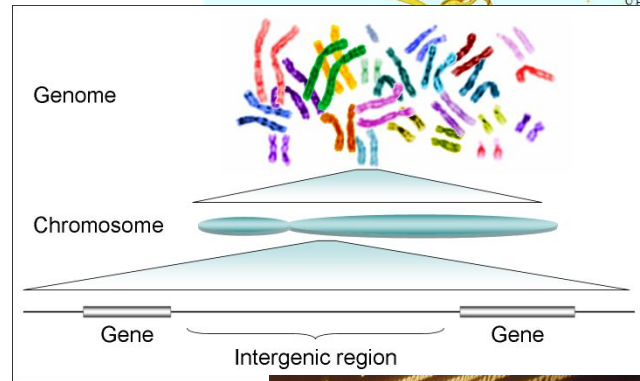
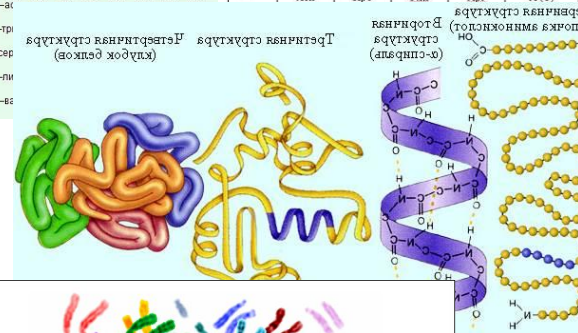


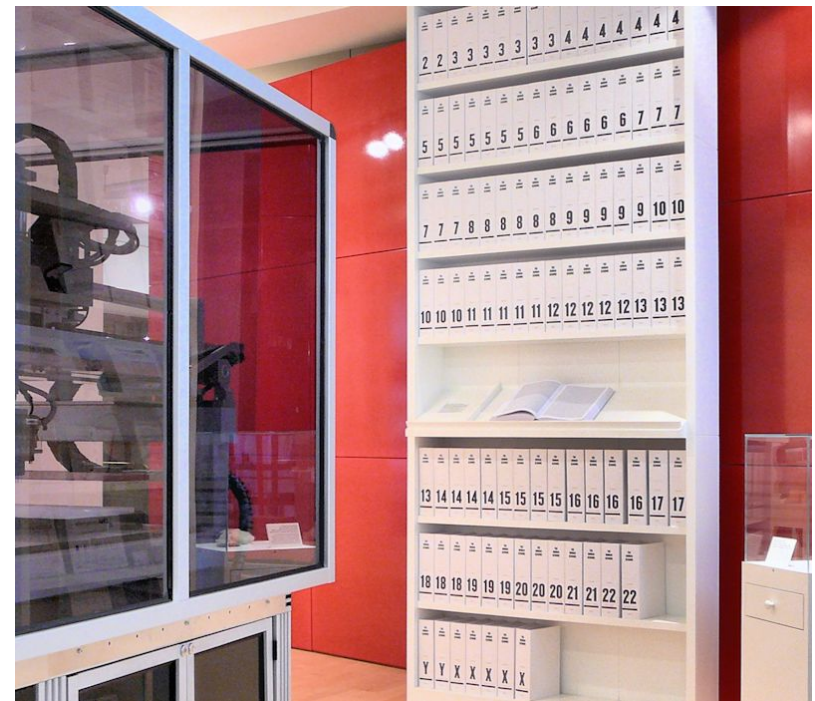
Таблица генетического кода

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	Гис	Арг	А(Т)

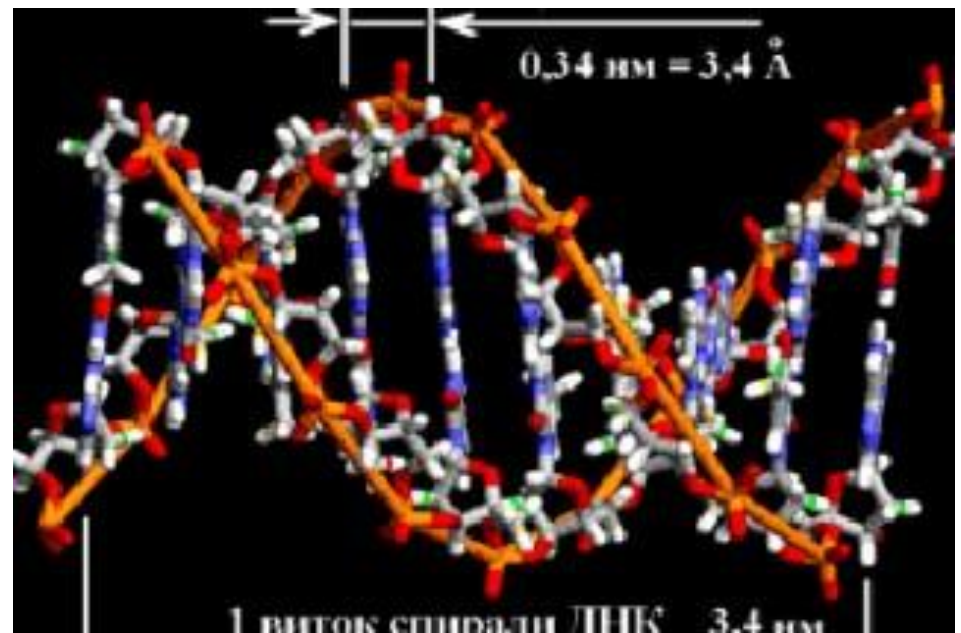
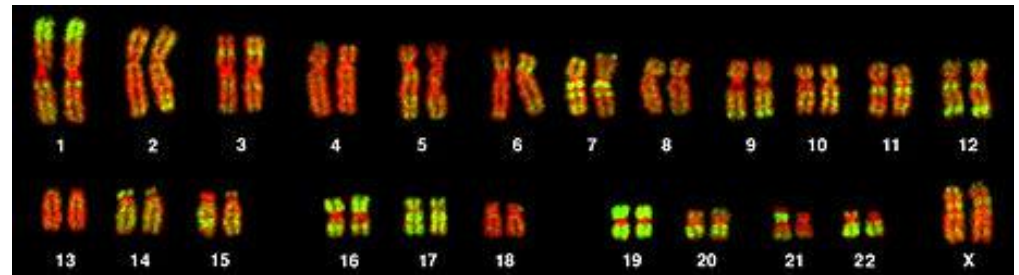
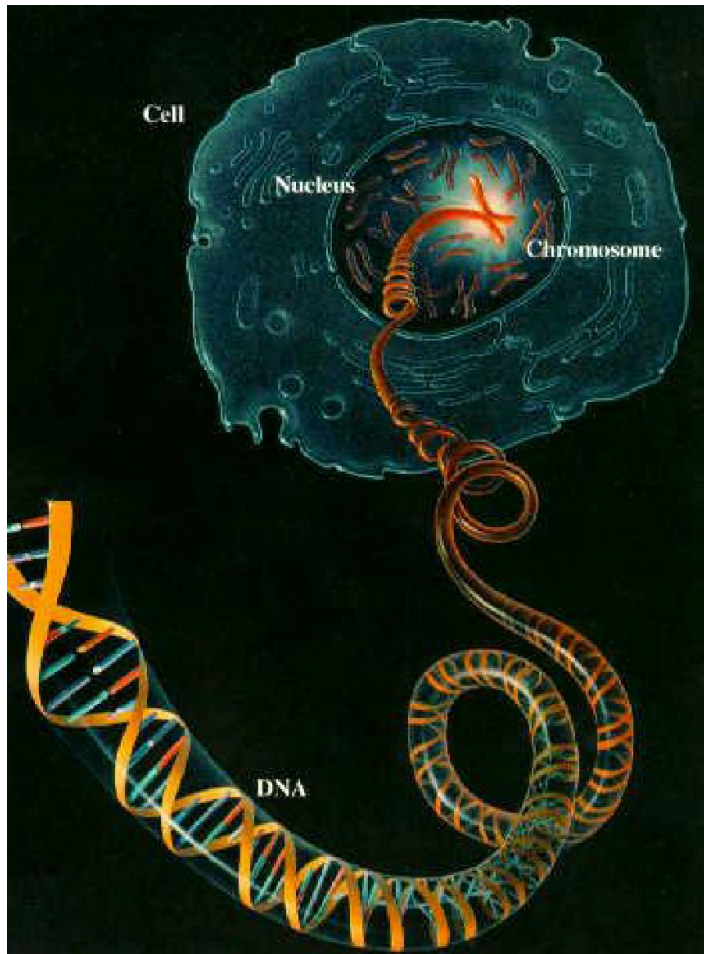


ATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCT  
TCGTTTCGCCCGCGCCGCTATACTTGGTAGATCAGAACTCAGGAGTGTCTTTT  
TTAATGCTAATTGGGAAAAAGACAGAACCAATCAAAGTGGTGTGAACCGTGT  
ATGGTGACAAAGATAAACGGCGGCATTGTTTTGCTACCTGGAAGAATATTTCTGG  
TTCCATTGAATAGTGAAACAAGGTTGTTGGCTGGATGATATCAACTGCTATGACA  
GGACTGATTGTGTAGAAAAAAGACAGCCCTGAAGTATATTTCTGTGTGCTGTA  
GGCAATATGTGTAATGAAAAGTTTTCTTATTTCCGGAGATGGAAGTCACACAG  
CCCACCTCAAATCCAGTTACACCTAAGCCACCCACCGTGGTGGAACTCACACAT  
GCCCACCGTGCCAGCACCTGAACCTCTGGGGGACCGTCAGTCTTCTCTTCCC  
CCCAAAACCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTG  
GTGGTGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGAC  
GGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAG  
CAGTACCGTGTGGTCAGCGTCTCACCGTCTGCACCAGGACTGGCTGAATGGC  
AAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGTCCCCATCGAGAAA  
ACCATCTCAAAGCCAAAGGGCAGCCCGAGAACACAGGTGTACACCCTGCC

- Если обратиться к роману Льва ТОЛСТОГО «Война и мир», то окажется, что писатель использовал примерно **три миллиона** букв. А у человека **три миллиарда** таких «букв» — нуклеотидов в геноме. То есть каждый из нас носит в себе условную **тысячу томов** известного романа. А в случае лиственницы сибирской это **четыре тысячи томов!**

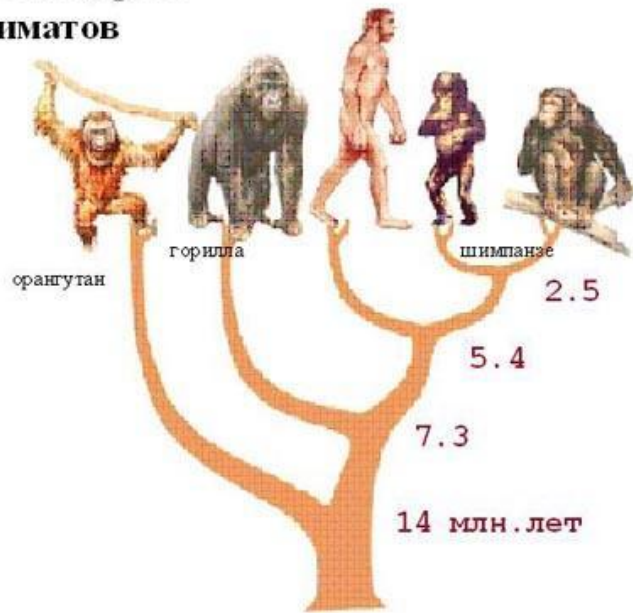


# Длина генома одной клетки человека (23 пары хромосом) около 2 м



# Все люди на 99% по составу ДНК одинаковые

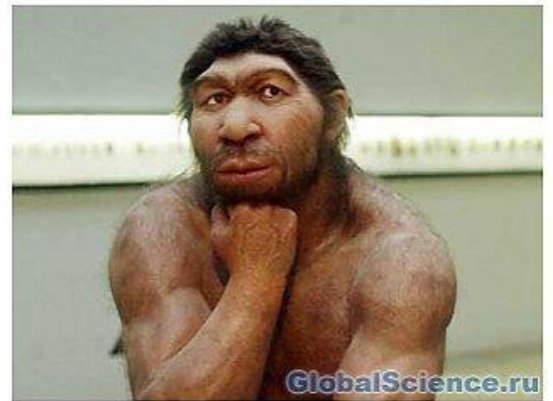
Филогенетическое древо  
высших приматов



Генетические различия на уровне ДНК

между людьми: 1 нуклеотид из 1000

между человеком и шимпанзе : 1 нукл. из 100

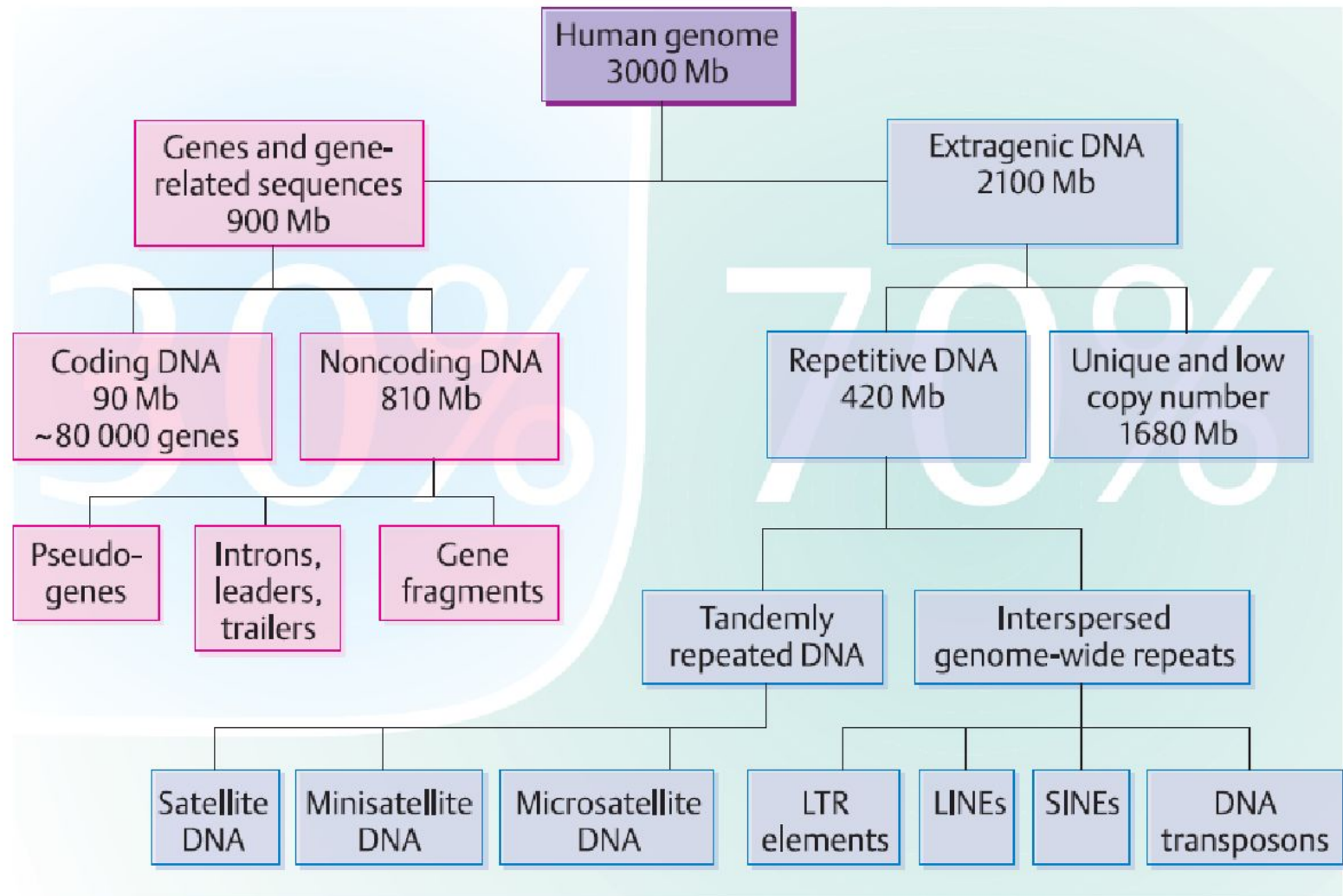


- В каждом европейце 1-4% генома неандертальца при совпадении геномов на 99,5% (человек-неандерталец)

