

ЭНШЕОЛОПЕДИЯ МЕТОДОВ NGS

Анастасия Жарикова – 10 марта 2021
azharikova89@gmail.com

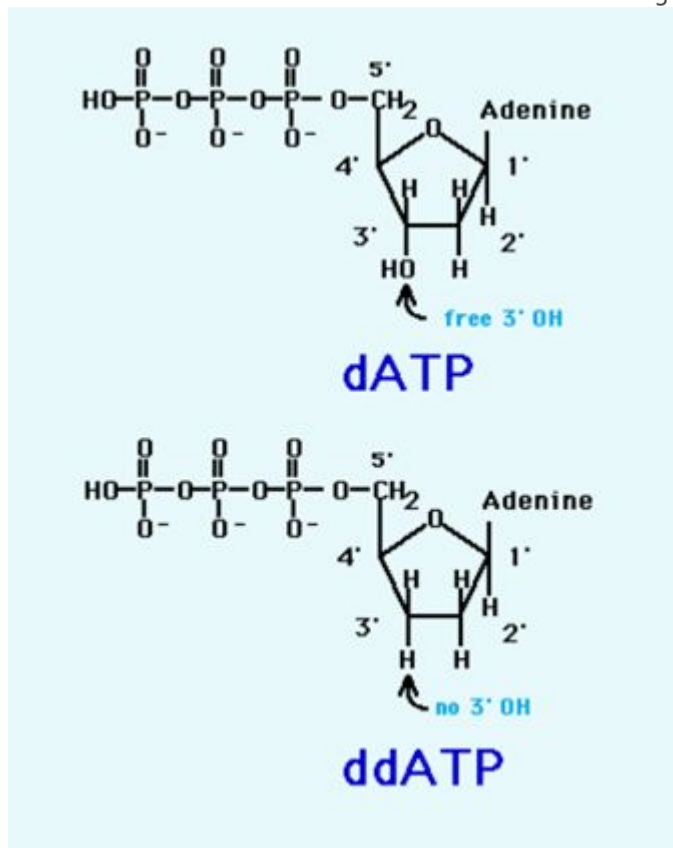
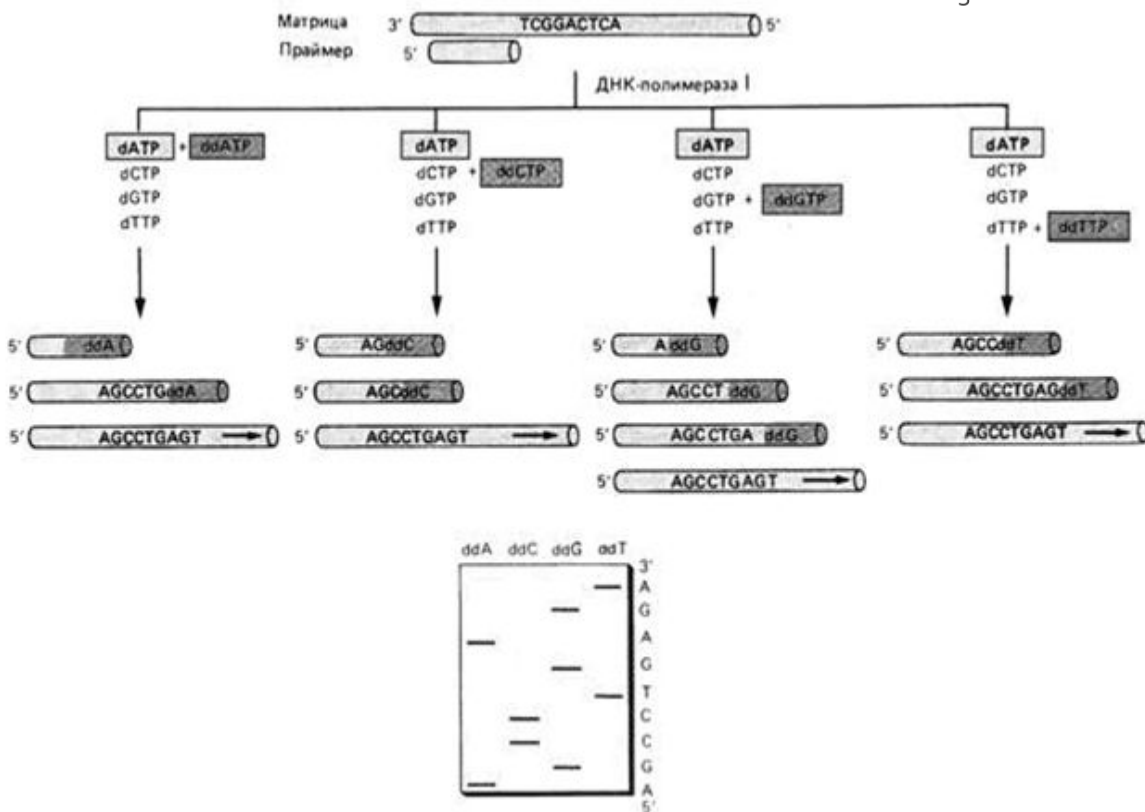
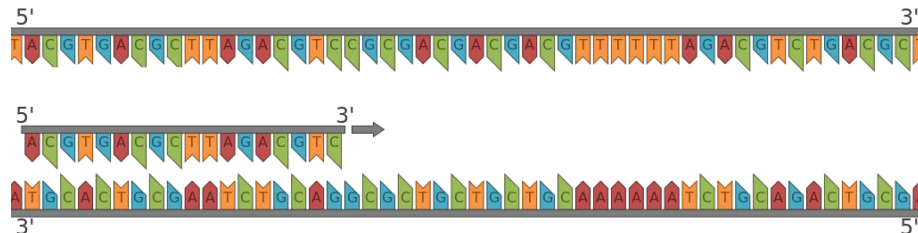
NGS

Next Generation Sequencing

Секвенирование следующего (нового) поколения

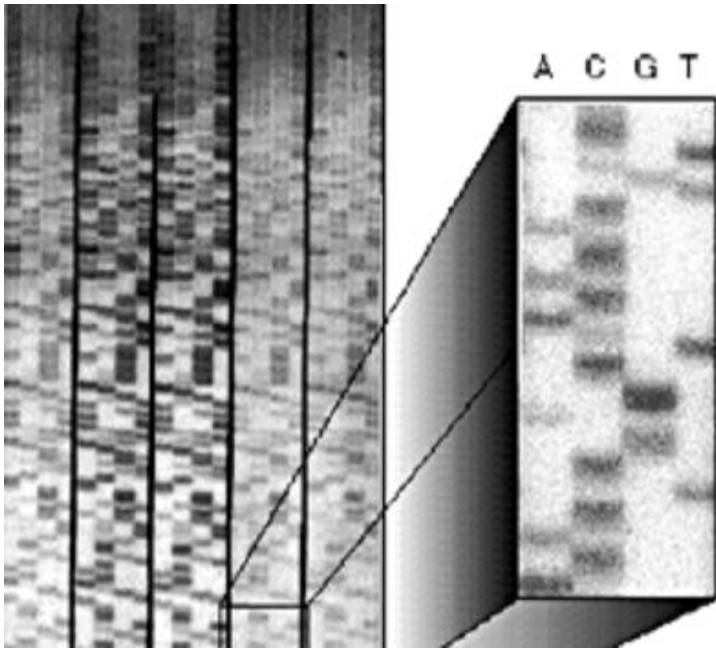
СЕКВЕНИРОВАНИЕ ДНК

Метод обрыва цепи



Фредерик Сэнгер - 1977 год
 Нобелевская премия по химии - 1980 год

МЕТОД ОБРЫВА ЦЕПИ - РЕЗУЛЬТАТ "ТОГДА"



