

INTRO TO BIOINFORMATICS

Евгений Пшеничный

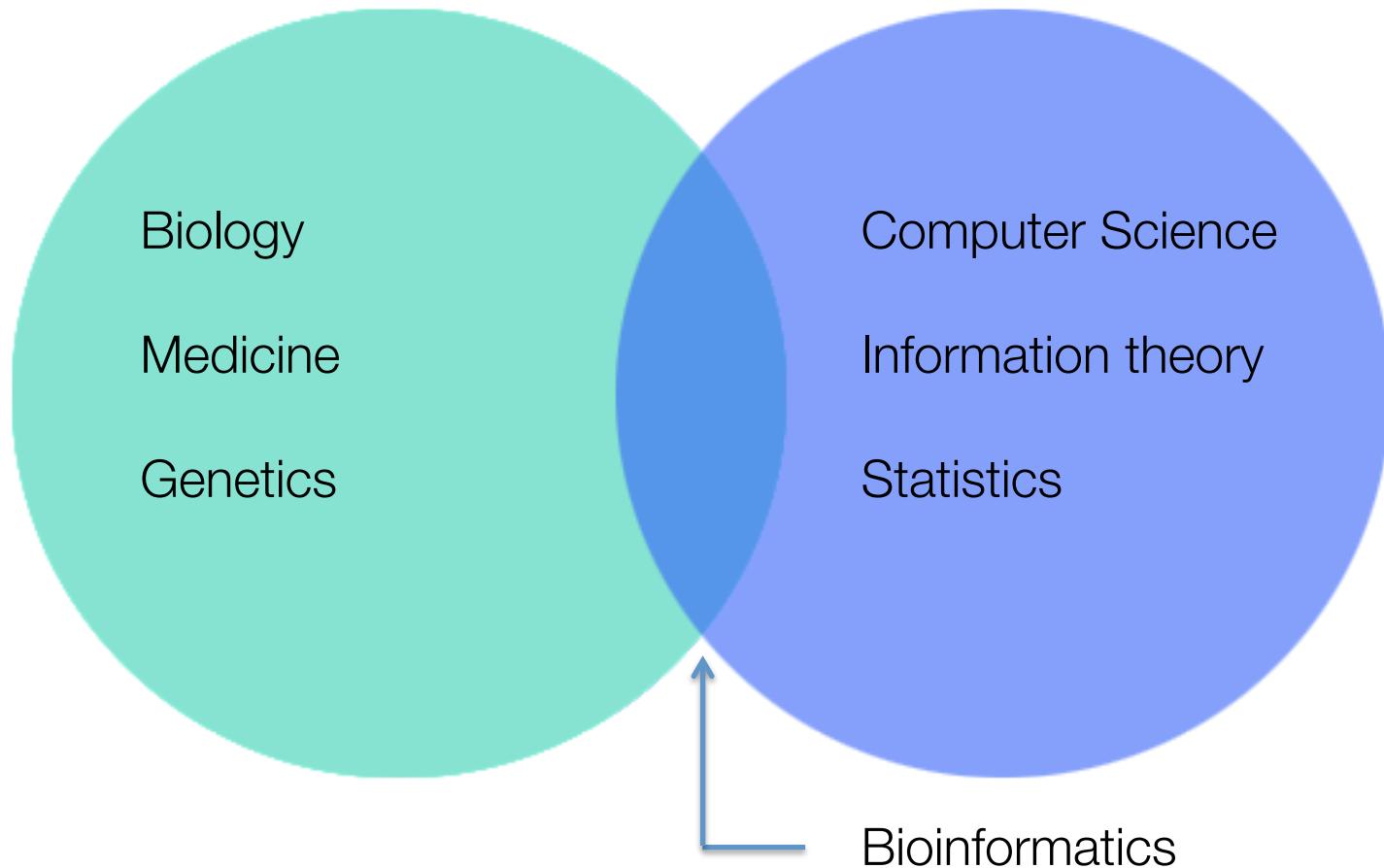
<http://app.ficus.io>

#kurilkarostov



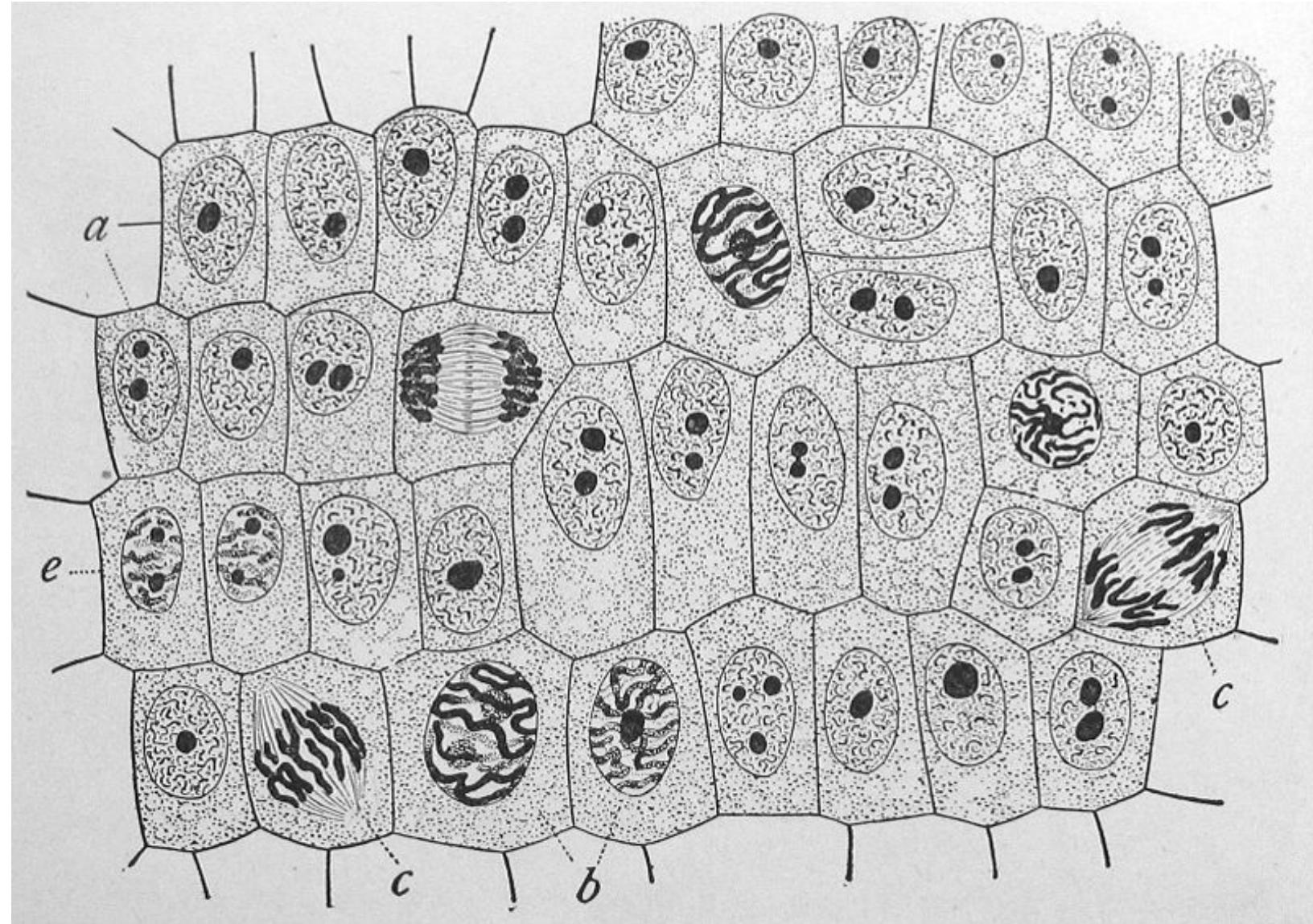


BIOINFORMATICS



ПРИМЕНЕНИЯ

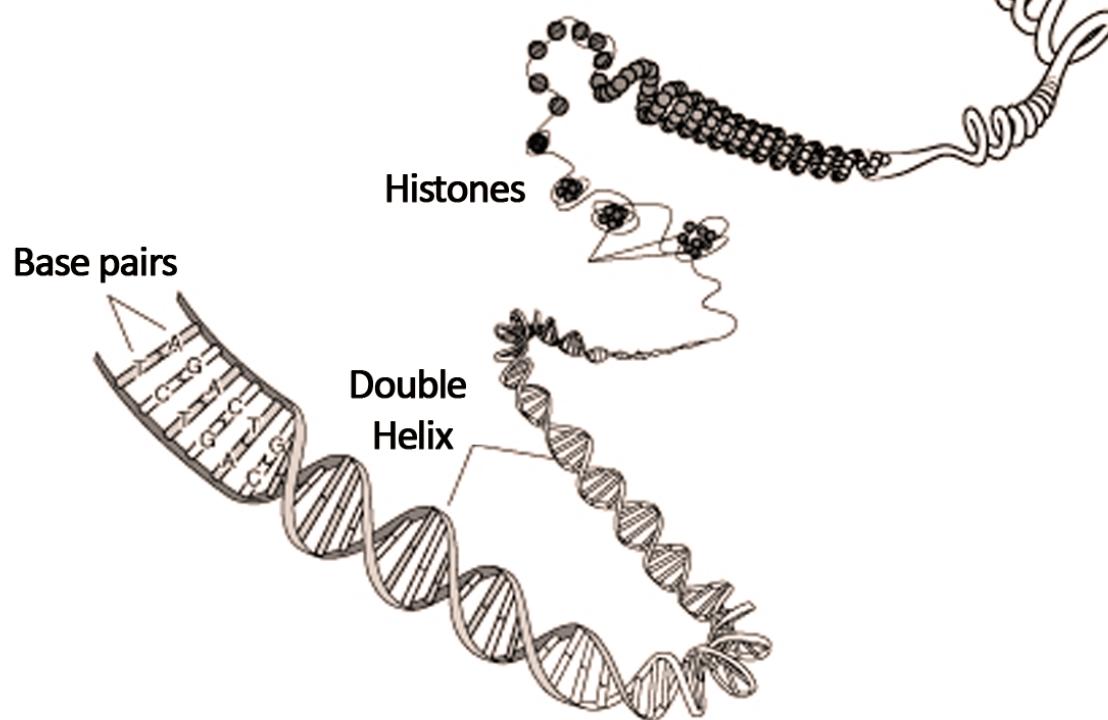
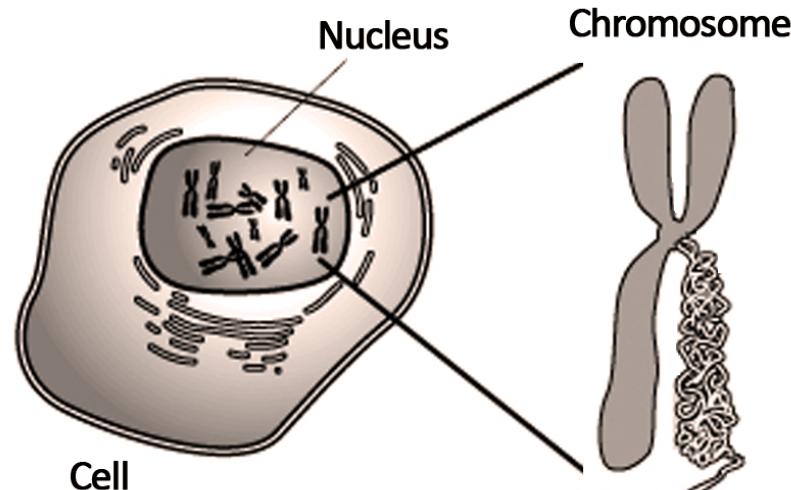
- Персонифицированная медицина
- Генная терапия
- Разработка лекарств
- Пренатальная диагностика