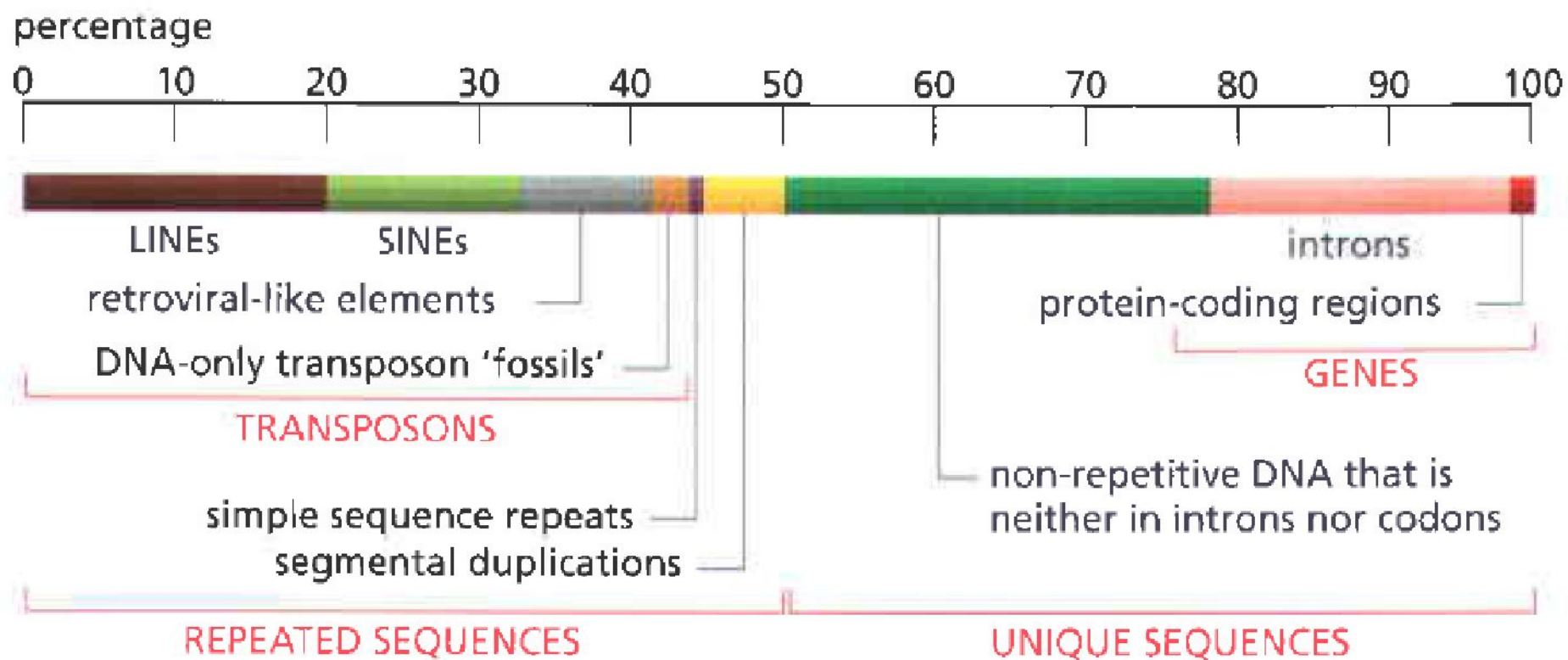
# Варианты геномов

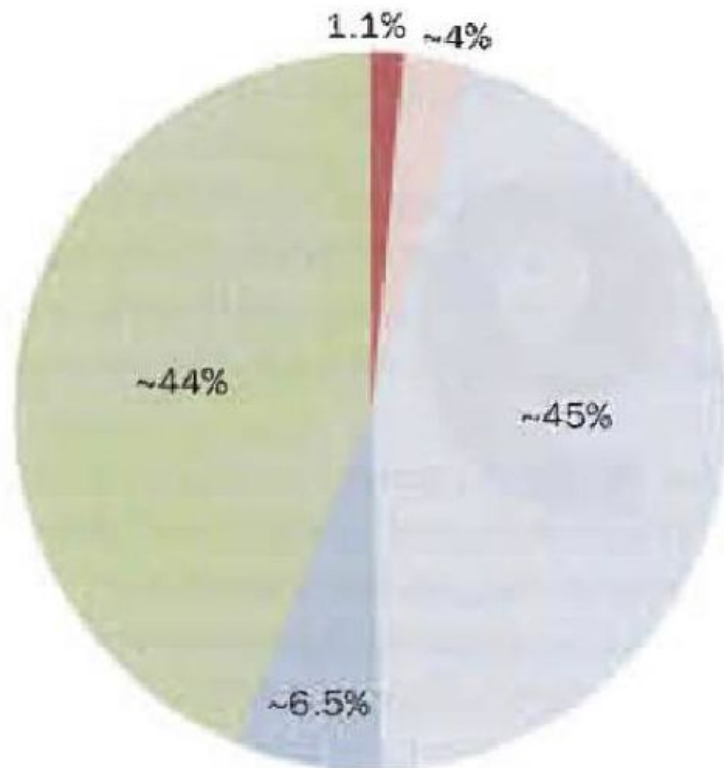


# Компоненты генома человека



## A. The components of the nuclear human genome

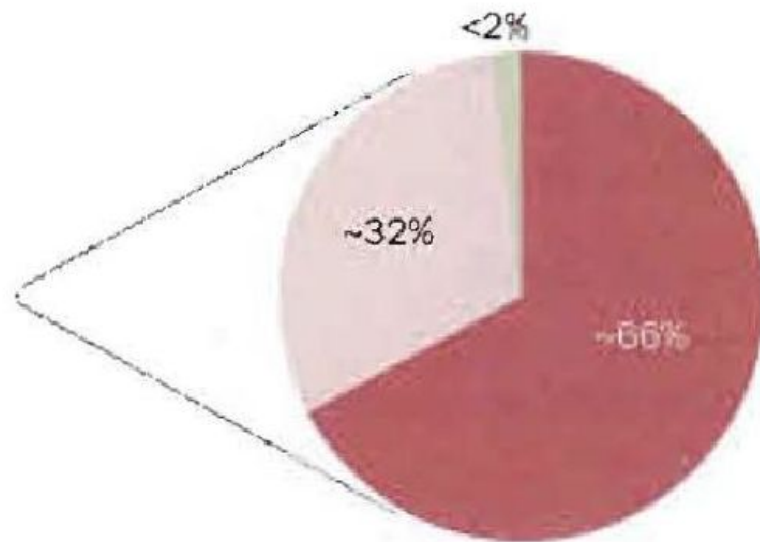




nuclear genome

highly conserved sequences

- protein-coding genes
- RNA genes, regulatory sequences



mitochondrial genome

poorly conserved sequences

- transposon-based repeats
- heterochromatin
- other sequences



# Особенности выделения ДНК из клеток растений!

- При экстракции ДНК из растительных объектов необходимо не только деактивировать клеточные ферменты, но и «удалить» запасные вещества, например, полисахариды и вторичные метаболиты, такие как алкалоиды, фенольные соединения, терпены, которые не просто мешают изолированию ДНК, но и отрицательно влияют на ее качество.

# Можно попробовать самим выделить свою ДНК

- 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)
- Ну или выделить ДНК лука  
<http://www.youtube.com/watch?v=ooZD3U62o50>

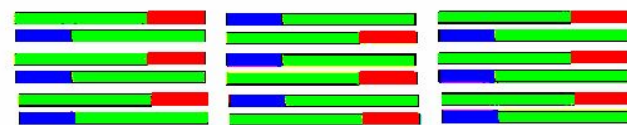
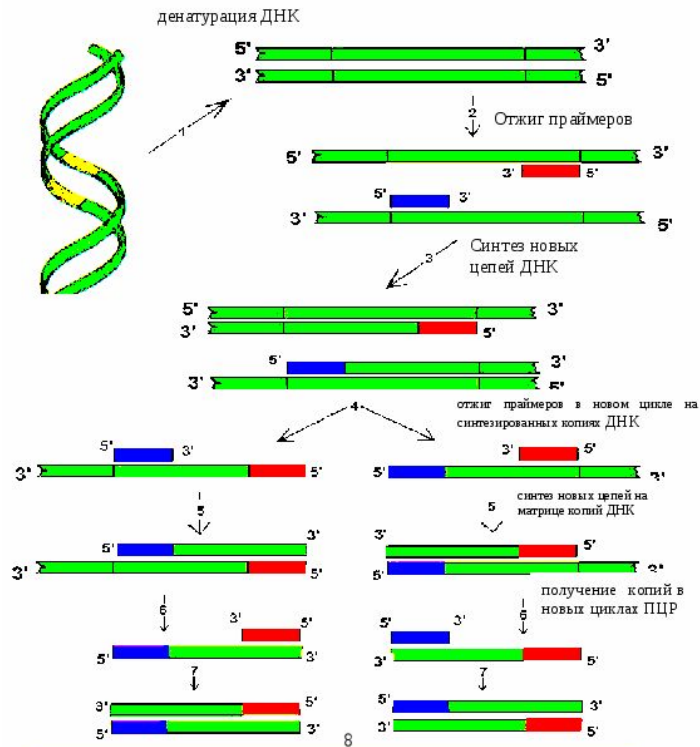
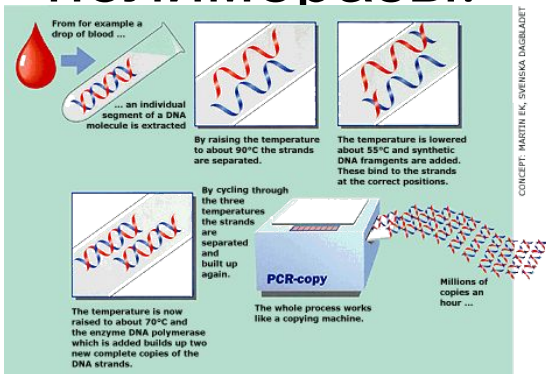
# Основные этапы выделения ДНК