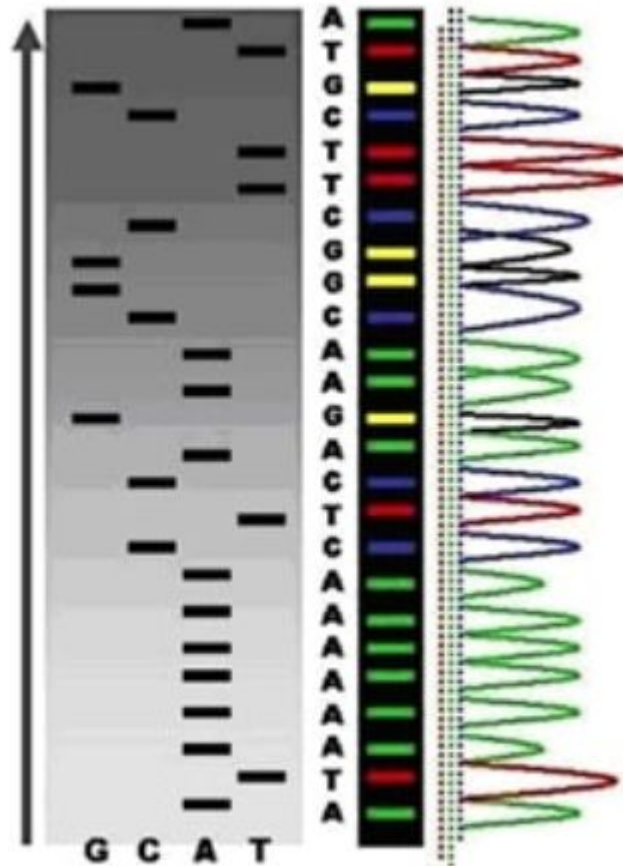
“-” / Длинные фрагменты / Конец посл-ти

“+” / Короткие фрагменты / Начало посл-ти

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

ПААГ

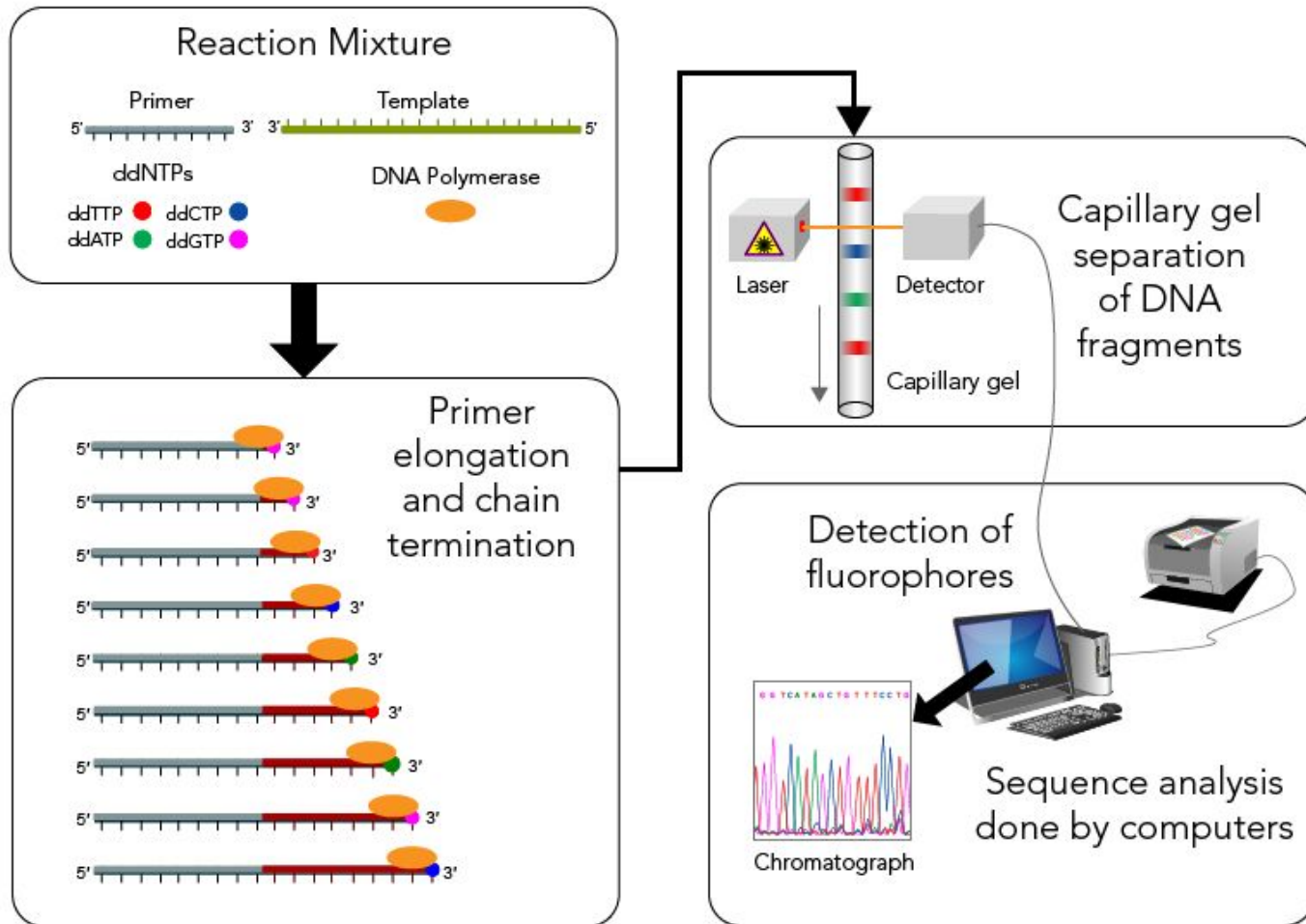
МЕТОД ОБРЫВА ЦЕПИ - РЕЗУЛЬТАТ "СЕЙЧАС"



~ 1000 п.н.

«Золотой стандарт»

АВТОМАТИЗАЦИЯ МЕТОДА ОБРЫВА ЦЕПИ

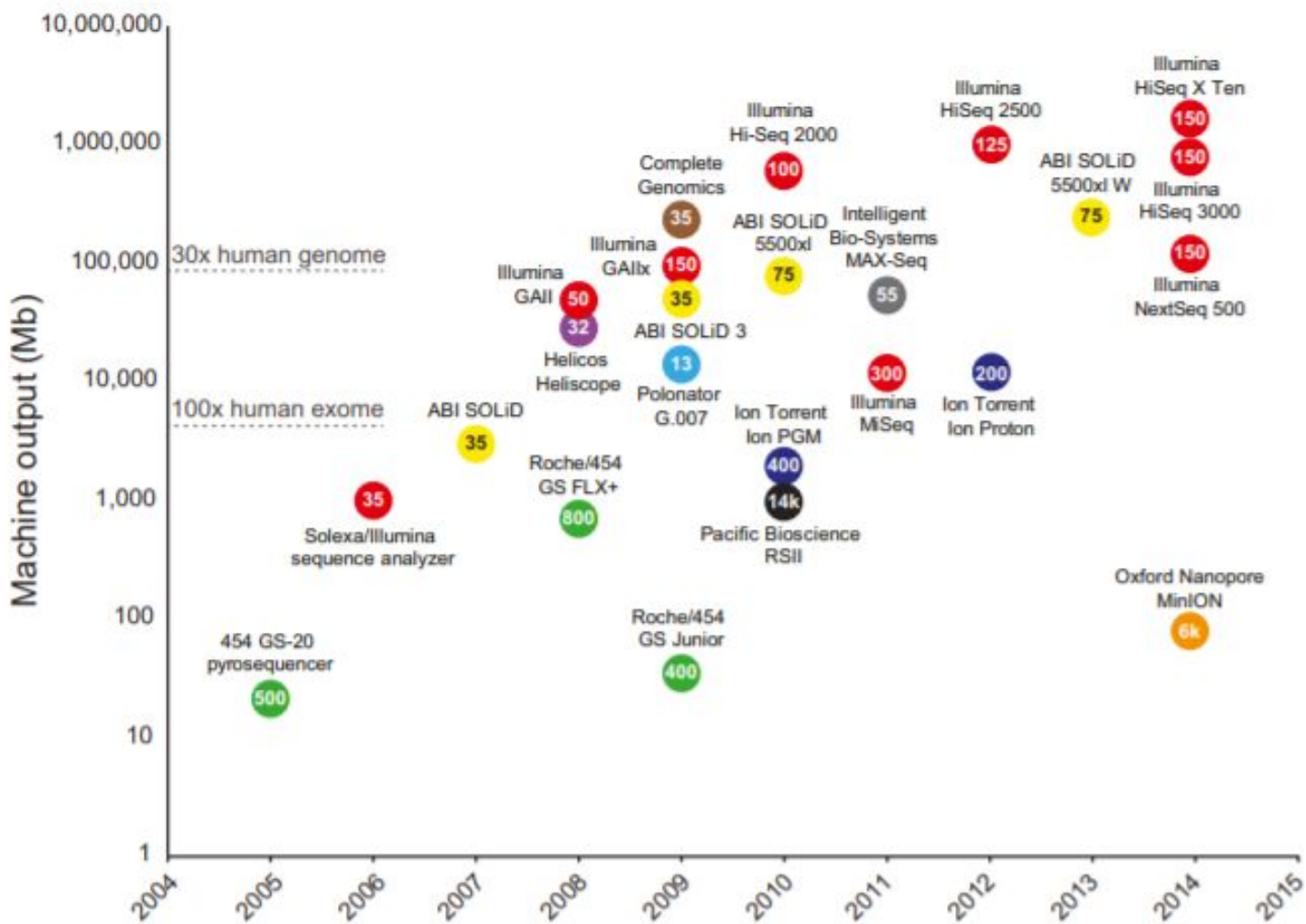


NGS

Метод обрыва цепи - 1 фрагмент за запуск!

NGS - сразу много РАЗНЫХ фрагментов за запуск!

