

# 2020. ХОБП, ч 1, Химическая биология (вт, чт 12-40 — 14-15 ЮХА)

## I Живое/жизнь как система

- 11. 02 Что такое живое/жизнь с точки зрения химии - 1
- 13. 02 Молекулы клетки. Вода. - 2
- 18. 02 Структура и функция белка - 3
- 20. 02 Биологические мембраны Обмен веществом. Преобразование энергии - 4
- 25. 02 Контрольная 1

## II Информационные потоки

- 27. 02 Структура нуклеиновых кислот, двойная спираль ДНК — 5. День откр двер
- 03. 03 Биосинтез нуклеиновых кислот — 6
- 05. 03 Упражнения с нуклеиновыми кислотами - 7
- 10.03 Биосинтез белка - 8
- 12. 03 Контрольная 2

## III Генотип и фенотип

- 17. 03 Регуляция экспрессии генов. Система передачи сигнала. Рак - 9
- 19. 03 Геном, плазмиды, вирусы. Грипп, ВИЧ — 10

## IV Инженерия и промышленность

- 24. 03 Генетическая инженерия - 11
- 26.03 Введение в биотехнологию. ГМО — 12
- 31.03 Контрольная 3

### Семинары по группам по отдельному расписанию

Разбор контрольной 1 с 02.03 Заболел семинарист 302 группы. Семинар переносится ровно на 1 неделю (след. Пт, 10.50)

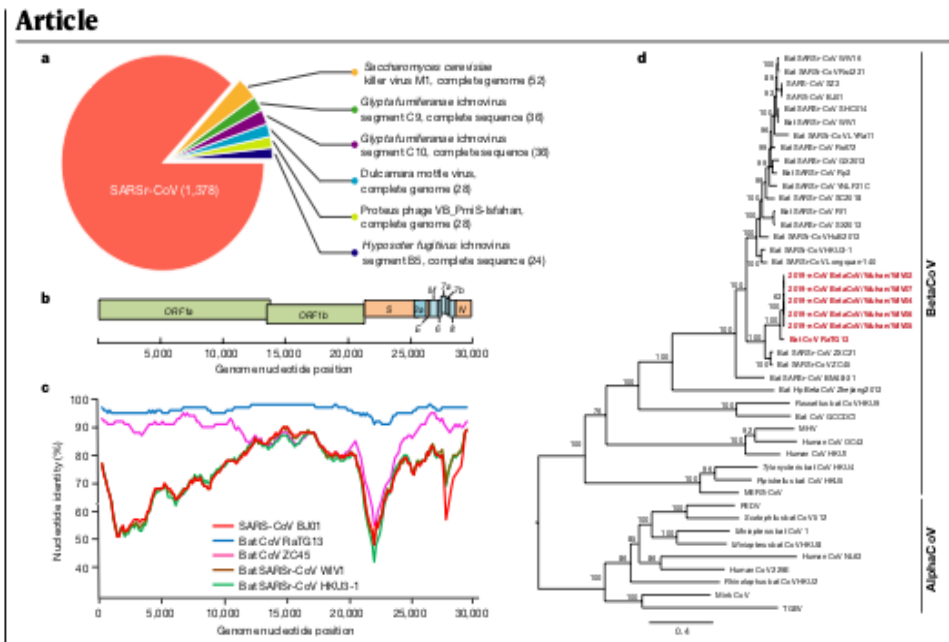
Разбор контрольной 2

Разбор контрольной 3

# A pneumonia outbreak associated with a new coronavirus of probable bat origin

For qPCR analysis, primers based on the S gene of 2019-nCoV were designed: RBD-qF1, 5'-CAATGGTTTAACAGGCACAGG-3'; RBD-qR1, 5'-CTCAAGTGTCTGTGGATCACG-3'. RNA extracted as described above was used for qPCR using the HiScript II One Step qRT-PCR SYBR Green Kit (Vazyme Biotech). Conventional PCRs were also performed using the following primer pairs: ND-CoVs-951F, 5'-TGT-KAGRTTYCCTAAYATTAC-3'; ND-CoVs-1805R, 5'-ACATCYTGATAN-ARAACAGC-3'. The 20- $\mu$ l qPCR reaction mix contained 10  $\mu$ l 2 $\times$  One Step SYBR Green mix, 1  $\mu$ l One Step SYBR Green Enzyme mix, 0.4  $\mu$ l 50 $\times$  ROX Reference Dye 1, 0.4  $\mu$ l of each primer (10  $\mu$ M) and 2  $\mu$ l template RNA. Amplification was performed as follows: 50  $^{\circ}$ C for 3 min, 95  $^{\circ}$ C for 30 s followed by 40 cycles consisting of 95  $^{\circ}$ C for 10 s and 60  $^{\circ}$ C for 30 s, and a default melting curve step in an ABI 7500 Real-time PCR machine.

High-throughput sequencing, pathogen screening and genome assembly Samples from patient BALF or from the supernatant of virus cultures were used for RNA extraction and next-generation sequencing (NGS) using BGI MGISEQ2000 and Illumina MiSeq 3000 sequencers. Metagenomic analysis was carried out mainly based on the bioinformatics platform MGmapper (PE\_2.24 and SE\_2.24). The raw NGS reads were first processed by Cutadapt (v.1.18) with minimum read length of 30 base pairs. BWA (v.0.7.12-r1039) was used to align reads to a local database with a filter hits parameter of 0.8 FMM ((match + mismatch)/read length  $\geq$  fraction) value and minimum alignment score of 30.

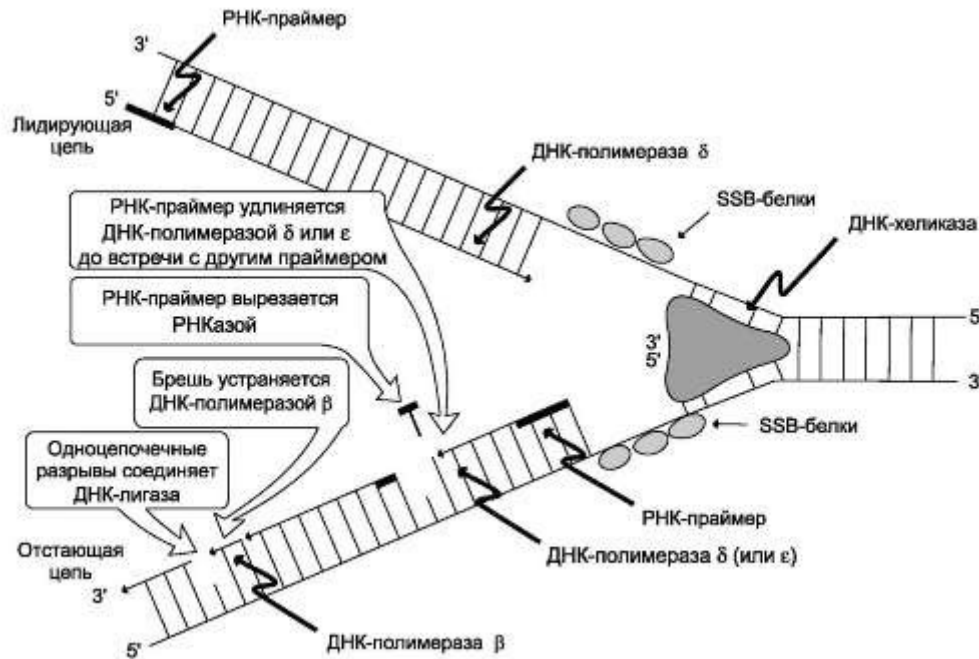


**Fig. 1 | Genome characterization of 2019-nCoV.** a, Metagenomic analysis of next-generation sequencing of BALF from patient ICU06. b, Genomic organization of 2019-nCoV WIV04. M, membrane. c, Similarity plot based on the full-length genome sequence of 2019-nCoV WIV04. Full-length genome sequences of SARS-CoV BJ01, bat SARS-CoV WIV1, bat coronavirus RaTG13 and ZC45 were used as reference sequences. d, Phylogenetic tree based on nucleotide sequences of complete genomes of coronavirus. MHV, murine hepatitis virus; PEDV, porcine epidemic diarrhoea virus; TGEV, porcine transmissible gastroenteritis virus. The scale bars represent 0.1 substitutions per nucleotide position. Descriptions of the settings and software that was used are included in the Methods.

# Упражнения с нуклеиновыми кислотами *в растворе и на твердой фазе*

1. Денатурация ДНК. Плавление, T<sub>пл</sub>.
2. Ренатурация ДНК, кинетика.
3. Гибридизация ДНК, зонды (FISH, микрочипы)
4. Полимеразная цепная реакция (+ в реальном времени, на носителе)
5. Определение первичной структуры ДНК - *секвенирование* (три поколения, метод Сэнгера, Illumina)

# Репликативная вилка в двух проекциях



Биоинженерия –  
как можно перенести процесс из живой клетки, *in vivo*,  
в пробирку, *in vitro*?

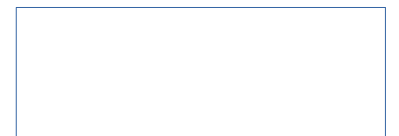
**ДЕНАТУРАЦИЯ** - («отход от природы»),  
**разрушение** исходной (нативной)  
вторичной/третичной структуры

Двойная спираль ДНК (N) ----- >  
два отдельных твжа ДНК как статистические  
клубки (D)

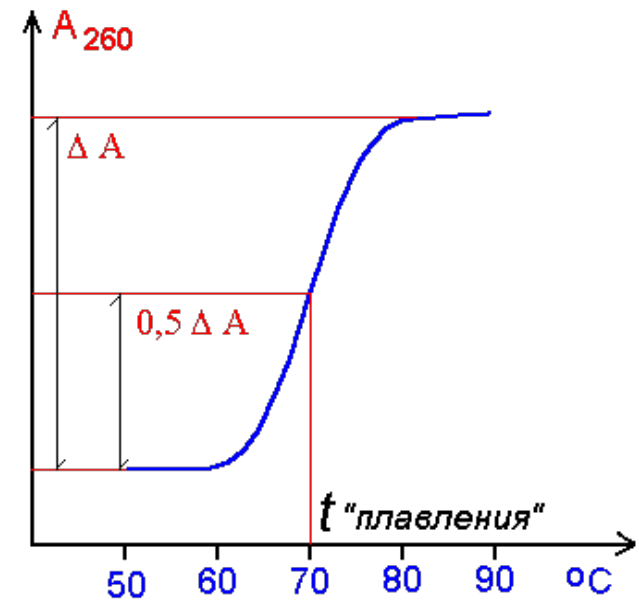
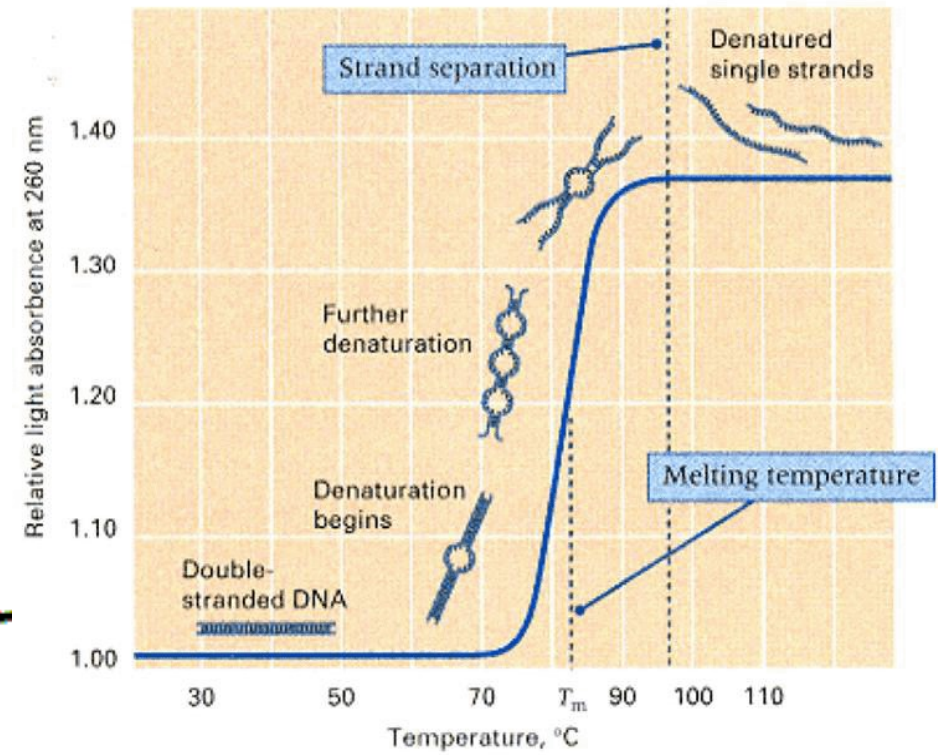
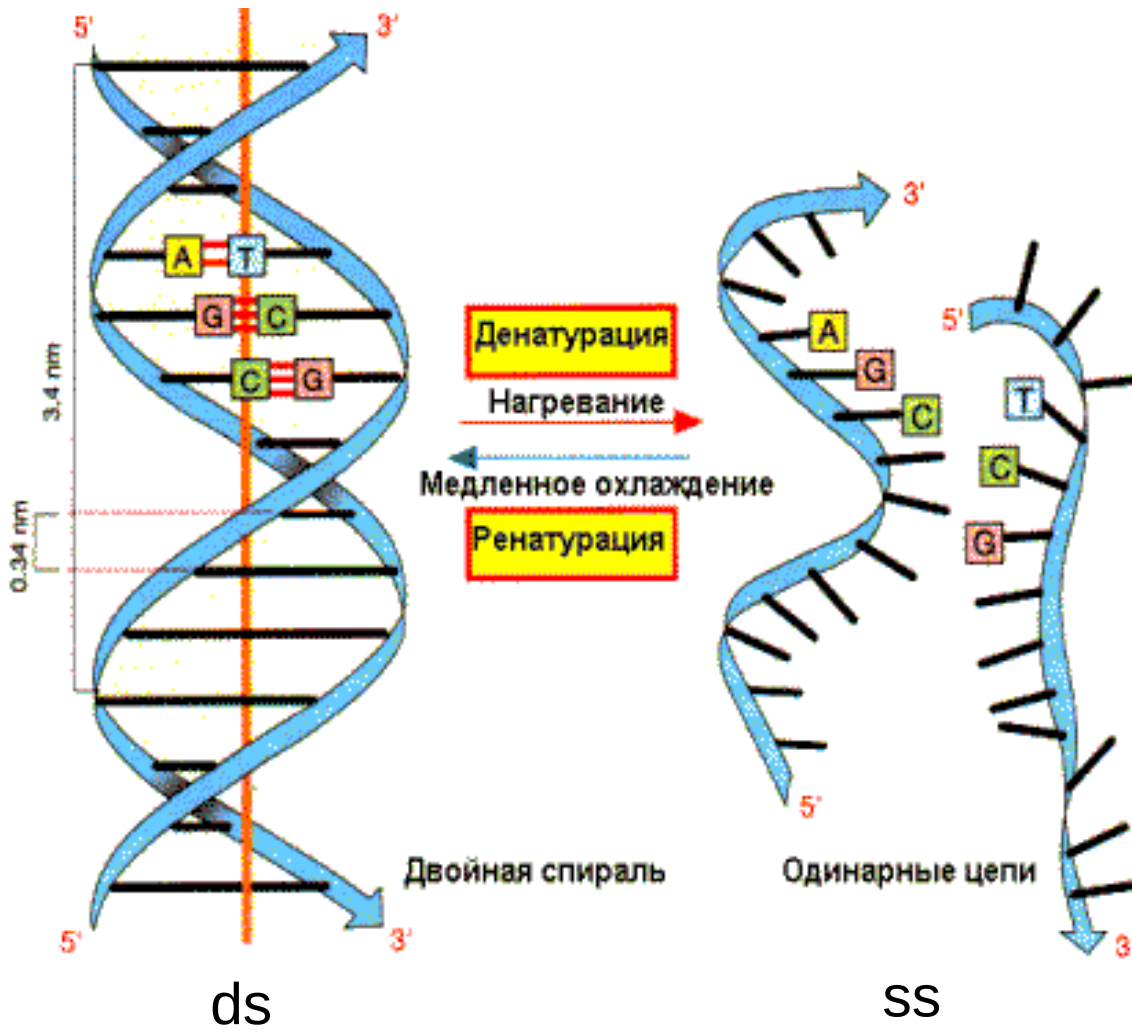
**РЕНАТУРАЦИЯ** - («возвращение к природе»),  
**восстановление** исходной (нативной)  
вторичной/третичной структуры

