

Биология и генетика, part 1

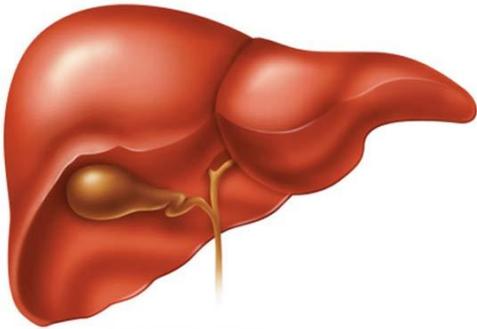
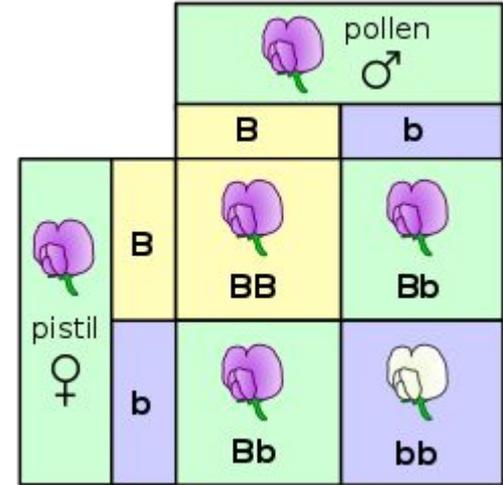
Ярослав Баранов
МНЛ«Компьютерные технологии», Университет ИТМО

Структура курса

- 8 или 10 лекций + домашние работы на stepik.org
- Цель курса - сделать нас умнее! :)
- Для успешного прохождения курса нужно выполнить не менее 70% домашних работ

1. Зачем вам биология и зачем биологи программируют.

Биология в школе



gettyimages.in

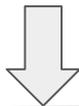


FFFFFFF
FFFFFFF
FFFFFFF
FFFUU
UUUU
UUUU
UUUU
UUUU
UUUU-

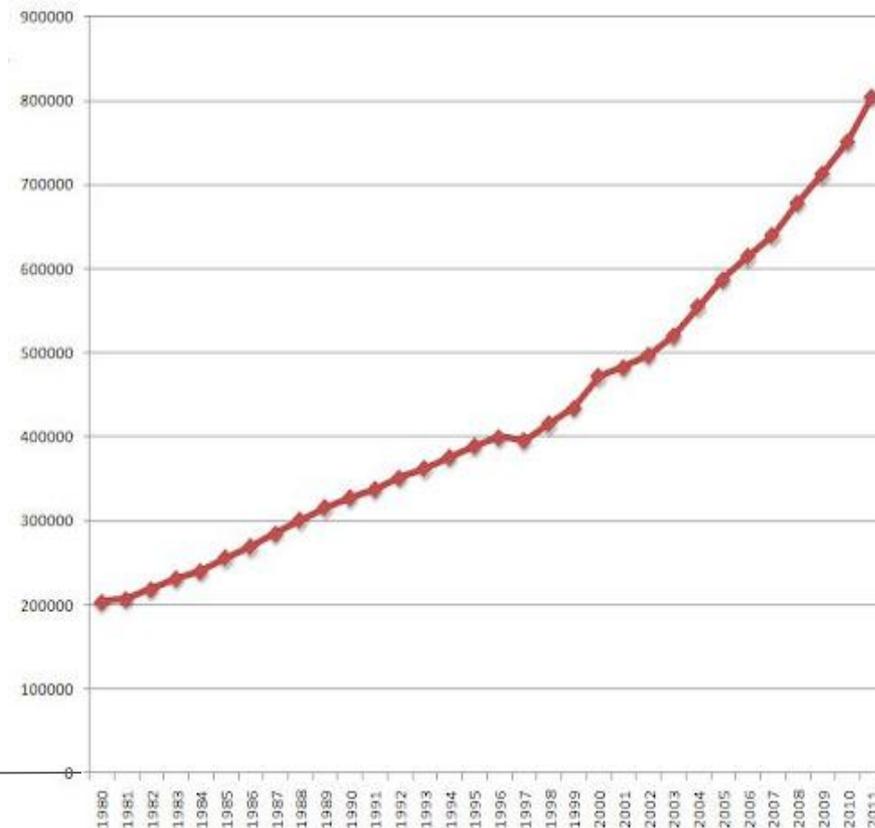
Но на самом деле биология гораздо интереснее



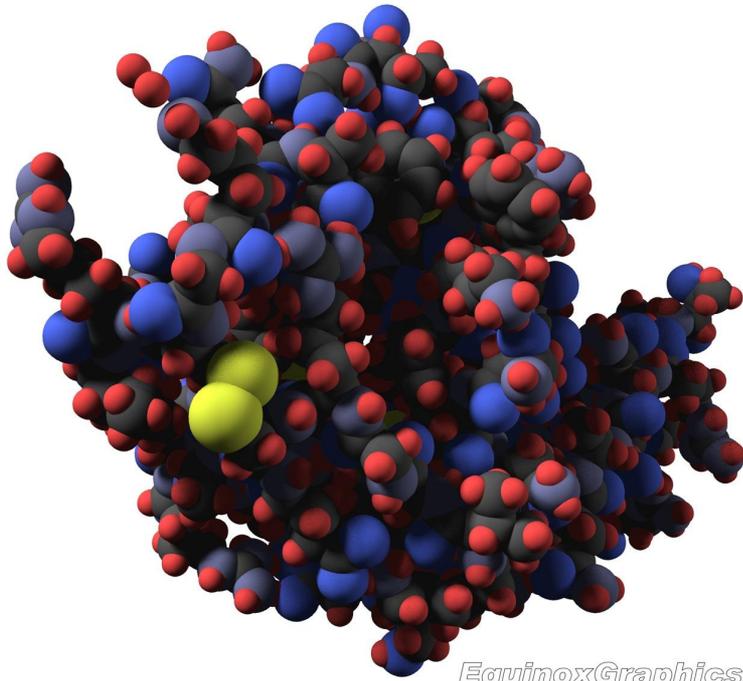
То, что изучают в школе



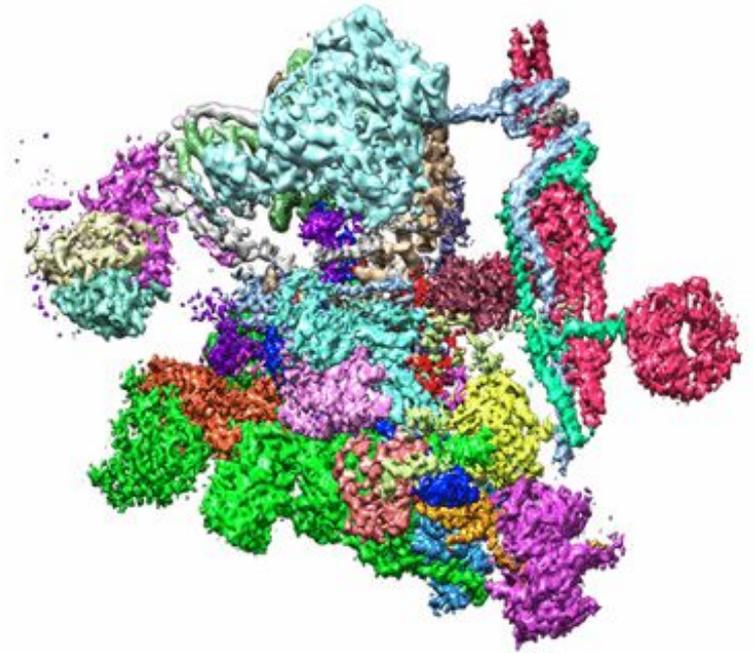
MEDLINE: English-language papers published per year 1980-2011



Биология это про молекулы



Лизоцим (белок)



Сплайсосома (комплекс белков и РНК)