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

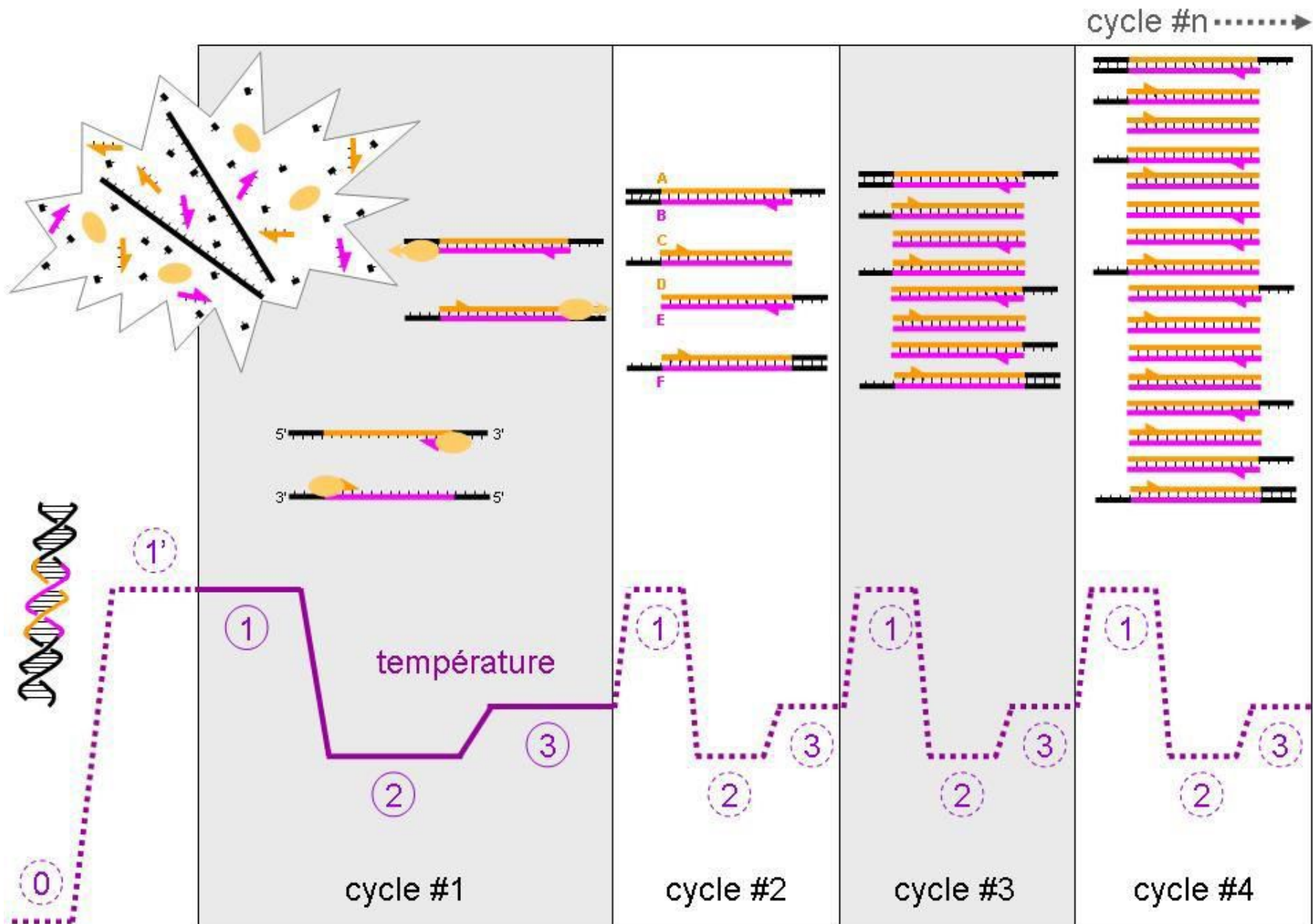
# Полимеразная цепная реакция

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

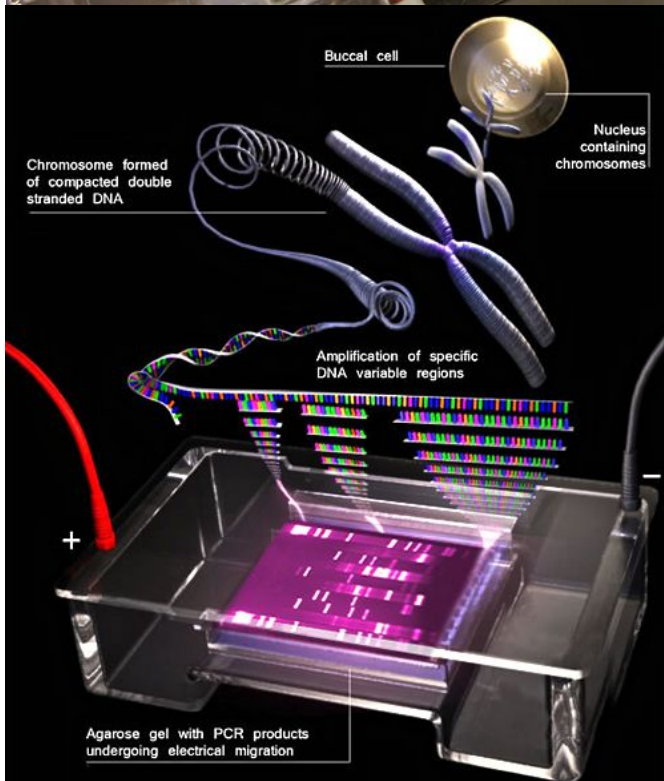
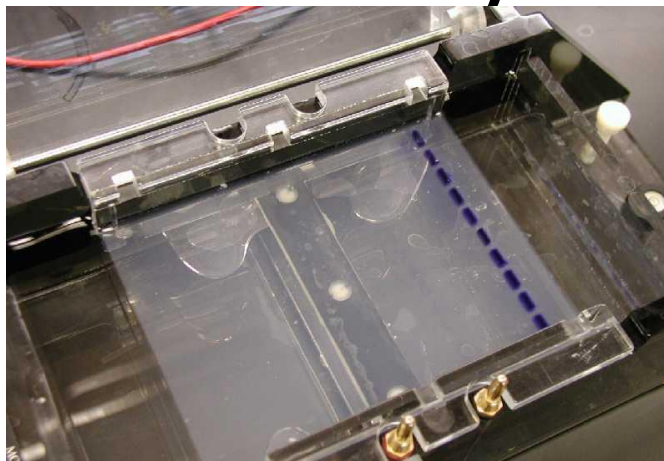


1 – денатурация двойных цепей ДНК; 2 – отжиг праймеров; 3 – синтез новых цепей на основе первоначальной матрицы ДНК; 4 – отжиг праймеров на синтезированных копиях в новом цикле; 5 – синтез новых цепей на матрице синтезированной копии ДНК; 6, 7 – в новом цикле отжиг праймеров и синтез копий на матрице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

- ПЦР очень мощный метод размножения фрагментов ДНК вне организмов (*in vitro*). С момента его открытия в 1985 году, этот метод используется сейчас практически во всех молекулярно-биологических лабораториях.
- Через 8 лет после этого за изобретение метода ПЦР Кэри Муллис получил Нобелевскую премию.
- После гибридизации матрицы с праймером (отжиг), последний служит затравкой для ДНК-полимеразы при синтезе комплементарной цепи матрицы.
- Важнейшая характеристика праймеров—температура плавления

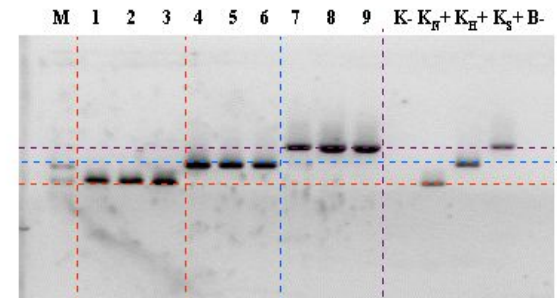


# Визуализация и контроль

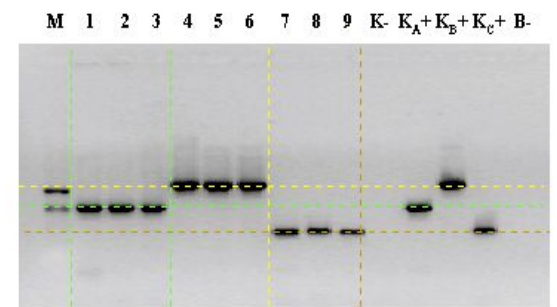


Для электрофоретического анализа ДНК обычно используют агарозные (для относительно длинных молекул ДНК) и полиакриламидные (для высокого разрешения коротких молекул ДНК, например, в случае секвенирования) гели.