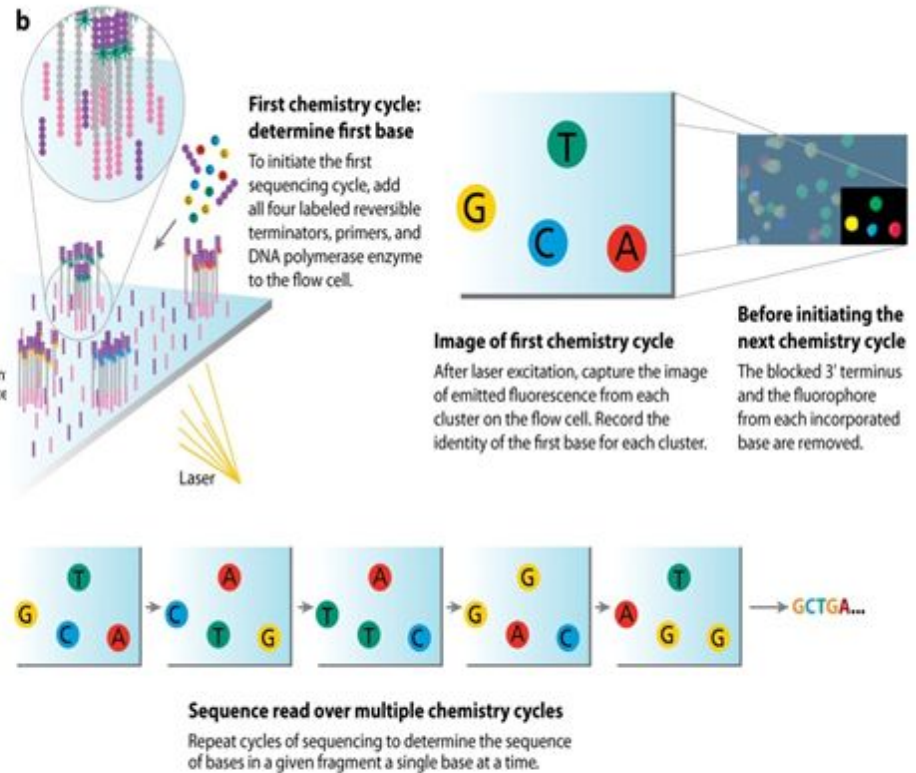
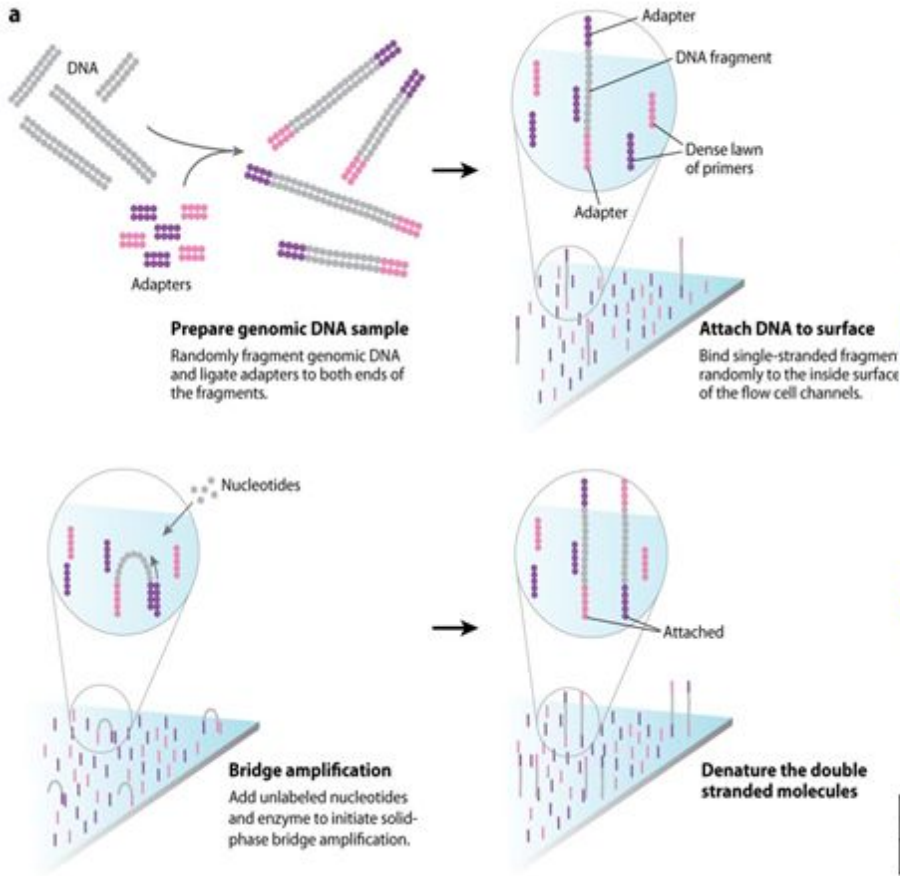
МЕТОДЫ



ГЕНОМНЫЕ ЦЕНТРЫ



ILLUMINA



AR Mardis ER. 2008.
Annu. Rev. Genomics Hum. Genet. 9:387–402

ПОСМОТРИТЕ!

Метод обрыва цепи: [The Sanger Method of DNA Sequencing](#)

ПЦР: [12 методов в картинках: полимеразная цепная реакция](#)

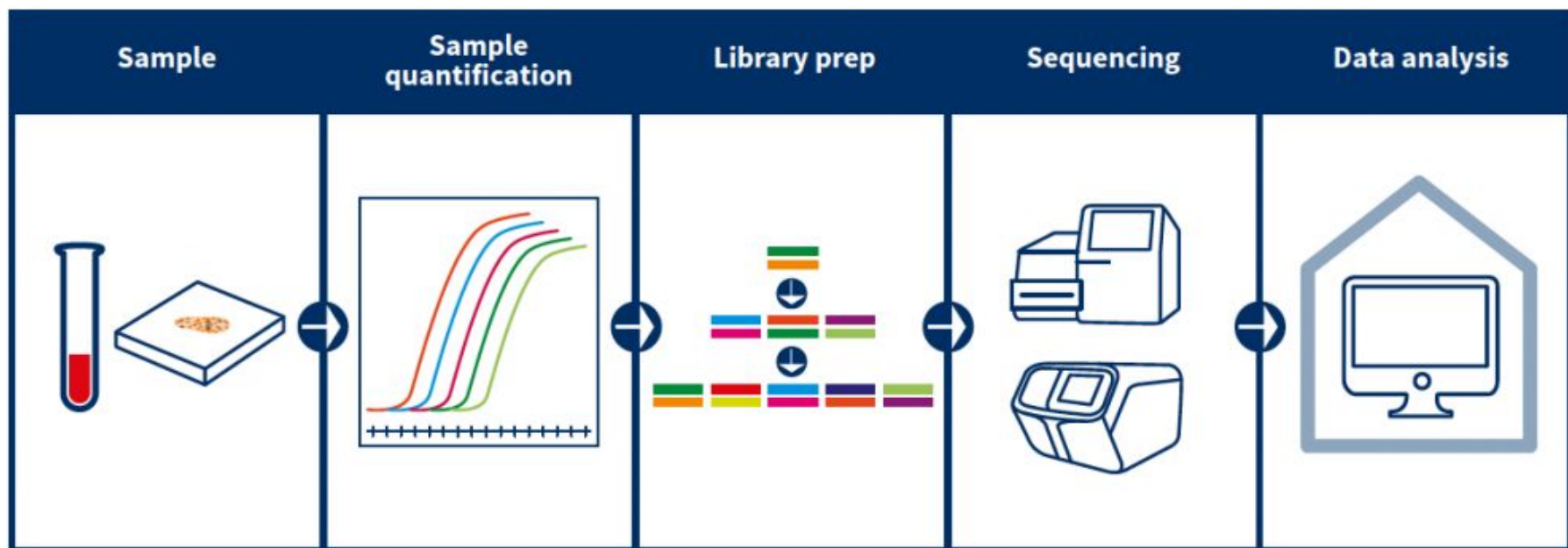
Секвенирование НК: [12 методов в картинках: секвенирование нуклеиновых кислот](#)

Illumina: [Intro to Sequencing by Synthesis: Industry-leading Data Quality](#)

[Milestones in Genomic Sequencing](#)

[Автоматизация омиксных технологий](#)

ОТ ОБРАЗЦА ДО РЕЗУЛТАТА



ЧТО МОЖНО СЕКВЕНИРОВАТЬ?

ДНК

Геном:

- организма с неизвестным геномом – секвенирование *de novo*



[Genome](#) > **Genome Information by Organism**

Overview (59870); [Eukaryotes \(15912\);](#) [Prokaryotes \(314195\);](#) [Viruses \(42364\);](#) [Plasmids \(28280\);](#) [Organelles \(18197\)](#)

РАЗМЕРЫ ГЕНОМОВ



ЧТО МОЖНО СЕКВЕНИРОВАТЬ?