Гибридизация ДНК

Олигонуклеотидные зонды



# Денатурация ДНК

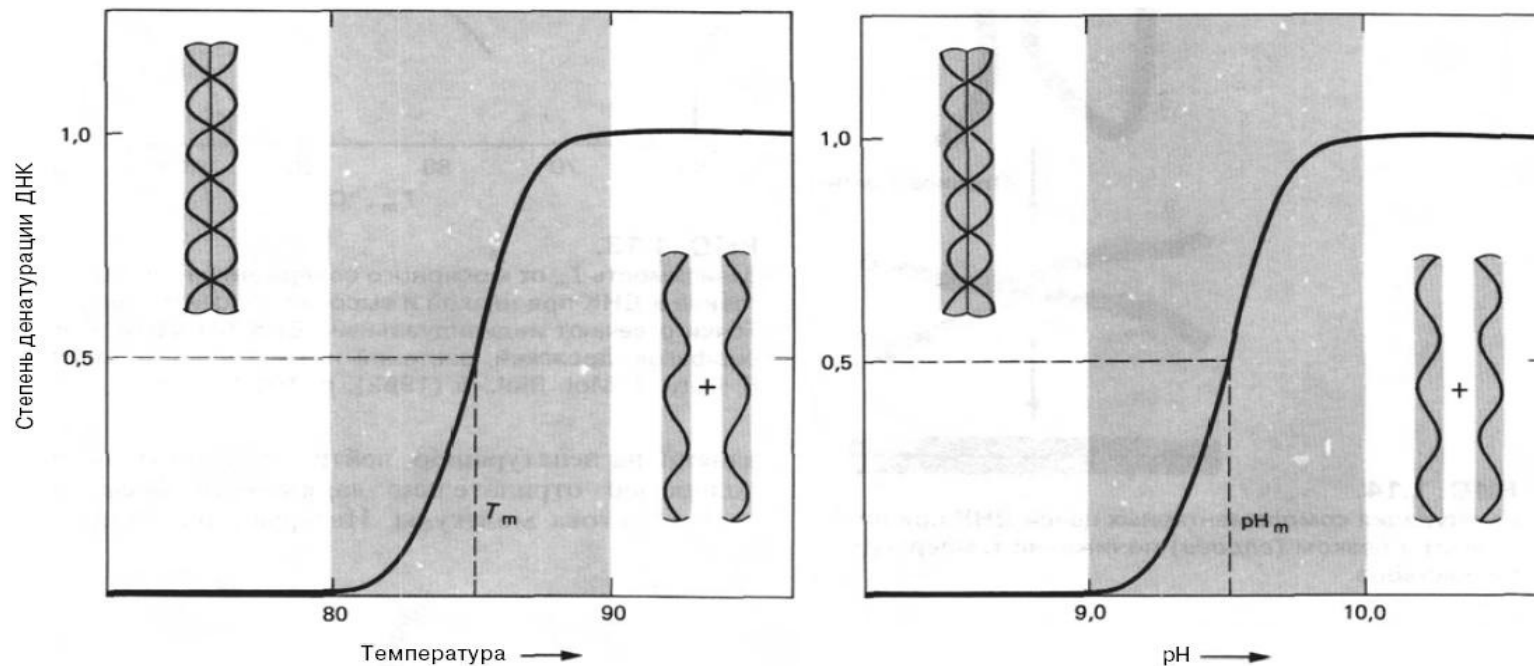


УФ: гиперхромный эффект

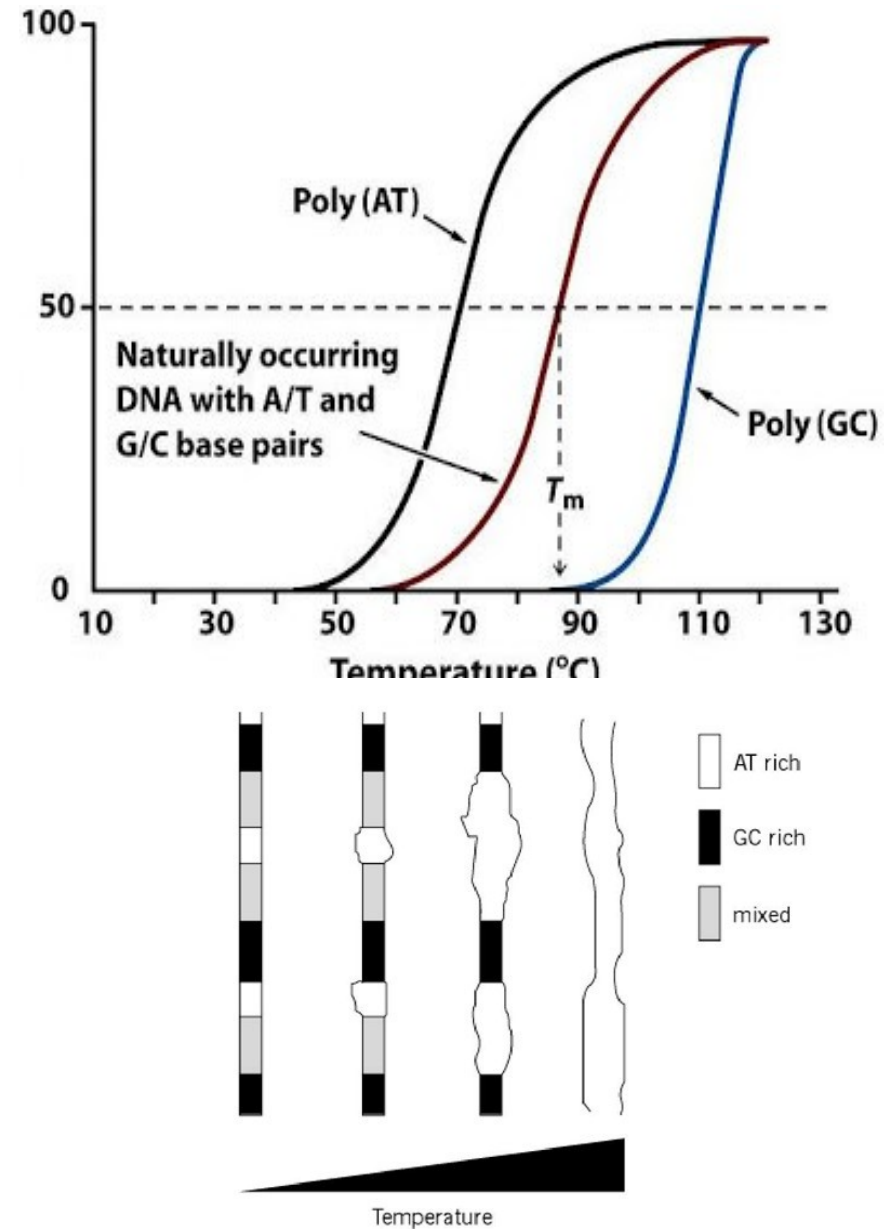
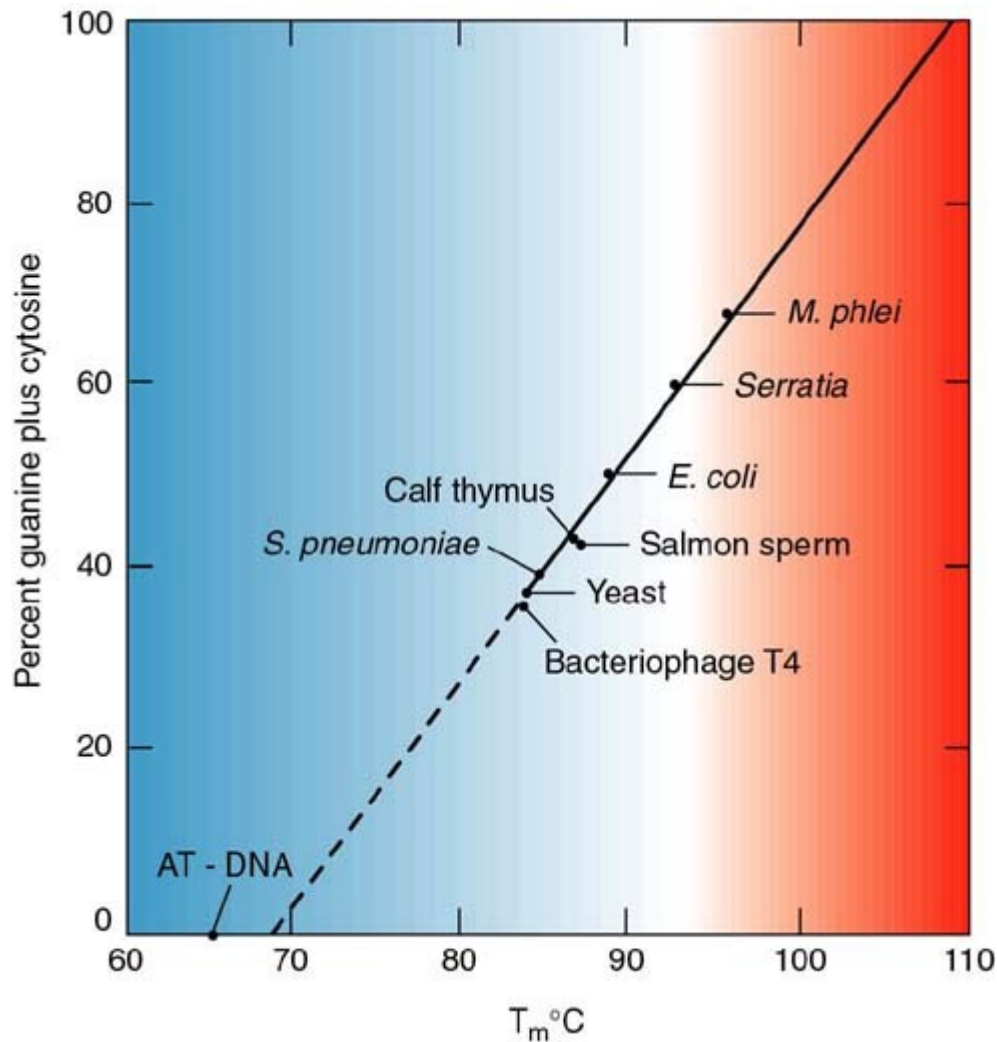


# Денатурация ДНК

*температура, химические вещества  
(рН, гуанидин хлорид, мочеви́на, формамид,  
диметилсульфоксид (ДМСО))*

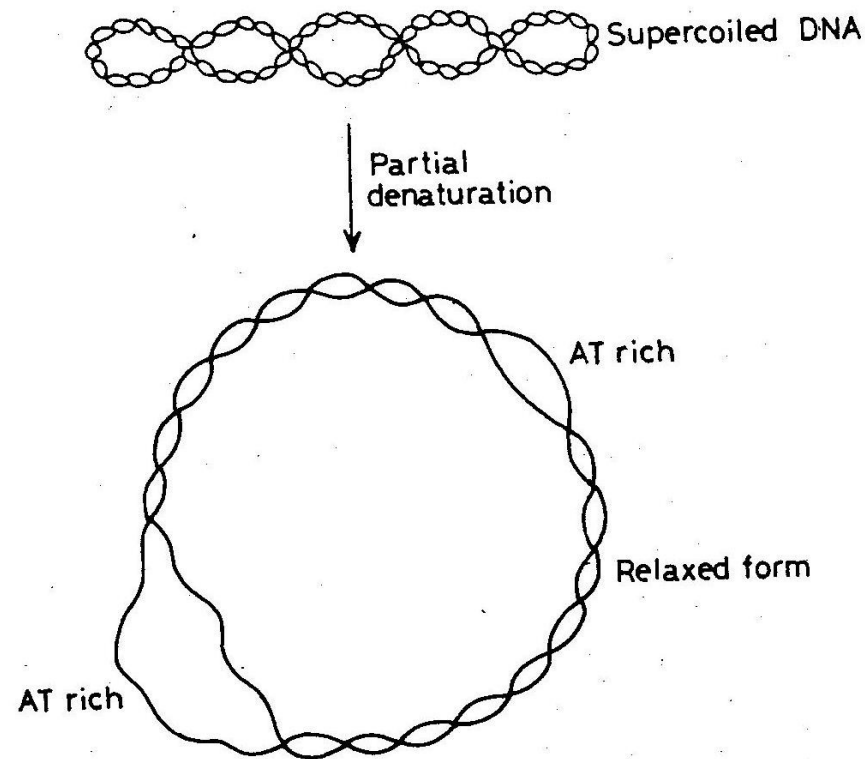


# Зависимость $T_m$ от состава А-Т, G-C





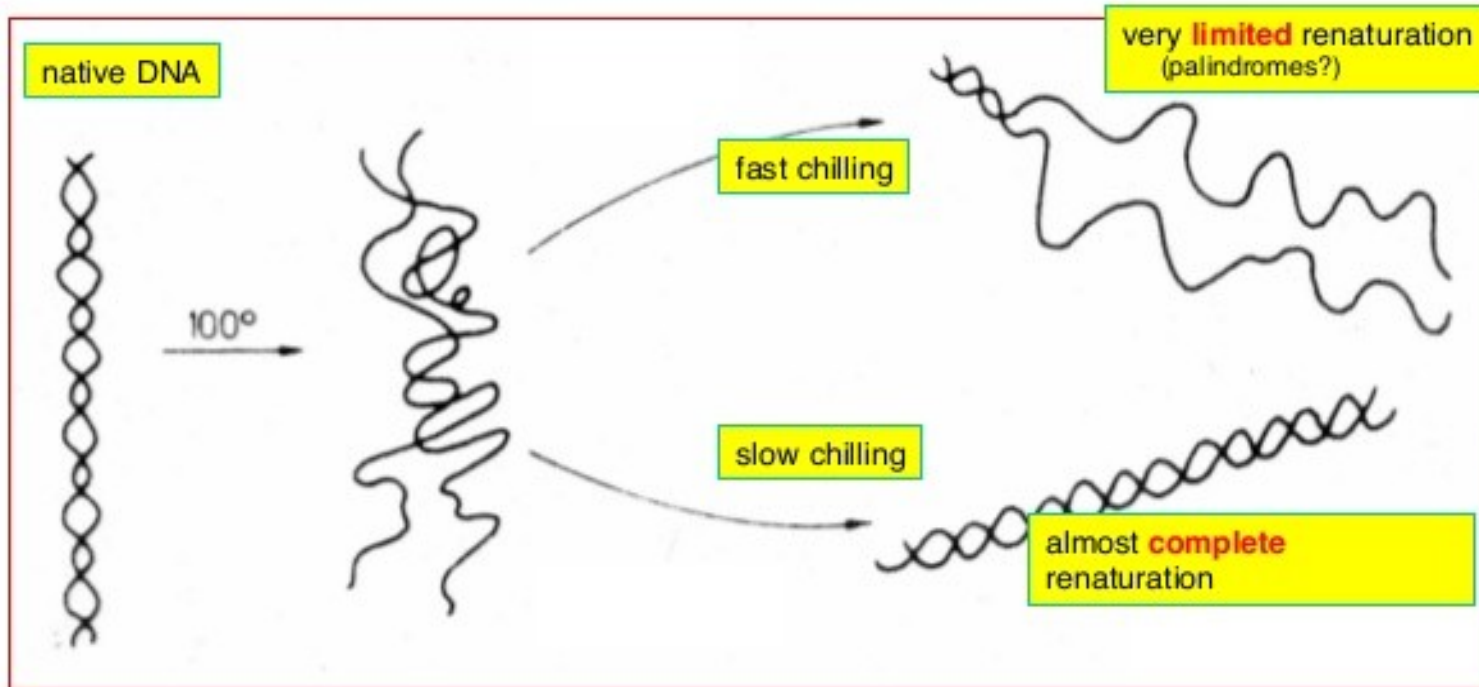
Обратимые процессы:  
суперспиральность стимулирует частичное  
раскручивание двойной спирали (денатурацию)



$$Lk = Tw + Wr$$

# Обратимые процессы: денатурация и ренатурация ДНК

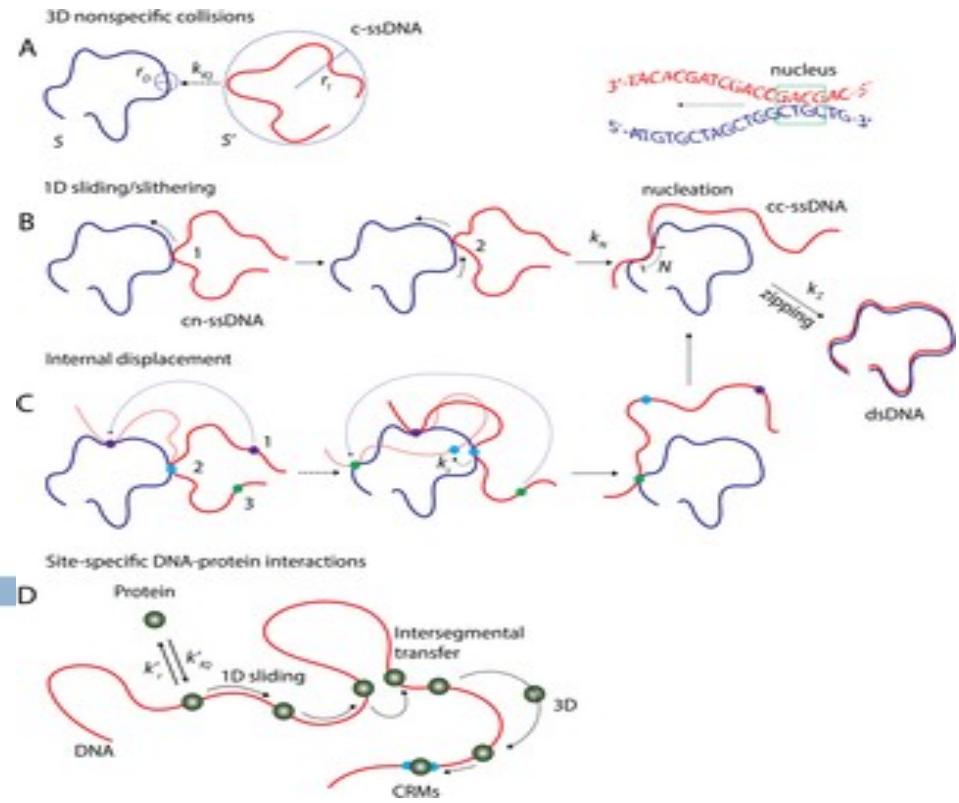
DNA denaturation and renaturation : strategic aspects



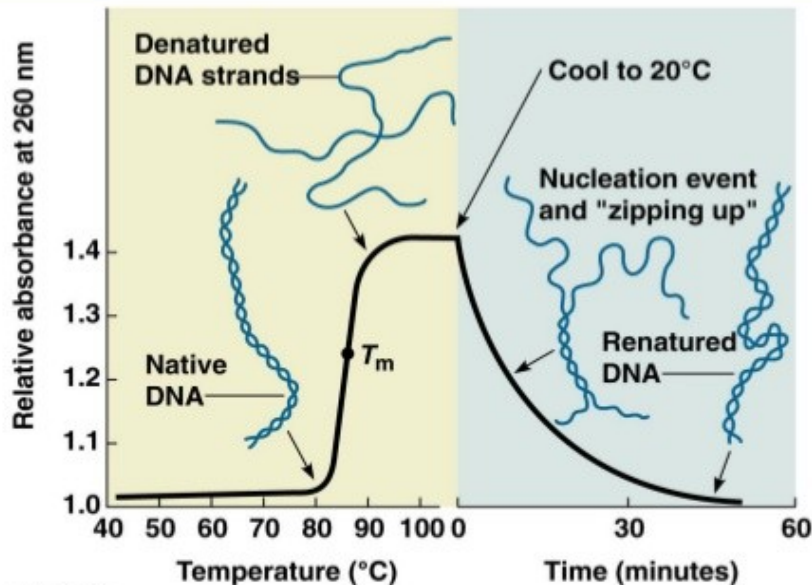
Hyperchromicity effect : disruption of the stacking  
=> 30 to 40 % increase of UV (260 nm) absorption

# Сложные механизмы ренатурации

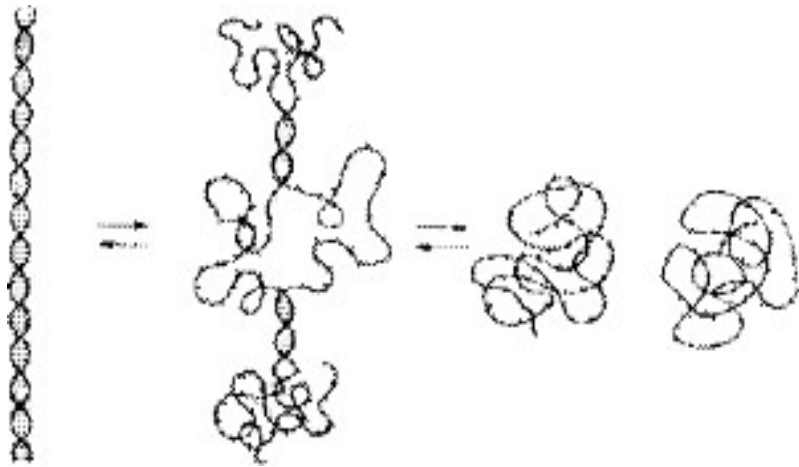
## Renaturation of DNA



22



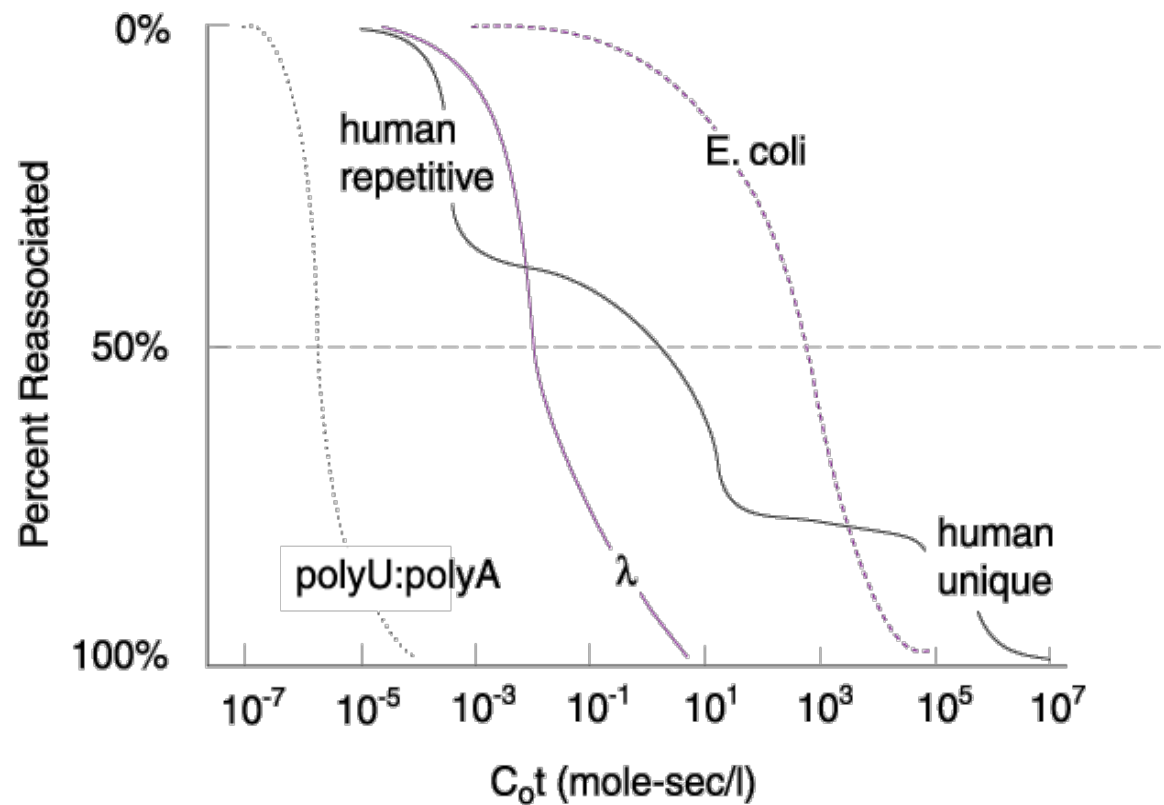
# Медленная кинетика ренатурации ДНК



$$C/C_0 = 1/(1+K_2 C_0 t)$$

Втор порядок

Линейн при  $K_2 = 1$



# Олигонуклеотидные зонды некомплементарная пара

ДНК GCTATTCAACTGAAGAGGGGCACAGC + GCTATTCAACTG<sup>G</sup>AGAGGGGCACAGC

олиг  
“зонд” CGATAAGTTGACTTCTCCCGTGTCG

CGATAAGTTGACTTCTCCCGTGTCG

$$T_m - 25^\circ \begin{array}{c} \uparrow \\ \downarrow \end{array} > T_m$$

$$T_m^* - 25^\circ \begin{array}{c} \uparrow \\ \downarrow \end{array} > T_m^*$$

GCTATTCAACTGAAGAGGGGCACAGC  
CGATAAGTTGACTTCTCCCGTGTCG

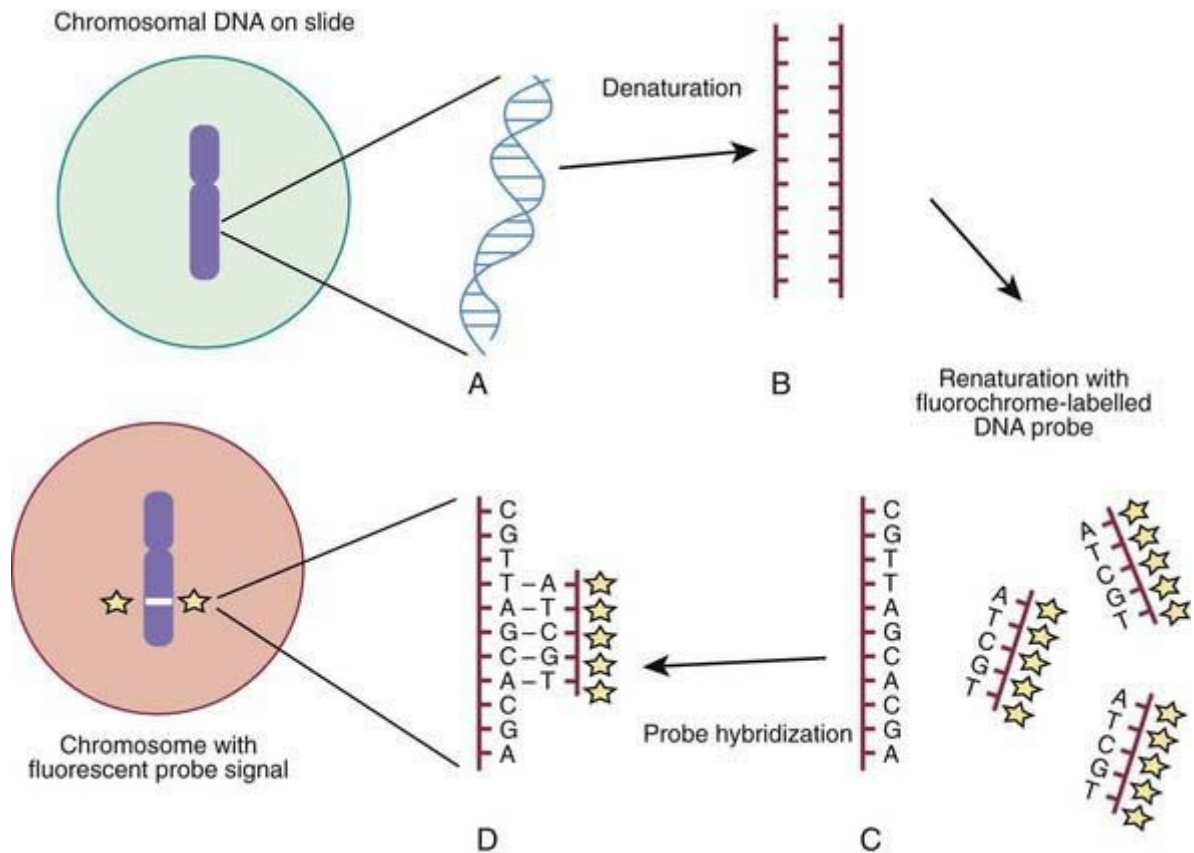
GCTATTCAACTG<sup>G</sup>AGAGGGGCACAGC  
CGATAAGTTGAC<sub>T</sub>TCTCCCGTGTCG