Biology это про эволюцию

*"Nothing in biology
makes sense except
in the light of
evolution"*

Ф. Г. Добжанский (с)

- [Опыты Ричарда Ленски](#)
- [Резистентность к антибиотикам](#)

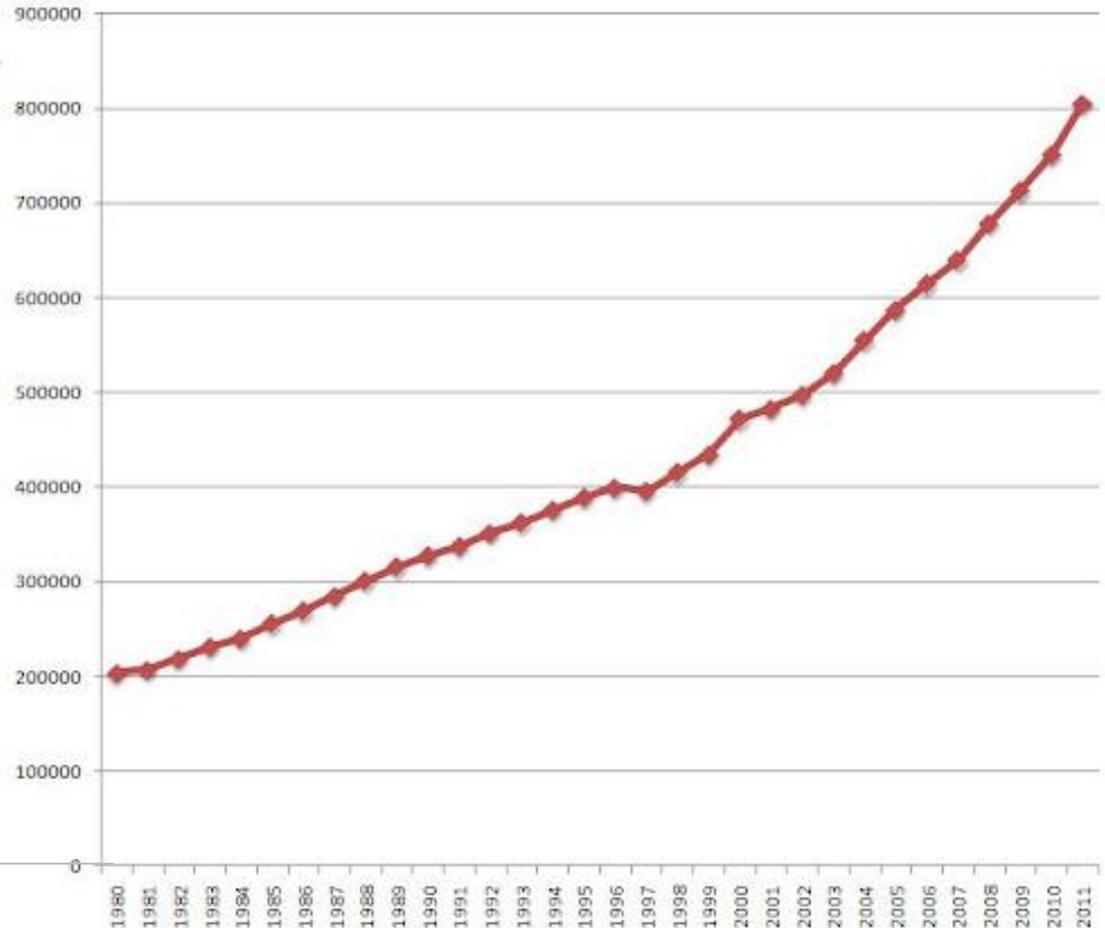


Biology was boring

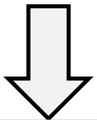
Число статей,
индексируемых в базе
PubMed/MEDLINE



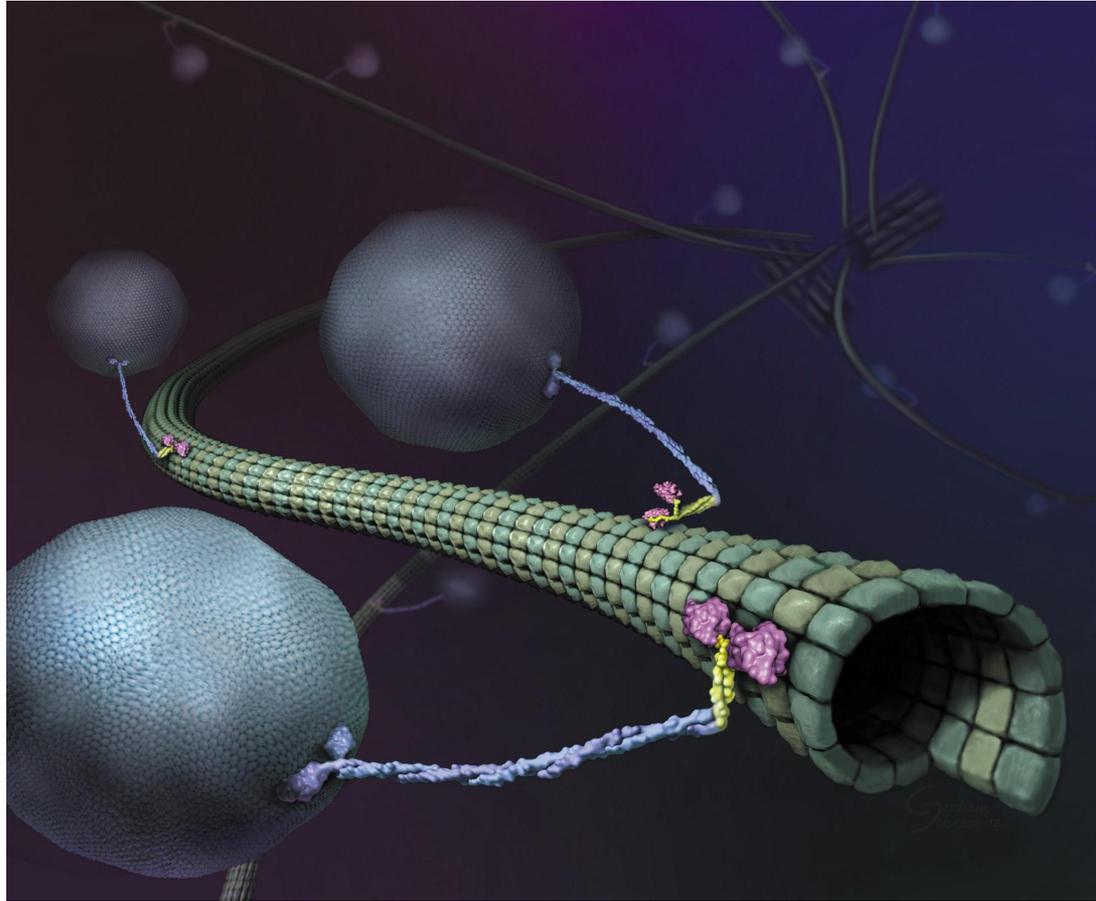
MEDLINE: English-language papers published per year 1980-2011