- Альтернативные источники энергии
- Улучшение качества питания
- Криминалистическая экспертиза
- Эволюционные исследования
- Исследования изменения климата
- Устойчивость к антибиотикам
- Утилизация отходов
- Защита от био-оружия
- Новые сорта пищевых культур

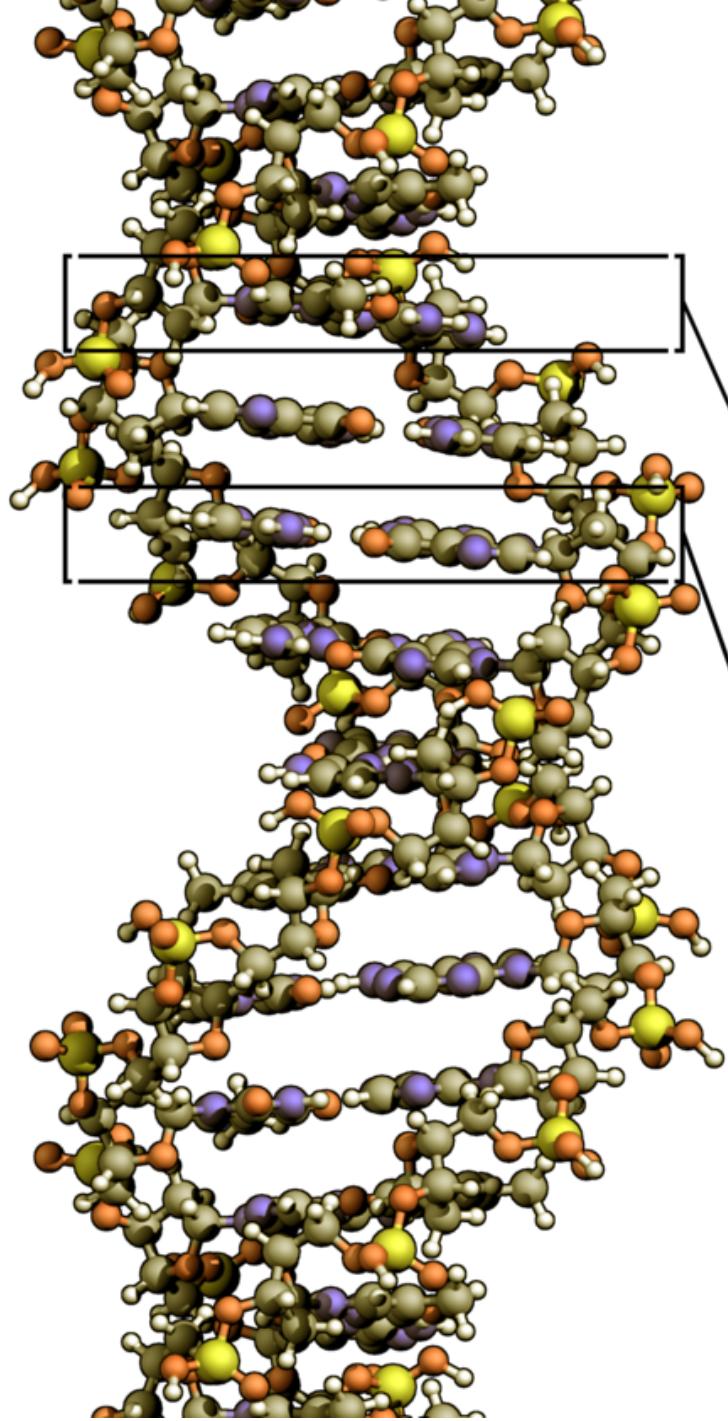


ДНК

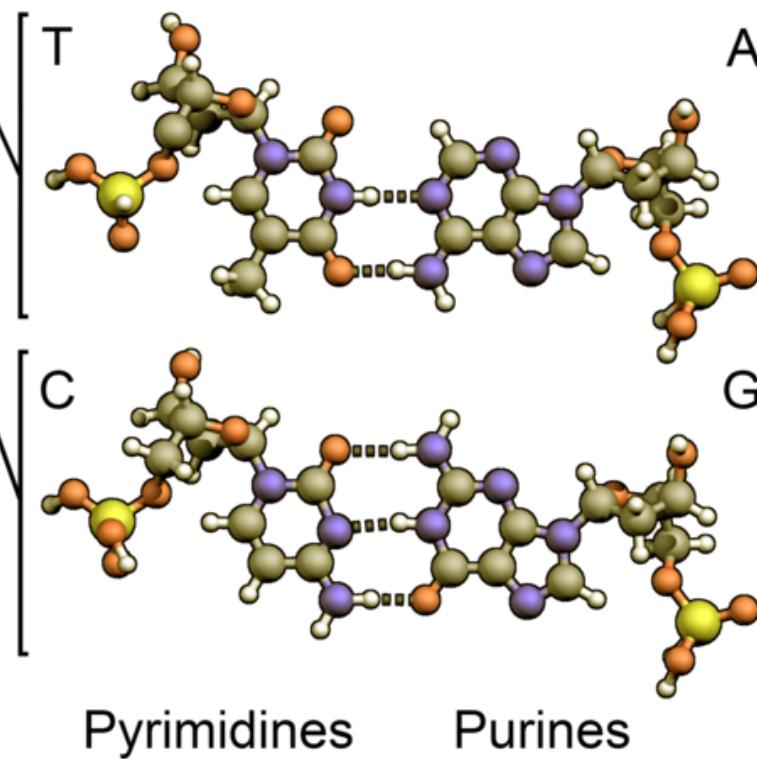
Двойная спираль,
нуклеотиды:

A T G C





Правило комплементарности



ДНК

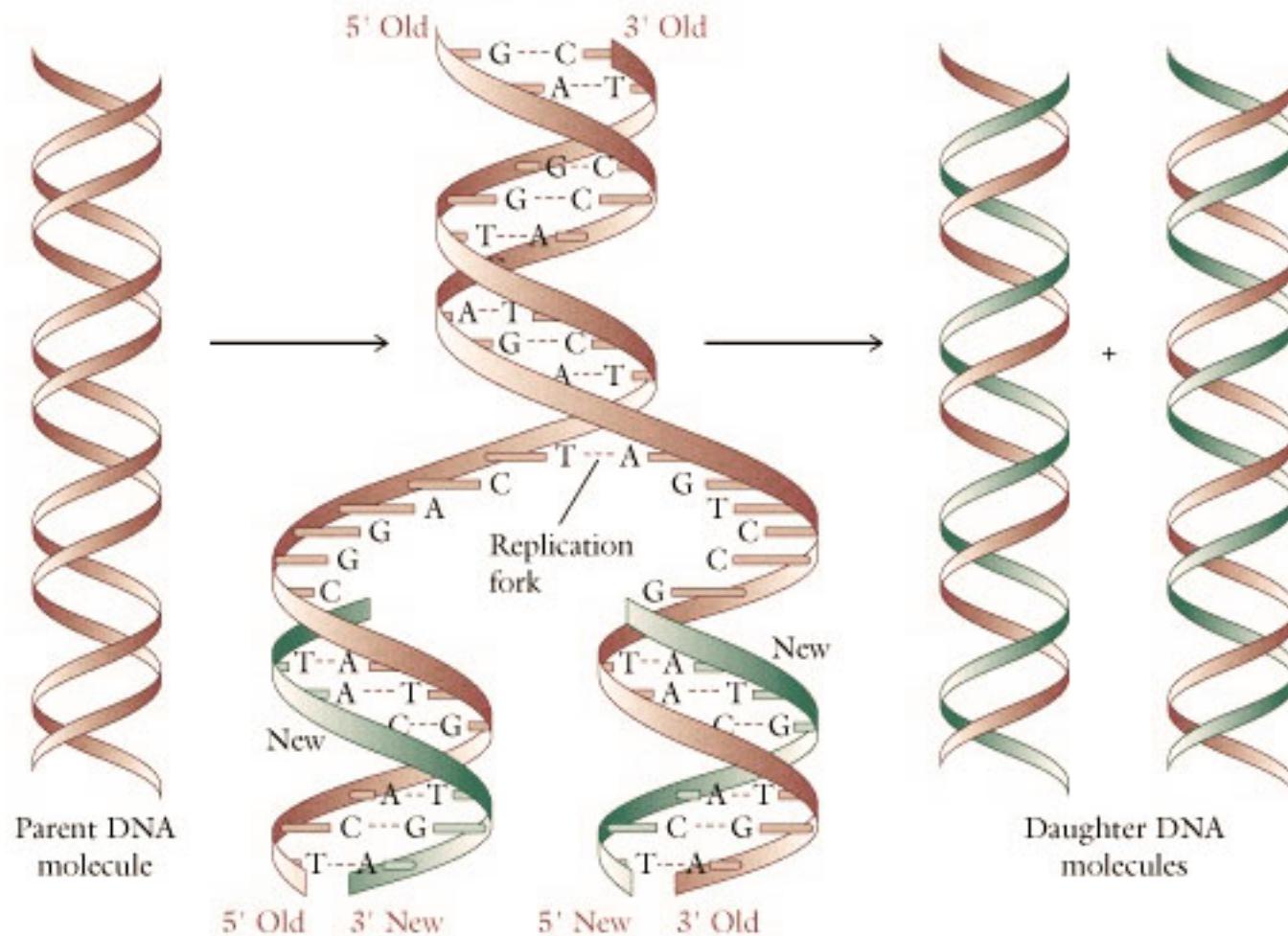
AGCTCCAGCTAGGCTTTACCTAGCTGTGT
TCGAGGGTCGATCCGAAATGGATCGACACA

ДНК



AGCTCCAGCTAGGCTTTACCTAGCTGTGT

Репликация



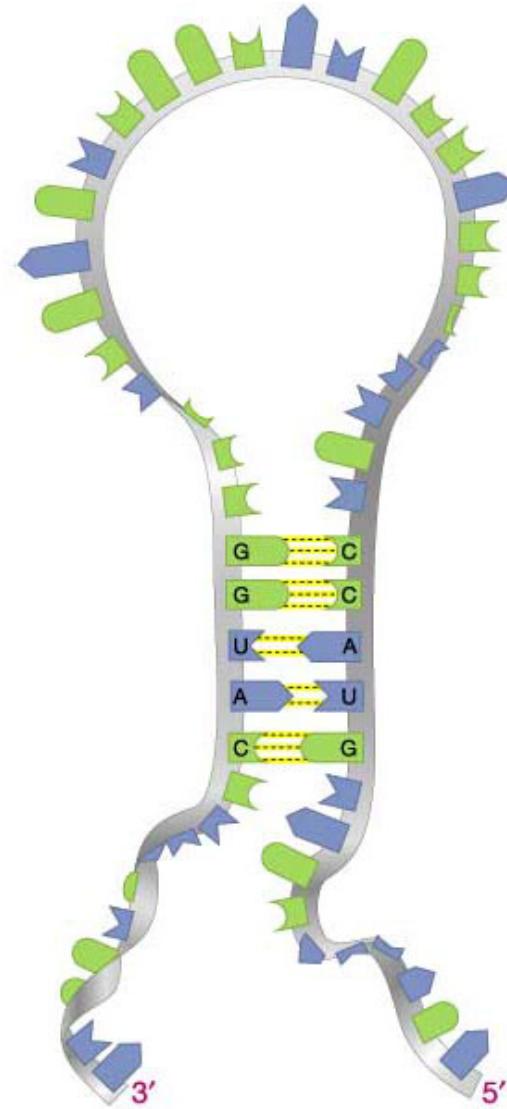
РНК

Одна цепь,
рибонуклеотиды:

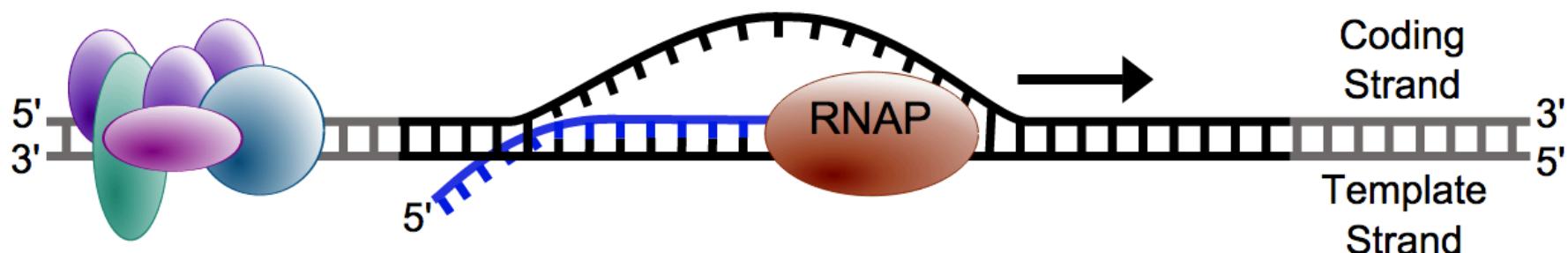
A U G C

Замена:

T -> U



Транскрипция



ДНК → РНК

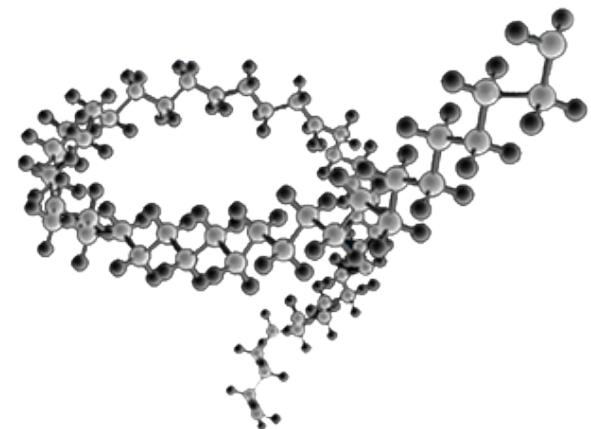
T → U

Белки

Аминокислоты (20):