эмпирика

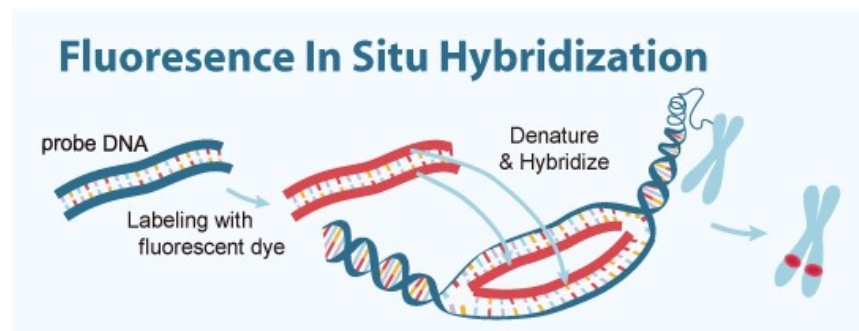
note:  $T_m^*$  is  $4^\circ$  lower than  $T_m$

(In general, there is a  $1^\circ$  drop for every 1% mismatch)

# Хромосомные зонды, FISH - флуоресцентная гибридизация

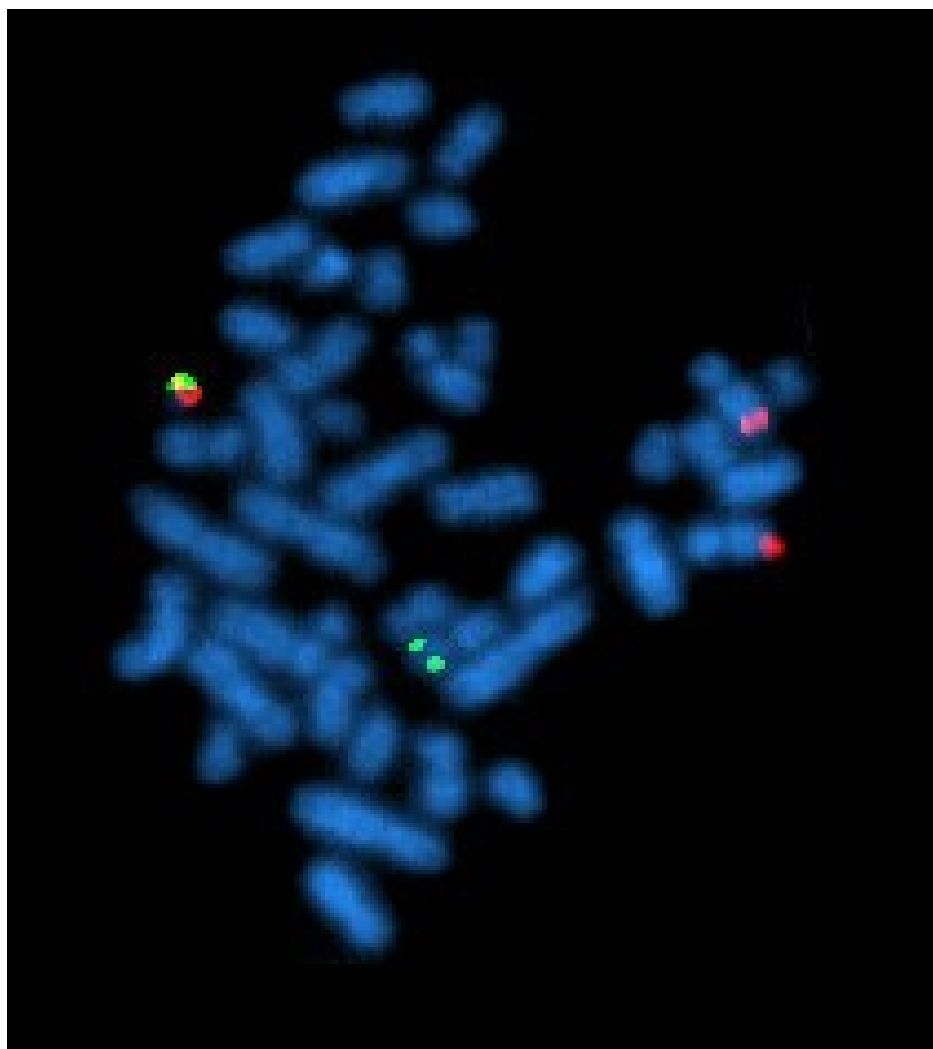


FISH





FISH. Метафазные клетки с “филадельфийской хромосомой” (перегруппировка *bcr/abl* связана с хронической миелогенной лейкемией). Хромосомы окрашены в синий цвет, локус *ABL1* — красный цвет, локус *BCR* — зелёный цвет. Вверху слева — хромосома с перестройкой, отмечена красно-зеленой точкой

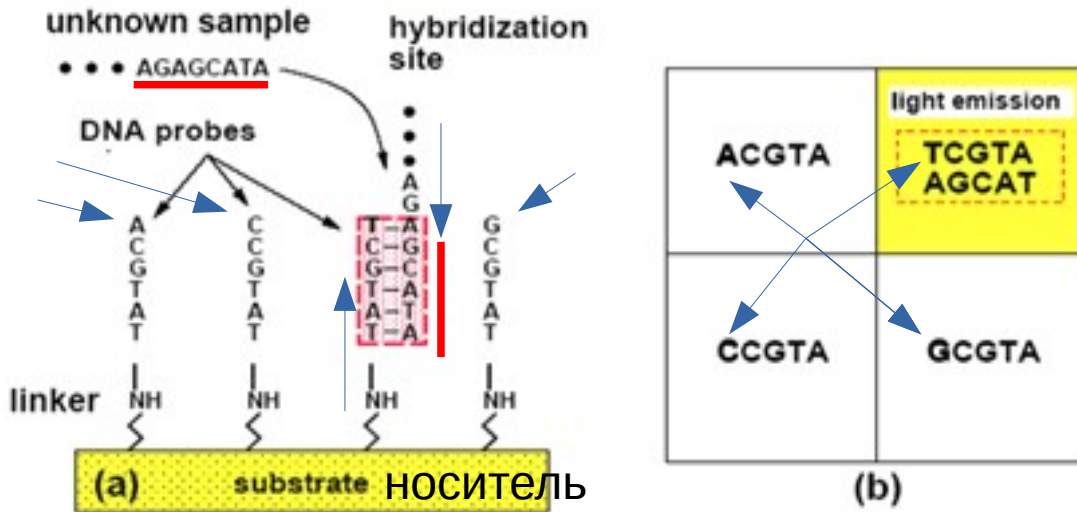


# Гибридизация на твердой фазе:

## ЧИПЫ



# Simple Example of DNA Microarrays

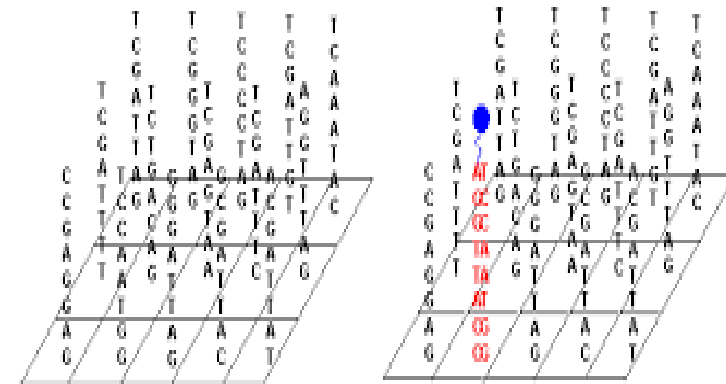
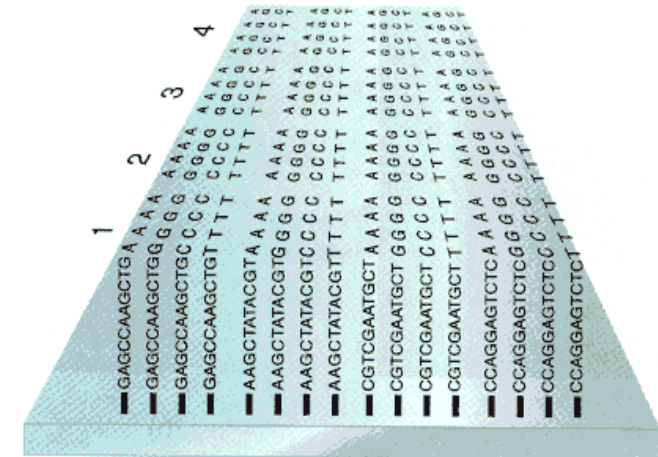


(a) Example immobilized DNA probes showing hybridization of unknown(target) to specific probe.

(b) Probes are arranged a planar arrays. The hybridized regions can be detected by the fluorescence of the duplex.

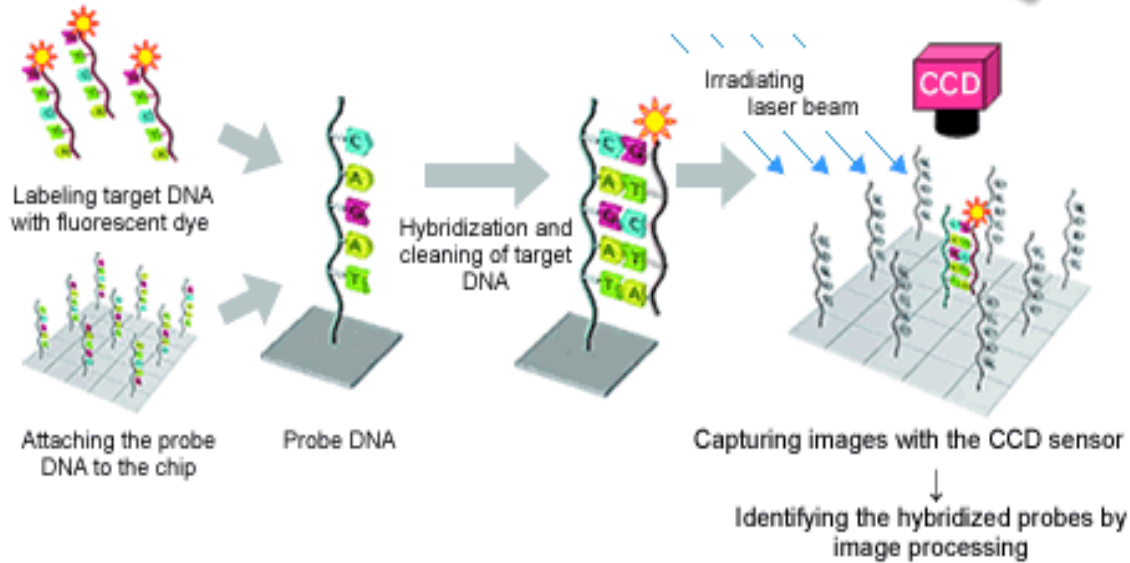
(Griffiths et al. 1999)

## Oligonucleotide array



# Fluorescence-Detection DNA Chip

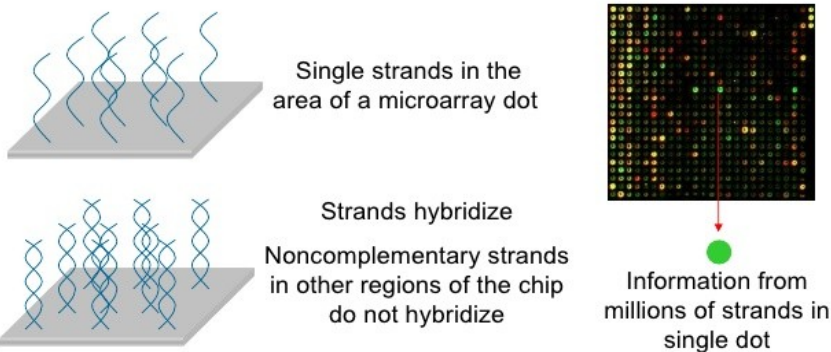
The four bases A, T, G, and C bind A to T or G to C. A target DNA sequence is analyzed by checking which bases the target DNA bases bind.



ПЗС – прибор с зарядовой связью  
 CCD – Charge-Coupled Device

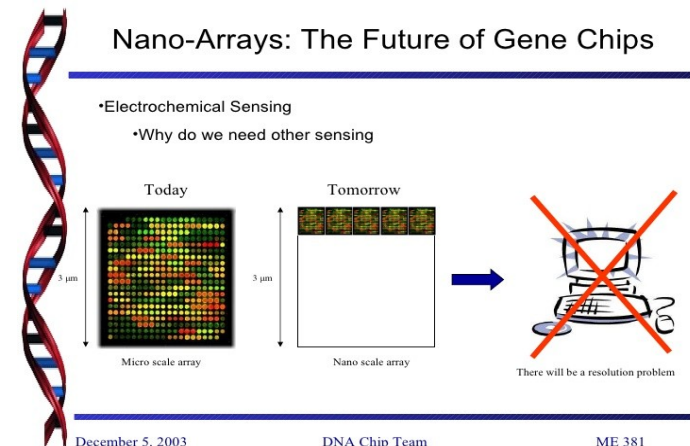
## Array Hybridization

- Single strand oligonucleotides stand on the chip
- Hybridization occurs in complementary strands
- Each microarray dot contains millions of identical strands



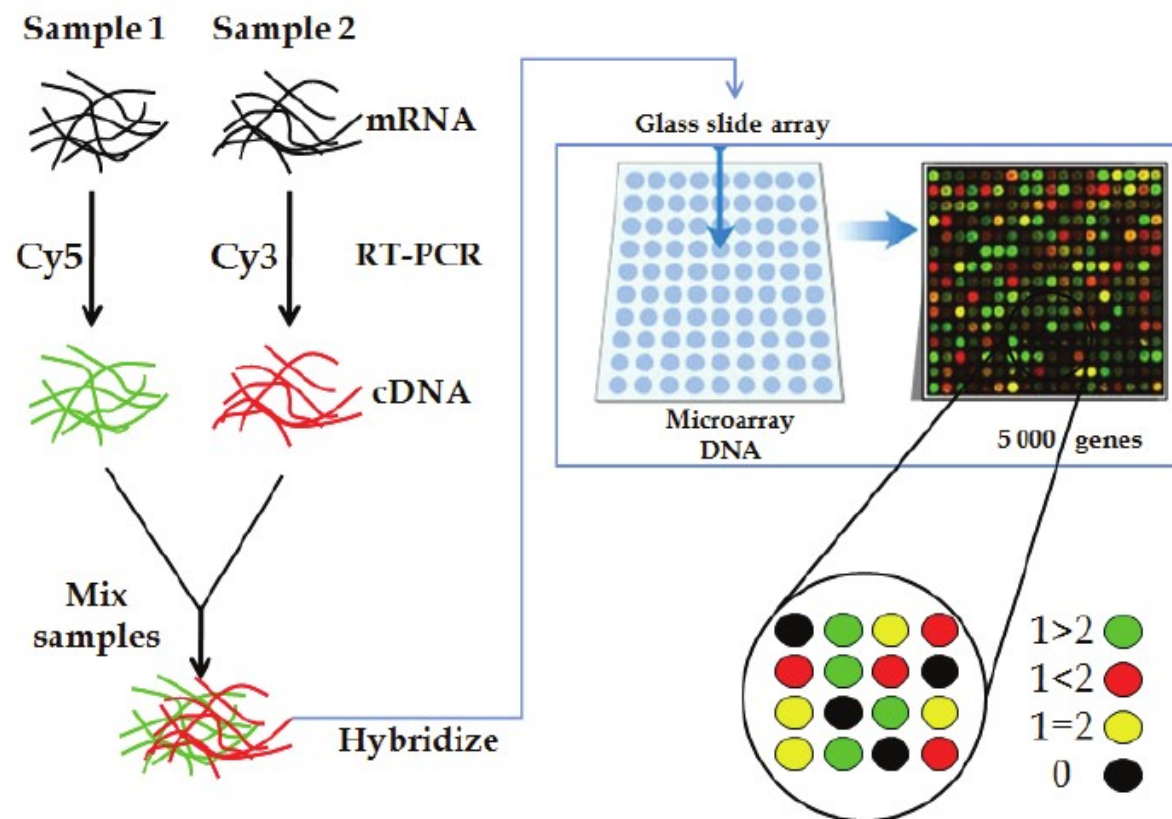
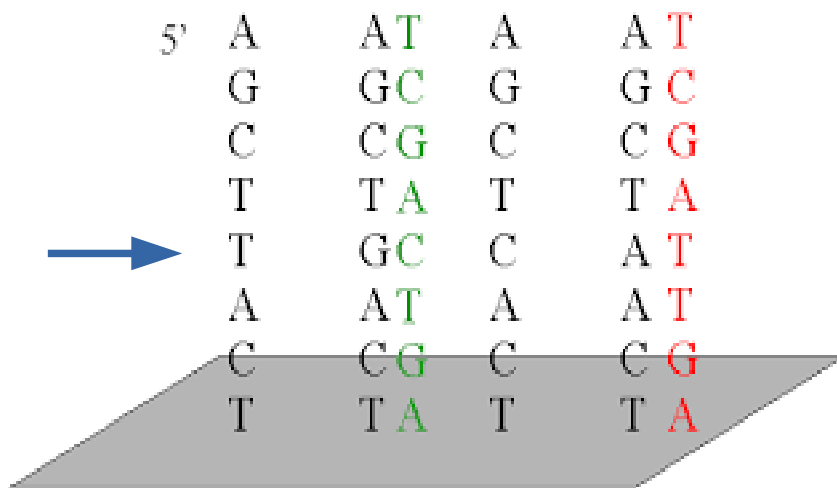
## Nano-Arrays: The Future of Gene Chips

- Electrochemical Sensing
- Why do we need other sensing



# Сравнительная гибридизация двух разных ДНК на чипах

## Несбалансированные изменения (перестройки)





# Мощь экспоненты

## Изобретение шахмат

$$T_{64} = 1 + 2 + 4 + \dots + 2^{63} = 2^{64} - 1 = 18\,446\,744\,073\,709\,551\,615 = 18E18$$

Количество зерна примерно в 1800 раз превышает мировой урожай пшеницы за год (в 2008-09 аграрном году урожай составил 686 млн тонн), то есть **превышает весь урожай пшеницы, собранный за всю историю человечества.**

В единицах массы: если принять, что одно зёрнышко пшеницы имеет массу 65 мг, тогда общая масса пшеницы на шахматной доске составит 1200 миллиардов тонн или 1,2 триллиона тонн:

$$18446744073709551615 \times 0,065 \text{ g} / (1000000 \text{ g/t}) \approx 1,2 \text{ E}12 \text{ t}$$

В то время как количество зёрен на первой половине доски велико, количество на второй половине многократно его превышает. Количество зёрен на **первой половине доски** составляет  $1 + 2 + 4 + \dots + 2^{147} = 2^{148} - 1 = 4\,294\,967\,295$  зёрен, или около **100 000 кг риса при массе одного зёрнышка 25 мг**

Количество зёрен составляет примерно  $3E-5$  числа Авогадро.