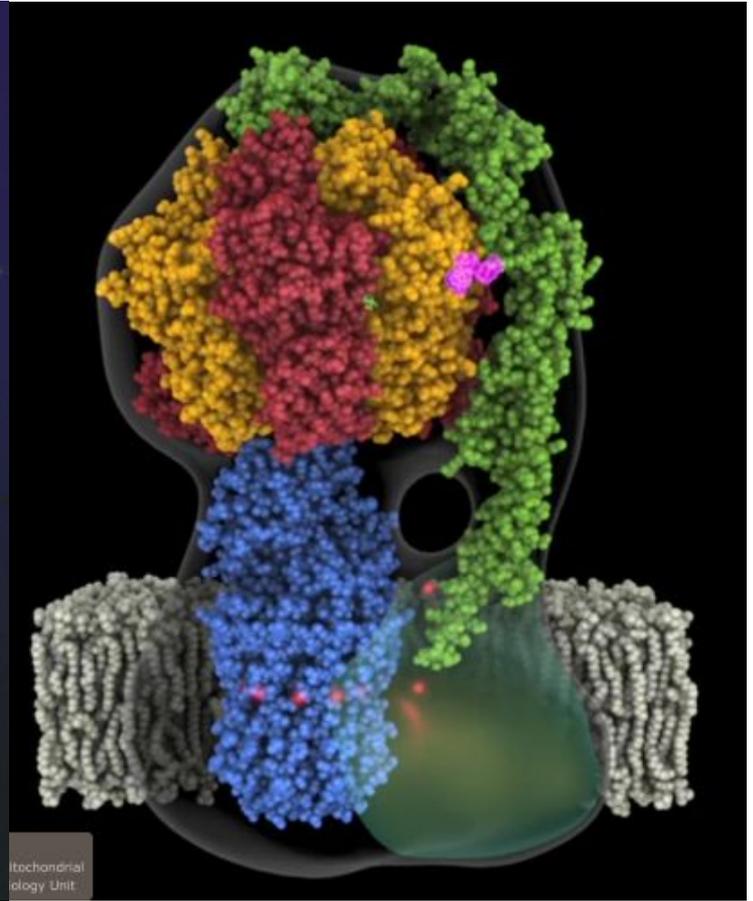
Биология
в школе



Биология это про машины!