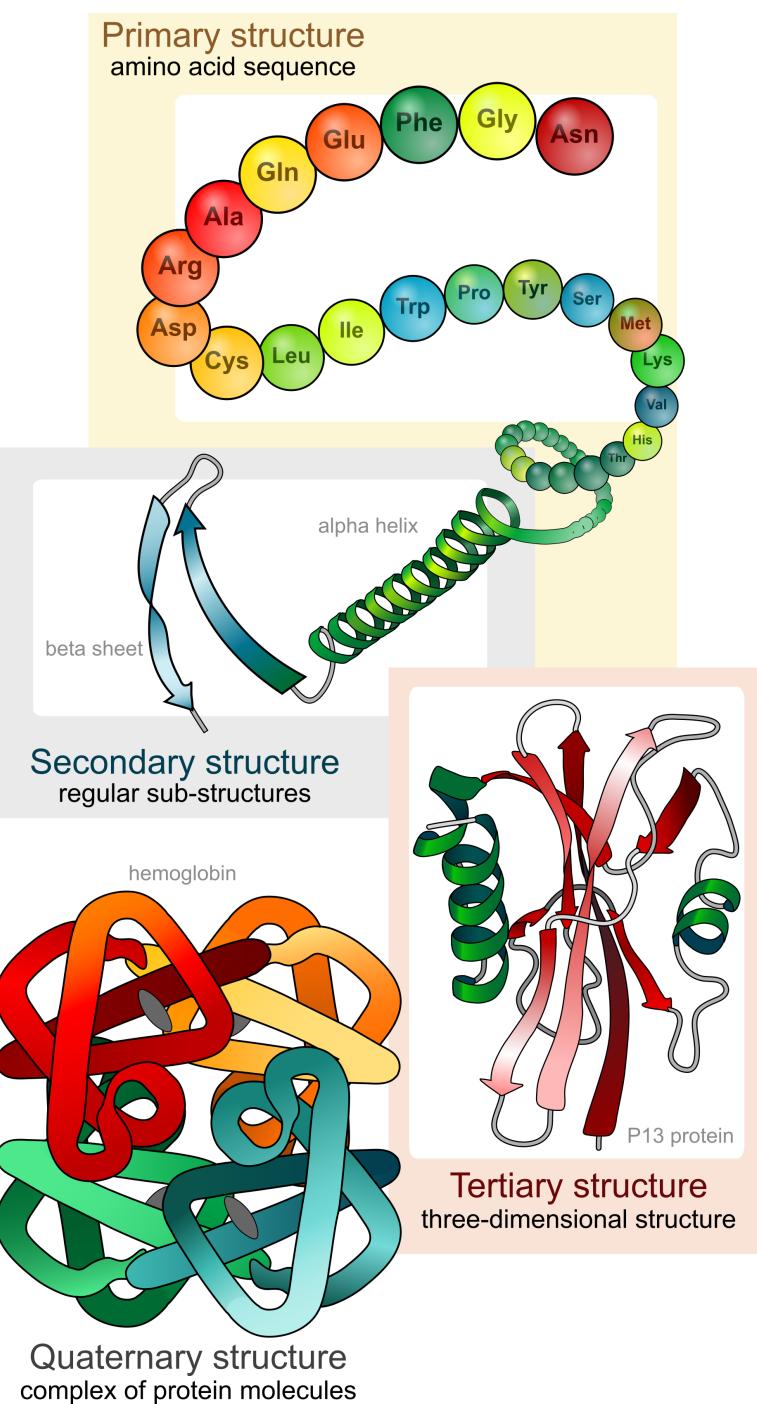
MATGSRTSLLLAFGLLCLEFGHIKLMDPW

Функции: структурная,
ферментативная,
рецепторная,
гормональная,
транспортная, защитная...

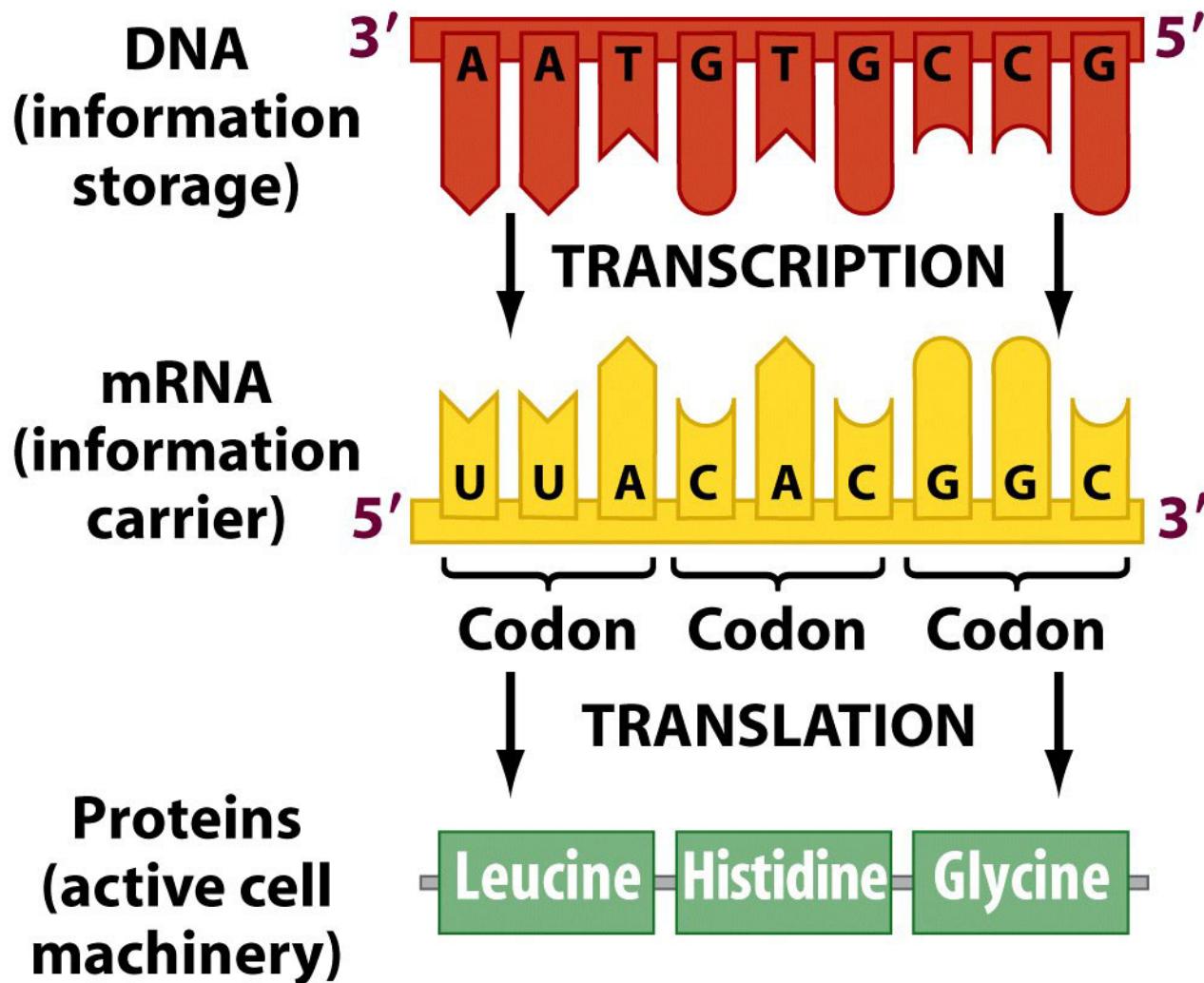


Структура белка

1. Последовательность
2. α – спираль / β - лист
3. Трехмерная структура
4. Комплекс



Центральная доктрина



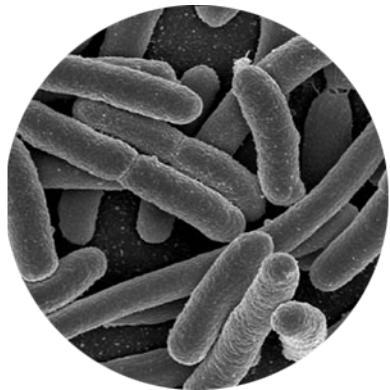
Трансляция

		Second base										
		U	C	A	G							
First base	U	UUU UUC UUA UUG	Phenylalanine Serine Leucine Leucine	F S L L	UCU UCC UCA UCG	Serine	UAU UAC UAA UAG	Tyrosine Stop codon Stop codon	Y	UGU UGC UGA UGG	Cysteine Stop codon Tryptophan	C C A G
	C	CUU CUC CUA CUG	Leucine	L	CCU CCC CCA CCG	Proline	CAU CAC CAA CAG	Histidine Glutamine	H Q	CGU CGC CGA CGG	Arginine	U C A G
	A	AUU AUC AUA AUG	Isoleucine Methionine start codon	I M	ACU ACC ACA ACG	Threonine	AAU AAC AAA AAG	Asparagine Lysine	N K	AGU AGC AGA AGG	Serine Arginine	U C A G
	G	GUU GUC GUA GUG	Valine	V	GCU GCC GCA GCG	Alanine	GAU GAC GAA GAG	Aspartic acid Glutamic acid	D E	GGU GGC GGA GGG	Glycine	U C A G
		Third base										