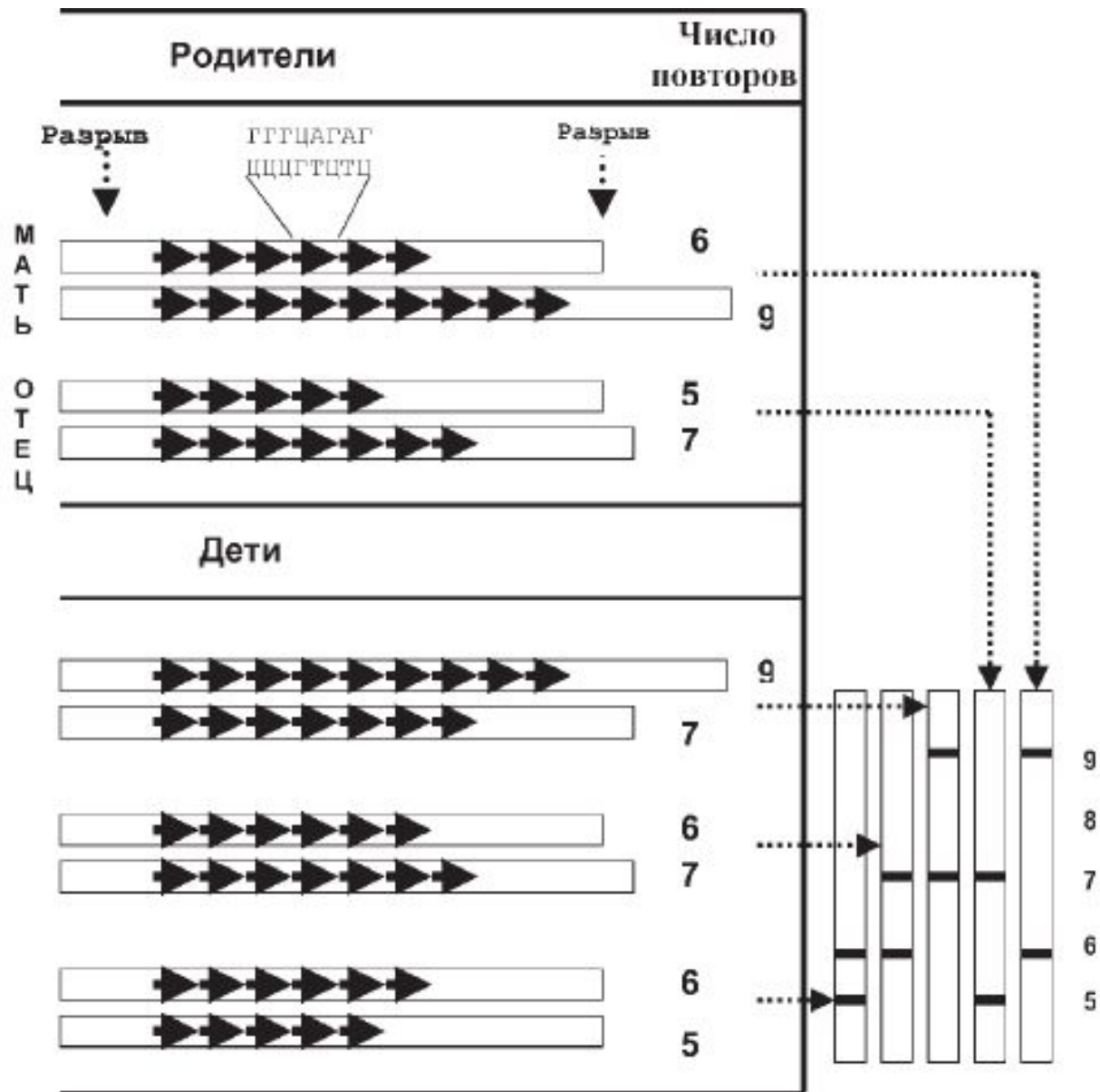
Рис. 1. Внешний вид результатов электрофореза продуктов ПЦР после амплификации с родоспецифичными (А) и серогрупповыми (В) праймерами. Направление электрофореза - снизу вверх (на рисунке)



Дорожки 1-9 - образцы СМЖ от 9 больных ГБМ  
 К- отрицательный контроль стадии 2 (ПЦР)  
 V- отрицательный контроль стадии 1 (выделение ДНК)  
 $K_N^+$ ,  $K_N^+$ ,  $K_S^+$  = ДНК *N.meningitidis*, *H.influenzae*, *S.pneumoniae*, то есть положительный контроль стадии 2 (ПЦР)  
 M - маркер длины фрагментов ДНК

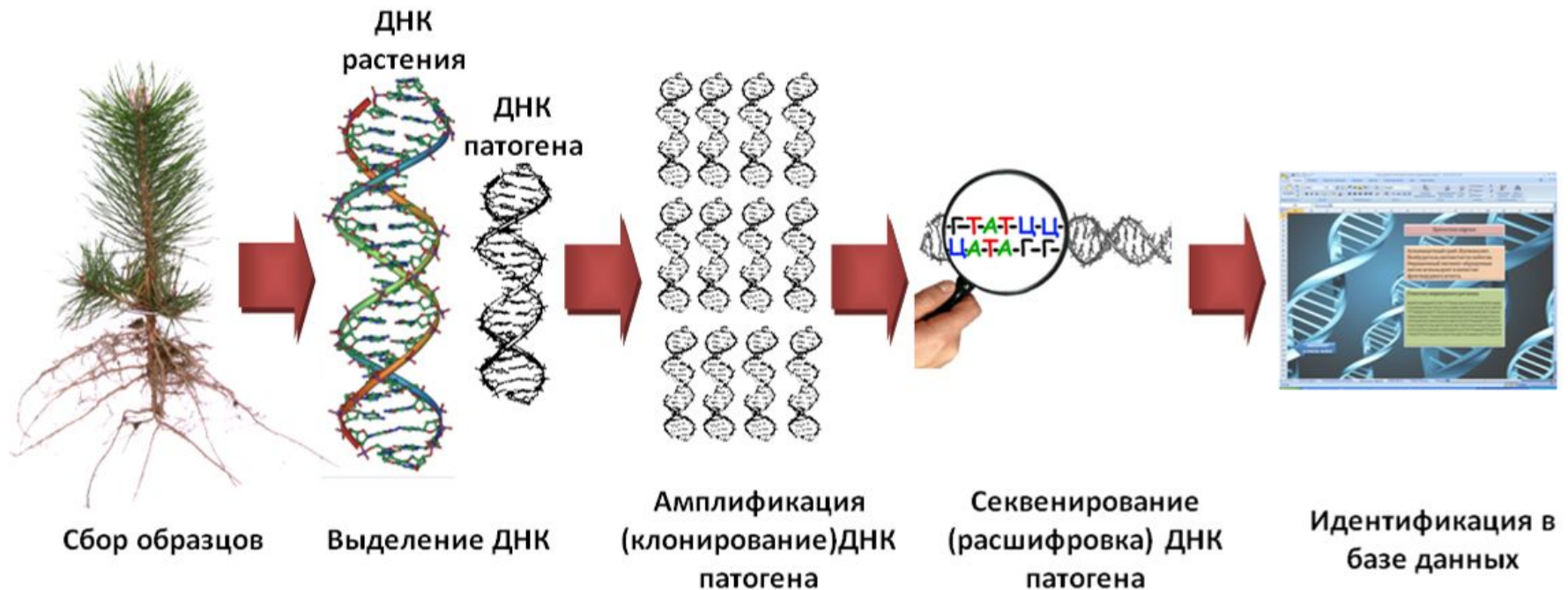


Дорожки 1-9 - образцы СМЖ от 9 больных менингококковым менингитом  
 К- отрицательный контроль стадии 2 (ПЦР)  
 V- отрицательный контроль стадии 1 (выделение ДНК)  
 $K_A^+$ ,  $K_B^+$ ,  $K_C^+$  = ДНК *N.meningitidis* серогрупп А, В, С, то есть положительный контроль стадии 2 (ПЦР)  
 M - маркер длины фрагментов ДНК





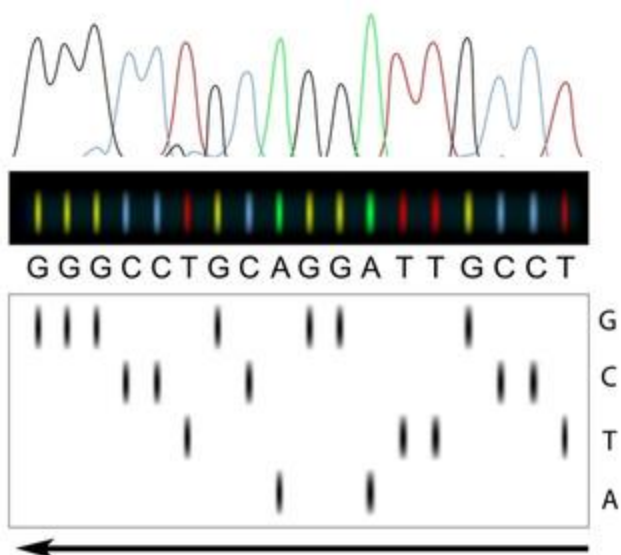
- **Секвенирование** биополимеров (белков и нуклеиновых кислот — ДНК и РНК) — определение их аминокислотной или нуклеотидной последовательности (от лат. *sequentum* — последовательность)



# 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?**

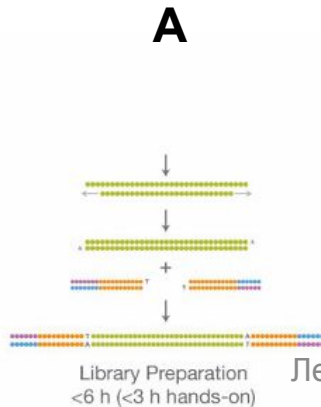
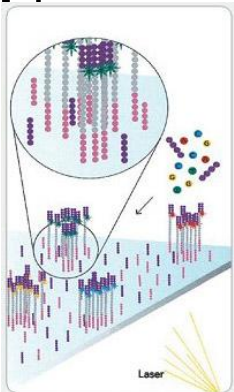
- Дидезоксинуклеотидный метод, или метод «обрыва цепи», был разработан **Ф. Сенгером в 1997** году и в настоящее время широко используется для определения нуклеотидной последовательности ДНК.



Лекция "Ребятам о геномах" 9 января  
2015 г. СФУ Дейч К.О.