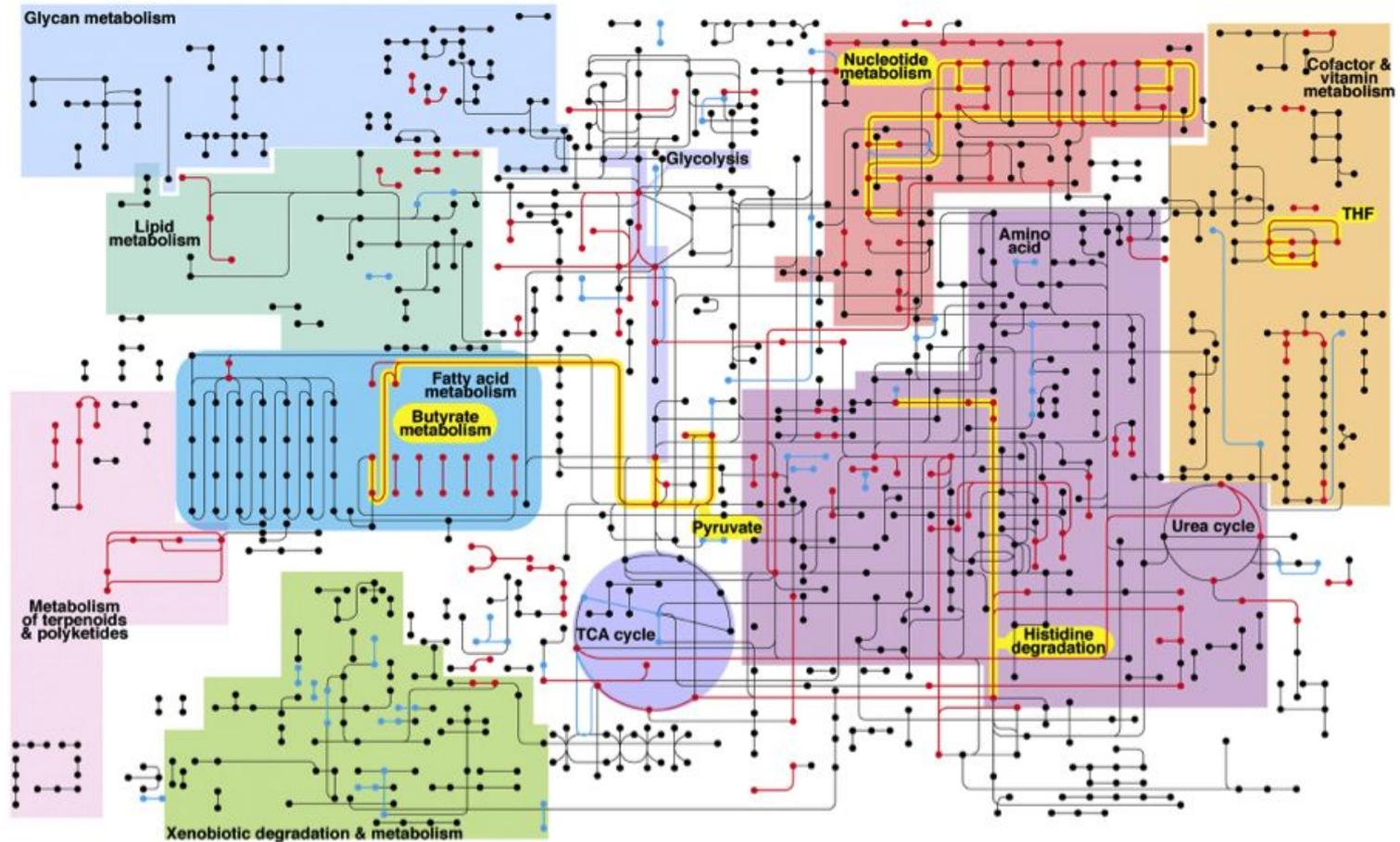


Кинезины и микротрубочки



АТФ-синтаза

Биология это про сложные системы



Метаболическая сеть *Escherichia coli*

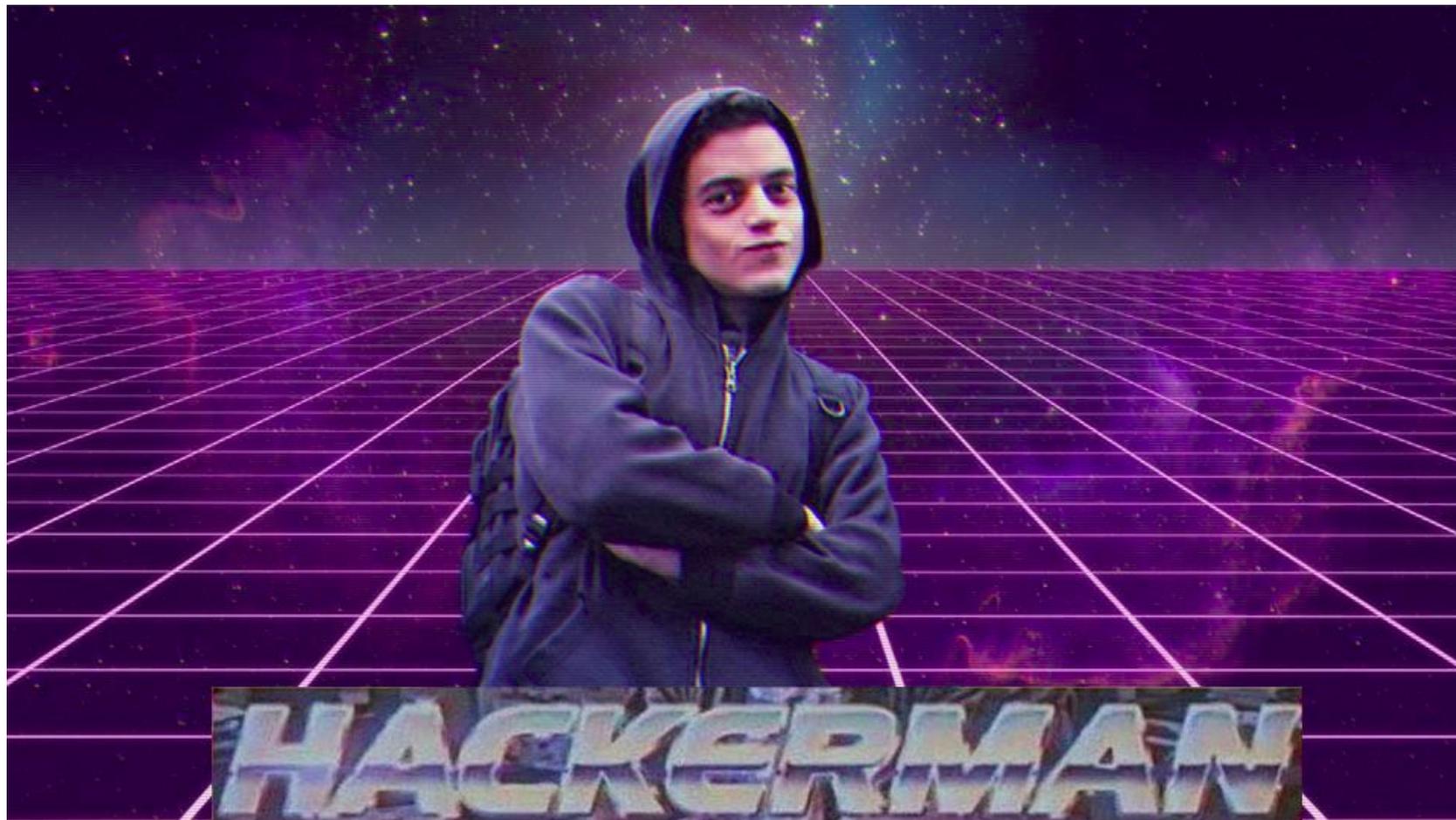
Biology is about Big Data

- A **single** human genome is approximately **120 GB**.
- The Broad is sequencing genomes at a rate of **1 every 10 minutes**.
- The Broad generates **17 TB** of new genomes **per day**.
- The Broad manages **45 PB** of scientific data.

- YouTube receives **24 TB** of new videos **per day**.
- YouTube stores about **1 billion videos** with an average size of 86 MB, for a total storage footprint of **86 PB**.

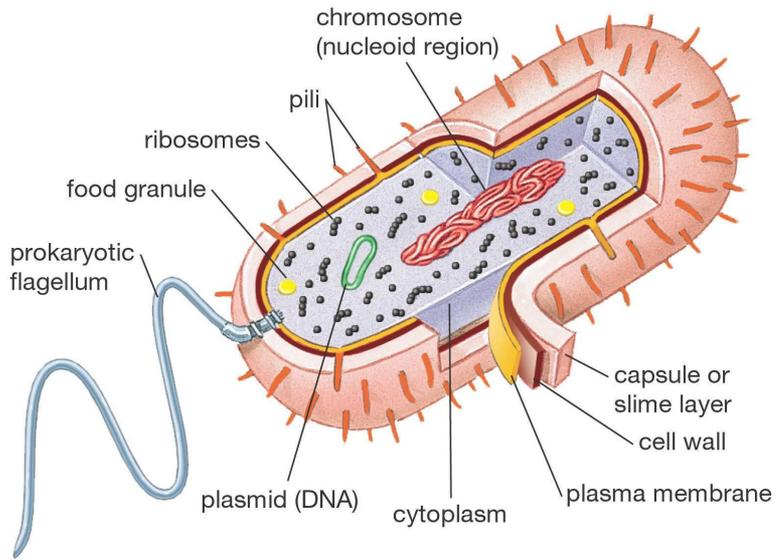


It needs to be hacked!



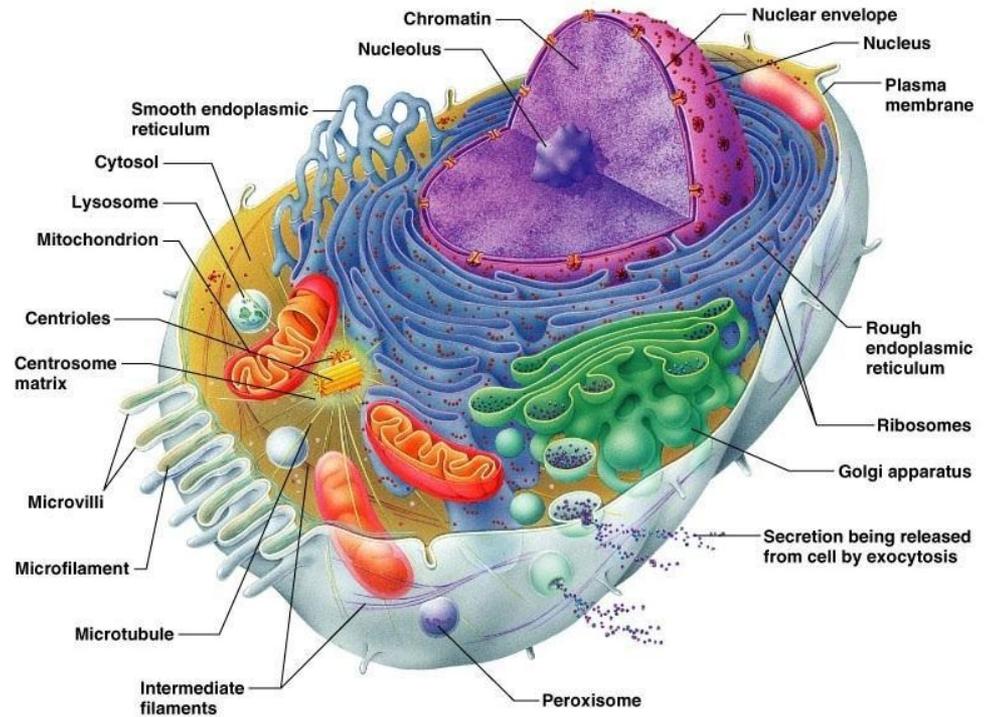
2. Основные понятия и концепции в биологии

Клетки



Copyright © 2005 Pearson Prentice Hall, Inc.

Прокариоты

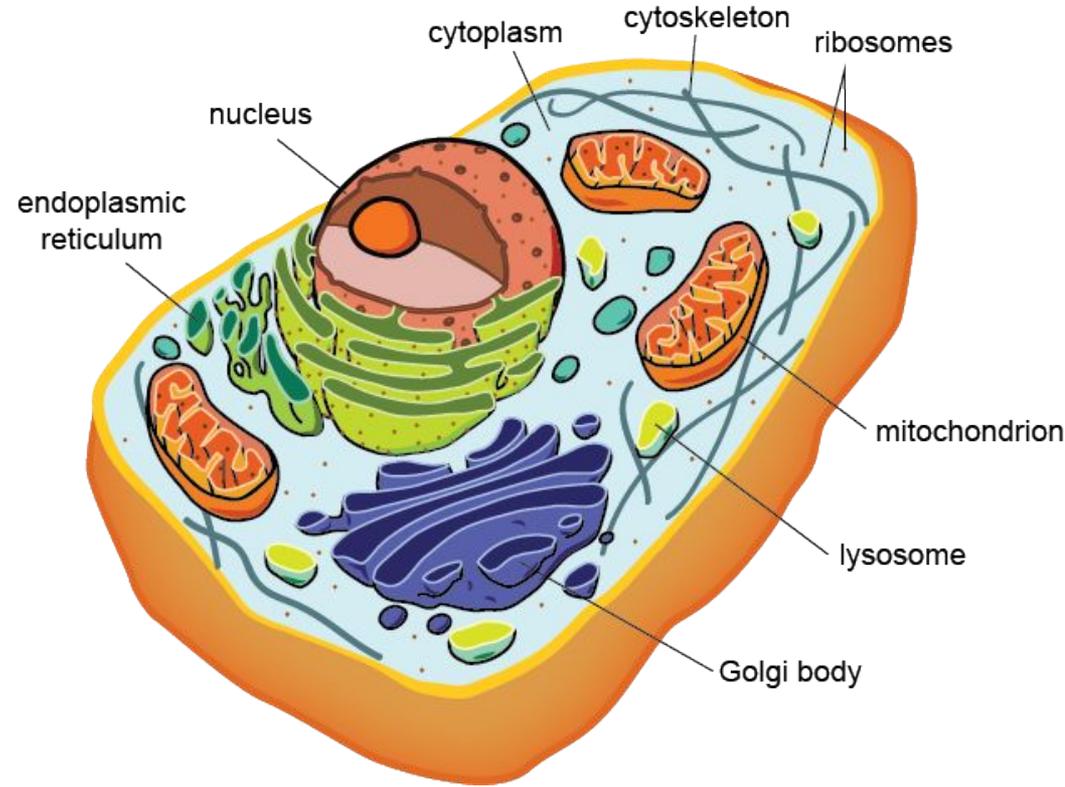


Copyright © 2004 Pearson Education, Inc., publishing as Benjamin Cummings.