РНК → Белок

3 нуклеотида (кодон) → 1 аминокислота

Размер геномов



E.coli

4,600,000bp



Drosophila

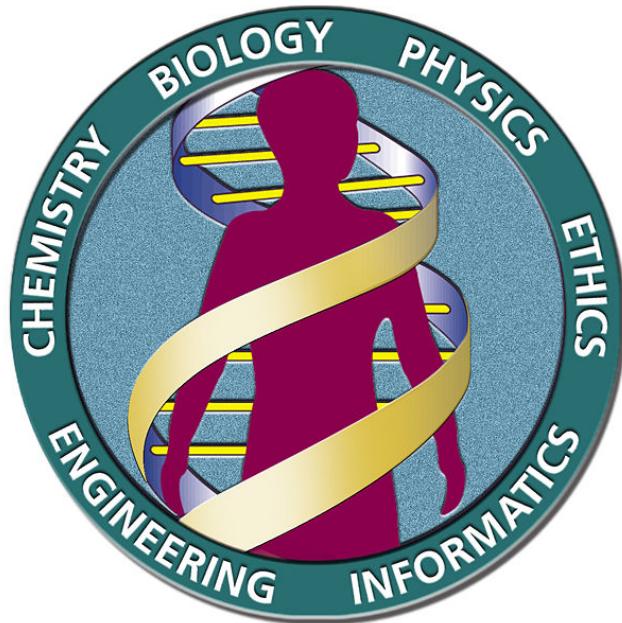
130,000,000bp



Homo sapiens

3,200,000,000bp

Геном человека



1990–2003
13 лет



0.01%

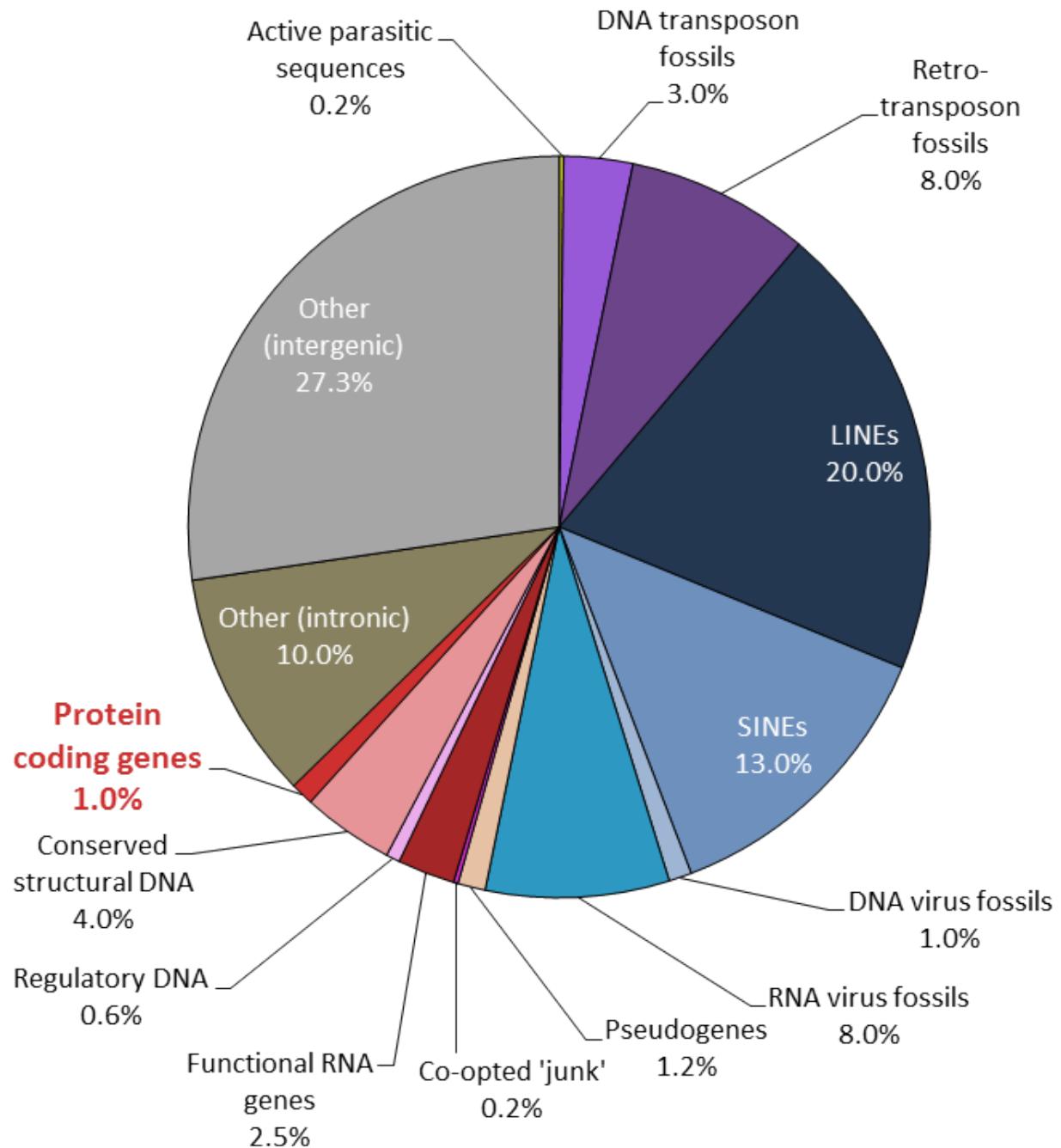
A G T A T A A G G T G C A T T C C C
A C A A A A A A G T A T A A A G G T G C A T T C
C C C A G A A G T C C C G A T A A A G T C G T C C C T C C G C A T T C C G G T
A A A G T A T A A G G T G C A T T C C C A G A G G G A T A A A G T C A T C
G A T C C C A T A A G G G T G C A T T C C C A G A T C
A C A A A A G T C A G G G A T A A A G T C A T C