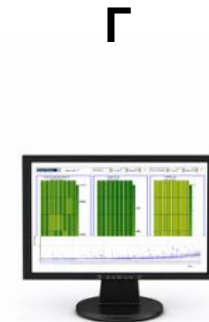
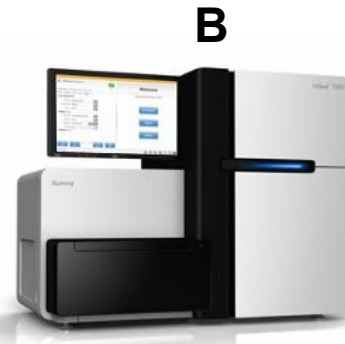
# Секвенирование с помощью HiSeq2000

- Основные этапы секвенирования включают:
- приготовление геномной библиотеки, состоящей из фрагментов ДНК нужного размера (**А**);
- амплифицирование фрагментов и создание из них кластеров для синтеза ДНК, используя автоматизированные системы типа cBot для создания кластеров (**Б**);
- секвенирование путём синтеза (**В**);
- компьютерный анализ полученных изображений и генерирования из них индивидуальных сиквенсов, используя Sequencing Control Software (SCS) и Consensus Assessment of Sequence and Variation (CASAVA) программы, которые анализируют цветовые сигналы и конвертируют графические изображения в нуклеотидные последовательности с оценкой их качества и вывода в виде текстовых файлов для дальнейшего анализа (**Г**).



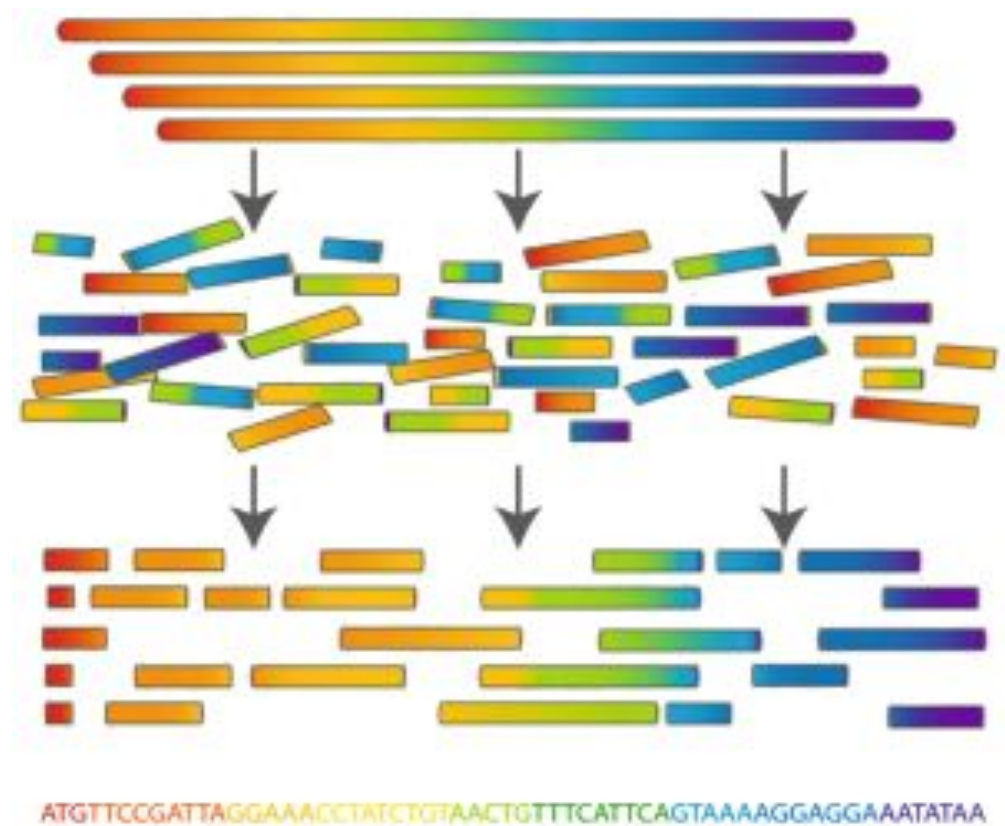
Cluster Generation  
<4 h (<10 min hands-on)

Лекция "Решаем о геномах" 9 января 2015г. СФУ ДИЧ К.О.

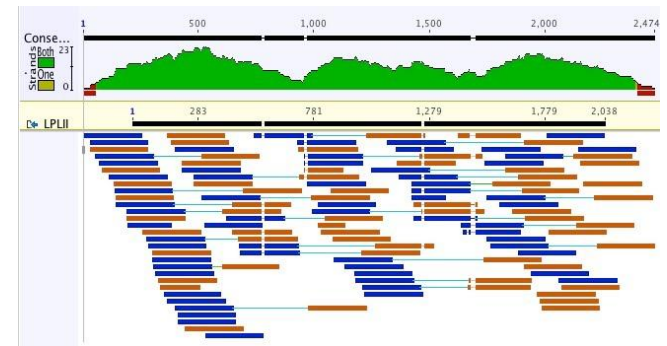
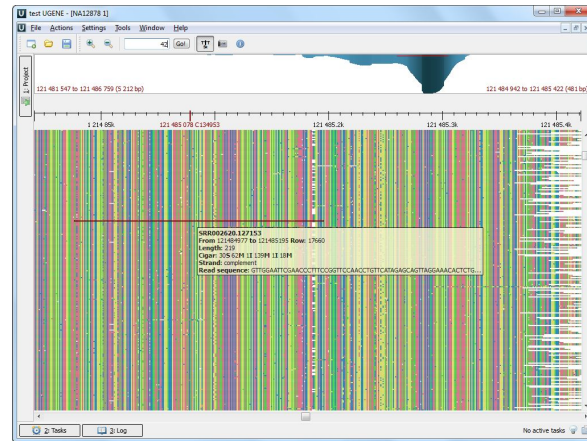
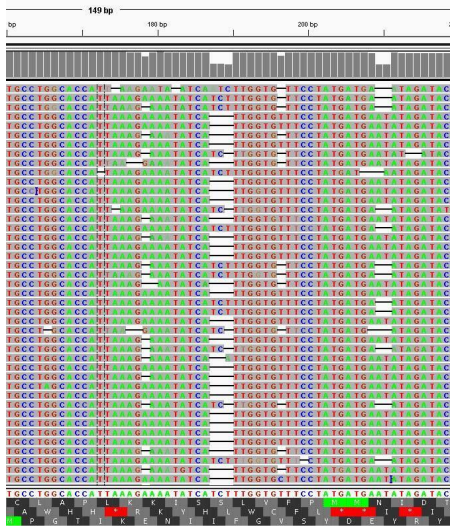
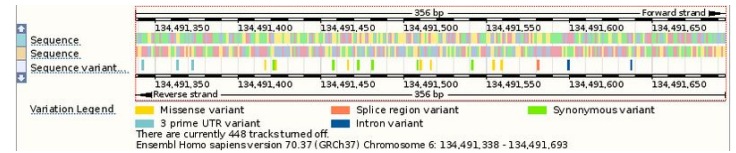
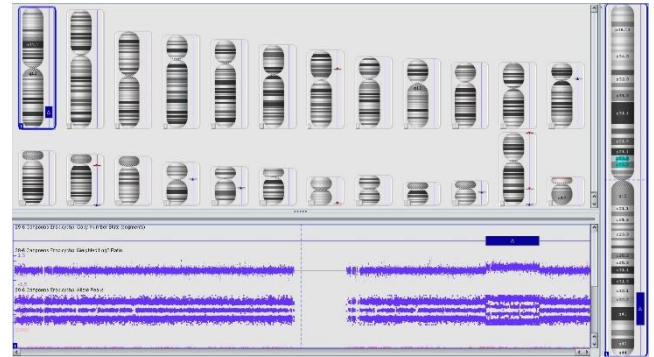
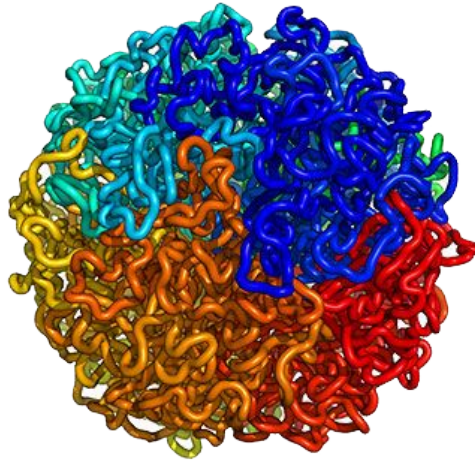
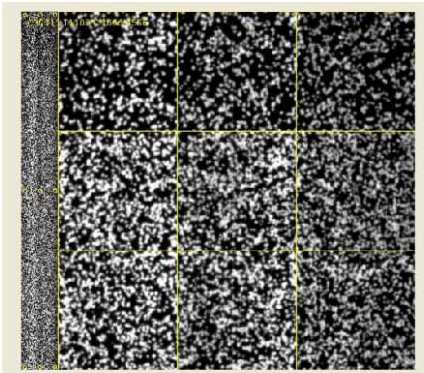