Эукариоты

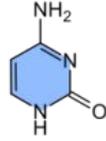
Химия жизни (из молекул состоит клетка)

- Вода
- Малые молекулы
- Ионы
- ДНК
- РНК
- Белки
- Углеводы (моносахариды, полисахариды)
- Липиды (триглицериды, холестерол, фосфолипиды)

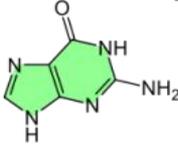


DNA and RNA

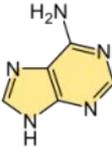
Cytosine **C**



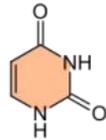
Guanine **G**



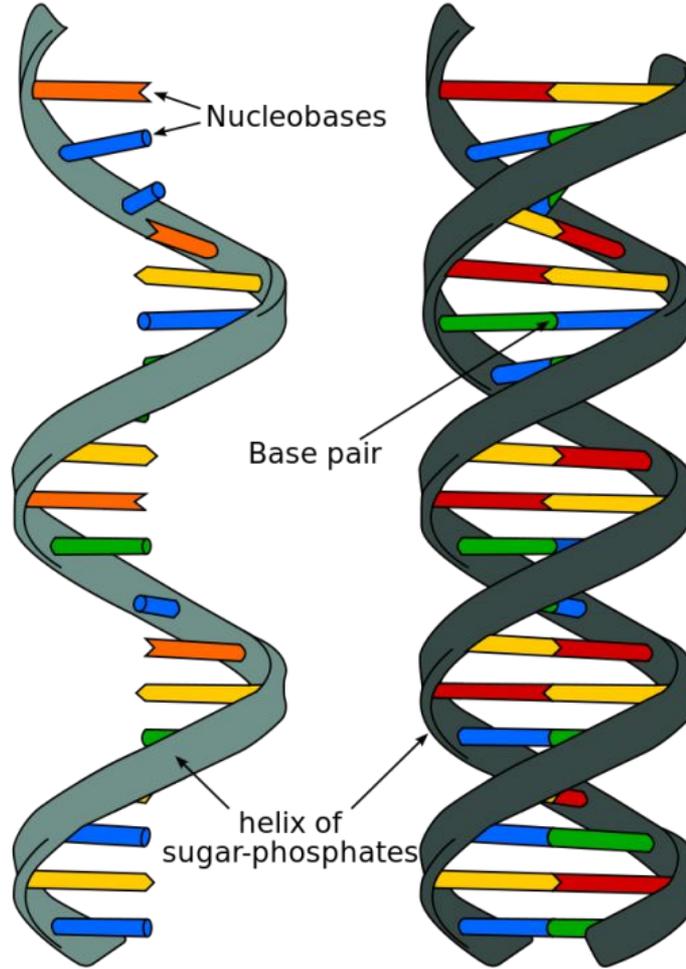
Adenine **A**



Uracil **U**



Nucleobases
of RNA



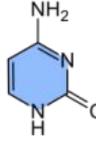
RNA

Ribonucleic acid

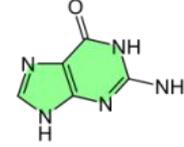
DNA

Deoxyribonucleic acid

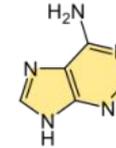
Cytosine **C**



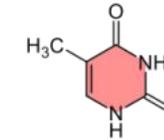
Guanine **G**



Adenine **A**



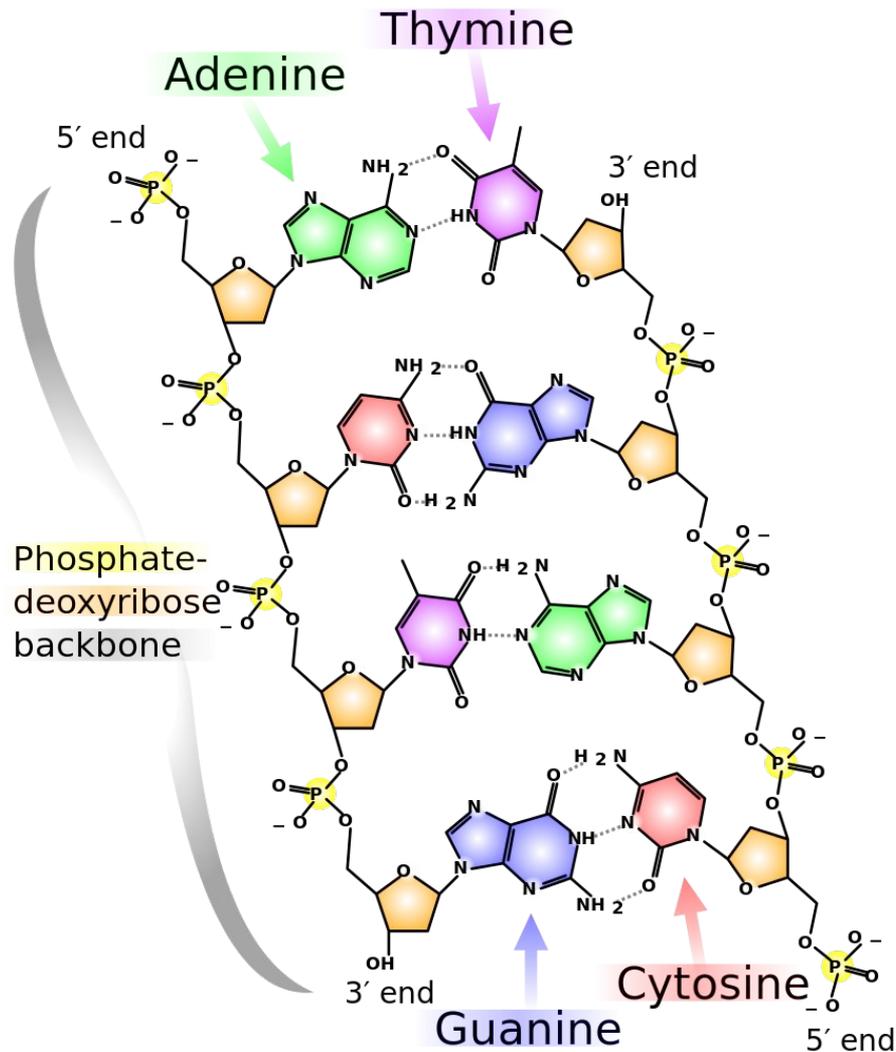
Thymine **T**



Nucleobases
of DNA

Структура ДНК

- Две цепи ДНК *антипараллельны*
- ДНК - правозакрученная спираль
- Есть 5'-конец, а есть 3'-конец (определяется по сахару)
- Цепи соединены водородными связями