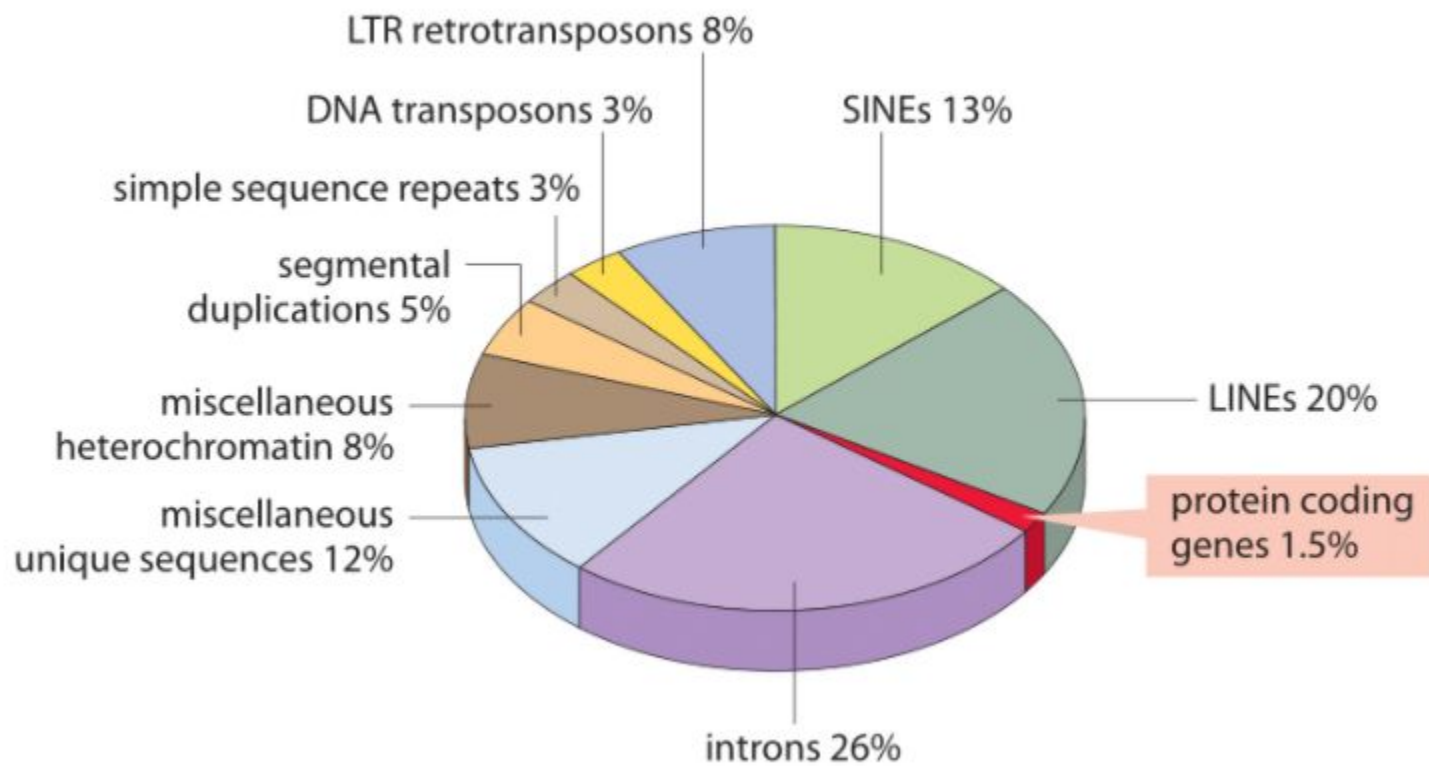
ДНК

Геном:

- организма с известным геномом – ресеквенирование

КОМПОЗИЦИЯ ГЕНОМА

main components of the human genome



МОЖНО РЕСЕКВЕНИРОВАТЬ ПОСЛЕДОВАТЕЛЬНОСТЬ:

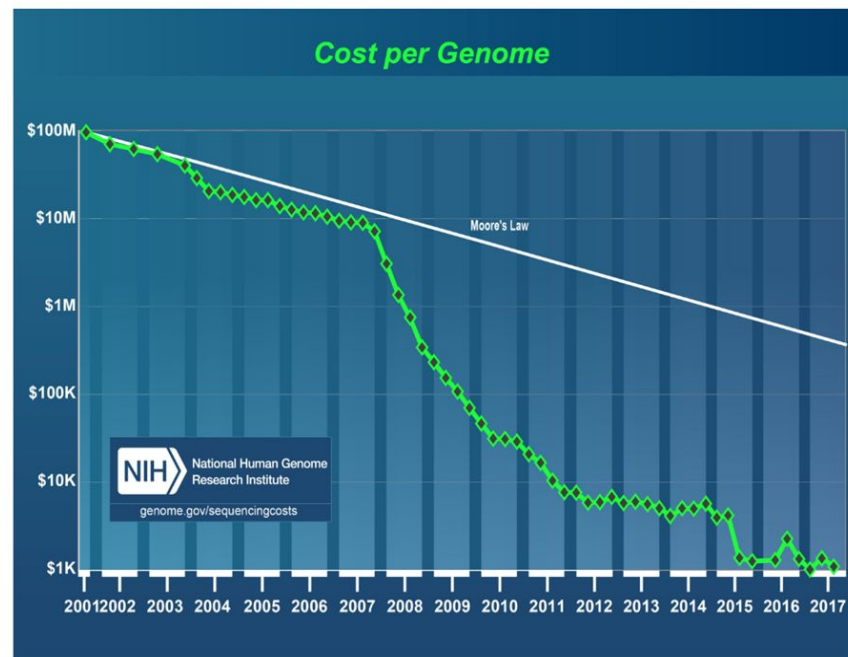
Полного генома

Экзома (кодирующую часть генома)

Отдельные таргетные гены или области

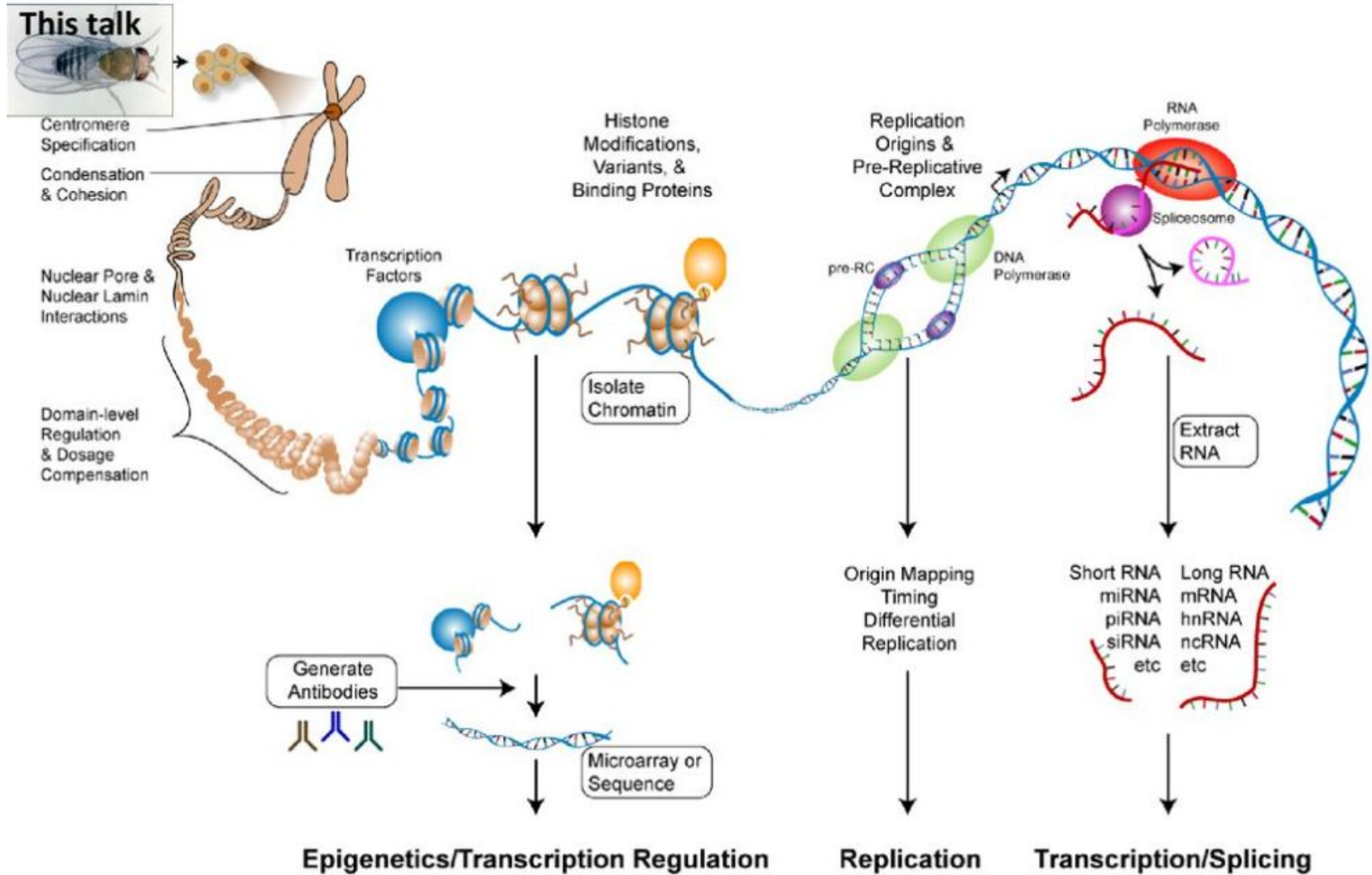
Выбор зависит от:

- бюджета
- целей исследования
- места для хранения данных
- вычислительных мощностей



Источник: Source: [DNA Sequencing Costs: Data](#)