1	2	4	8	16	32	64	128
256	512	1024	2048	4096	8192	16384	32768
65K	131K	262K	524K	1M	2M	4M	8M
16M	33M	67M	134M	268M	536M	1G	2G
4G	8G	17G	34G	68G	137G	274G	549G



## ДНК из мумий

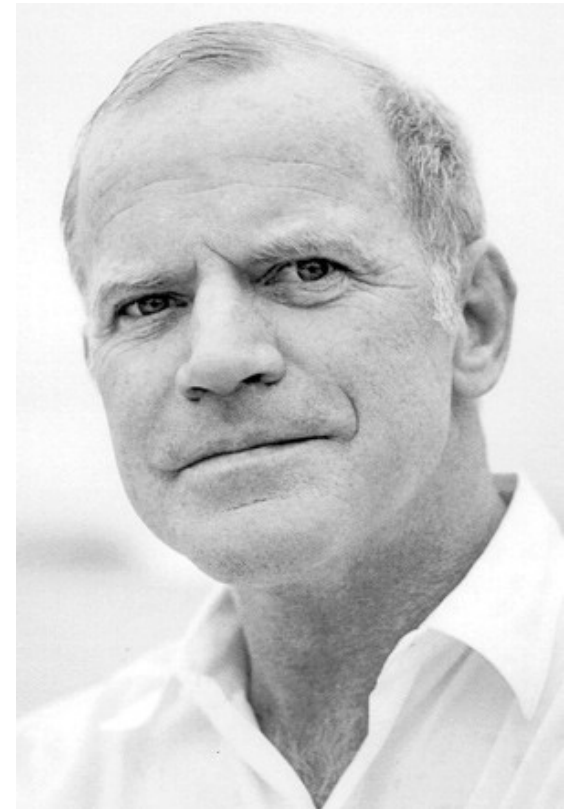
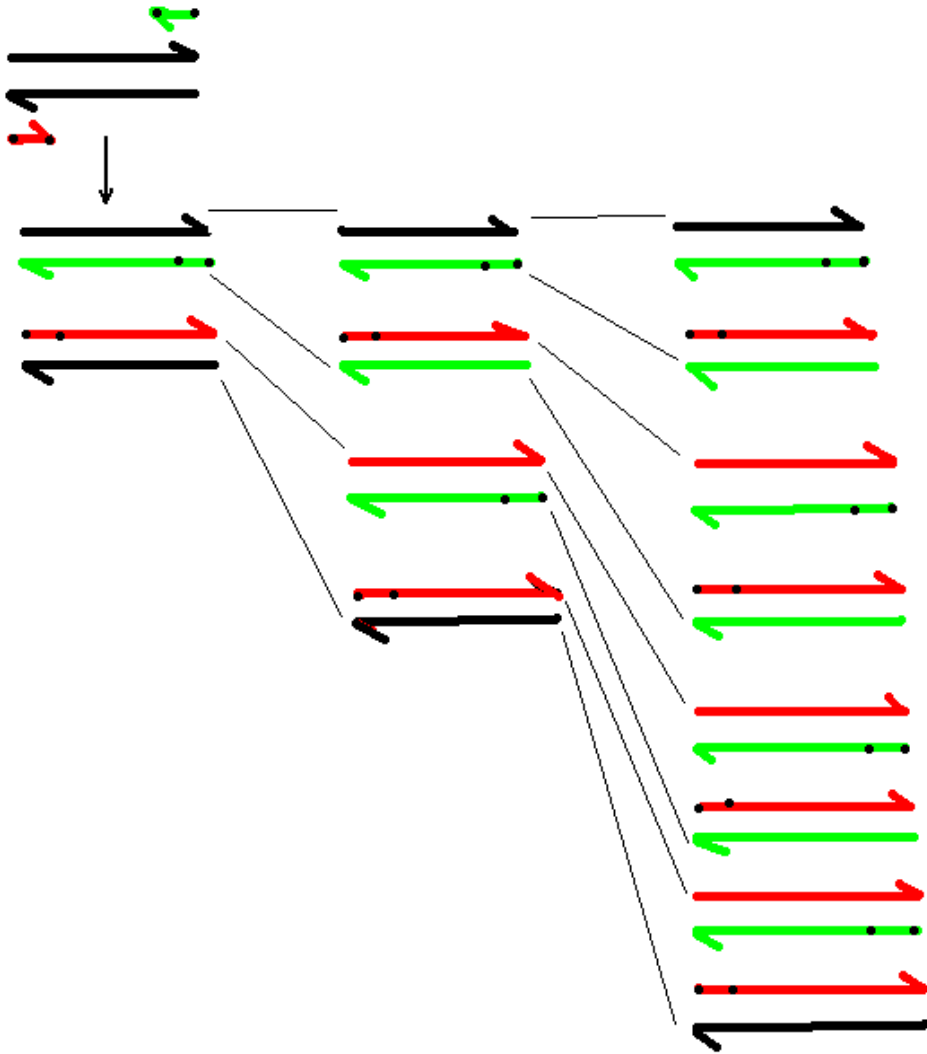
Тутанхамон был сыном Эхнатона (Аменхотепа) и его неидентифицированной сестры

Тутанхамон был женат на Анхесенамон, третьей дочери Эхнатона,

Ancestry and Pathology in King Tutankhamun's Family // JAMA: the Journal of the American Medical Association. — 2010.  
DOI:10.1001/jama.2010.121. — PMID 20159872.

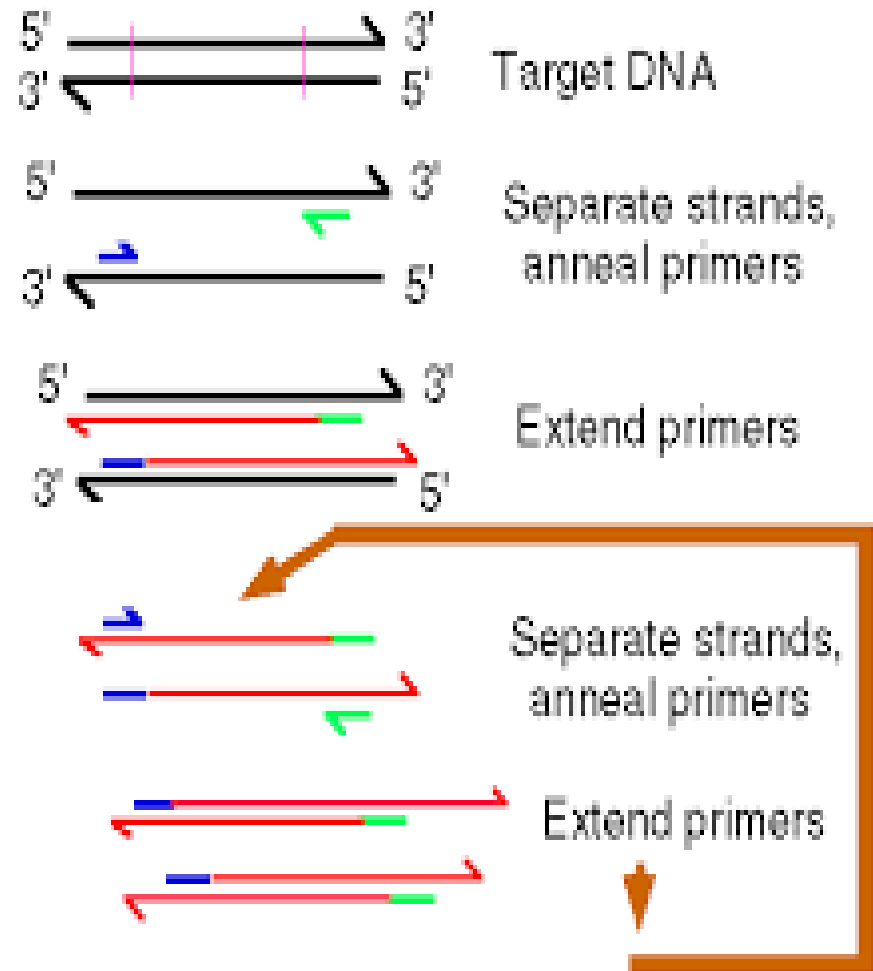
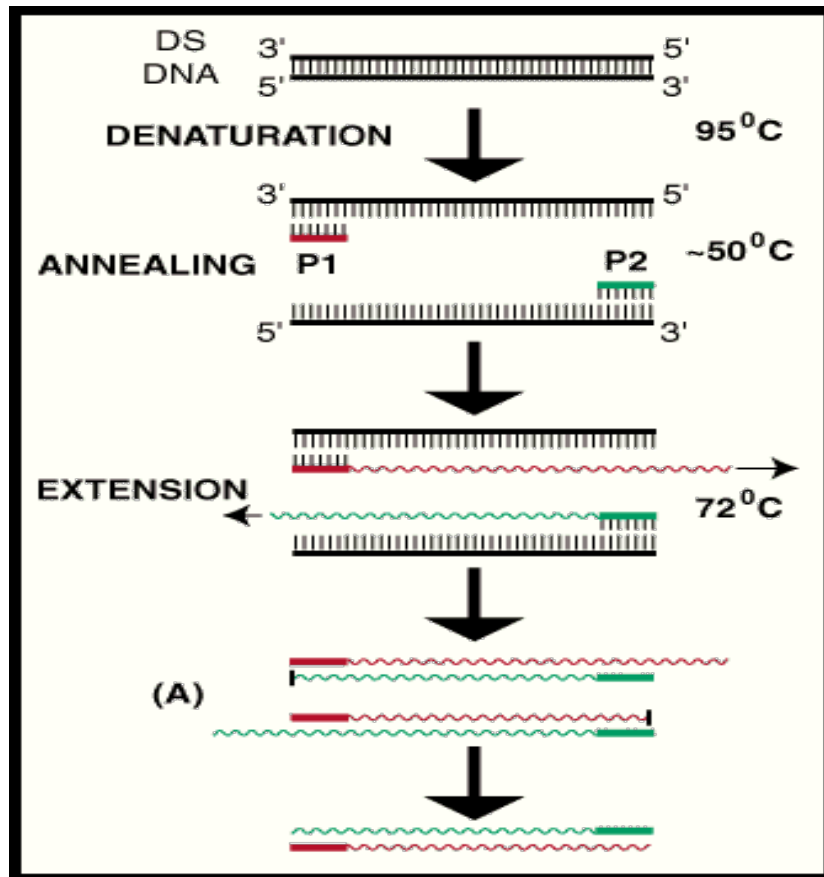


# Полимеразная цепная реакция (ПЦР)



Кэрри Муллис (1983-5)  
Нобелевская премия  
по химии 1993 г.

# Что нужно было сделать для полимеразной цепной реакции?



Each cycle doubles the target DNA copy number

## Состав для ПЦР:

ДНК-матрица

Два праймера, комплементарные противоположным концам разных цепей требуемого фрагмента ДНК.

Термостабильная ДНК-полимераза из термофилов: *Thermus aquaticus* (Taq-полимераза), *Pyrococcus furiosus* (Pfu-полимераза), *Pyrococcus woesei* (Pwo-полимераза), *Thermus thermophilus* (Tth-полимераза) и др.

Дезоксирибонуклеозидтрифосфаты (dATP, dGTP, dCTP, dTTP).

Ионы  $Mg^{2+}$  для работы полимеразы.

Буферный раствор, обеспечивающий pH, ионную силу. Содержит соли, бычий сывороточный альбумин.

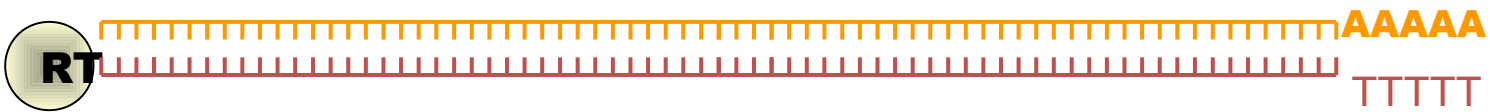
Чтобы избежать испарения реакционной смеси, на поверхность жидкости в пробирке наслаивают высококипящее масло (вазелиновое). Для амплификатора с подогревающейся крышкой это не делают.

Пирофосфатаза увеличивает выход ПЦР-реакции, поскольку смещает равновесие.

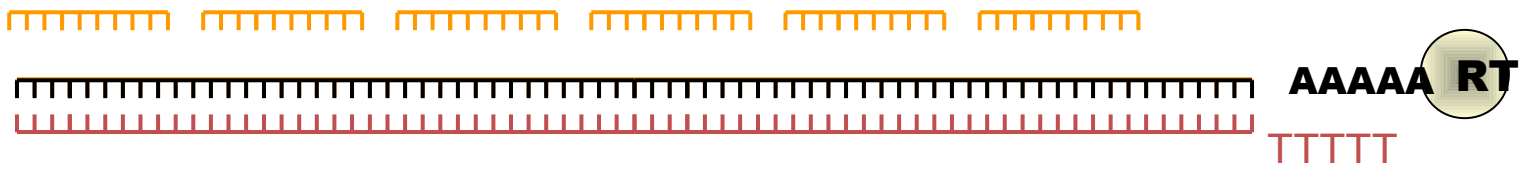
Для РНК-матрицы нужна РНК-зависимая ДНК-полимераза, чтобы получить ДНК-матрицу



Oligo dT primer is bound to mRNA



Reverse transcriptase (RT) copies first cDNA strand

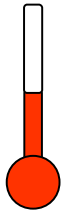


Reverse transcriptase digests and displaces mRNA and copies second strand of cDNA

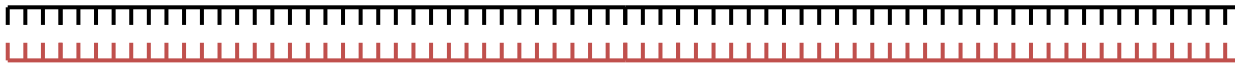


Double strand cDNA

Копирование мРНК в кДНК обратной транскриптазой



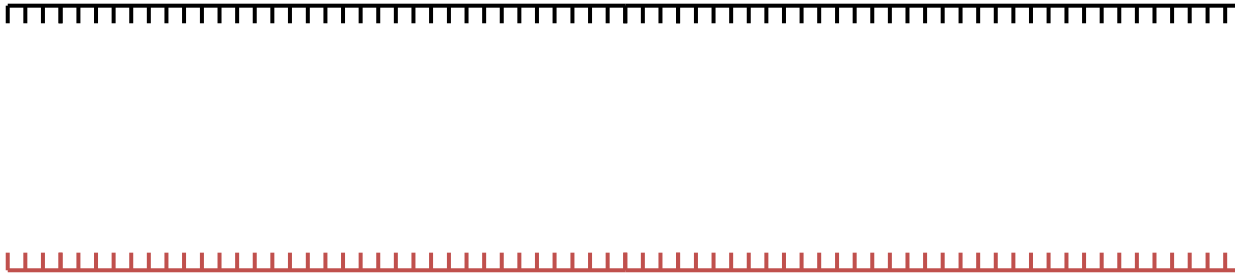
50°



A. Double  
strand  
DNA



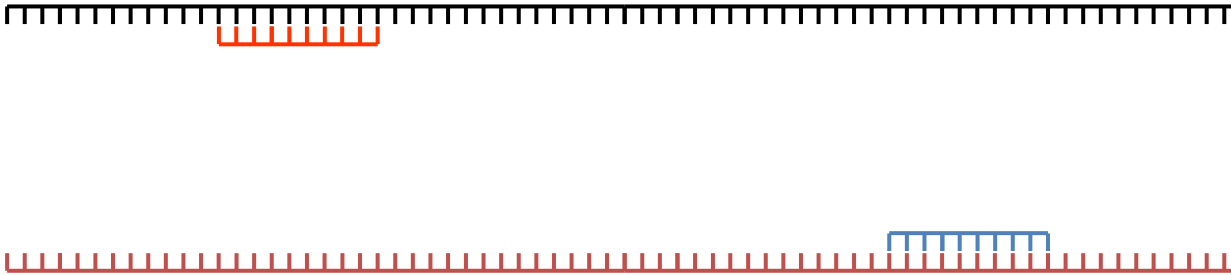
96°



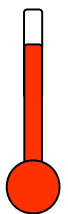
B.  
Denature



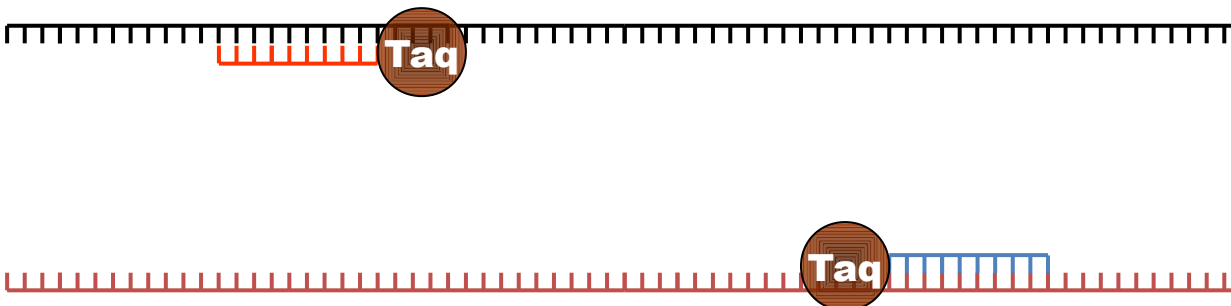
50°



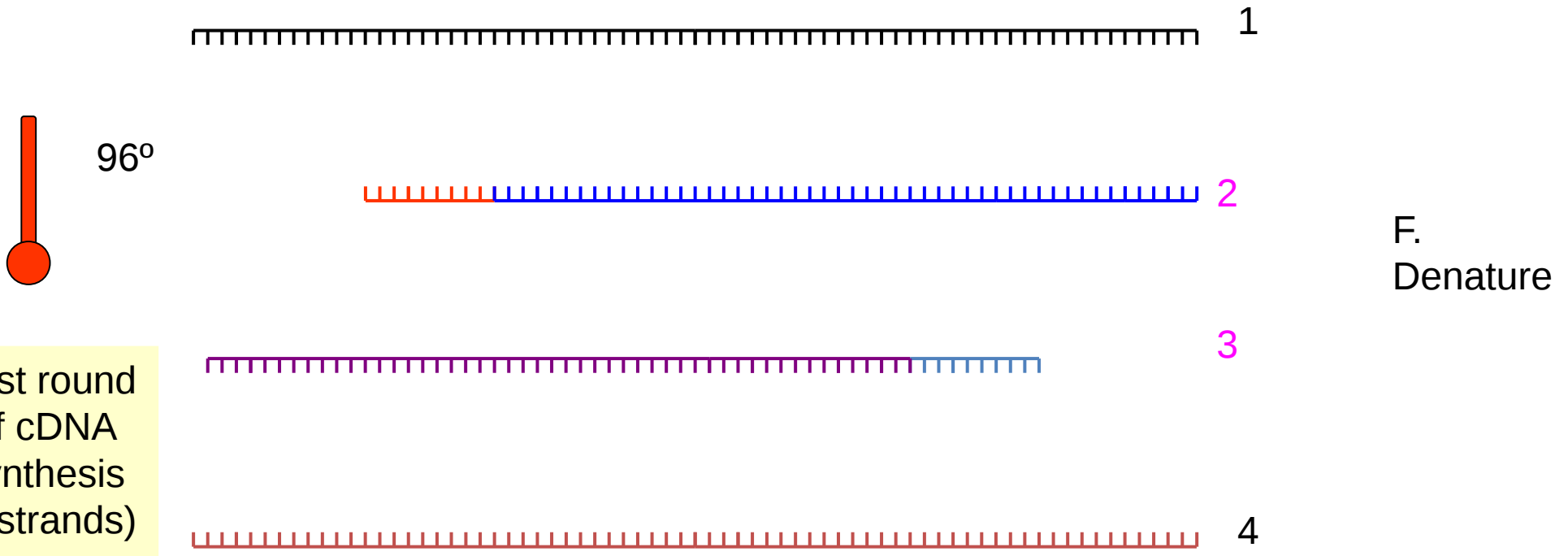
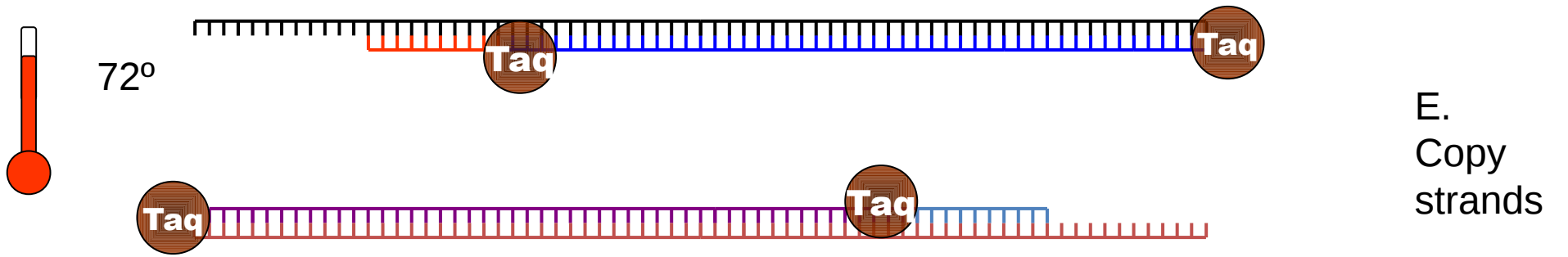
C. Anneal  
primers



72°



D.  
Polymerase  
binds

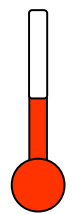


First round  
of cDNA  
synthesis  
(4 strands)