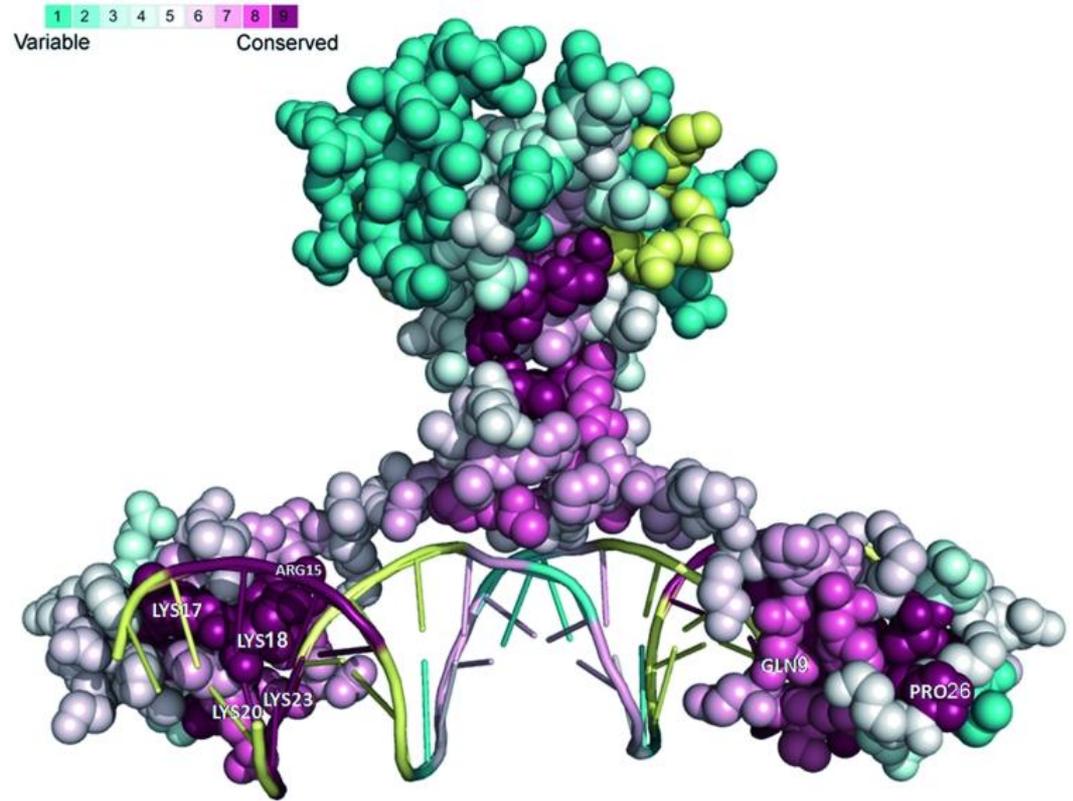


DNA vs. RNA

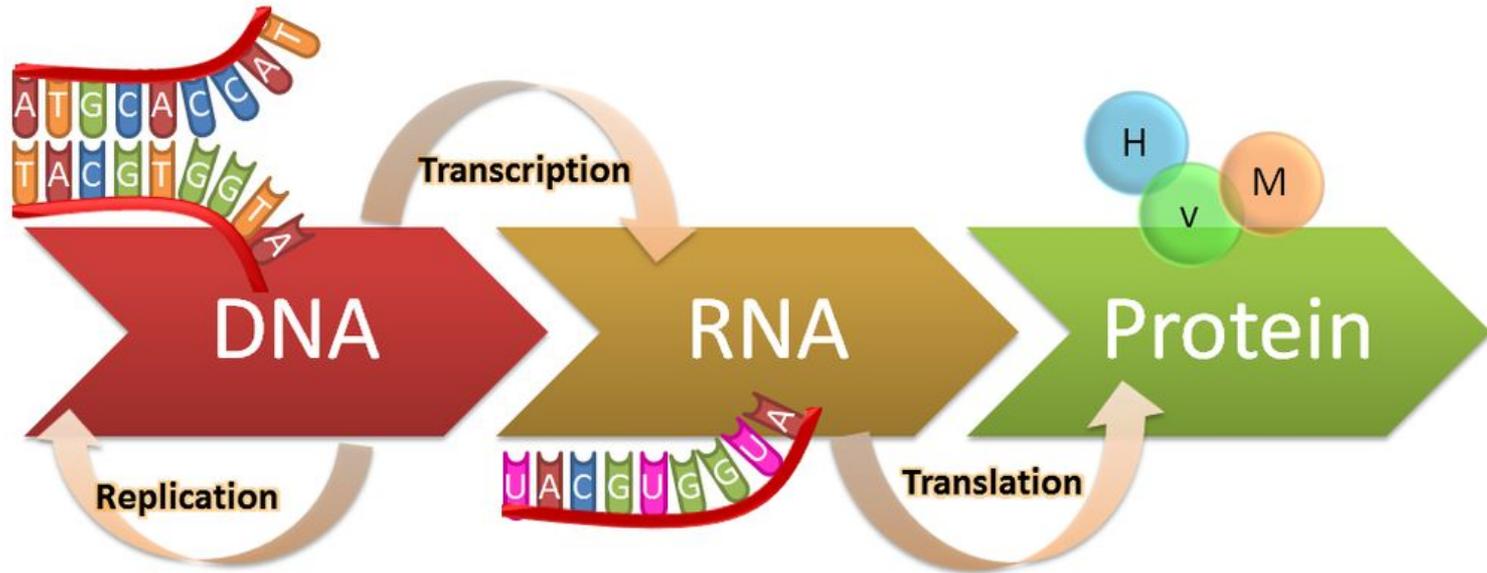
Элемент	ДНК	РНК
Мономер	Нуклеотид	Нуклеотид
Количество цепей	2	1
Основания	А, Т, Г, Ц (A, T, G, C)	А, У, Г, Ц (A, U, G, C)
Стабильность	Высокая	Низкая (свободные -ОН группы)
Размер молекулы	Средний или большой	Малый или средний, редко большой
Функционал	Хранение и передача информации	Посредническая функция, иногда ферментативная и еще очень много....

Протеины

- Полимеры аминокислот
- Образуют множество 3D-структур
- Мотивы, домены
- Барьерные, структурные, моторные, каналы, рецепторы, ДНК-связывающие...

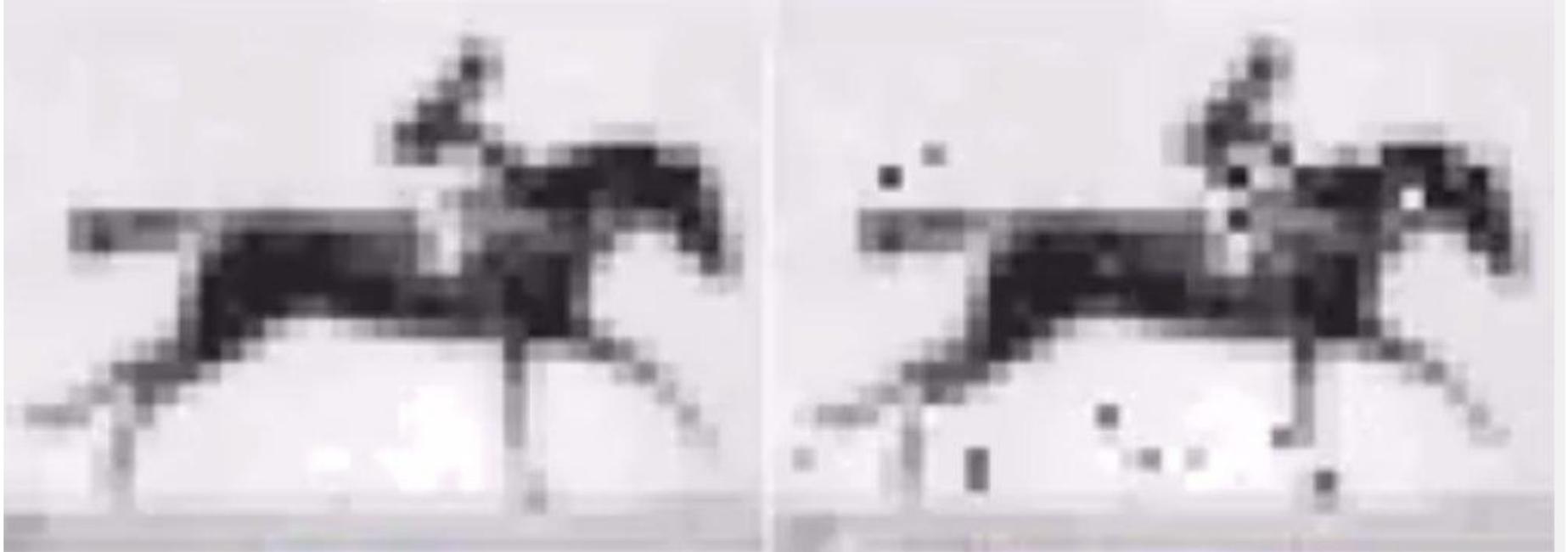


Центральная догма молекулярной биологии



- Генетическая информация реализуется за счет серии матричных процессов
- Основные матричные процессы - репликация, транскрипция, трансляция
- ДНК - HDD, РНК - RAM, белок - вывод

Why to use polymers?