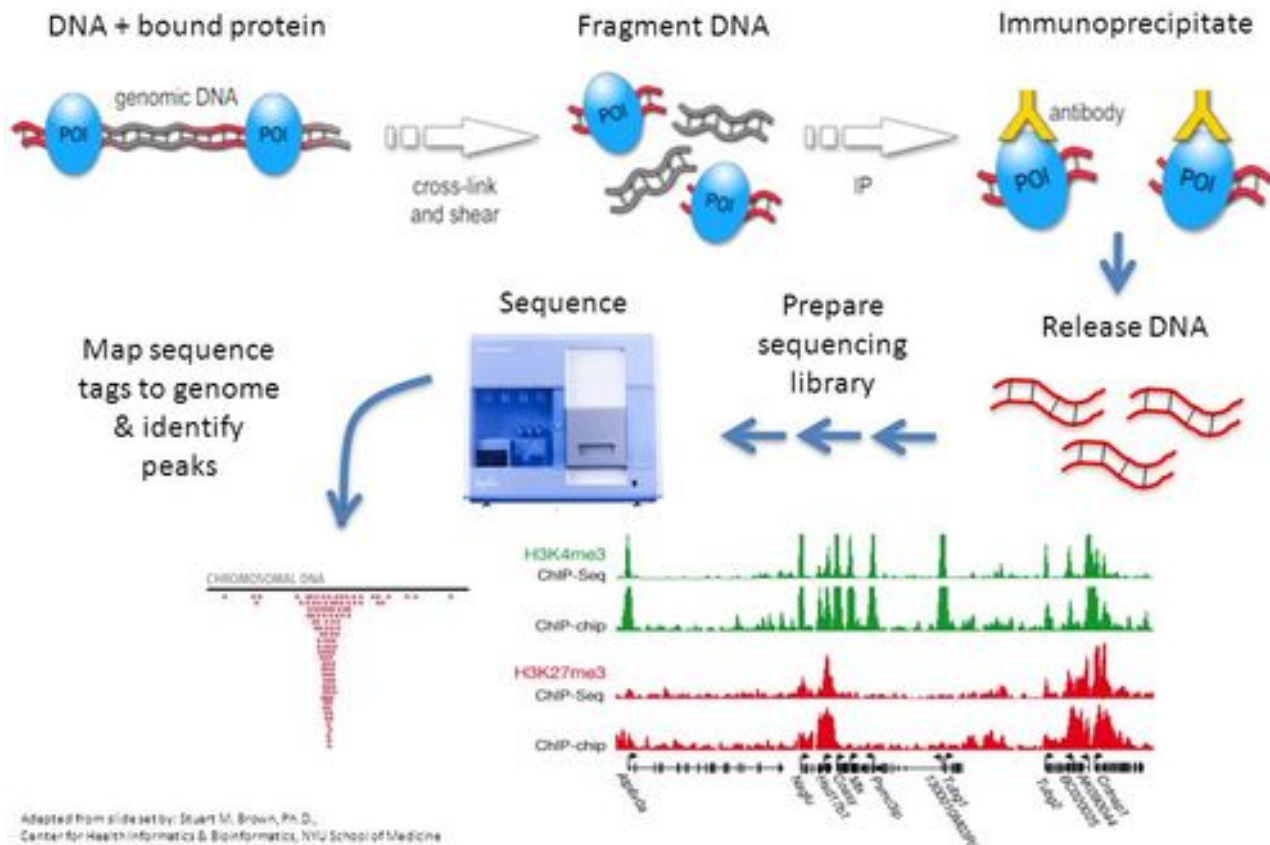
ДНК + ...



CHIP-SEQ

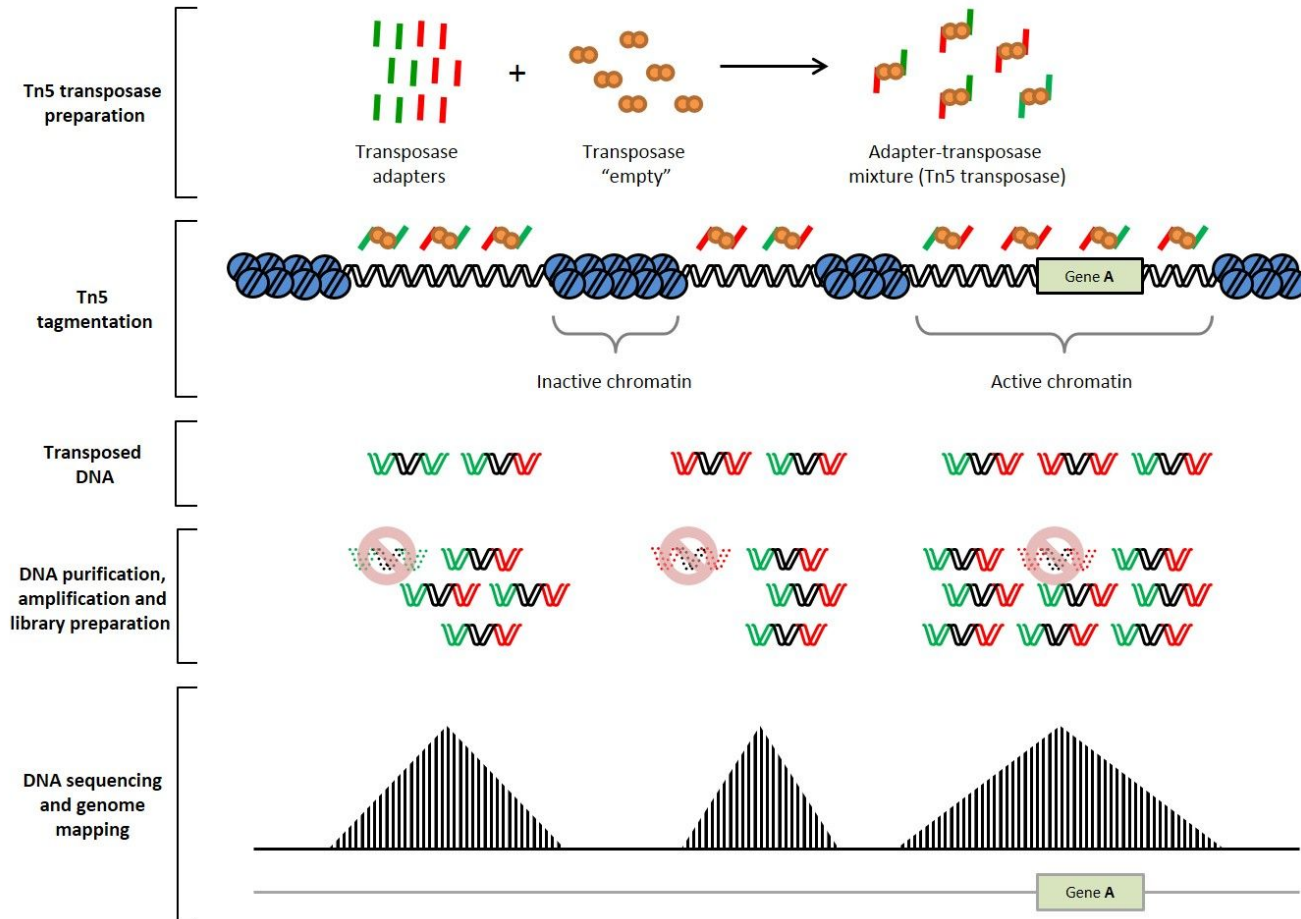
Chromatin immunoprecipitation sequencing