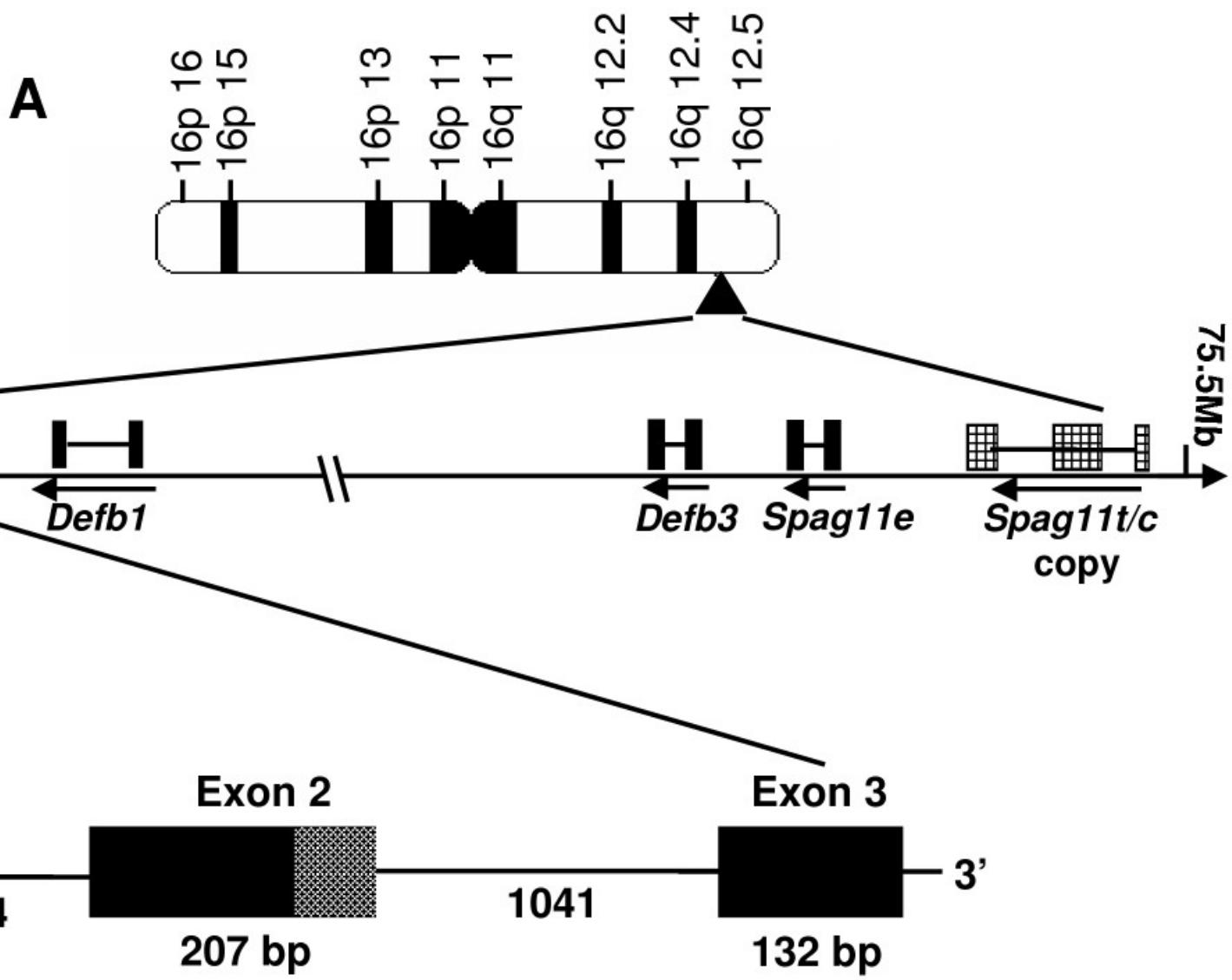


A G A G T A G A A G G C G T C T C C G C A T A C C C G G T C
A T C C C A G T A T A A G G T C A T T C C C A G A G G G A T A A A G T C E G
C G A T C C C A T A A G G G T G C A T T C C C A G A T C G
C A C C A A A G T C A G G G A T A T T A G G G T C C G C A T T C C C G G G T
A A A G T A T A A G G T C A G G G A T A A A G T C A T C

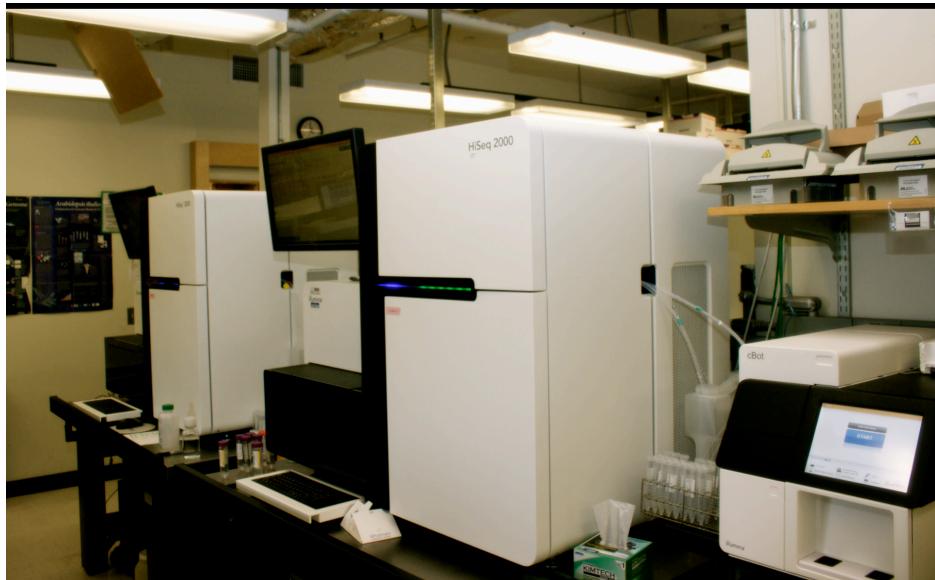


G A T G A T A A G T C A A T A A A G T A T A A G G G T G C A T T C C C A G A T A T C
A C T T A A G C C A G T A T A A G G T C C A T T C C C G A A G T C C C C C G A T A G
C A T T C C C A A G T C C C C C A G A T C T C A T T C C C E E E G T C C
G G C A T T A A G T A T A A G G G T G C A T T G G G T G C A T T G G G T G C A C C C G G
A C A A A A G T C A G G G A T A