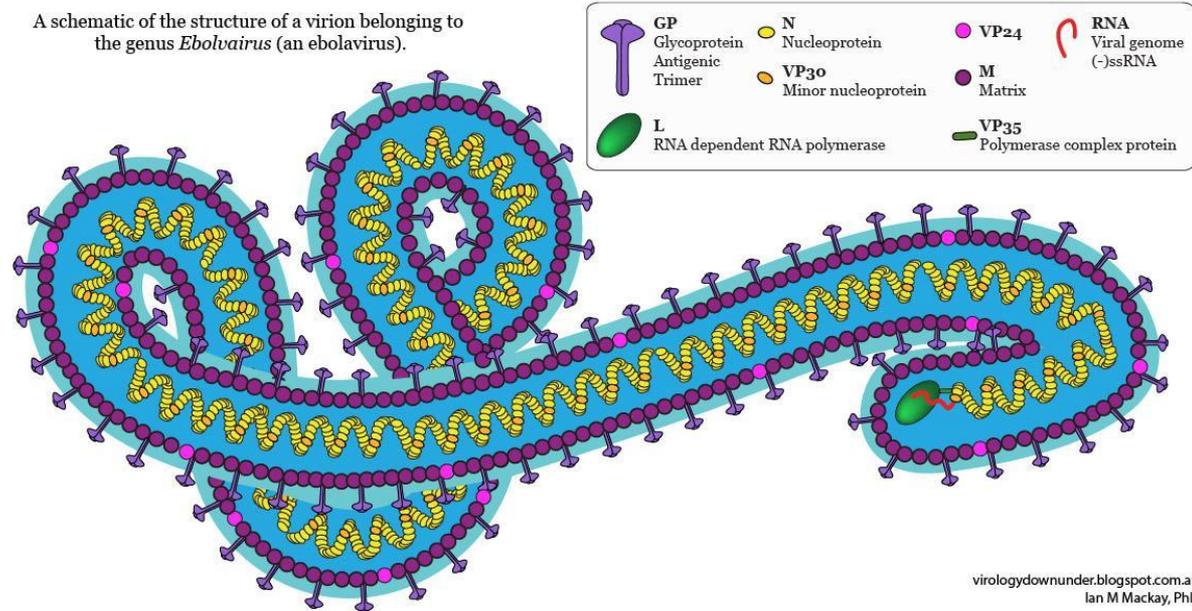


CRISPR–Cas encoding of a digital movie into the genomes of a population of living bacteria
Shipman *et al.*, 2017, Nature

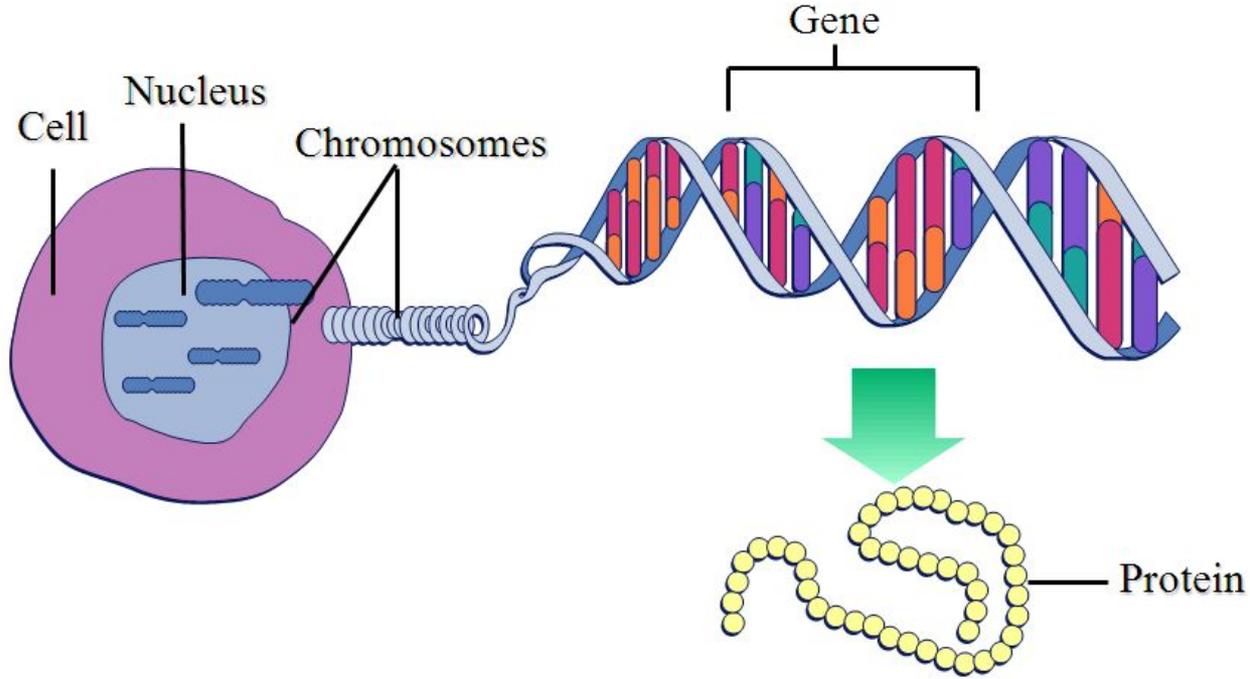
Живое существо

- Почти все живое состоит из клеток
- Все живое так или иначе способно к самовоспроизведению

A schematic of the structure of a virion belonging to the genus *Ebolavirus* (an ebolavirus).