RTA v1.7, CASAVA v1.7  
2 days (30 min hands-on)

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

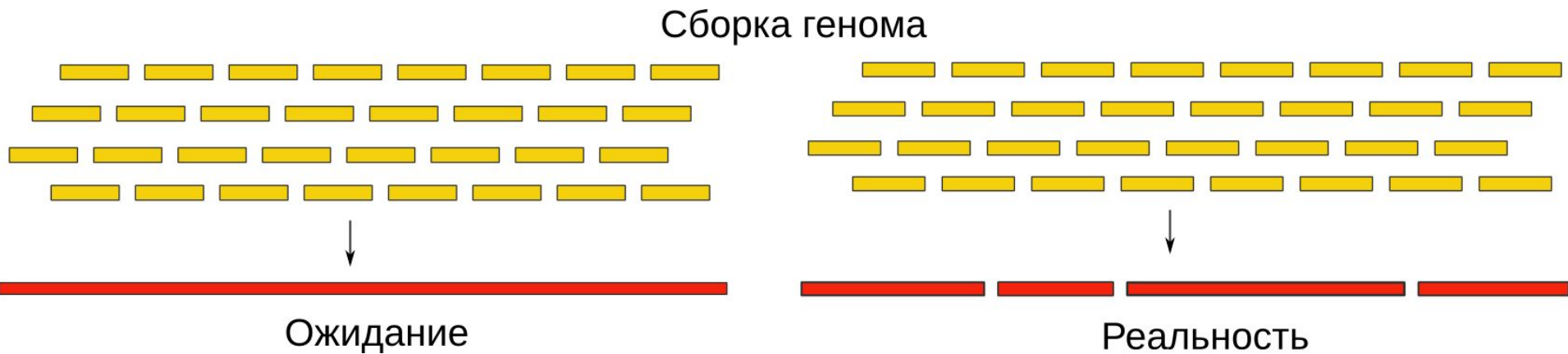


# Обработка данных



# Идеальной сборке мешает:

- большое количество повторов в геномах
- несовершенство программ сборки
- ошибки ПЦР и секвенирования



Референсная  
последовательность ДНК

короткие прочтения

```
AAACGTGAACTCGTCGCAGAAGCAGCATTTCCTCAAGCCCATAGCGCTTGCAATTGTAATTGTTGCCACAAGCAGGCCAGCCAAATC
ACGTGAACTCGTCGGAGAAGCAGCA TTCCAAGCCCATAGCGCTTGTAATTG TGTTGGCACAAGCAGGCCAGCCAAA
CGTGAACTCGTCGCAGAAGCAGCAT TCCAAGCCCATAGCGCTTGTAATTGT GTTGGCACAAGCAGGCCAGCCTAAT
GTGAACTCGTCGCAGAAGCAGCATT CCAAGCCCATAGCGCTTGTAATTGTA TTGGCACAAGCAGGCCAGCCTAATC
TGAACTCGTCGCAGAAGCAGCATT CCAAGCCCATAGCGCTTGTAATTGTA
TGAACTCGTCGGAGAAGCAGCATT CAAGCCCATAGCGCTTGTAATTGTAA
TGAACTCGTCGGAGAAGCAGCATT CGCTTGTAATTGTAATTGTTGGCAC
TGAACTCGTCGGAGAAGCAGCATT GCTTGTAATTGTAATTGTTGGCACA
GAACTCGTCGCAGAAGCAGCATT GCTTGTAATTGTAATTGTTGGCACA
AACTCGTCGGAGAAGCAGCATTCC GCTTGTAATTGTAATTGTTGGCACA
ACTCGTCGCAGAAGCAGCATTTC CTTGTAATTGTAATTGTTGGCACAA
CTCGTCGGAGAAGCAGCATTACCA TTGTAATTGTAATTGTTGGCACAAG
```

↑  
SNP  
(гетерозиготный)

↑  
ошибка  
секвенирования

↑  
SNP  
(гомозиготные)

↑  
SNP  
ошибка  
секвенирования



- АТ- богатые регионы – промоторы
- GC- богатые регионы – шпильки
- GC состав больше 40% (42-48%)- митохондриальная ДНК
- Если в организме экспрессируется много повторов (в транскриптом) – это первая реакция на развитие рака
- Вставки ретровирусов в геном антисмысловые

- 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 website interface. At the top, there is a navigation bar with 'NCBI Resources' and 'How To' links, and a 'Sign in to NCBI' button. Below this is a search bar with a dropdown menu set to 'All Databases' and a 'Search' button. The main content area is divided into three columns. The left column contains a vertical navigation menu with items like 'NCBI Home', 'Resource List (A-Z)', 'All Resources', 'Chemicals & Bioassays', 'Data & Software', 'DNA & RNA', 'Domains & Structures', 'Genes & Expression', 'Genetics & Medicine', 'Genomes & Maps', 'Homology', 'Literature', 'Proteins', 'Sequence Analysis', 'Taxonomy', 'Training & Tutorials', and 'Variation'. The middle column features a 'Welcome to NCBI' section with a brief description of the center's mission and links to 'About the NCBI', 'Mission', 'Organization', 'Research', and 'NCBI News'. Below this is a 'Get Started' section with a list of links: 'Tools', 'Downloads', 'How To's', and 'Submissions'. A featured article titled 'Genotypes and Phenotypes' is also visible, with a sub-headline 'Data from Genome Wide Association studies that link genes and diseases. See study variables, protocols, and analysis.' The right column contains 'Popular Resources' with links to PubMed, Bookshelf, PubMed Central, PubMed Health, BLAST, Nucleotide, Genome, SNP, Gene, Protein, and PubChem. At the bottom of the right column is an 'NCBI Announcements' section with a notice about GenBank release 205.0 being available via FTP, dated Dec 16, 2014.

# 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  
GTCCCGGGCАТСТGTCАТТGTACCTGCAATGATC  
GGTCАТАСААТТGCTGTTCAАТААТGGGAGGGAA  
САТТТАСССАТТТАТGТААСAGATCGTATGGTAG  
АССАСАААТТGGGGGAАТТТGCACСТАСТСТАСТ  
ТТТТСАGGGACATGCGAGAAACGATAAGAAATC  
TCGTCGTТАА

# Пример первого участка

**complement(1..1062) gene=«psbA»**

**product="photosystem II protein D1**

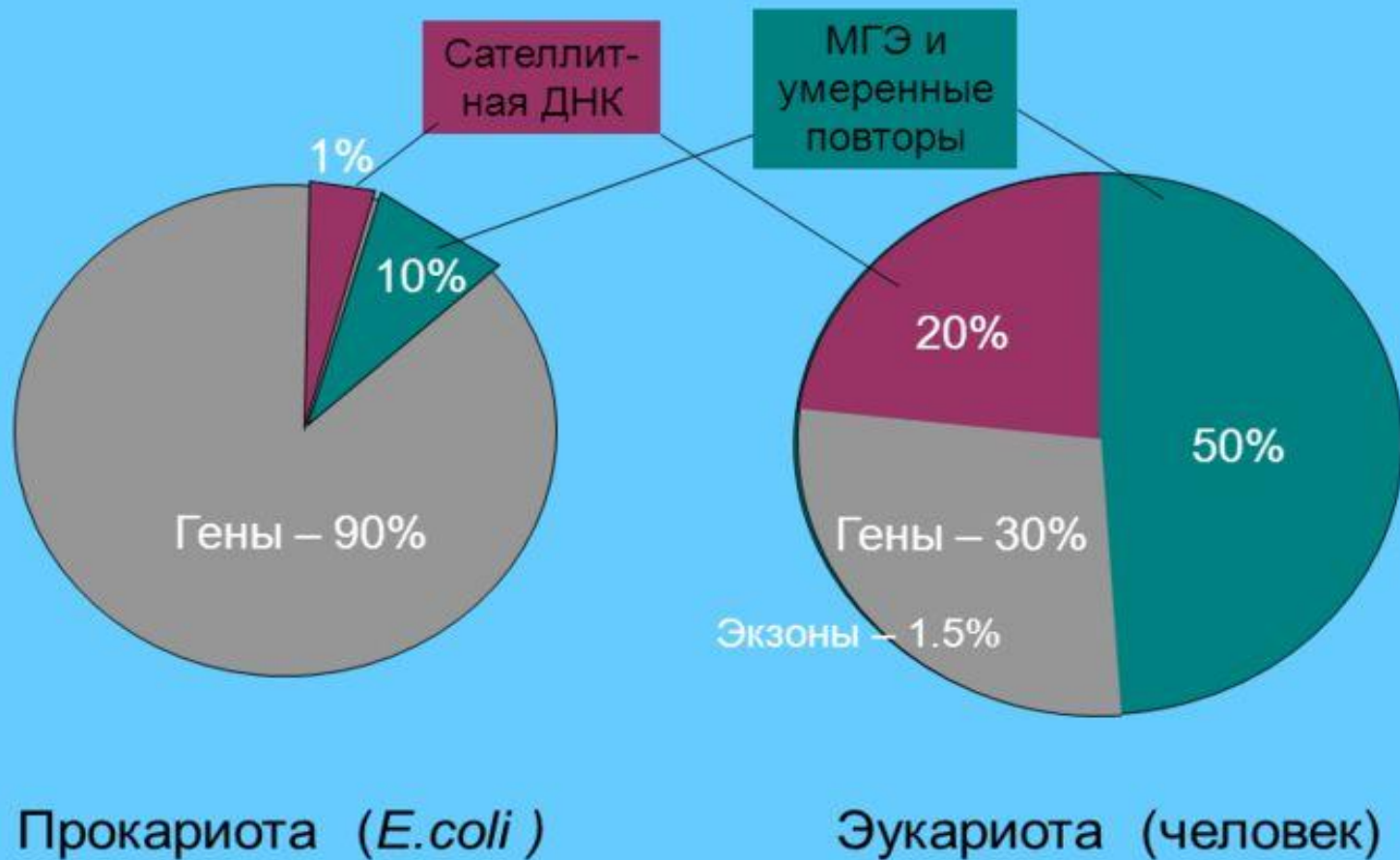
- /translation="MTAIIERRESANLWGRFCDWITSTENR  
LYIGWFGVLMIPTLLTASVFIIAFIAAPPVDIDGIREPVSG  
SLLYGNNIISGAIPTSAAGLHFYPIWEAASVDEWLYNG  
GPYELIVLHFLLGVACYMGREWELSFRLGMRPWIAVA  
YSAPVAAASAVFLIYPIGQGSFSDGMPLGISGTFNFMIV  
FQAEHNILMHPFHMLGVAGVFGGSLFSAMHGSLVTSS  
LIRETTENQSANAGYKFGQEEETYNIVAAGHYFGRLIFQ  
YASFNNSRSLHFFLAAWPVAGIWFTALGISTMAFNLN  
GFNFNQSVVDSQGRVINTWADIINRANLGMMEVMHER  
NAHNFPLDLAAVESISIGG"