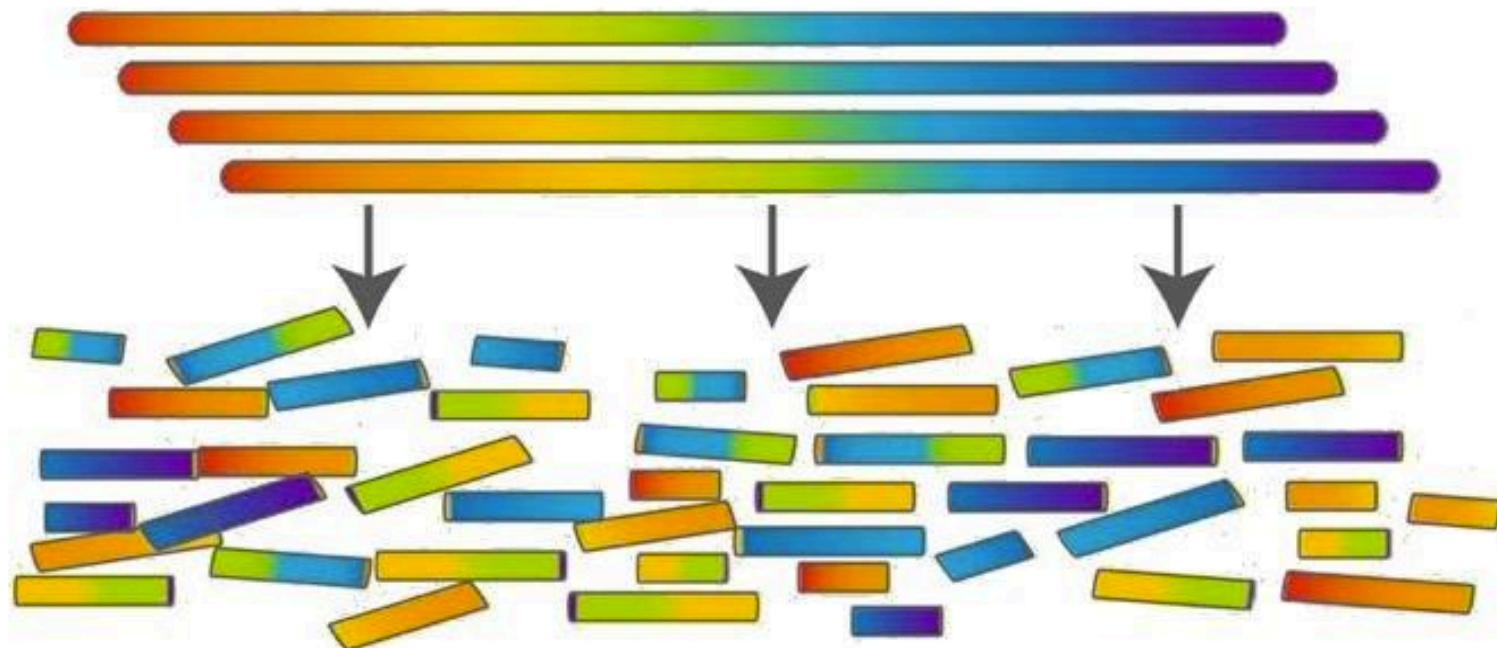


Так выглядит секвенатор



Метод дробовика

- Разделение генома на фрагменты
- Прочтение коротких фрагментов
- Сборка генома



A C A A A G T C A G G G A T A C A A A G T C A G G G A T A

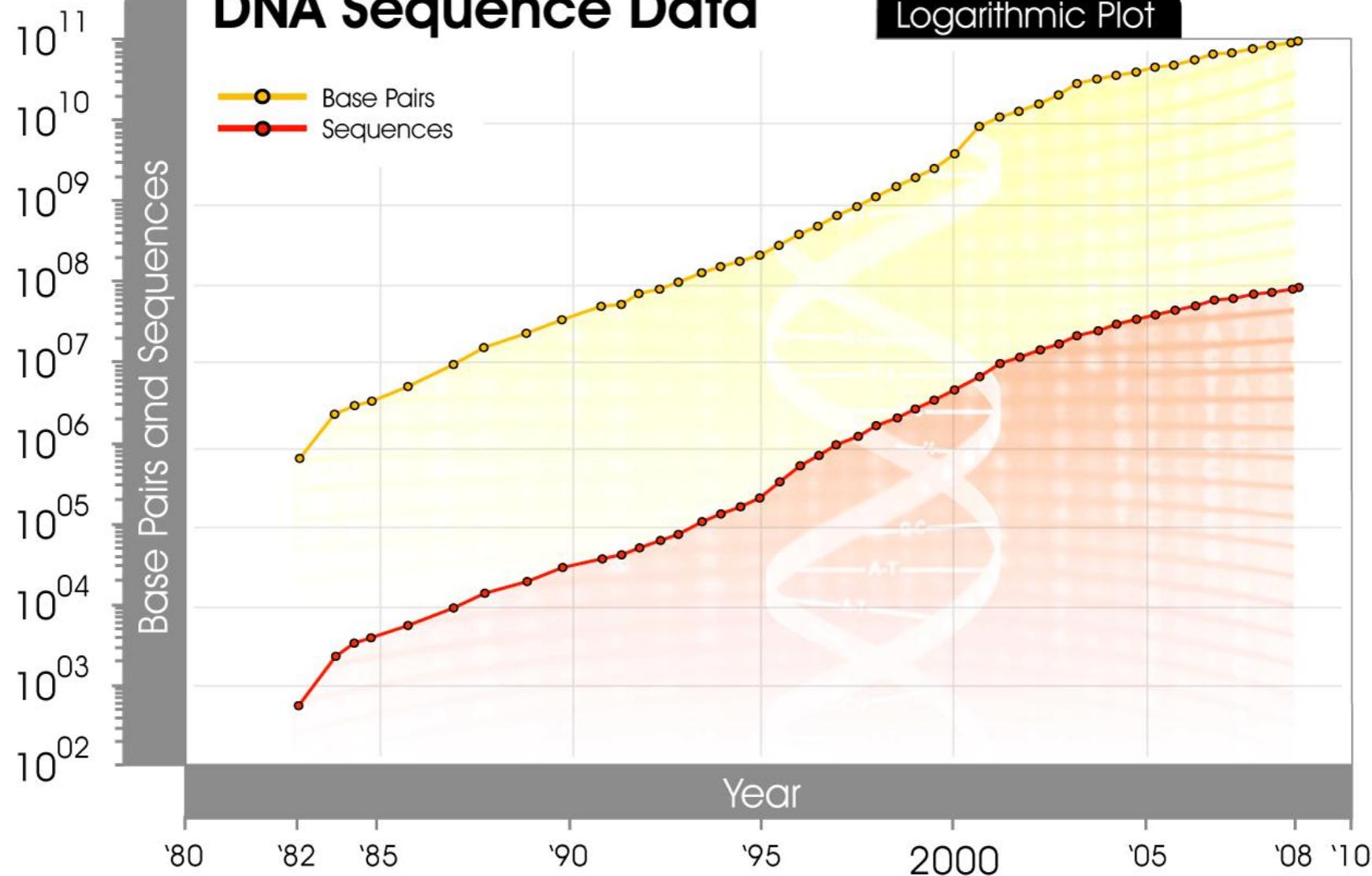
A wide-angle photograph of a massive data center. The floor is filled with rows of server racks, many of which are illuminated from within, casting a warm glow. Above, a complex network of steel beams, pipes, and cables forms a grid-like ceiling structure. The lighting is dramatic, with strong highlights and deep shadows, emphasizing the scale and complexity of the facility.