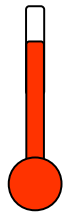




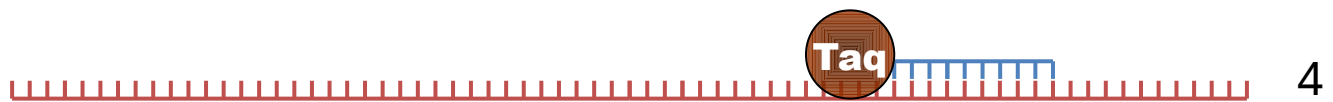
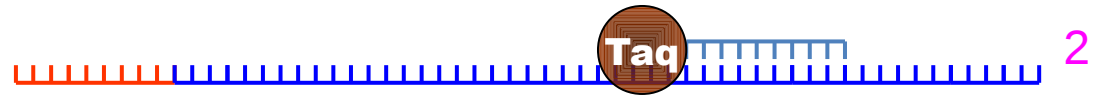
50°



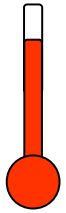
G.  
Anneal  
primers



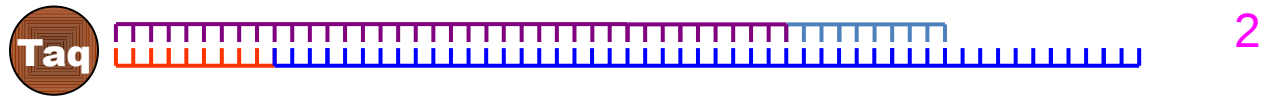
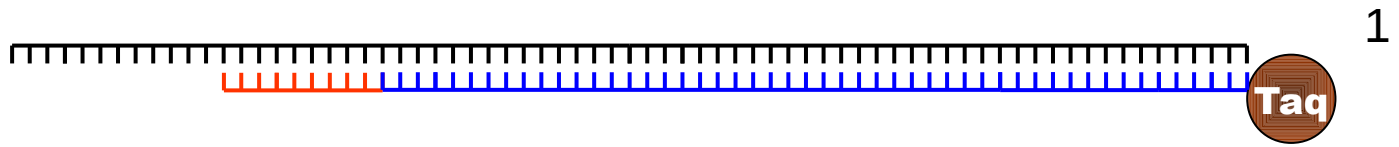
72°



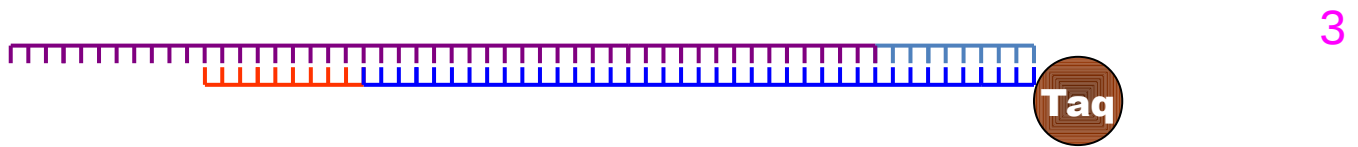
H.  
Polymeras  
e binds



72°

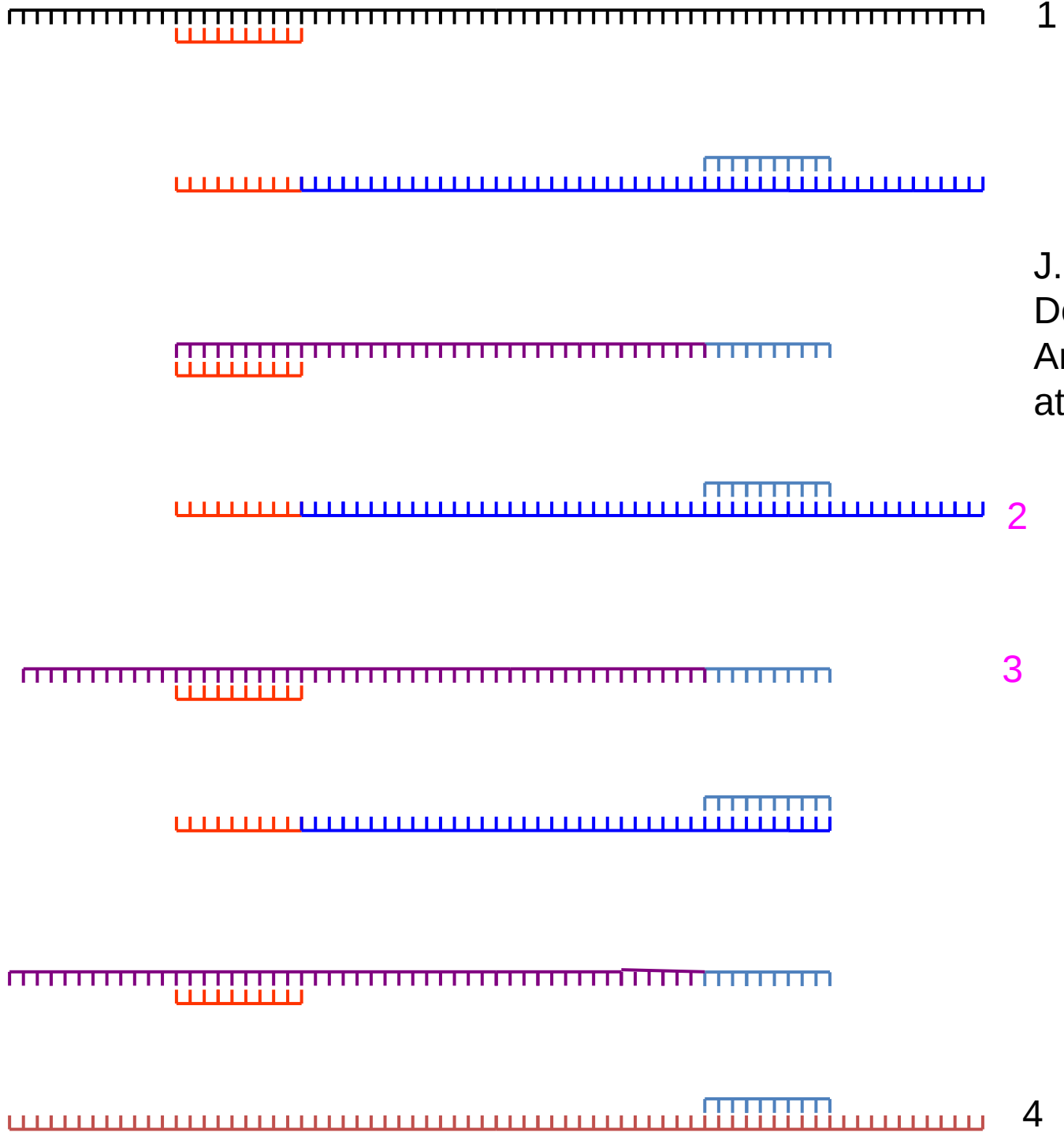
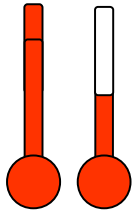


I. Copy strands

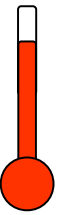


Second round of cDNA synthesis (8 strands)





1



72°

K. Bind polymerase (not shown) and copy strands

2



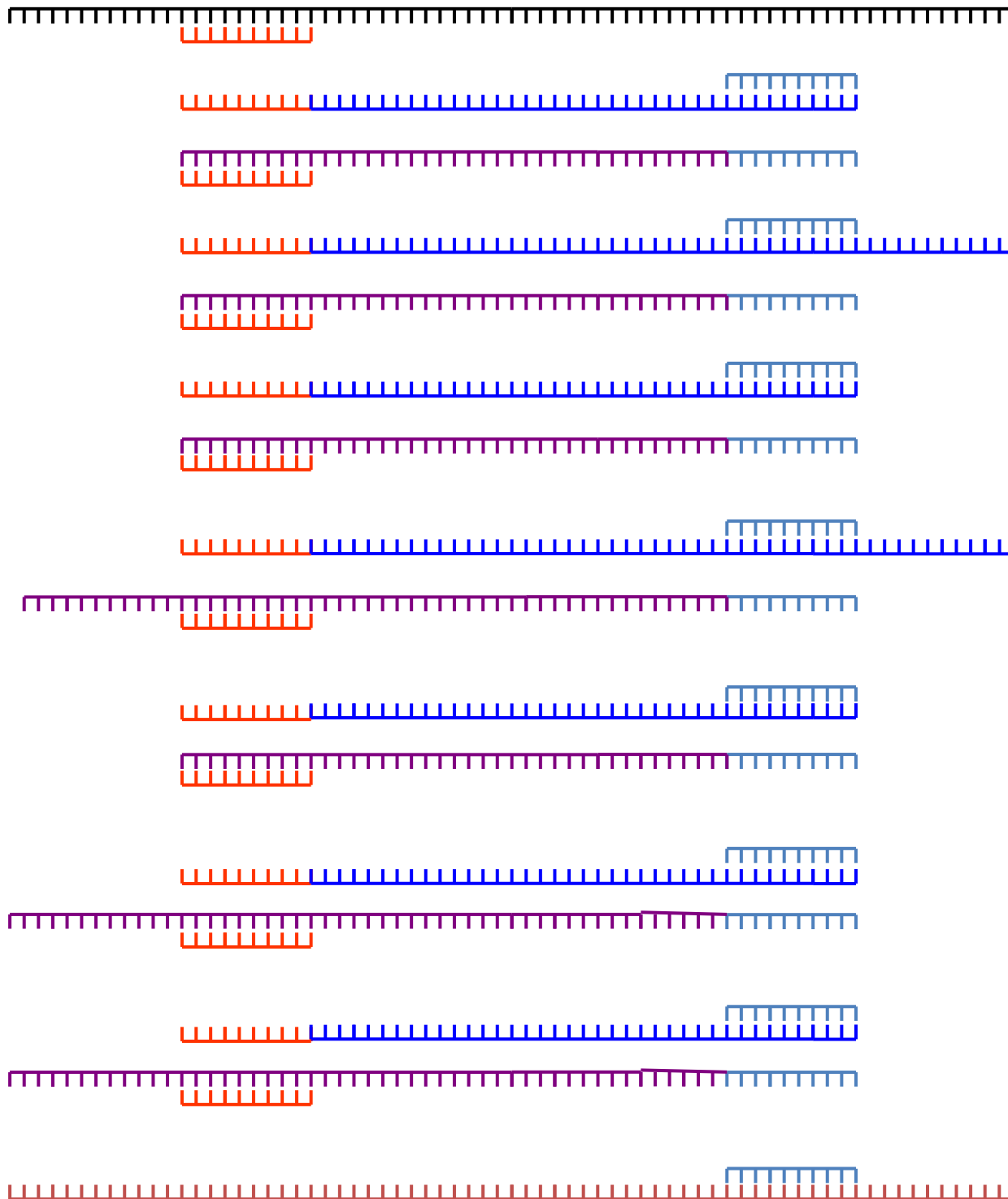
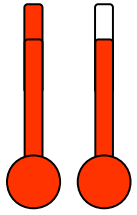
3



4



Third round of cDNA synthesis (16 strands)



1

L.  
Denature at 96°  
Anneal primers  
at 50°

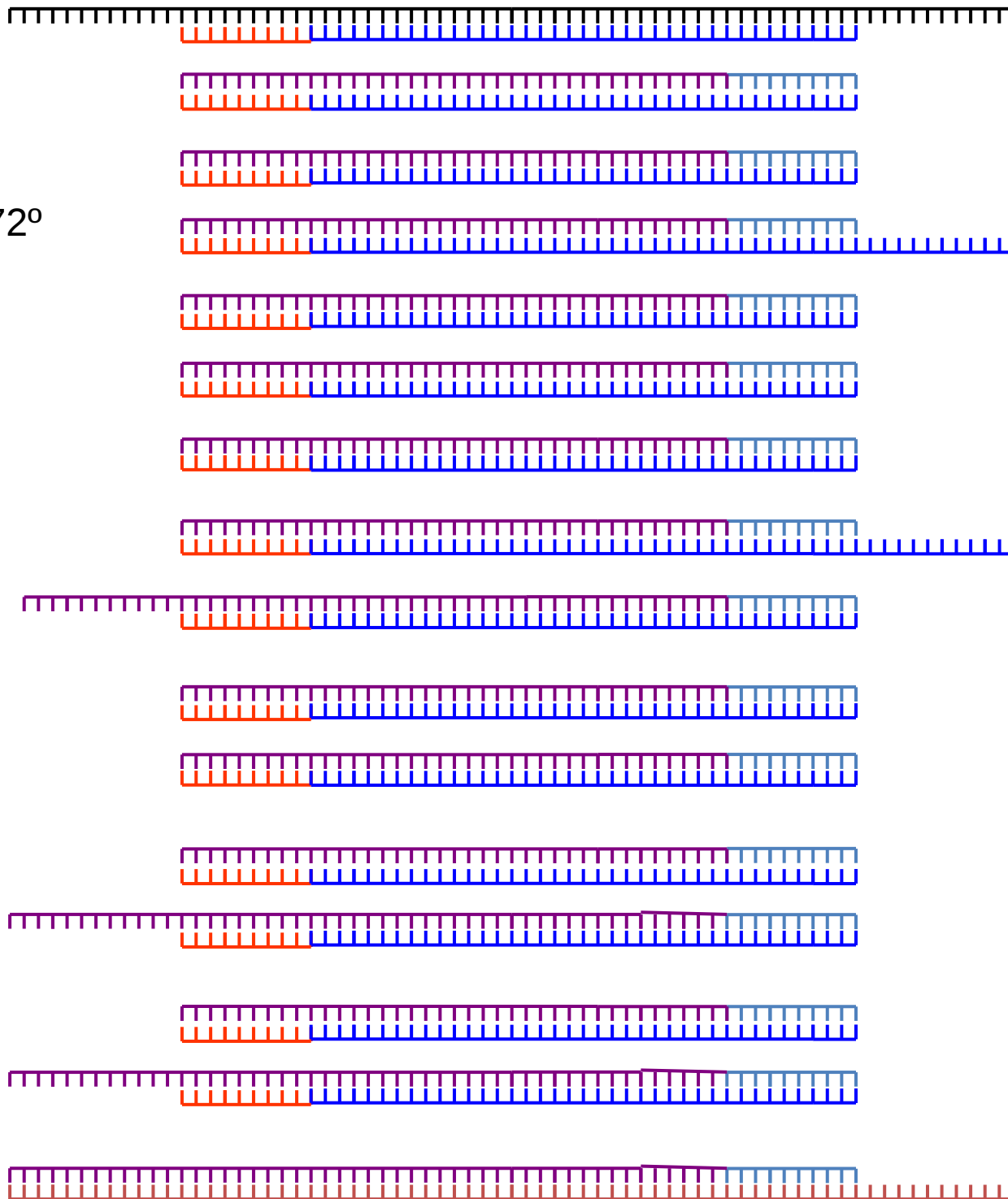
2

3

4



72°



1

M.  
Copy strands at  
72°

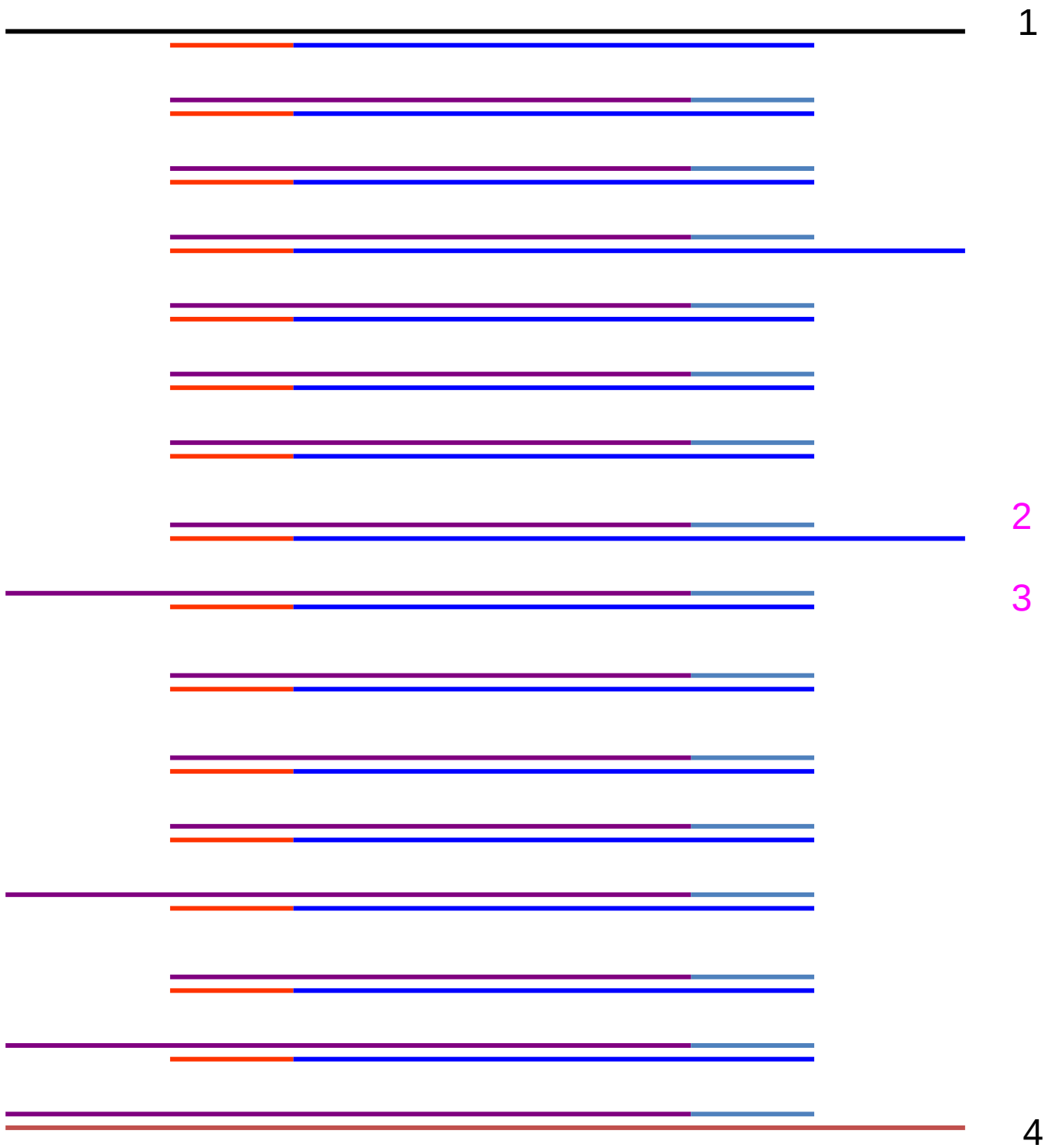
2

3

4

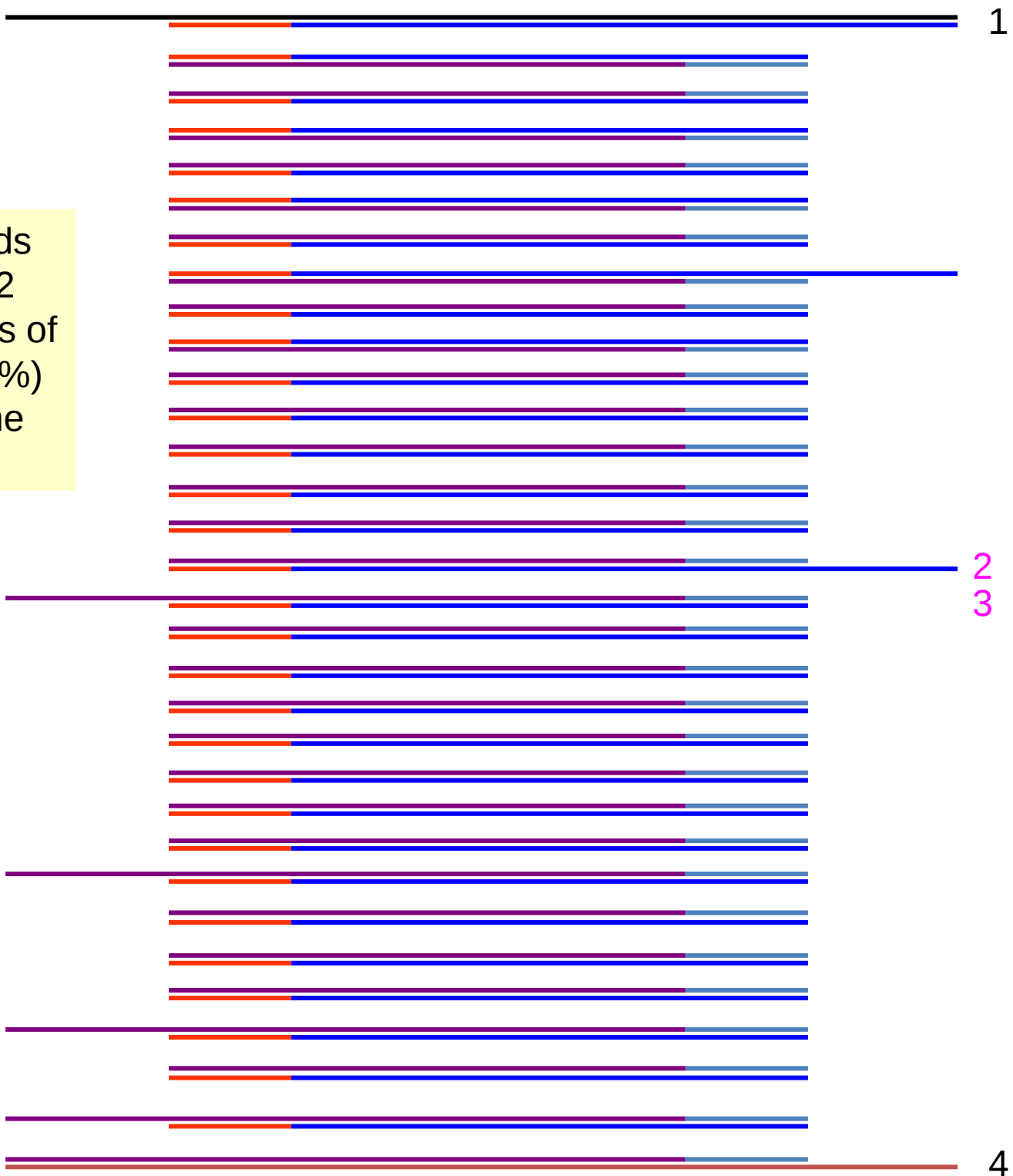
Fourth  
round of  
cDNA  
synthesis (32  
strands)

cDNA strands (32) are now shown as lines





After 5 rounds  
there are 32  
double strands of  
which 24 (75%)  
are are same  
size



1

2  
3

4

# Мощь экспоненты

## Изобретение шахмат

$$T_{64} = 1 + 2 + 4 + \dots + 2^{63} = 2^{64} - 1 = 18\,446\,744\,073\,709\,551\,615 = 18E18$$

Количество зерна примерно в 1800 раз превышает мировой урожай пшеницы за год (в 2008-09 аграрном году урожай составил 686 млн тонн), то есть **превышает весь урожай пшеницы, собранный за всю историю человечества.**

В единицах массы: если принять, что одно зёрнышко пшеницы имеет массу 65 мг, тогда общая масса пшеницы на шахматной доске составит 1200 миллиардов тонн или 1,2 триллиона тонн:

$$18446744073709551615 \times 0,065 \text{ g} / (1000000 \text{ g/t}) \approx 1,2 \text{ E}12 \text{ t}$$

В то время как количество зёрен на первой половине доски велико, количество на второй половине многократно его превышает. Количество зёрен на **первой половине доски** составляет  $1 + 2 + 4 + \dots + 2^{147} = 2^{148} - 1 = 4\,294\,967\,295$  зёрен, или около **100 000 кг риса при массе одного зёрнышка 25 мг**

Количество зёрен составляет примерно  $3E-5$  числа Авогадро.