Фрагменты ДНК, связанные с белком

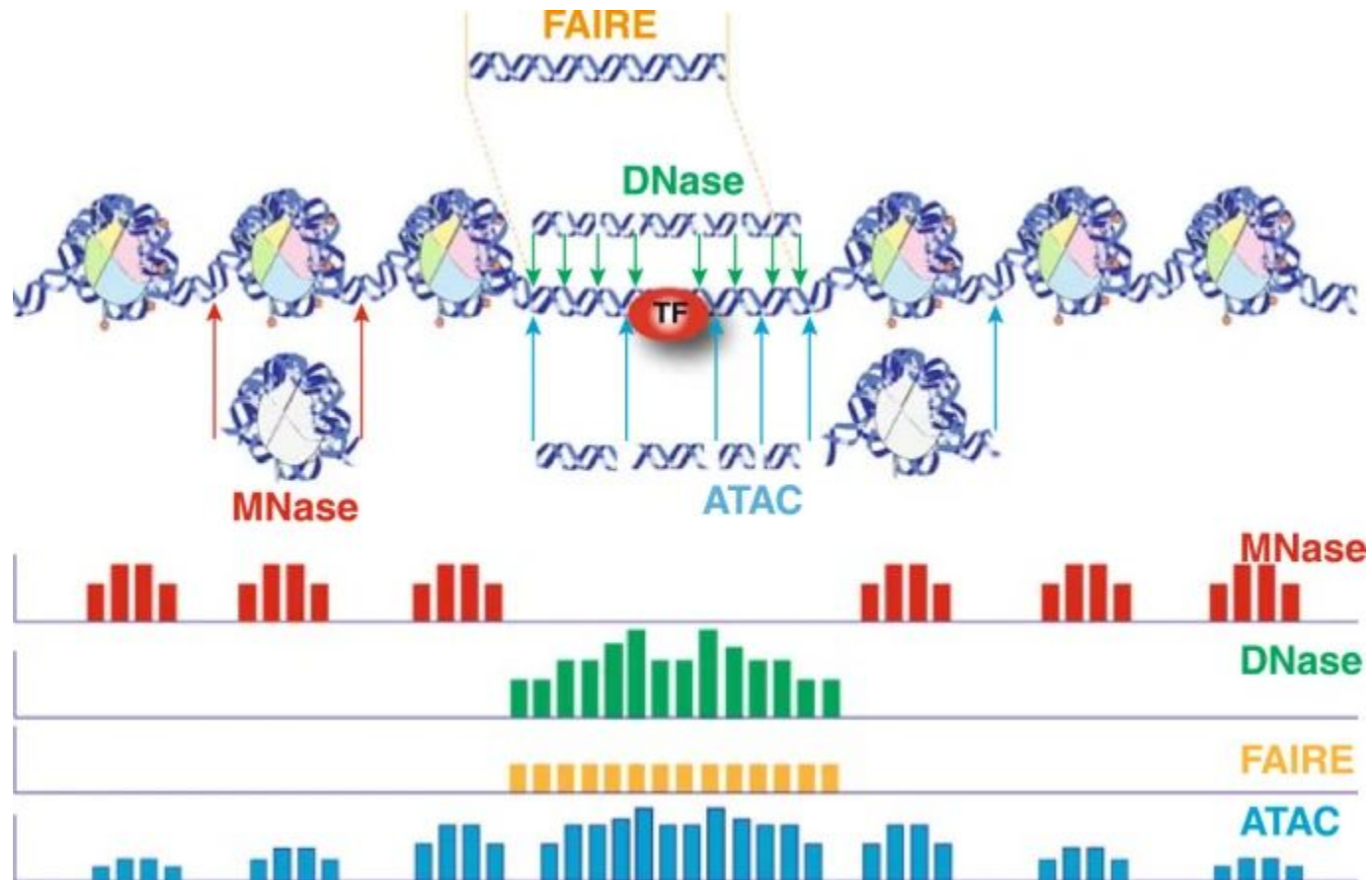


ATAC-SEQ

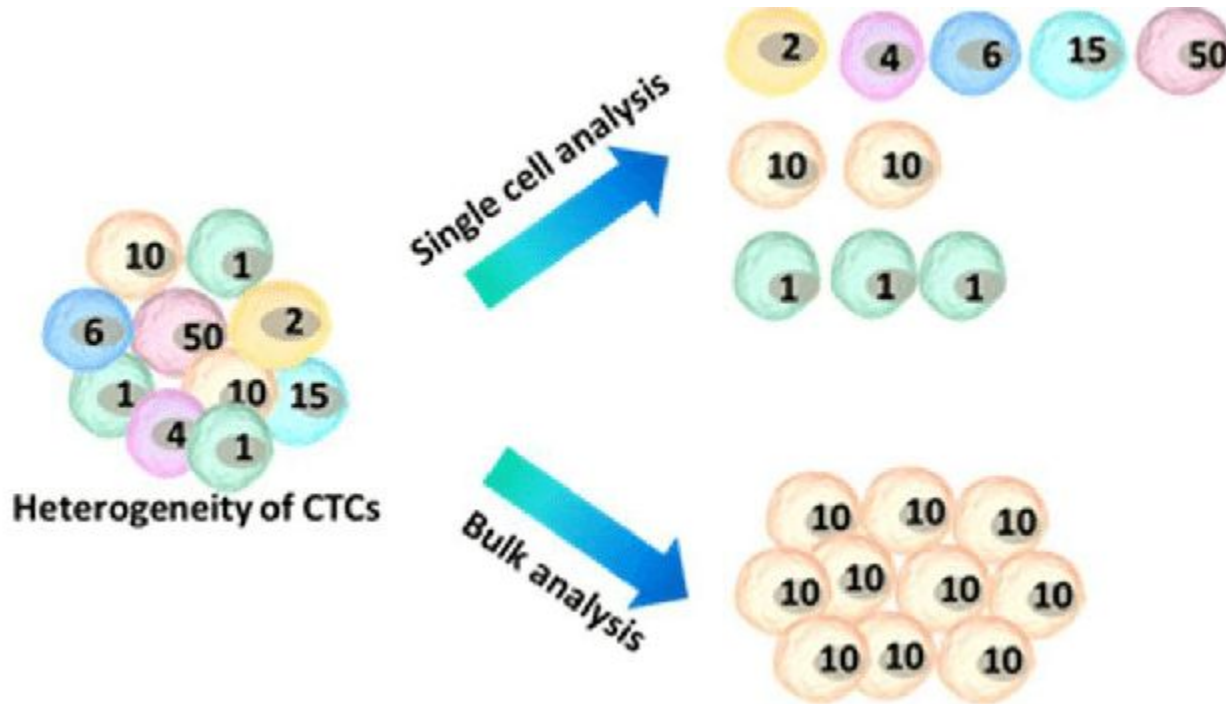
Assay for Transposase-Accessible Chromatin using sequencing



У МЕТОДОВ ЕСТЬ АНАЛОГИ И МОДИФИКАЦИИ



... И SINGLE-CELL ВАРИАЦИИ



 1026

Labs mapped

 386

NGS methods

RNA Sequencing Methods

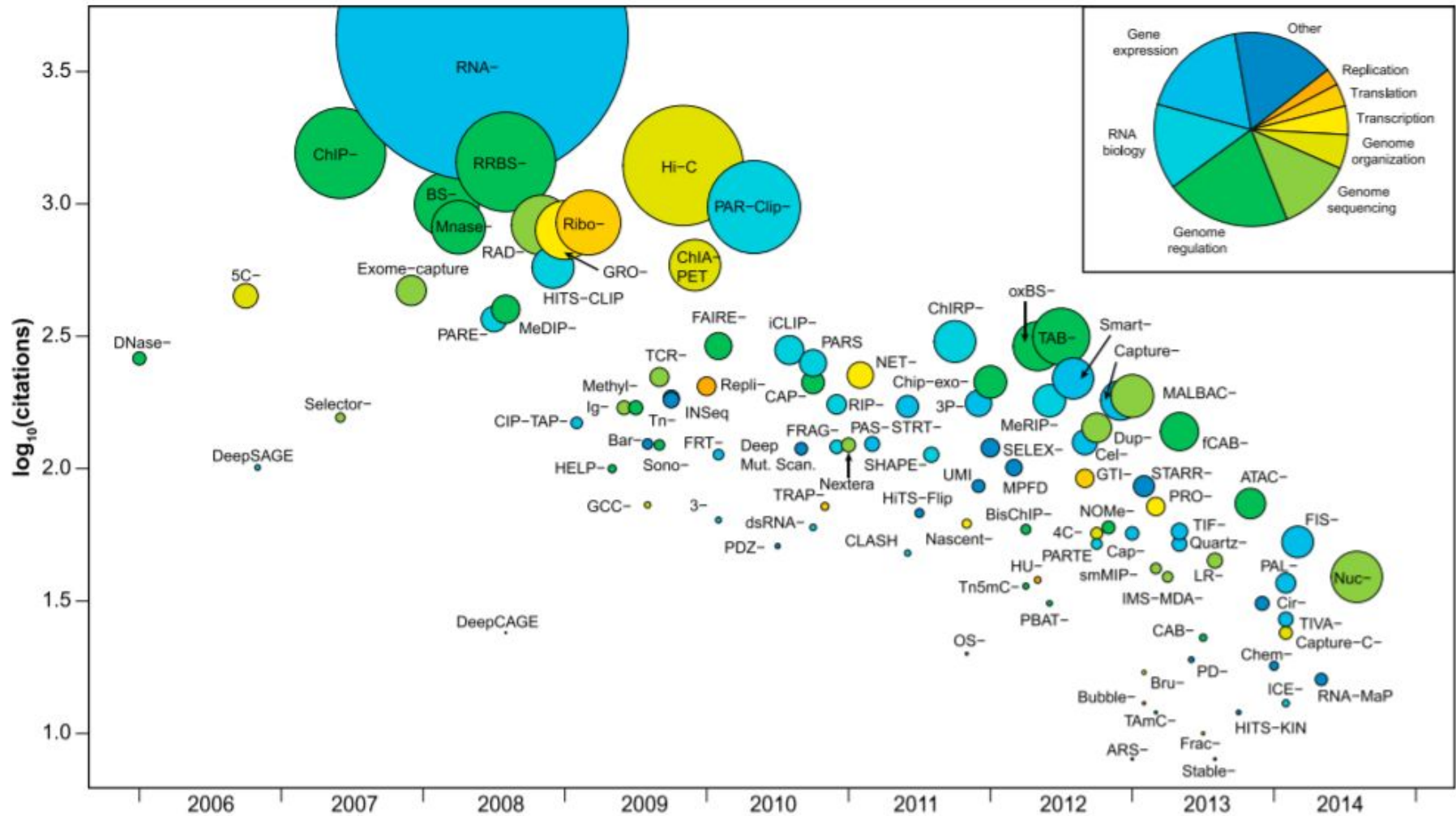
- Low-Level RNA Detection
 - CEL-Seq
 - CirSeq
 - CLaP
 - CytoSeq
 - Digital RNA Sequencing
 - DP-Seq
 - Drop-Seq
 - Hi-SCL
 - InDrop
 - MARS-Seq
 - Nuc-Seq
 - PAIR