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)		: GKRIELDELTSNVPKTRAGGQ	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)	: IDIEINPADLRIDVYRASGAGGQ	HVNKTESAVRITHMP	---	: 267	
	2 Helicobacter pylori (pro)	: IDIEIDEKDVRYDYRSGAGGQ	HVNKTESAVRITHFP	---	: 266	
	2 Mycobacterium tuberculosis (pro)	: DHIDIPEGDVYRVDVYRSGGPGGQ	SVNTTDSAVRLTHIP	---	: 275	
	2 Salmonella typhimurium (pro)	: IDIDINPADLRIDVYRASGAGGQ	HVNRTESAVRITHIP	---	: 267	
	2 Streptomyces coelicolor (pro)	: DHIEIDSELRVDVYRSGGPGGQ	GVNTTDSAVRLTHIP	---	: 266	
	Mitoch.	mit1 Saccharomyces cerevisiae	: YERTFKPEIRVDIMRASGKGGQ	HVNTTDSAVRLTHIP	---	: 302
		mit1 C. elegans	: VSVVVPDSVKDEAMRASGPGGQ	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

- Индивидуальная идентификация организмов;
- Установление родства между организмами;
- Поиск участков (маркеров) ответственных за болезни;
- Картирование (поиск местоположения) генов на хромосомах;
- Изучение эволюции и коэволюции генов, организмов, целых популяций;
- Определение однонуклеотидного полиморфизма (проект 1000 геномов);
- Предсказание функции генов (точнее, кодируемых ими белков) для разработки новых препаратов;
- Предсказание структуры белков и РНК, и т.п. ;
- Регуляция работы гена (в каких условиях он включается или выключается), с какими другими генами он взаимодействует и т.д.

# Геном прокариот и эукариот





# ОСНОВНЫЕ КЛАССЫ ПОВТОРОВ В ГЕНОМЕ ЧЕЛОВЕКА

## Диспергированные повторы



## Тандемные повторы



## Обращенные повторы



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

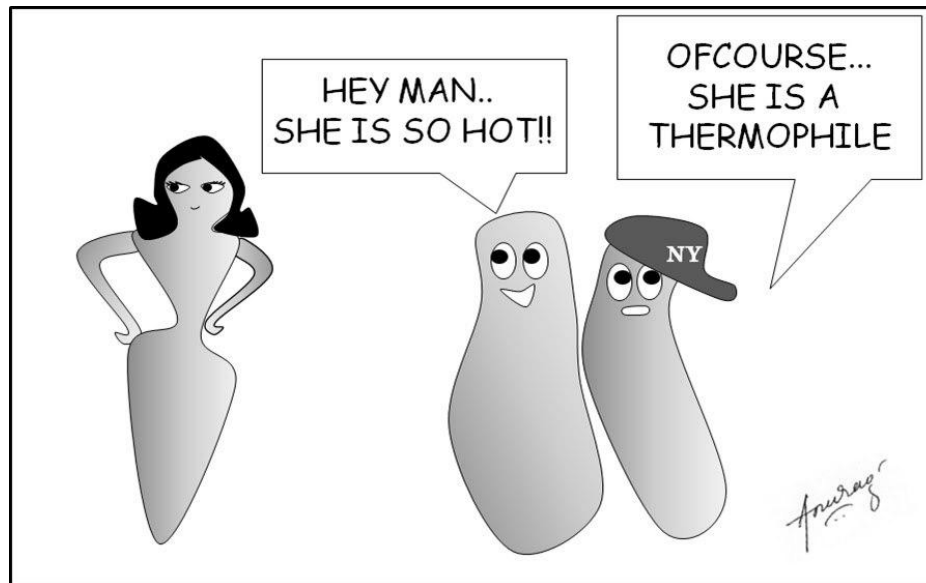
# «Митохондриальная Ева»

- 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.

- 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.

- **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 (ГП) от архей
- У *E.coli* – из 4289 генов 755 получены в результате ГП от бактерий других видов.

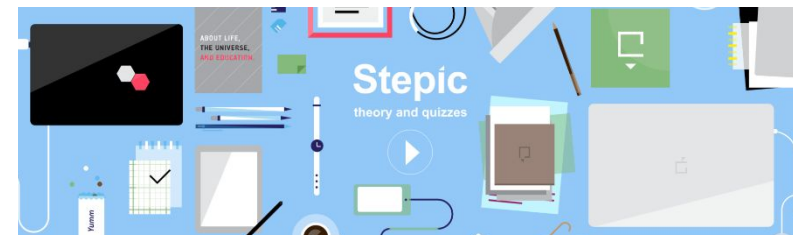
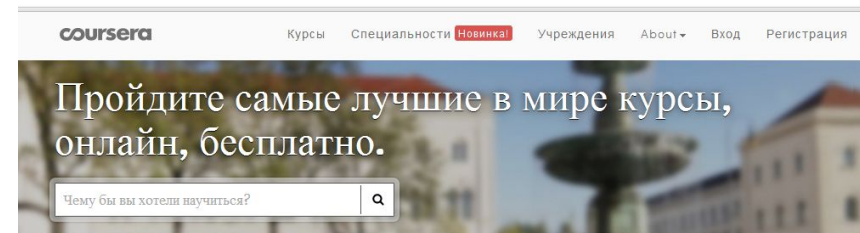
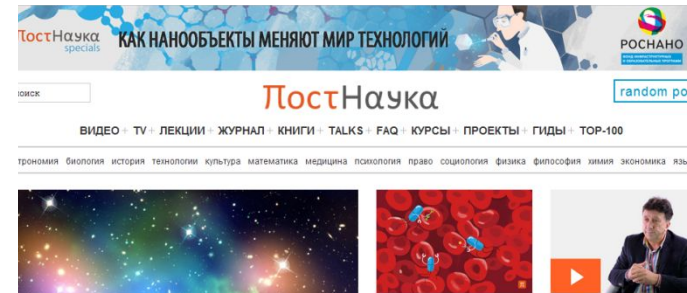


# Старение ДНК

- Благодаря кольцевому строению ДНК нуклеоида и плазмид клетки прокариот не стареют в череде поколений

# Образовательные курсы и лекции

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





Лекция "Ребятам о геномах" 9 января  
2015 г. СФУ Дейч К.О.

# Сайт НОЦ

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

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

Поиск

Центр геномных исследований  
Научно-образовательный центр СФУ

English

**Новости**

- Главная
- Лаборатория лесной геномики
- Учебная лаборатория
- Нам помогают
- Партнеры
- Контакты
- СМИ о нас
- Форум

**Пятница!!!**

12.12.2014 в 15:00  
Александр Цыбин  
"Предварительные результаты поиска повторов в геномах хвойных"  
Серафима Новикова


**Секвенирование геномов 17 людей старше 110 лет пока не смогло объяснить феномен долгожительства**  
14.11.2014

Исследователи из США под руководством Стюарта Кима (Stuart Kim) прочитали геномы 17 долгожителей. К сожалению, тайна долгой жизни остается тайной на данный момент - проведенный анализ геномов не выявил каких-либо редких мутаций, которые могли бы быть ответственны за долгожительство.

[Подробнее...](#)

**В СФУ рассказали о клонировании мамонта**  
06.11.2014

Максим Чепрасов, ученый из Северо-восточного федерального университета принял участие в Фестивале



Лекция "Ребятам о геномах" 9 января  
2015 г. СФУ Дейч К.О.