DATA SCIENCE

Growth in Genbank

DNA Sequence Data

Logarithmic Plot



Базы данных

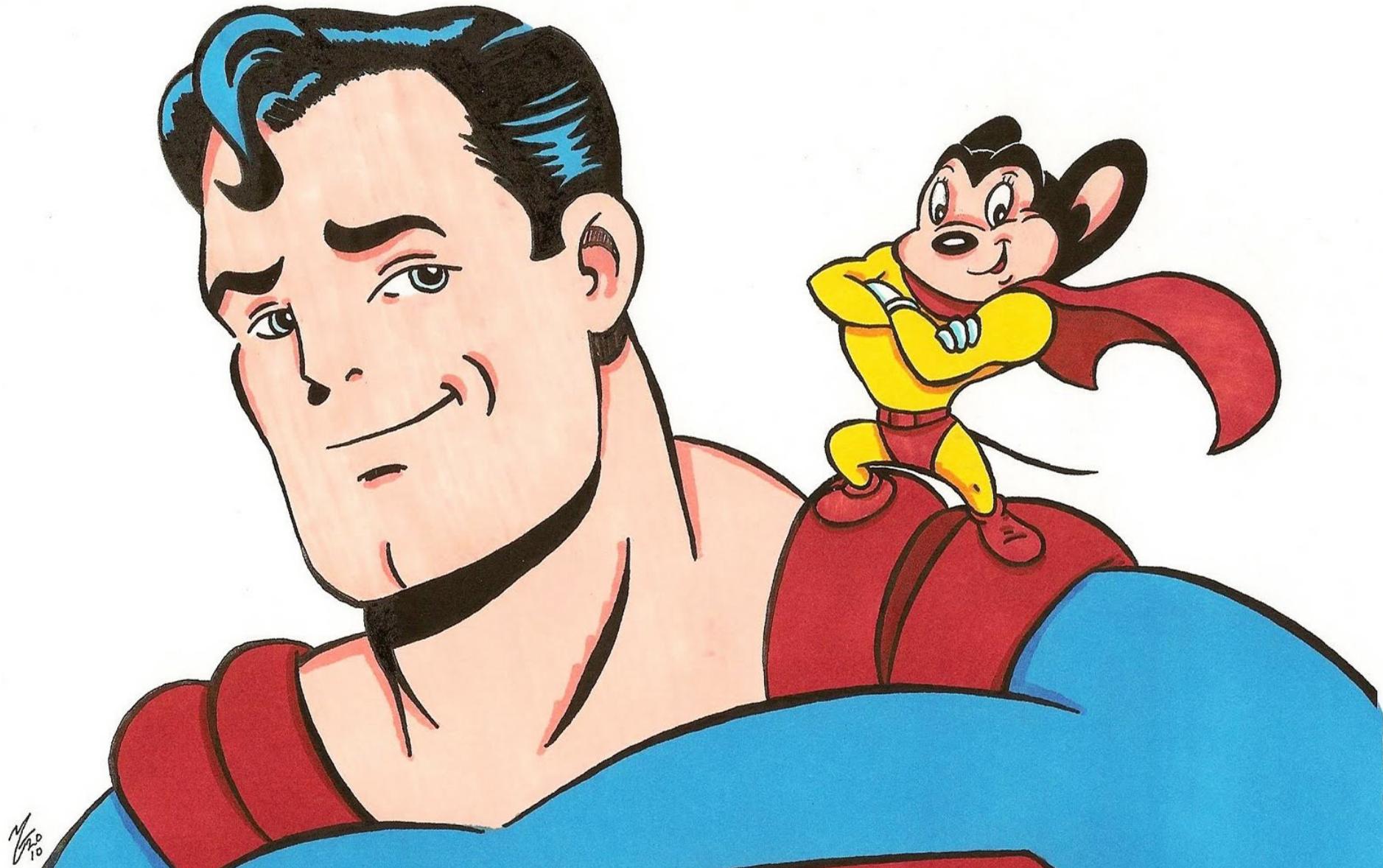
e! Ensembl





In silico

**БЕЛКИ МЛЕКОПИТАЮЩИХ
СХОЖИ НА 90%**

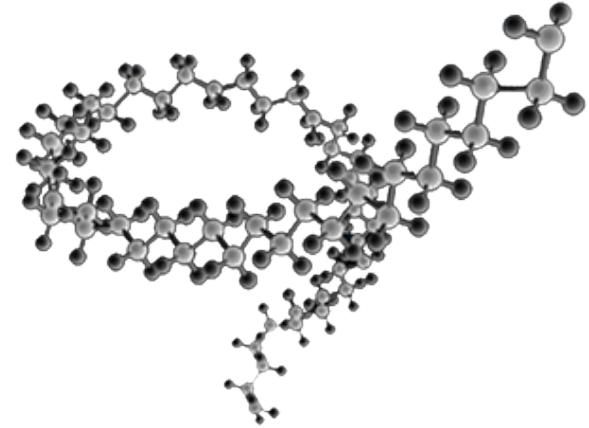




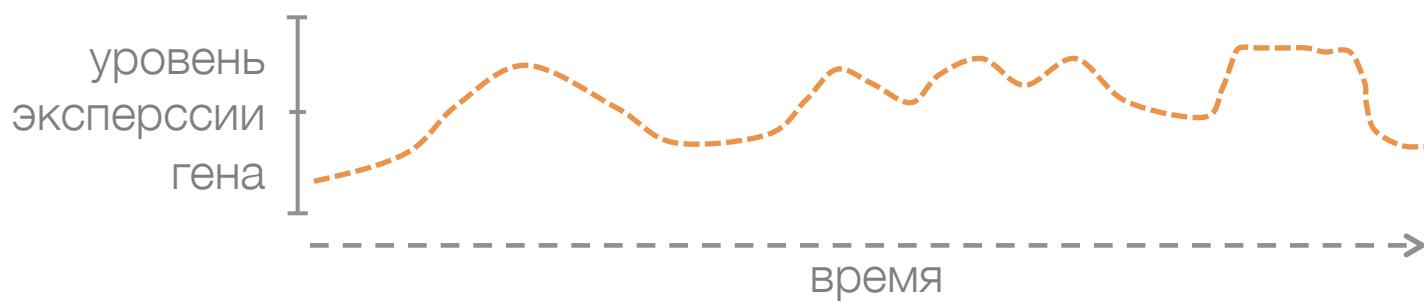
WE NEED TO GO DEEPER

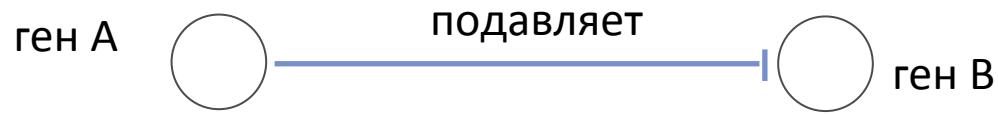


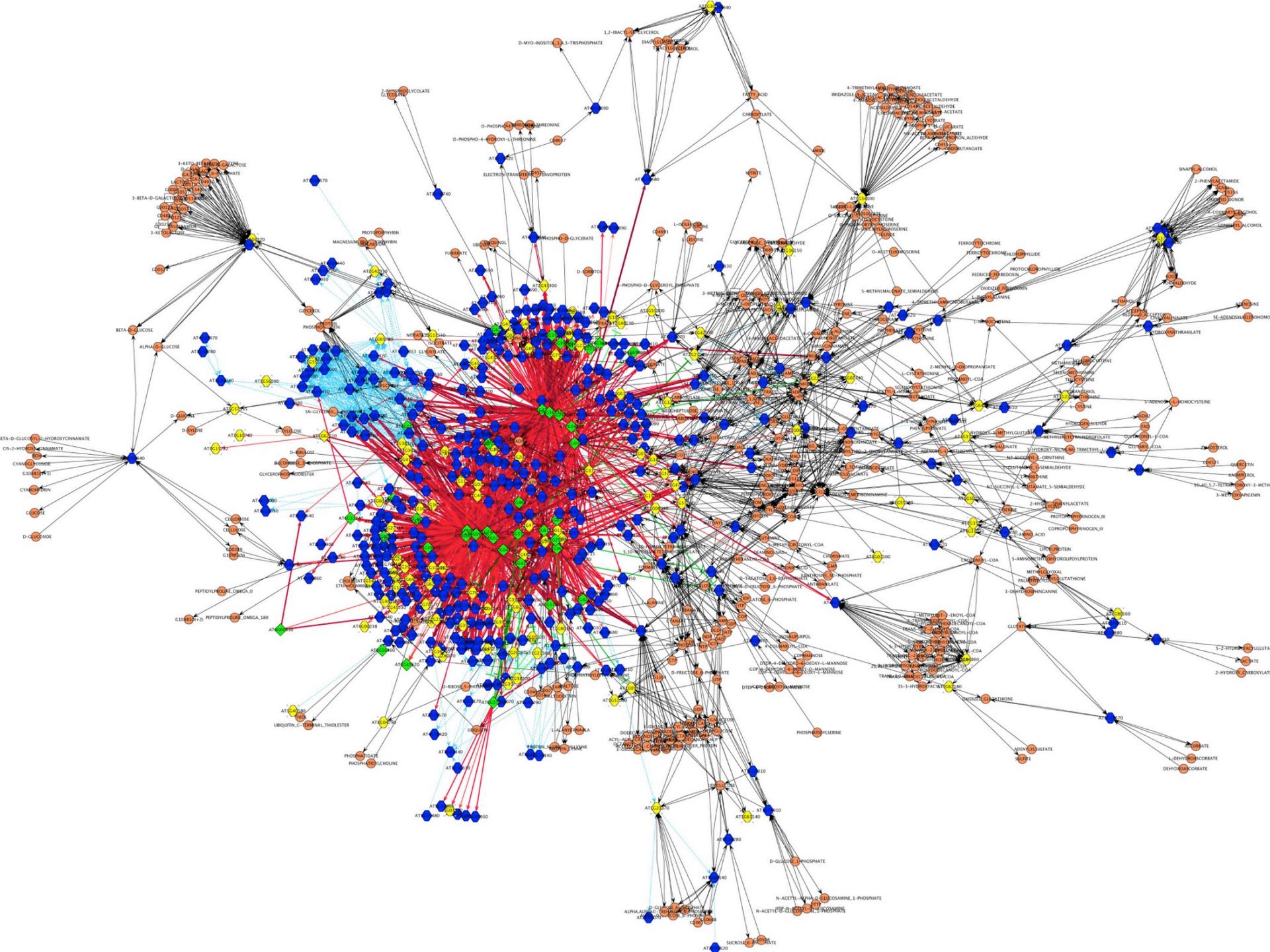
Преобразование
генов в белки



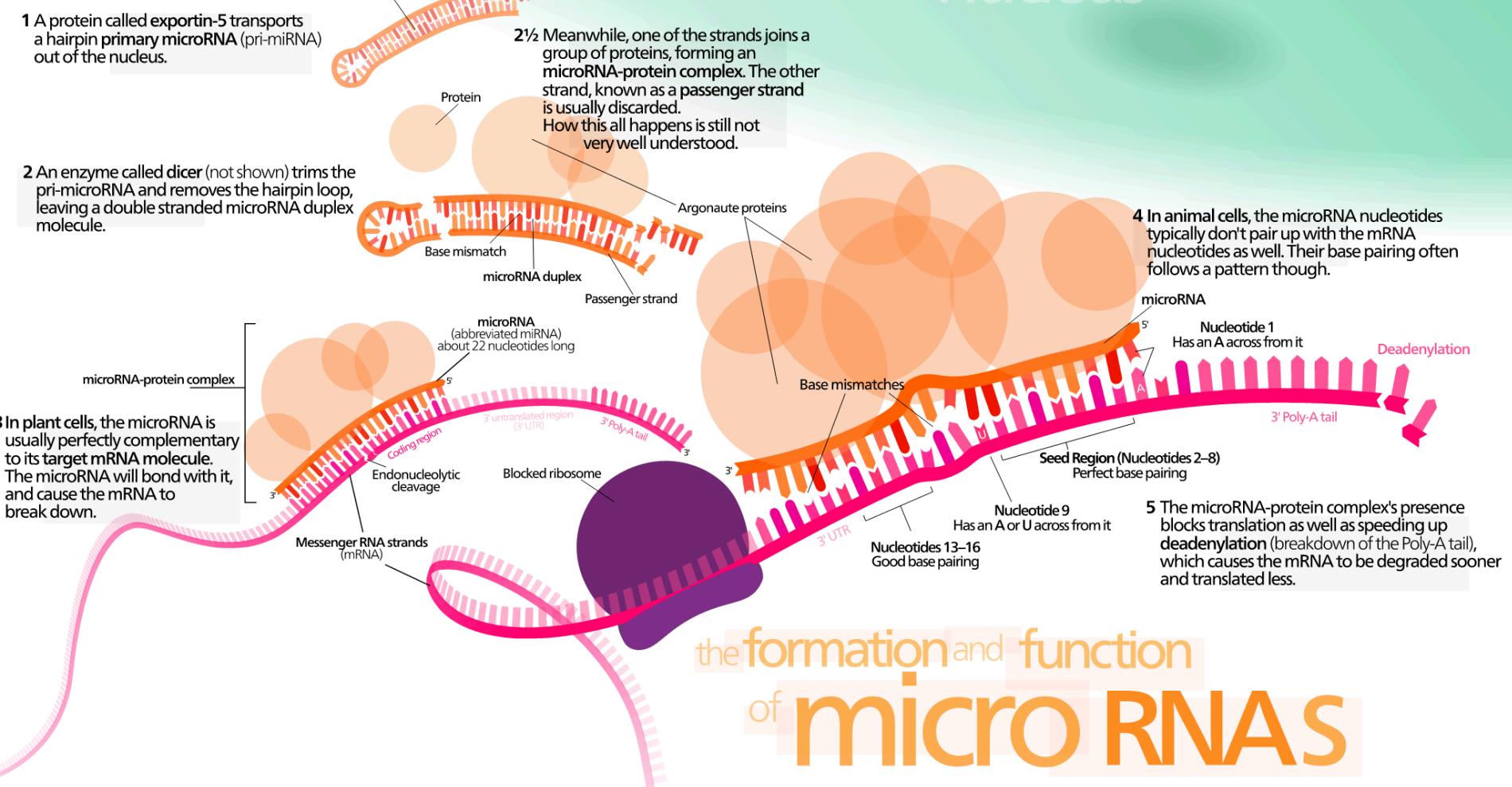
Уровень экспрессии определяет скорость
«производства» белков













Чего ждать в будущем?









Have a questions?