Генетич анализ – 3 цикл

1	2	4	8	16	32	64	128
256	512	1024	2048	4096	8192	16384	32768
65K	131K	262K	524K	1M	2M	4M	8M
16M	33M	67M	134M	268M	536M	1G	2G
4G	8G	17G	34G	68G	137G	274G	549G



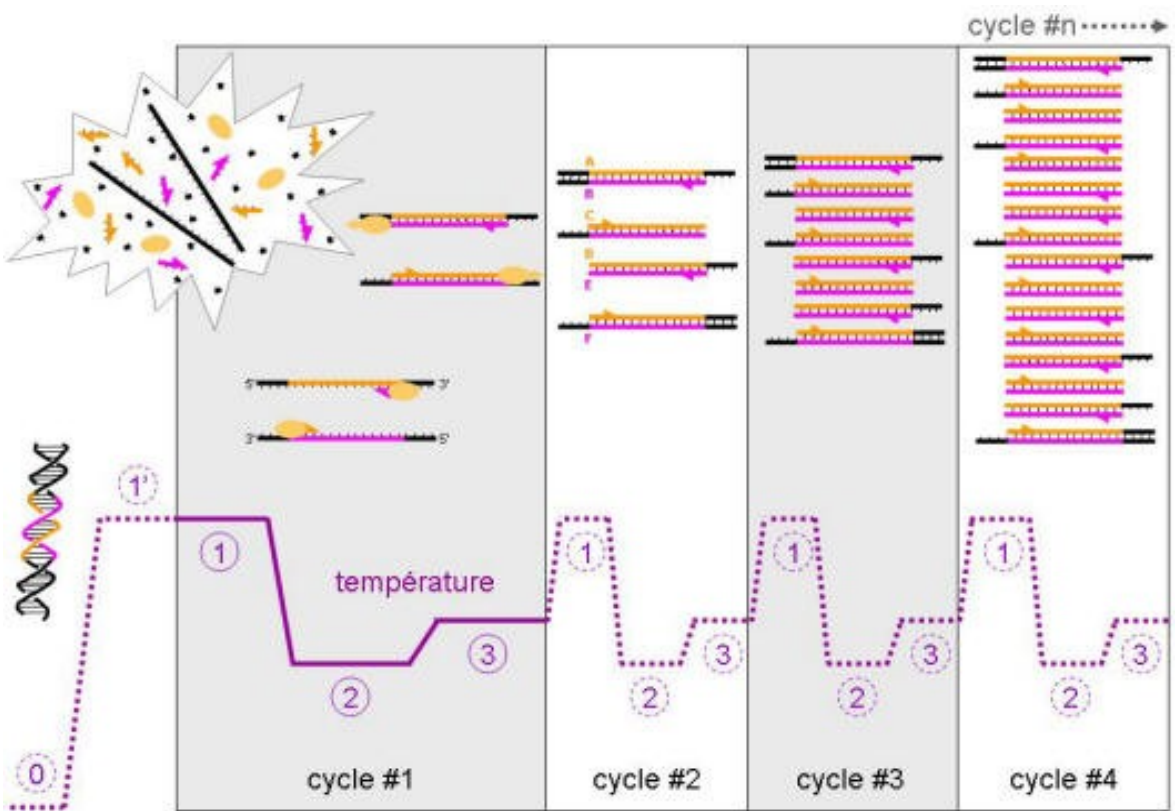
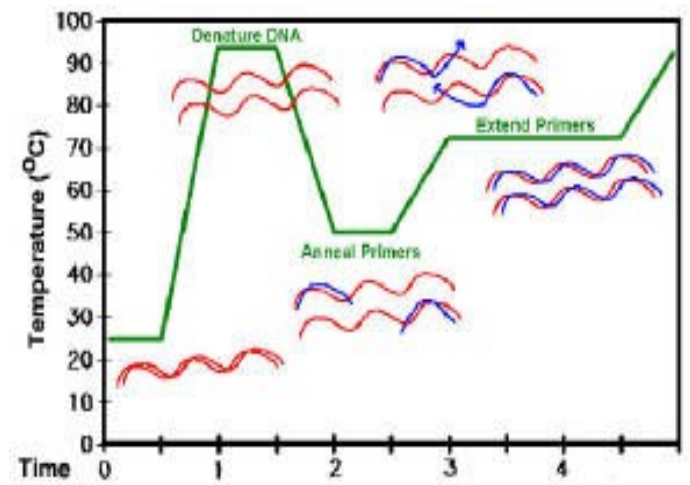
## ДНК из мумий

Тутанхамон был сыном Эхнатона (Аменхотепа) и его неидентифицированной сестры

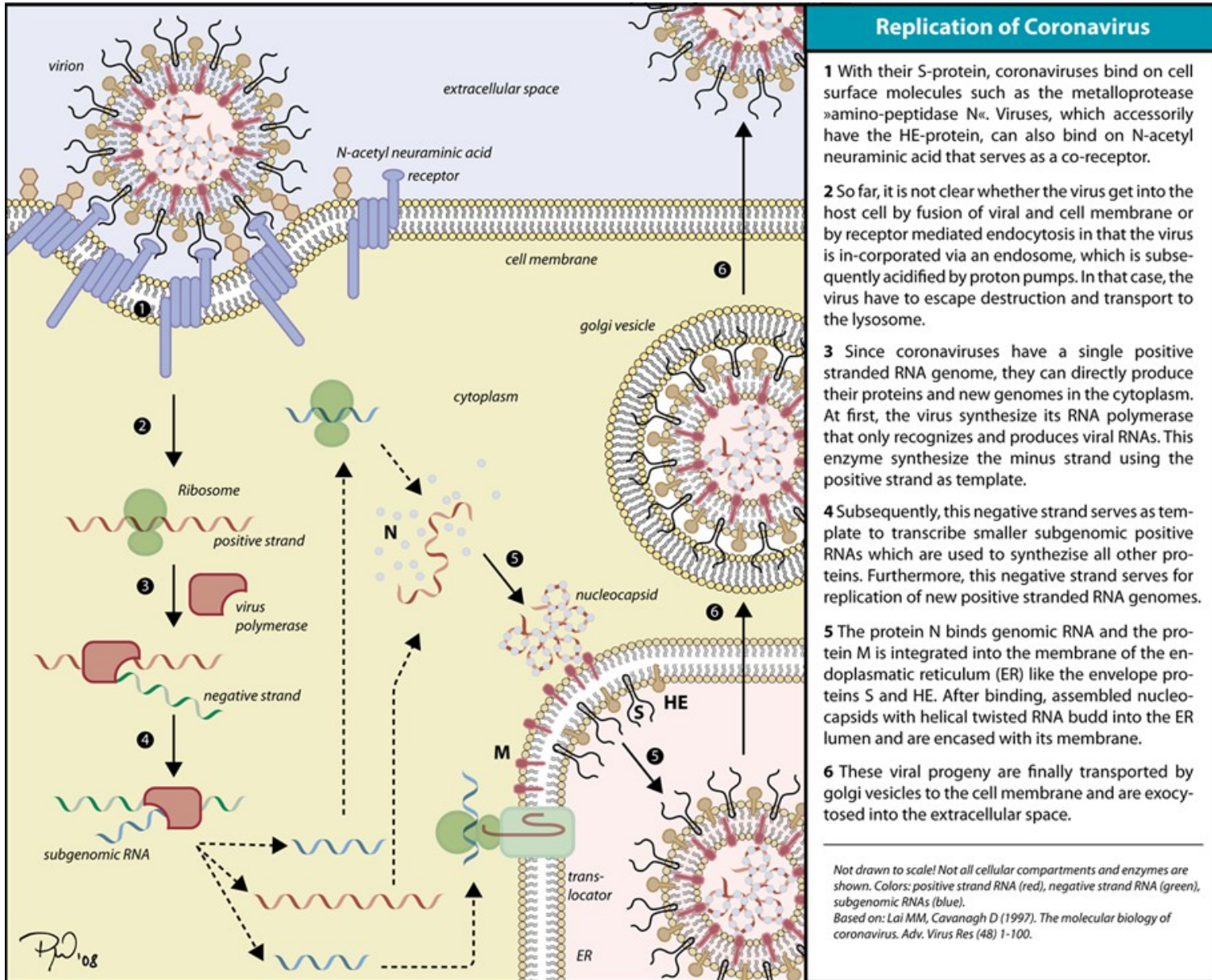
Тутанхамон был женат на Анхесенамон, третьей дочери Эхнатона,

Ancestry and Pathology in King Tutankhamun's Family // JAMA: the Journal of the American Medical Association. — 2010.  
DOI:10.1001/jama.2010.121. — PMID 20159872.

# Циклирование температурного режима



# Репликация коронавируса





# Как “увидеть” продукт ПЦР и следить на его накопением в режиме реального времени?

## Системы детекции

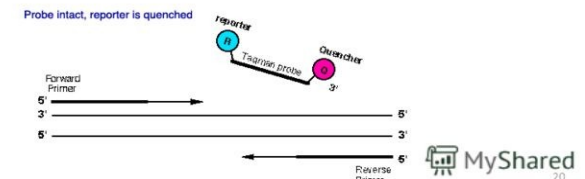
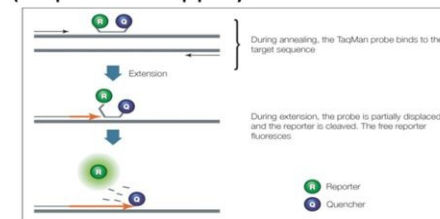
### 1. Неспецифичные

- интеркалирующие красители (SYBR Green и др.)

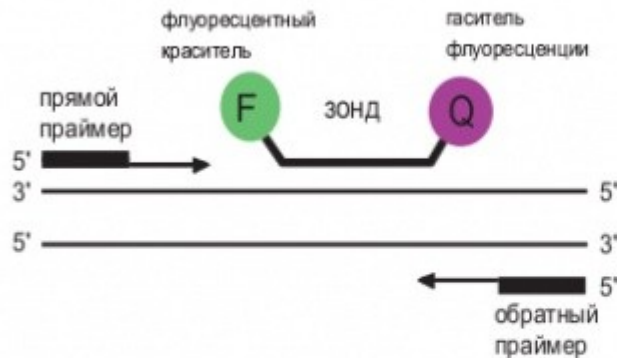
### 2. Специфичные

- линейные разрушаемые пробы (TaqMan)
- «молекулярные маячки» (beacons)
- примыкающие пробы

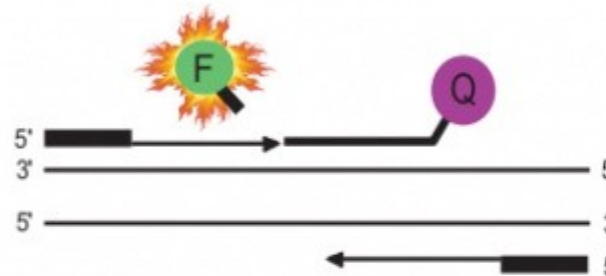
## Линейные разрушаемые пробы (TaqMan = Taq-polymerase + Pac-Man)



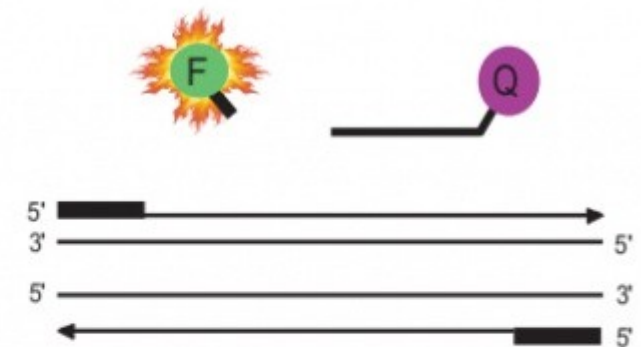
MyShared 19



1 - зонд отжигается на одной из цепей между прямым и обратным праймерами



2 - Taq ДНК полимераза расщепляет зонд, и флуоресценция разгорается



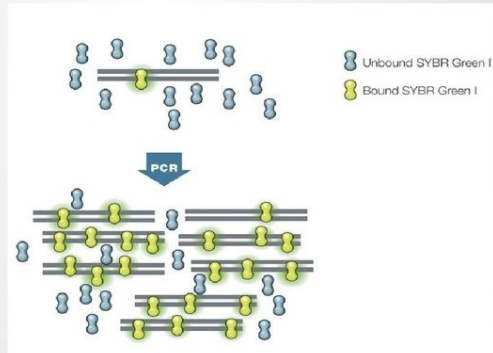
3 - Удвоение ампликона вызывает разгорание одного флуорофора

MyShared 20

## 1. Применение интеркалирующих флуоресцентных агентов. Принцип SYBR Green.

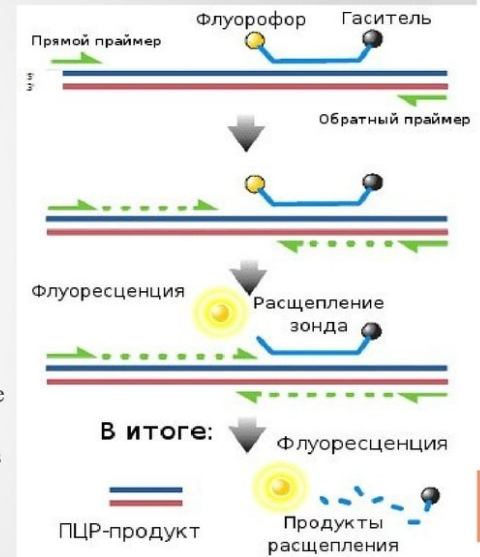
В качестве интеркалирующего красителя наиболее часто используют SYBR Green.

Особенностью красителя является его способность образовывать комплекс только с двухцепочечной молекулой ДНК, которая существует в процессе полимеразной цепной реакции только на стадии отжига (праймер-матрица ДНК) и элонгации (ампликон).



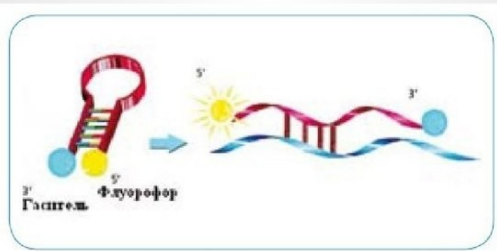
## 2. Использование меченых лигонуклеотидных проб. TaqMan протокол.

TaqMan PCR основан на использовании 5'-экзонуклеазной активности полимеразы. В реакцию добавляют ДНК-пробы, меченные на 5'-конце флуоресцентным красителем, а на 3'-конце - фосфатной группой и гасителем флуоресценции. Пробы комплементарны участку амплифицируемой области. Гаситель поглощает испускаемое флуоресцентной меткой излучение, а фосфатная группа в 3'-положении блокирует полимеразу.

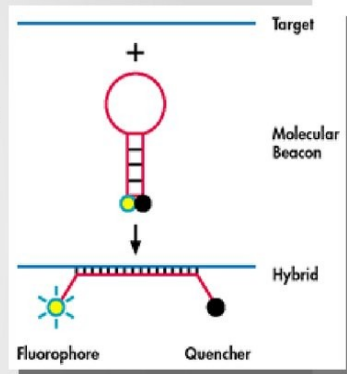


## Метод Molecular Beacons.

Molecular Beacons отличается от TaqMan тем, что концы пробы комплементарны друг другу. В свободном состоянии молекулярные маяки образуют «шпильку» (структуру, когда 5' и 3'-концы зонда гибридизуются между собой).

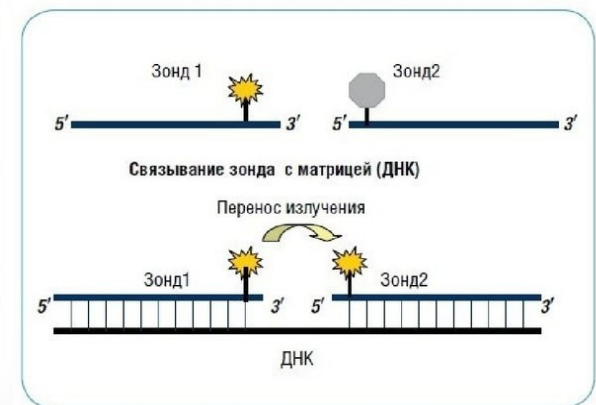


Флуорогенная шпилька - небольшая одноцепочечная молекула ДНК, которая в свободном состоянии способна образовывать вторичную структуру. При гибридизации пробы с матрицей вторичная структура разрушается



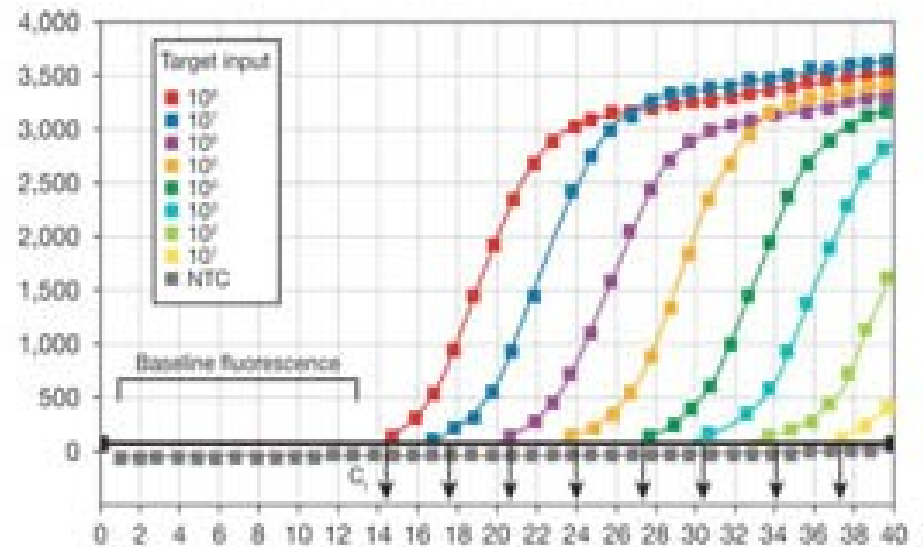
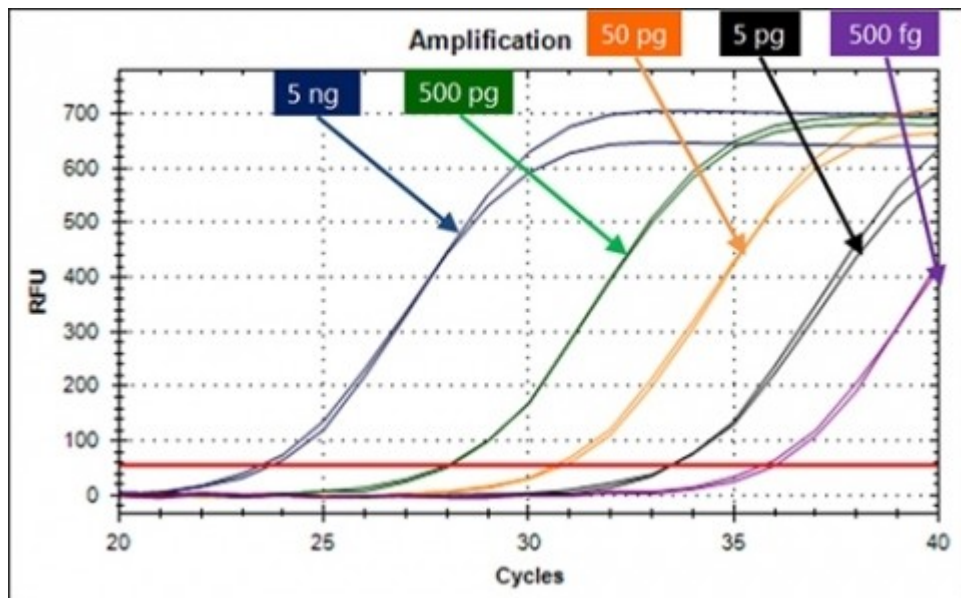
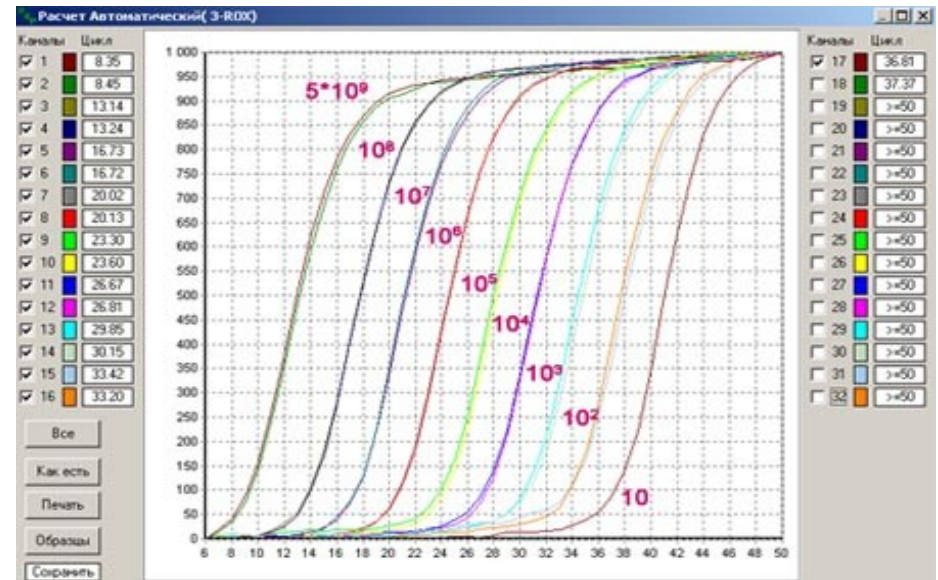
## Применение 2-х зондов с резонансным переносом энергии. Метод LightCycler.