DNA Sequencing Methods

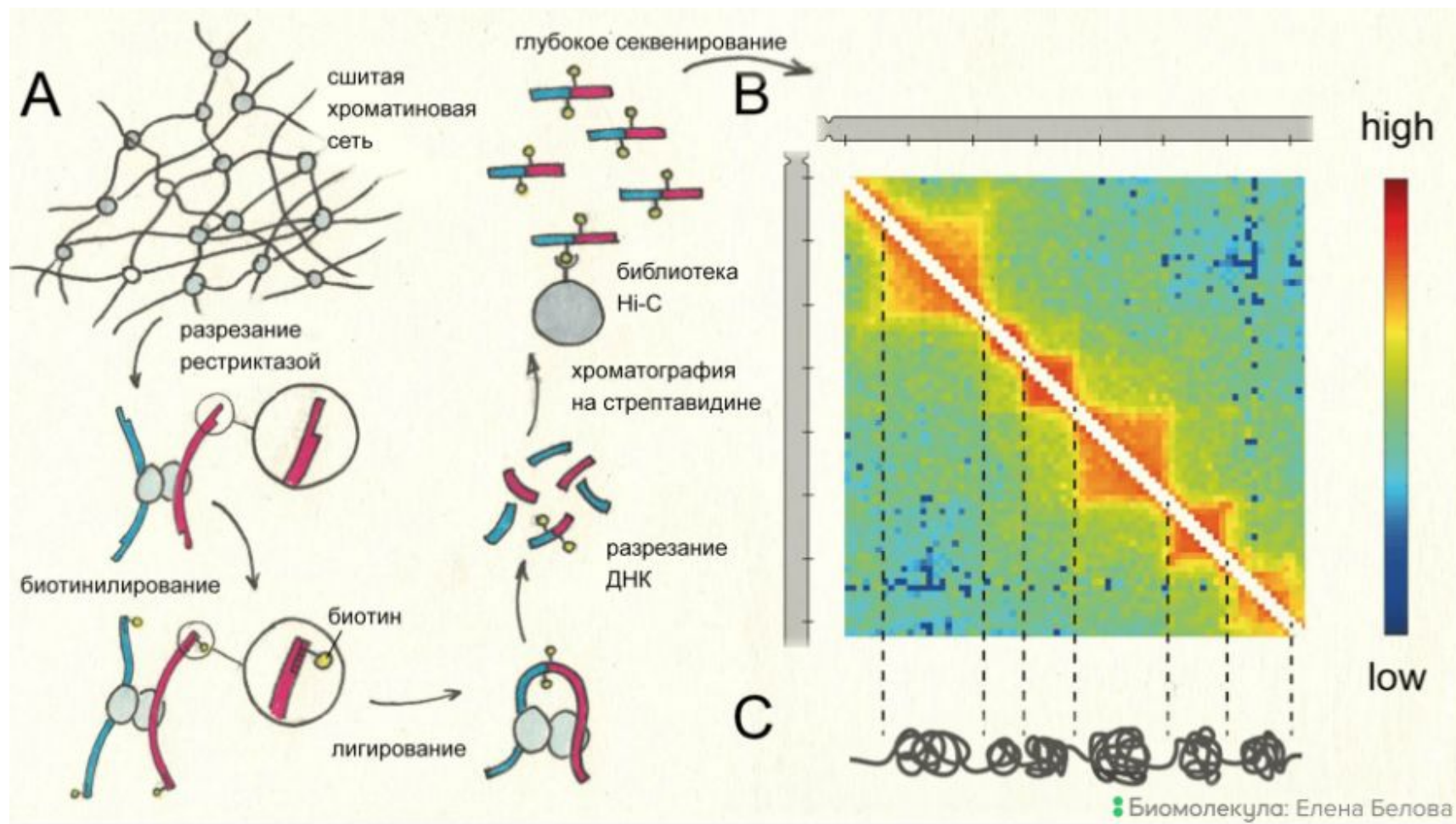
- Protein-Protein Interaction
 - PD-Seq
 - ProP-PD/PDZ-Seq
- Sequence Rearrangements
 - 2b-RAD
 - CPT-seq
 - ddRADseq
 - Digenome-seq
 - EC-seq
 - hyRAD
 - RAD-Seq
 - Rapture
 - RC-Seq
 - Repli-Seq

И ЭТО ЕЩЕ НЕ ВСЕ!