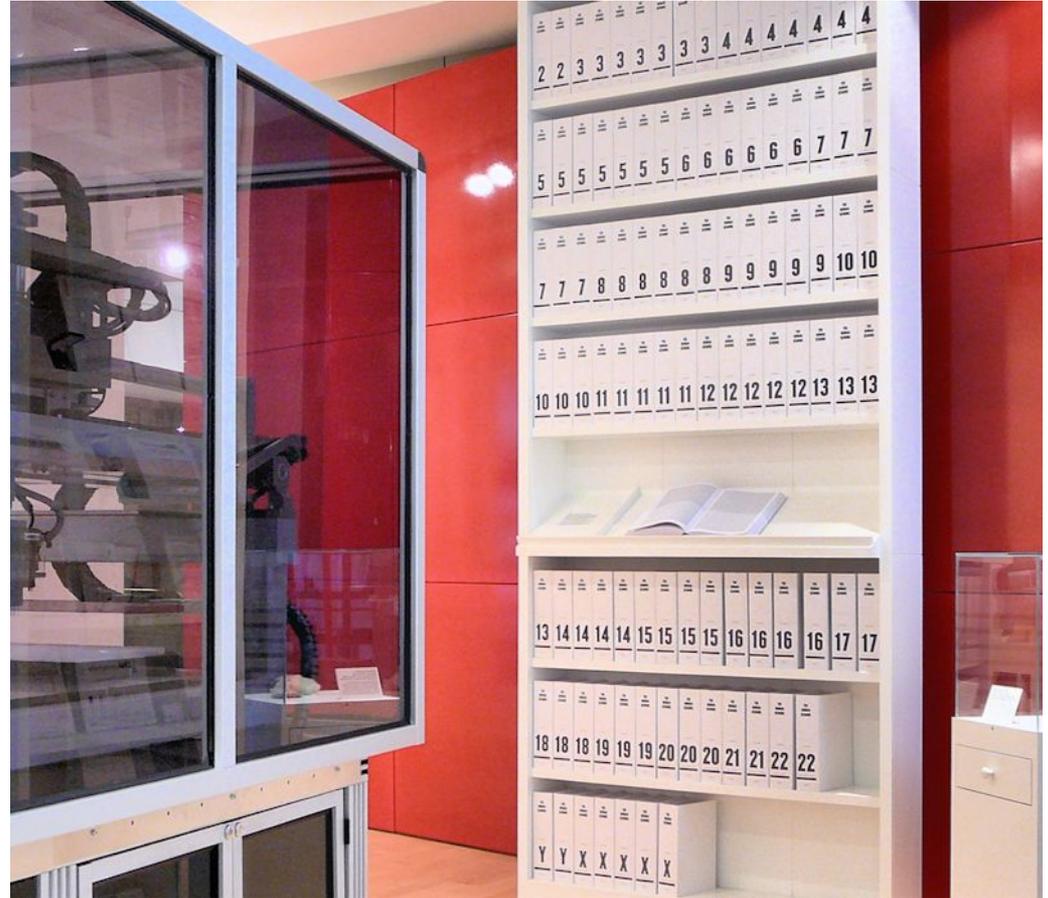
Ген



- Ген - единица наследования
- Ген - участок молекулы ДНК
- Ген способен транскрибироваться (кодировать что-то: белок или РНК)

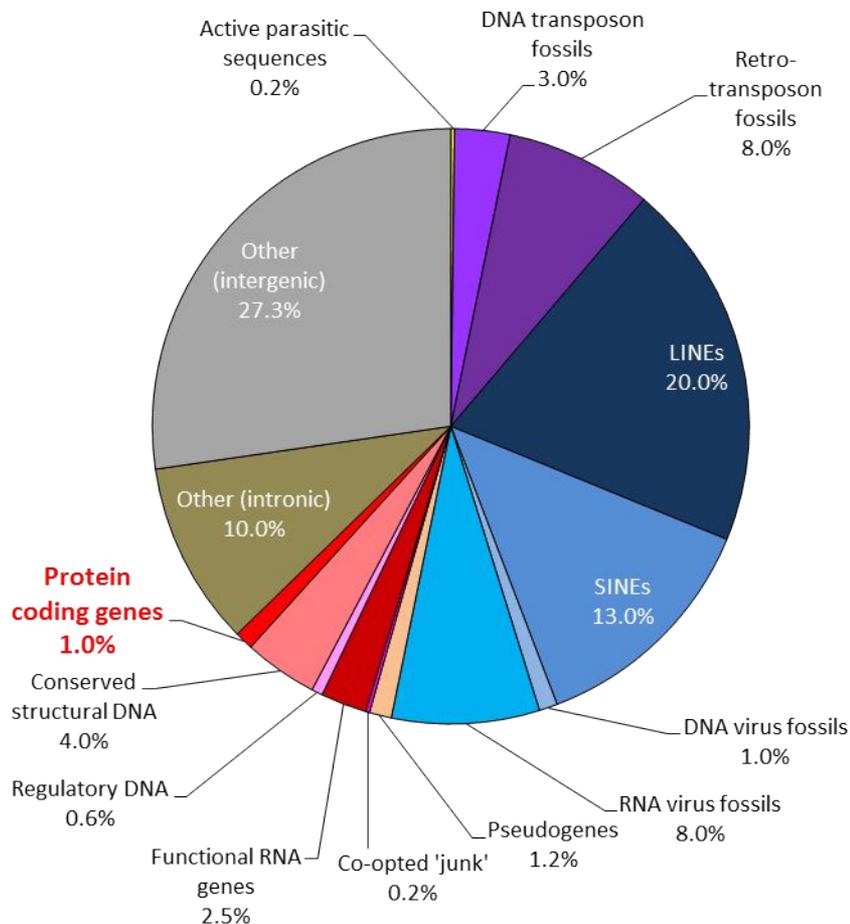
Геном

- Геном - совокупность всего генетического материала организма (клетки)
- Геномы различаются по размеру, составу, количеству частей (хромосом)
- Первый геном клеточного организма полностью прочитан в 1995 году (*Haemophilus influenzae*)



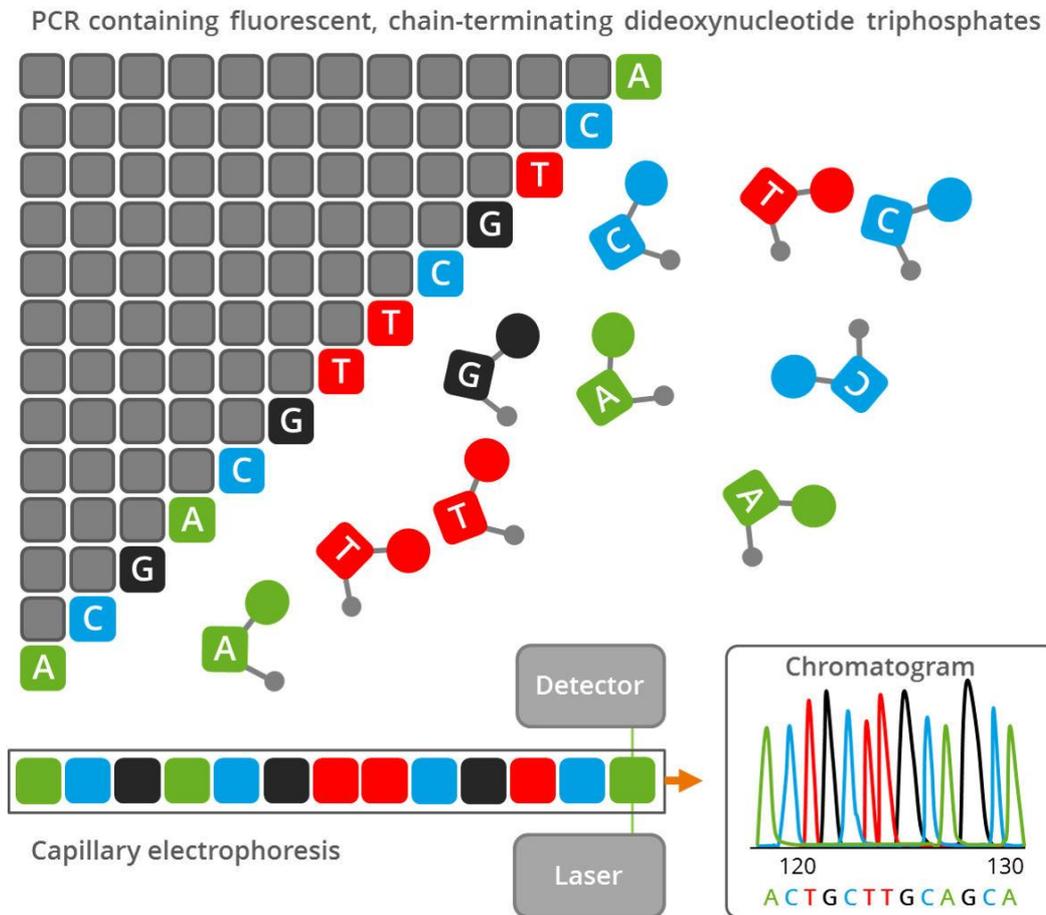
Что такое геном?

- Далеко не все в геноме что-то кодирует
- Гены:
 - белок-кодирующие
 - кодирующие РНК
- Повторы:
- Регуляторные элементы
- ...



Секвенирование

- Секвенирование - определение последовательности
- Sanger sequencing - первый широко распространенный метод
- Высокая точность, длина до 1,5 т.н. (kb), низкая скорость работы



Спасибо за внимание!