Принцип метода заключен в переносе энергии от одного флуорофора, находящегося на 3'-конце первого зонда, ко второму, находящемуся на 5'-конце второго зонда, который происходит, если расстояние между двумя флуорофорами составляет 1-3 нуклеотида.



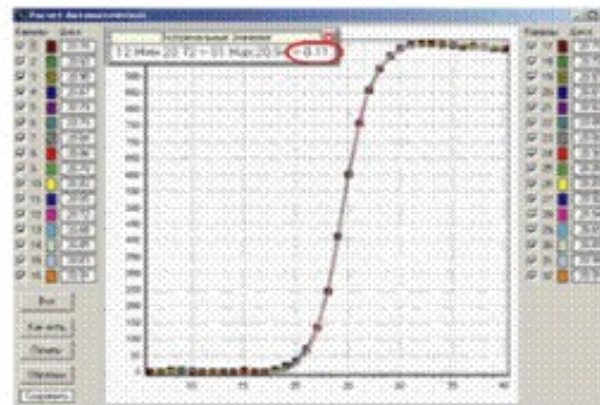


# Кинетика амплификации ДНК

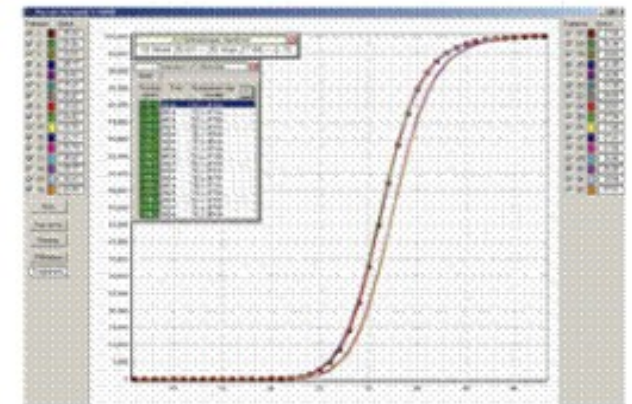


# Кинетика амплификации ДНК

КИНЕТИЧЕСКИЕ КРИВЫЕ АМПЛИФИКАЦИИ ФРАГМЕНТА ДНК ГЕНА ИММУНОГЛОБУЛИНА G ЧЕЛОВЕКА

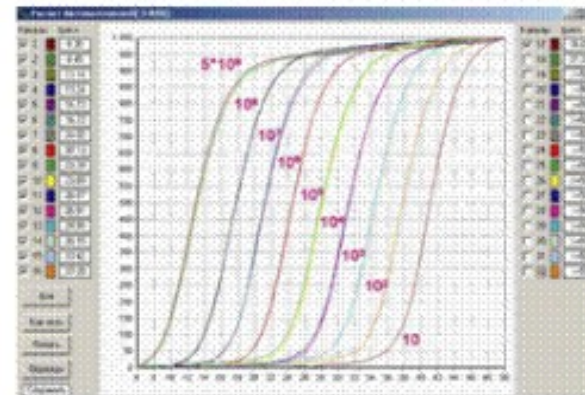


Высокая точность и воспроизводимость результатов. Разброс значений пороговых циклов для 32 одинаковых образцов составляет  $\pm 0.11$



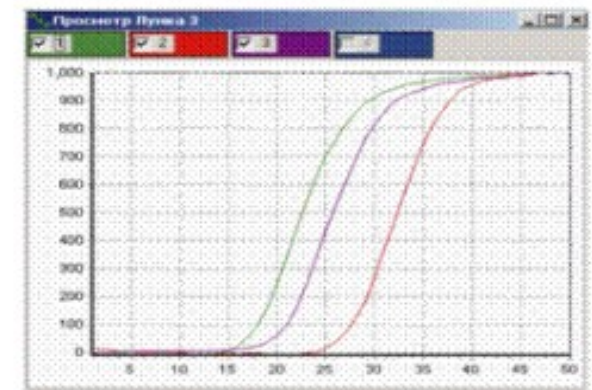
Возможность идентифицировать незначительные различия в исходном количестве ДНК (в данном случае количество ДНК в образцах – 5 и 10 тысяч копий)

ТЕСТ-СИСТЕМА ДЛЯ ОБНАРУЖЕНИЯ ДНК ГМИ (35 S -ПРОСМОТР)



Широкий динамический диапазон количественного определения ДНК от единичных до  $10^9$  копий

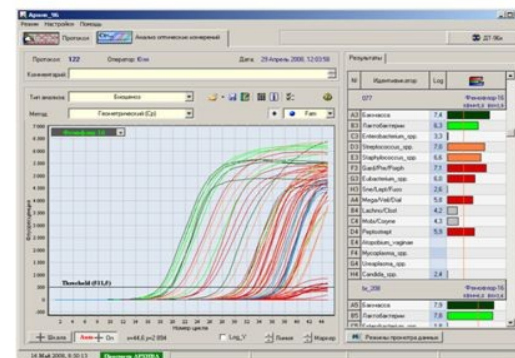
ТЕСТ-СИСТЕМА ДЛЯ ОБНАРУЖЕНИЯ ДНК M. tuberculosis, ВОЗМОЖНОСТЬ ПОСТАНОВКИ МУЛЬТИПЛЕКСНОЙ ПЦР



Возможность постановки мультиплексной ПЦР. Первый цветной канал – обнаружение фрагмента ДНК IS6110; второй – внутренний положительный контроль; третий – обнаружение фрагмента ДНК *senX3-regX3* региона генома M. Tuberculosis



## Оборудование и программное обеспечение



Детектирующий термоциклер ДТ-96, «ДНК-Технология»

28 ПЦР - это ДНК-Технология



14 октября 2012 г.

# ПЦР В РЕАЛЬНОМ ВРЕМЕНИ

Издательство **БИНОМ**  
ozon.ru

НПФ ДНК-Технология  
2009

### ТЕСТ-СИСТЕМА ДЛЯ ОБНАРУЖЕНИЯ ВИРУСА ПАРАГРИППА-3 МЕТОДОМ ПОЛИМЕРАЗНОЙ ЦЕПНОЙ РЕАКЦИИ В РЕАЛЬНОМ ВРЕМЕНИ

«Для животных»  
Хранение набора I при температуре от 2°C до 6°C,  
набора II – при температуре от минус 18°C до минус 20°C

Количество анализов: **50**

серия №: \_\_\_\_\_ изготовлена: \_\_\_\_\_  
контроль №: \_\_\_\_\_ годна до: \_\_\_\_\_

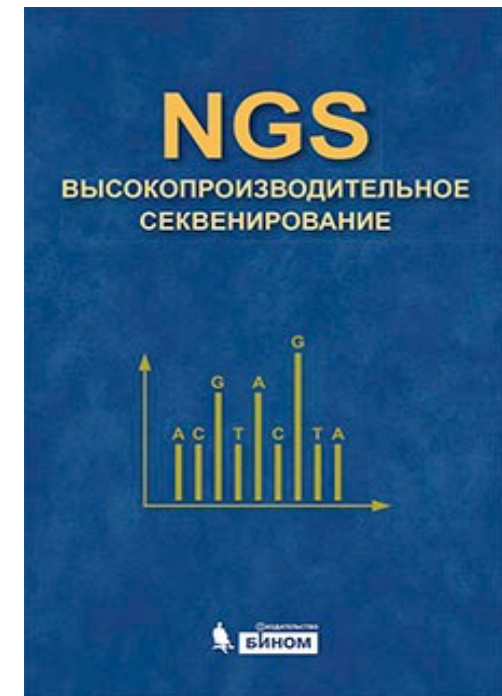
Производитель – ООО «ВЕТБИОХИМ», Россия, г. Москва

Медицина, ветеринария

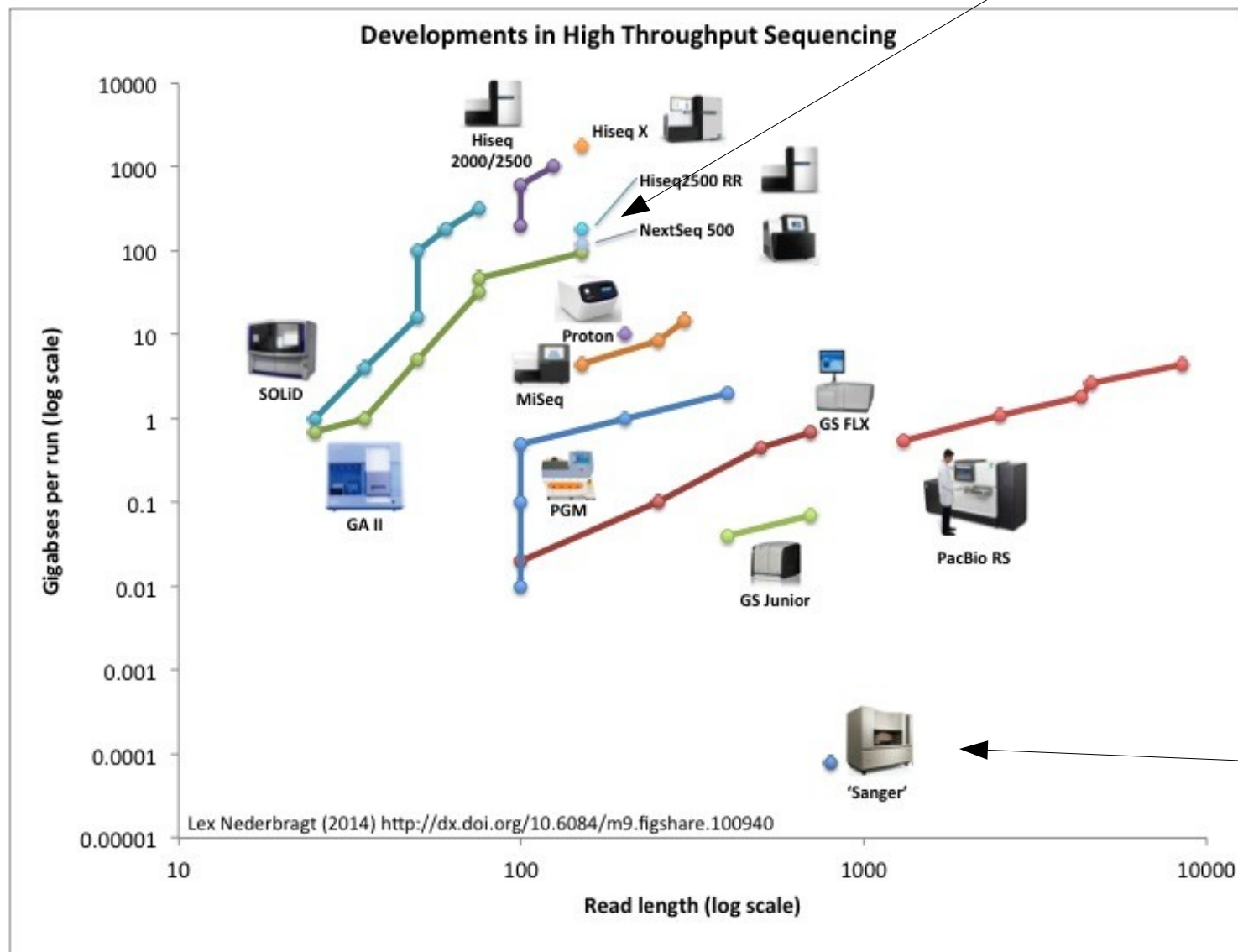
Определение первичной структуры ДНК -  
последовательности нуклеотидов – это  
секвенирование (sequencing)  
~~“первичная последовательность”~~

Три поколения методов секвенирования:

- классика, метод Сэнгера
- высокопроизводительное  
 (“следующего поколения”, NGS)
- единичные молекулы

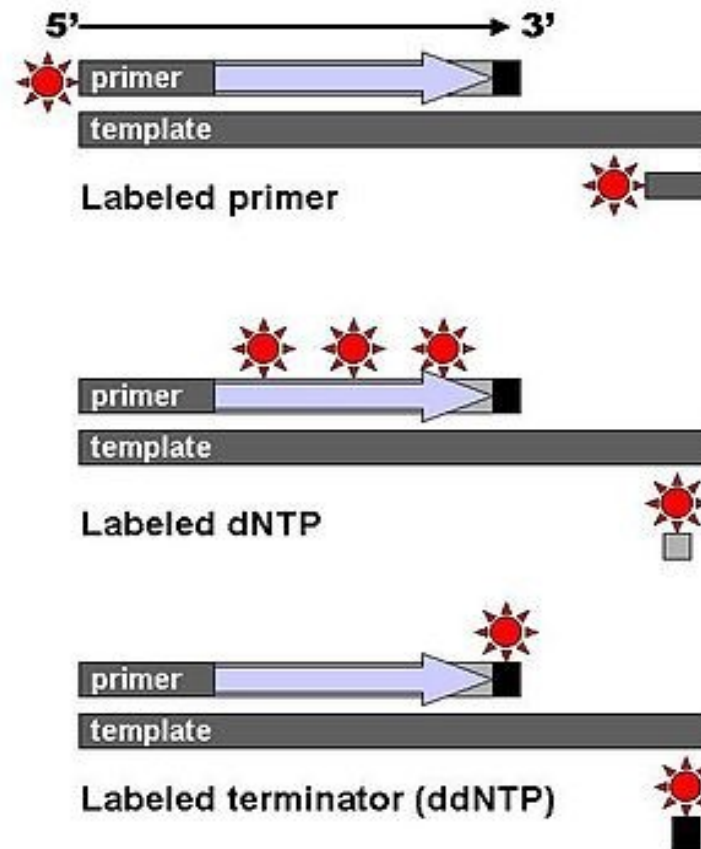


# Динамика развития высокопроизводительного секвенирования



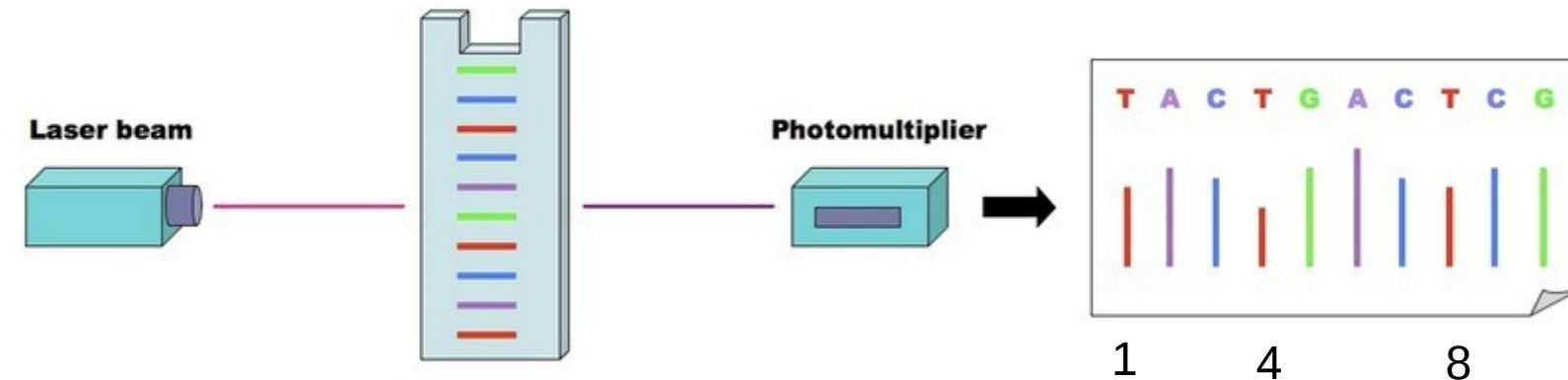
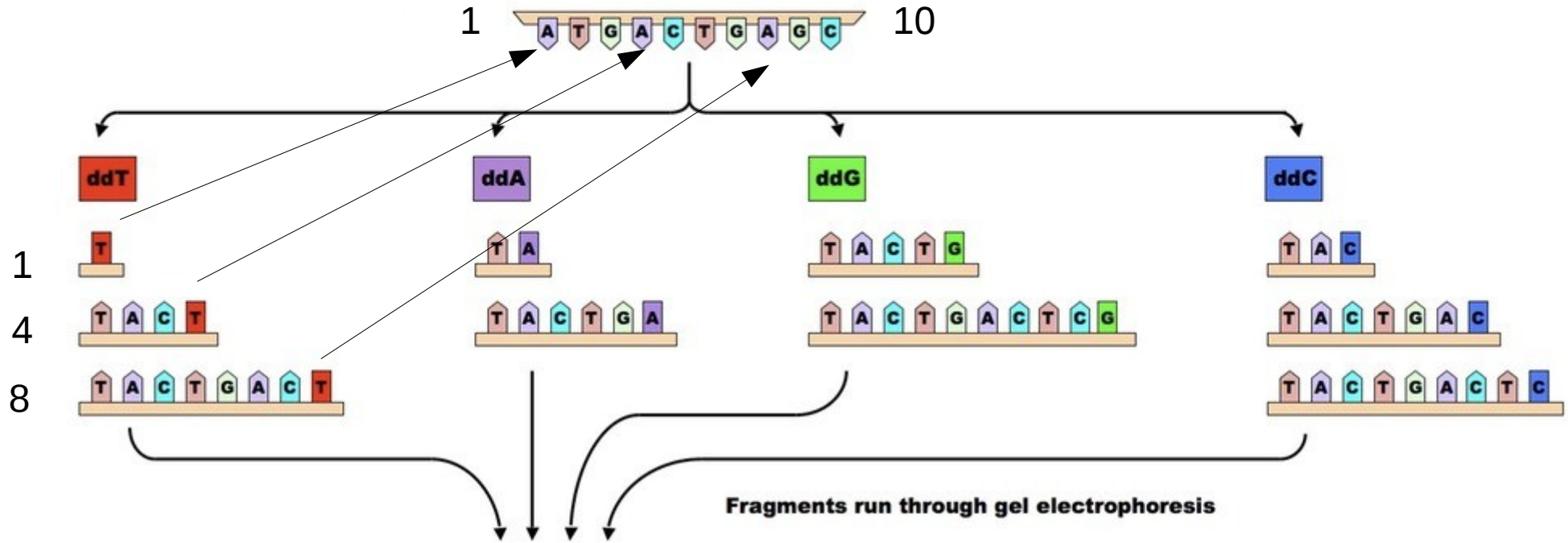
# ДНК-полимеразное секвенирование

## Три варианта введения метки в новую цепь



# Получение фрагментов в помощью флуоресцентных терминаторов (вспомните ацикловир)

PCR in presence of fluorescent, chain-terminating nucleotides



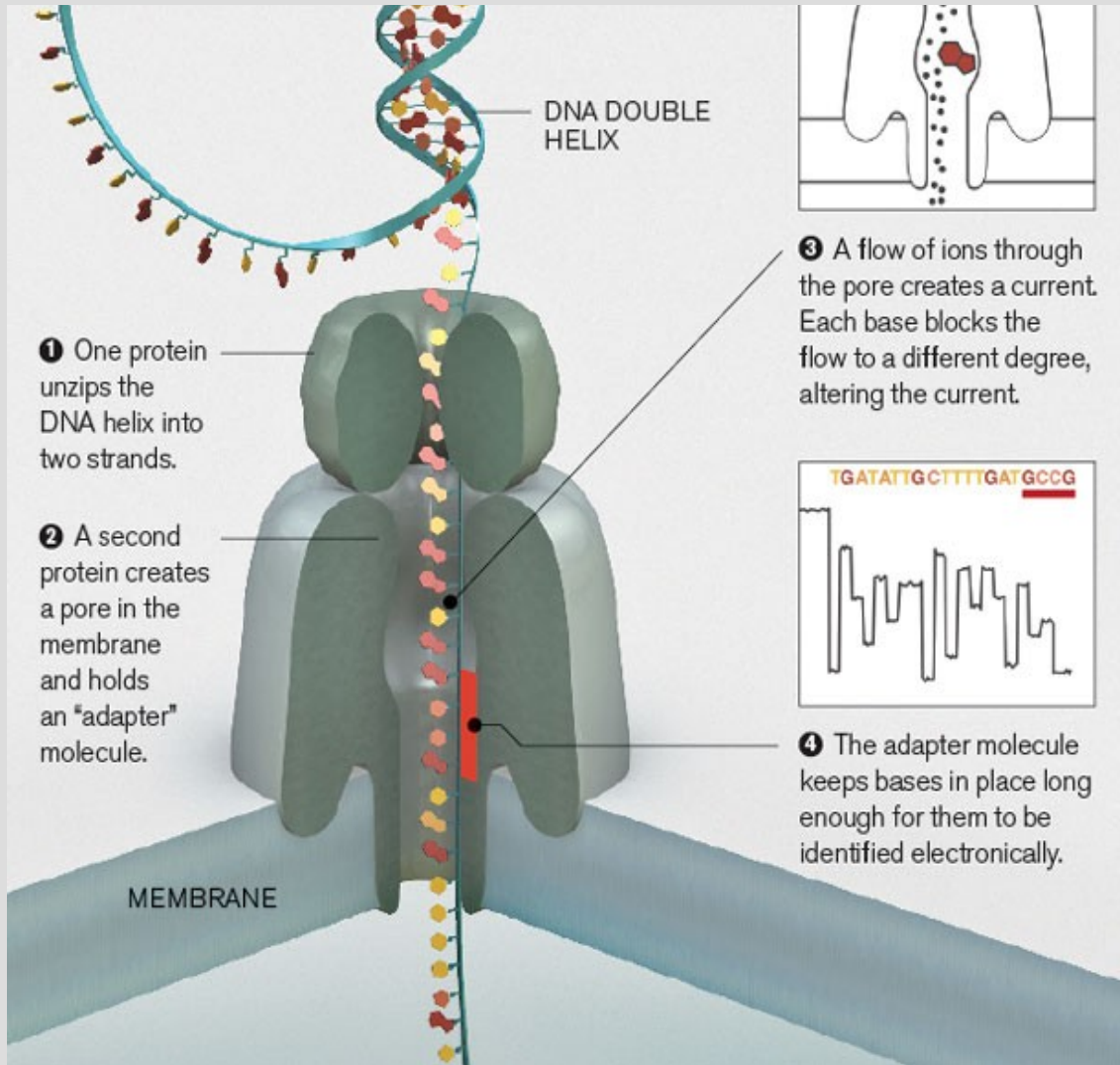
Fluorescent fragments detected by laser and represented on a chromatogram



# Секвенирование единичных молекул.

## Oxford Nanopore

Принцип основан на использовании мембран с белковыми нанопорами, через которые протягивается молекула ДНК. Секвенатор размером с USB-диск. Высокая скорость, очень высокая частота ошибок, низкая производительность (пока!)





# Секвенирующие центры

Институт Сэнгера, Великобритания. 1992



## BGI



Филиал МГУ?

# Области применения NGS

- **секвенирование геномов и транскриптомов de novo**

- начало большинства молекулярно-биологических и генетических исследований на немодельных объектах,

- поиск крупных геномных перестроек

- **полногеномное ресеквенирование**

поиск мутаций, ассоциированных с болезнями,

картирование генов,

геномы отдельных типов клеток,

анализ древней ДНК

- **направленное ресеквенирование**

биомедицина: скрининг мутаций с известной ролью в развитии болезней и поиск новых мутаций

- **анализ транскриптома**

сравнение уровней экспрессии, поиск новых генов и изоформ, аннотация de novo секвенированных геномов

- **ДНК-белковые и ДНК-ДНКовые взаимодействия**

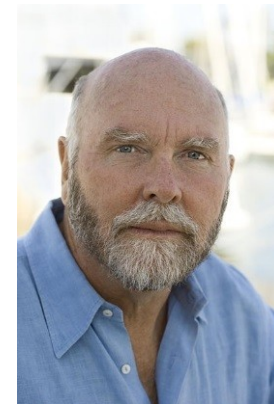
поиск сайтов связывания транскрипционных факторов, изучение пространственной организации хроматина

- **метагеномика**

анализ разнообразия микробных сообществ



# Определена первичная структура многих геномов. Зачем много секвенировать и копировать «массивы»? («Слишком много букв» - из интернета)

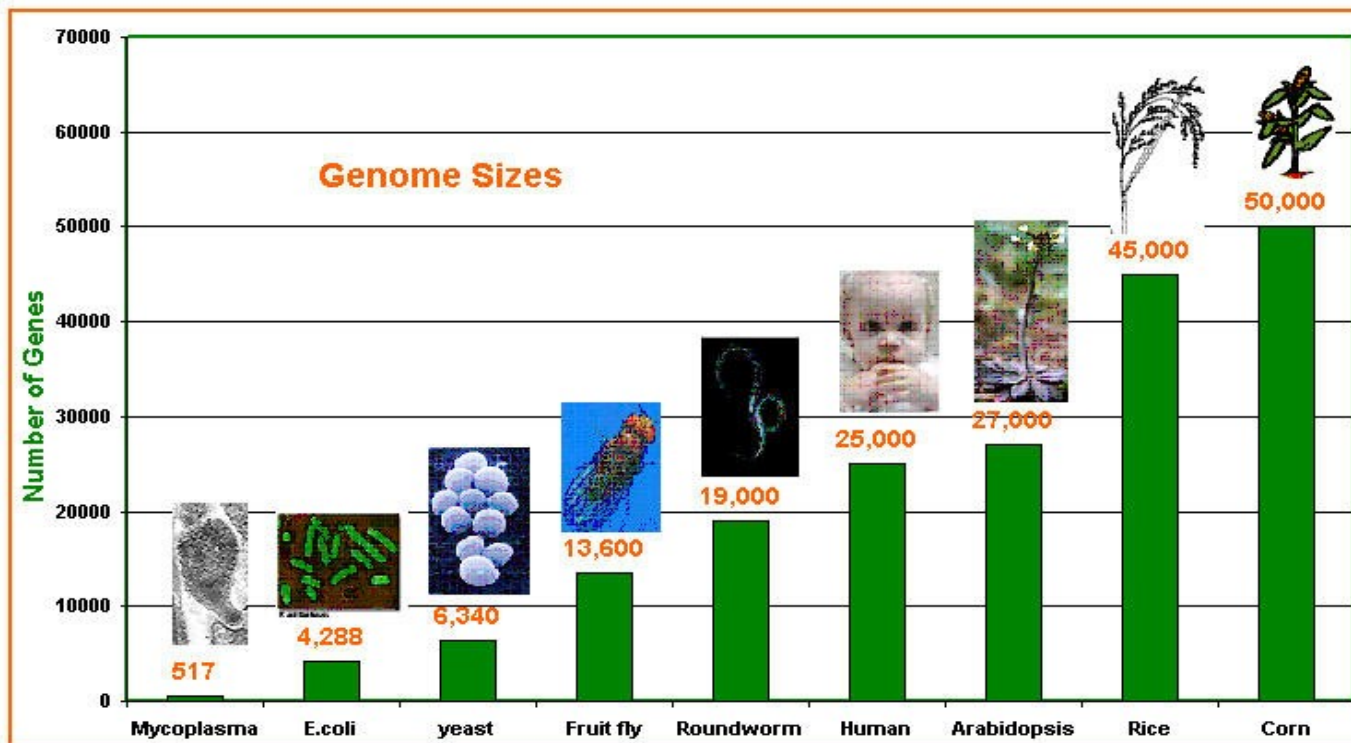


<http://www.jcvi.org/cms/home/>

Крейг Вентер  
Институт геномный исследований

In January 2012, Life Technologies introduced a sequencer to decode a human genome in one day for \$1,000

22-23 мая 2019 NGS-2019  
РНИМУ Пирогова



**Вирусы:** грипп, ВИЧ, гепатит, проч.

**Патогенные бактерии:** чума, холера, туберкулез, сифилис, язва  
(Проблема модельных бактерий: «Голден Стаф» Гузель)

**Модельные и основные эукариоты**

**Животные:** нематода, дрозфила, мышь. Человек.

**Растения:** арабидопсис, рис, кукуруза

Рынок 2018 – 3 млрд  
2024 – 22 млрд

1

# Разнообразие видов ЖИВОГО

«Эдем» Ван Кессель  
1626-1679



Dividing *Saccharomyces cerevisiae* (baker's yeast) cells.

4 μm



....aaatgcttgggca...

3

4 ?

Многочелюстные

Одноклеточные

(99%)

синтетические

• видимые глазом

видимые инструментом

невидимые и только читаемые

1665, 1674

Гук, Левенгук

2014

Вентер

2010

Вентер

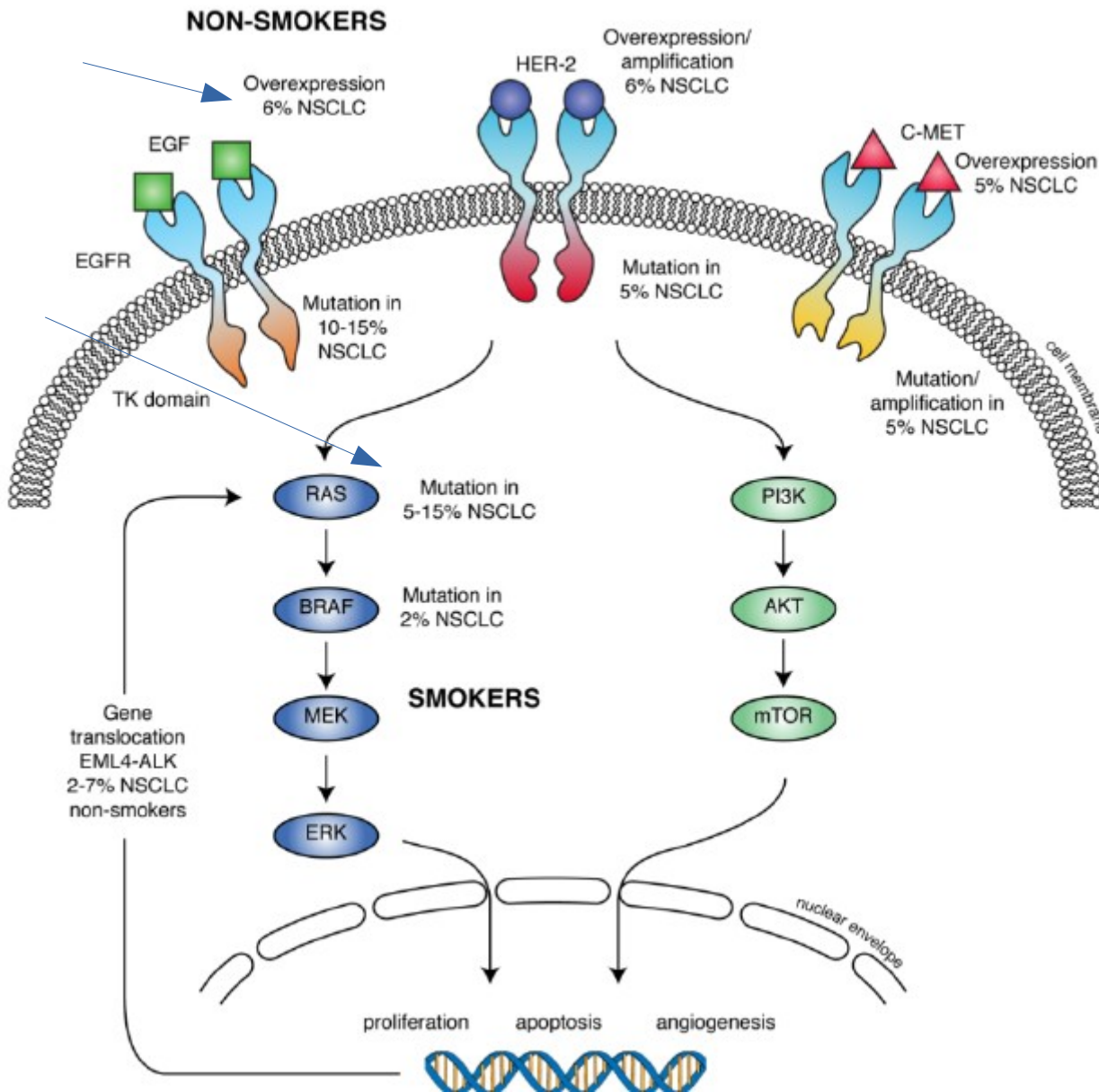
5e

Виртуальная e-клетка  
1996, Томита (127 генов)



wholecell.stanford.edu  
все 525 генов

# Немелкоклеточный рак легкого секвенирование и индивидуальная медицина?



Lung cancer remains as one of the most aggressive cancer types with nearly 1.6 million new cases worldwide each year. There are an estimated 222,520 new cases and 157,300 deaths from lung cancer in the United States in 2010 [1]. Non-small cell lung cancer (NSCLC) subtype represents 85% of all cases of lung cancer, while small cell lung cancer (SCLC) subtype comprises 15%

Mutations in at least one of the established lung cancer driver genes including egfr, kras, braf, her2, akt1, nras, pik3ca, mek1, eml4-alk and met amplification are found in approximately 60% of tumor specimens, and **greater than 90% were "exclusive": only one mutation was found in a particular tumor** [2].

Epidermal growth factor receptor (EGFR) exhibits overexpression or aberrant activation by mutations in 50 to 90% of NSCLC

EGFR pathway in NSCLC. Mutations, amplification or overexpression of growth factors receptors such as EGFR, HER-2 and C-MET are most frequent in NSCLC tumors from non-smokers patients. All these genetic alterations have been observed commonly in adenocarcinomas, women and Asiatic ethnicity. EML4/ALK fusion gene is associated to NSCLC from young and non-smokers patients. KRAS mutations and signaling pathway depending to KRAS are most frequent in smoker patients. PI3K signaling pathway modifications are most frequently observed in squamous cell carcinomas



# • Секвенирование COVID-19

## PubMed – 34 ref 05/03/202

Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL.

1 CAS Key Laboratory of Special Pathogens, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, People's Republic of China.

2 Wuhan Jinyintan Hospital, Wuhan, China.

3 University of Chinese Academy of Sciences, Beijing, People's Republic of China.

4 Hubei Provincial Center for Disease Control and Prevention, Wuhan, People's Republic of China.

5 CAS Key Laboratory of Special Pathogens, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, Wuhan, People's Republic of China. zlshi@wh.iov.cn.

### A pneumonia outbreak associated with a new coronavirus of probable bat origin.

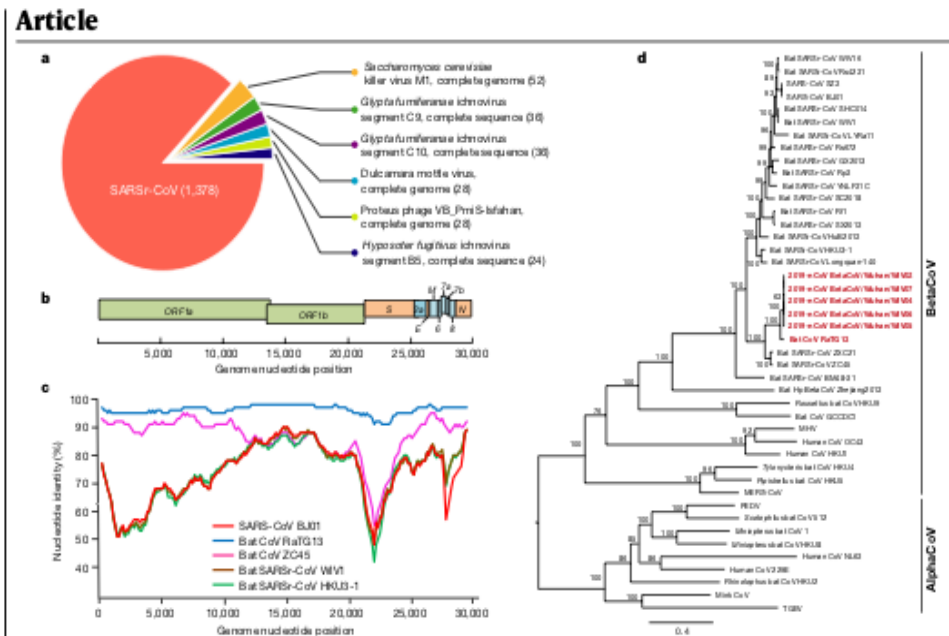
Nature. 2020 Feb 3. doi: 10.1038/s41586-020-2012-7. Epub ahead of print. PMID: 32015507.

Since the SARS outbreak 18 years ago, a large number of severe acute respiratory syndrome-related coronaviruses (SARSr-CoV) have been discovered in their natural reservoir host, bats<sup>1-4</sup>. Previous studies indicated that some of those bat SARSr-CoVs have the potential to infect humans. Here we report the identification and characterization of a novel coronavirus (2019-nCoV) which caused an epidemic of acute respiratory syndrome in humans in Wuhan, China. The epidemic, which started from 12 December 2019, has caused 2,050 laboratory-confirmed infections with 56 fatal cases by 26 January 2020. Full-length genome sequences were obtained from five patients at the early stage of the outbreak. They are almost identical to each other and share 79.5% sequence identity to SARS-CoV. Furthermore, it was found that 2019-nCoV is 96% identical at the whole-genome level to a bat coronavirus. The pairwise protein sequence analysis of seven conserved non-structural proteins show that this virus belongs to the species of SARSr-CoV. The 2019-nCoV virus was then isolated from the bronchoalveolar lavage fluid of a critically ill patient, which can be neutralized by sera from several patients. Importantly, we have confirmed that this novel CoV uses the same cell entry receptor, ACE2, as SARS-CoV.

# A pneumonia outbreak associated with a new coronavirus of probable bat origin

For qPCR analysis, primers based on the S gene of 2019-nCoV were designed: RBD-qF1, 5'-CAATGGTTTAACAGGCACAGG-3'; RBD-qR1, 5'-CTCAAGTGTCTGTGGATCACG-3'. RNA extracted as described above was used for qPCR using the HiScript II One Step qRT-PCR SYBR Green Kit (Vazyme Biotech). Conventional PCRs were also performed using the following primer pairs: ND-CoVs-951F, 5'-TGT-KAGRTTYCCTAAYATTAC-3'; ND-CoVs-1805R, 5'-ACATCYTGATAN-ARAACAGC-3'. The 20- $\mu$ l qPCR reaction mix contained 10  $\mu$ l 2 $\times$  One Step SYBR Green mix, 1  $\mu$ l One Step SYBR Green Enzyme mix, 0.4  $\mu$ l 50 $\times$  ROX Reference Dye 1, 0.4  $\mu$ l of each primer (10  $\mu$ M) and 2  $\mu$ l template RNA. Amplification was performed as follows: 50  $^{\circ}$ C for 3 min, 95  $^{\circ}$ C for 30 s followed by 40 cycles consisting of 95  $^{\circ}$ C for 10 s and 60  $^{\circ}$ C for 30 s, and a default melting curve step in an ABI 7500 Real-time PCR machine.

High-throughput sequencing, pathogen screening and genome assembly Samples from patient BALF or from the supernatant of virus cultures were used for RNA extraction and next-generation sequencing (NGS) using BGI MGISEQ2000 and Illumina MiSeq 3000 sequencers. Metagenomic analysis was carried out mainly based on the bioinformatics platform MGmapper (PE\_2.24 and SE\_2.24). The raw NGS reads were first processed by Cutadapt (v.1.18) with minimum read length of 30 base pairs. BWA (v.0.7.12-r1039) was used to align reads to a local database with a filter hits parameter of 0.8 FMM ((match + mismatch)/read length  $\geq$  fraction) value and minimum alignment score of 30.



**Fig. 1 | Genome characterization of 2019-nCoV.** a, Metagenomic analysis of next-generation sequencing of BALF from patient ICU06. b, Genomic organization of 2019-nCoV WIV04. M, membrane. c, Similarity plot based on the full-length genome sequence of 2019-nCoV WIV04. Full-length genome sequences of SARS-CoV BJ01, bat SARS-CoV WIV1, bat coronavirus RaTG13 and ZC45 were used as reference sequences. d, Phylogenetic tree based on nucleotide sequences of complete genomes of coronaviruses. MHV, murine hepatitis virus; PEDV, porcine epidemic diarrhoea virus; TGEV, porcine transmissible gastroenteritis virus. The scale bars represent 0.1 substitutions per nucleotide position. Descriptions of the settings and software that was used are included in the Methods.