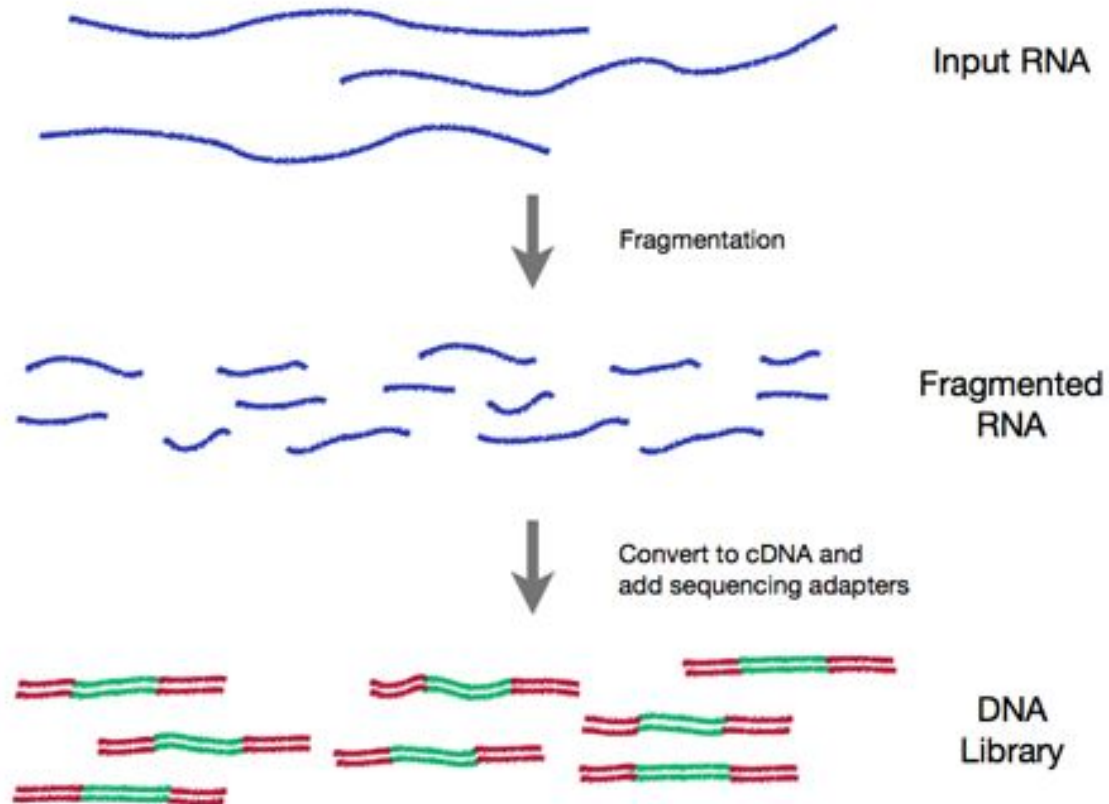
МЕТОДЫ NGS



Hi-C

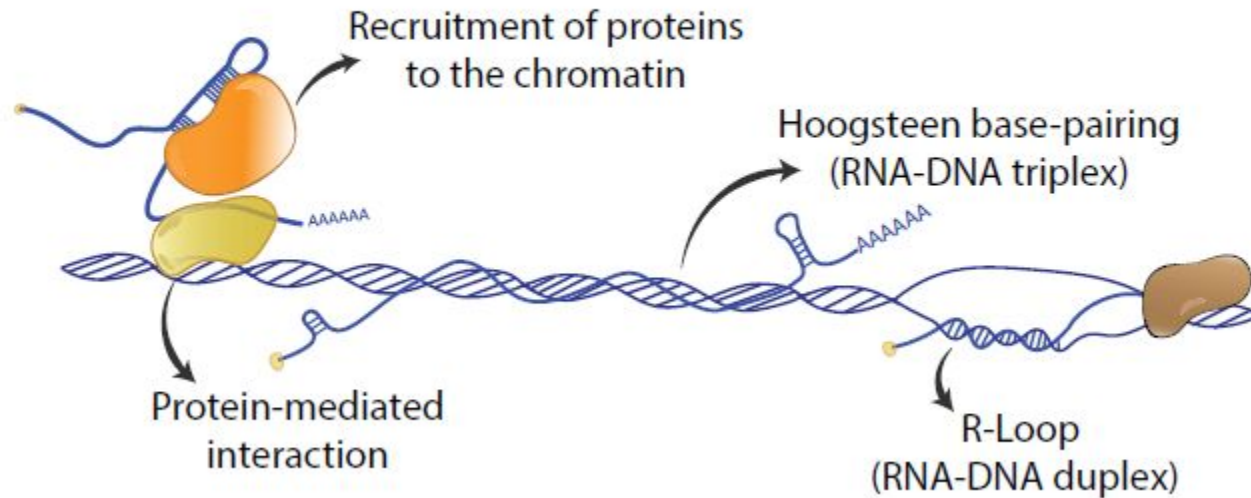


RNA-SEQ



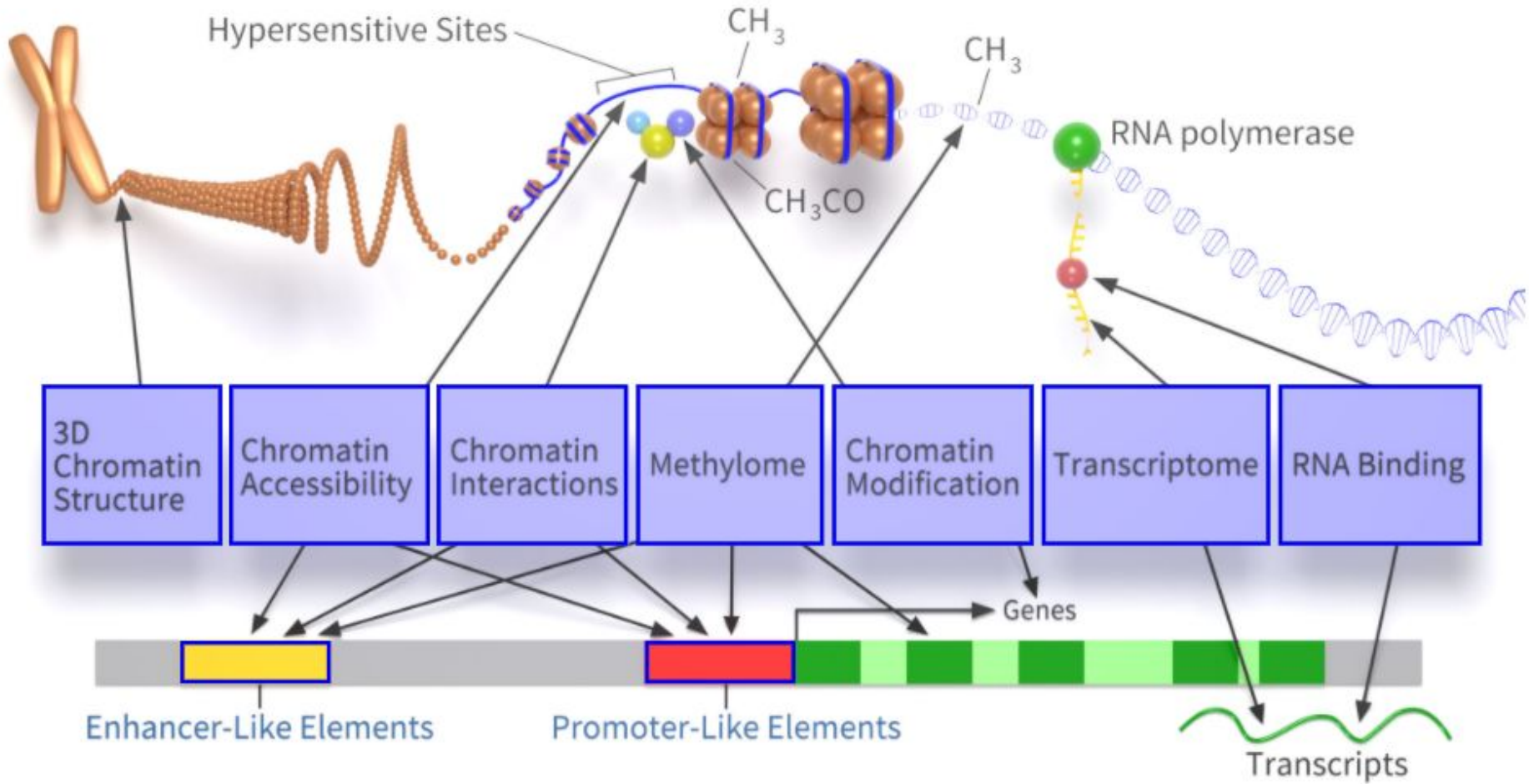
RNA-DNA INTERACTOME

RNA interactions with chromatin



GRID-seq
RADICL
Red-C
MARGI

ENCODE

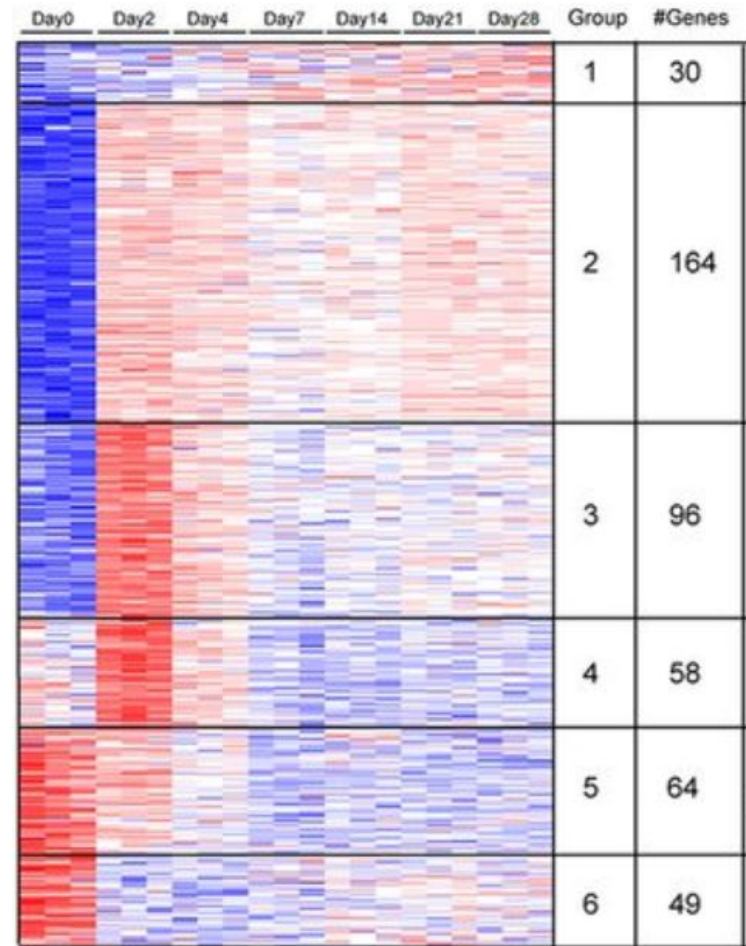
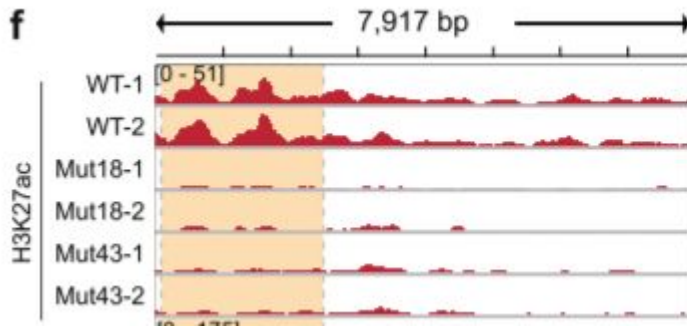


ENCODE

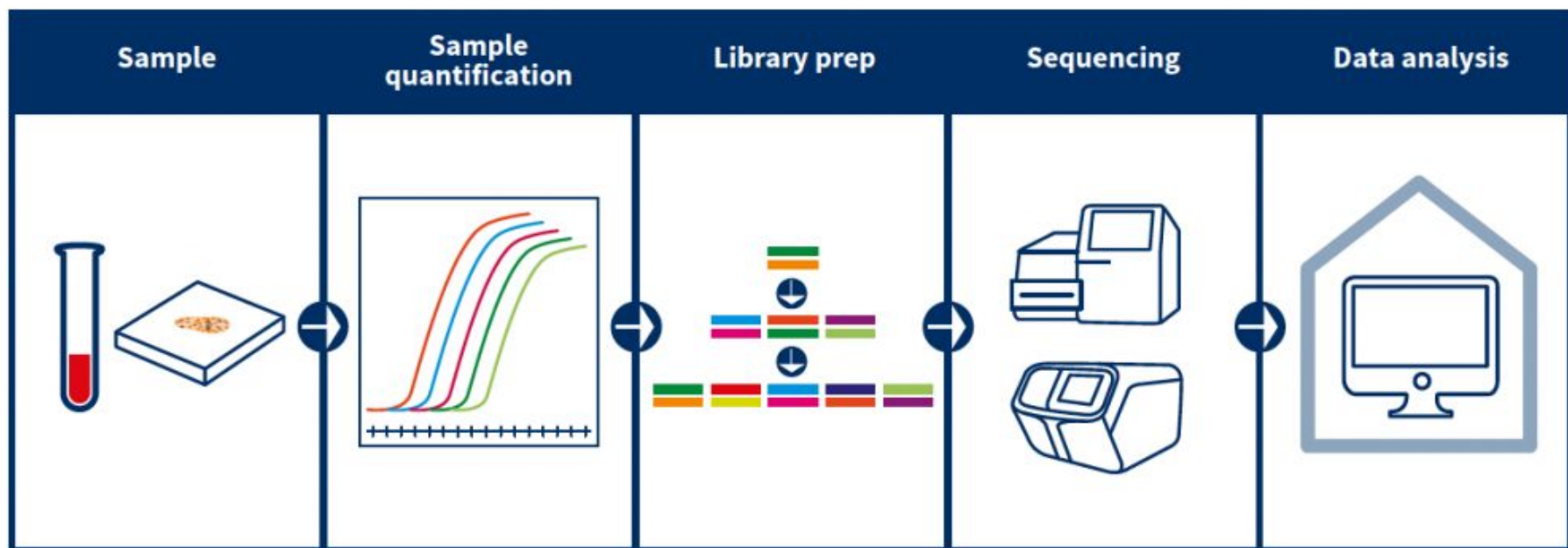
← BIOS

	TF ChIP-seq	Histone ChIP-seq	Control ChIP-seq	DNase-seq	scRNA-seq	total RNA-seq	polyA plus RNA-seq	shRNA RNA-seq	ATAC-seq	Mint-ChIP-seq	microRNA-seq	DNase array	CRISPR RNA-seq	Control eCLIP	eCLIP	small RNA-seq	WGBS	RNA Bind-n-Seq	RAMPAGE	RNA microarray	genotyping array	CAGE	microRNA counts
cell line	2657	750	703	210	2	115	185	553	123	16	26	91	252	232	223	111	18		30	73	75	53	8
K562	634	19	186	7		13	19	285	2		2	3	140	125	120	7	1		1	10	2	9	1
HepG2	654	15	77	2		6	11	268	2		2	3	112	107	103	3	2			6	2	6	1
A549	243	86	77	14			27		6			2				9	1			2	2	3	
GM12878	188	15	27	1	2	5	13		2		2	3				6	1		1	7	2	6	1
MCF-7	148	18	34	7		1	4		1		2	2				7				4	2	3	1
tissue	332	1793	567	703	14	205	397		181		180	122		2	2	67	162		104	2	7	17	101
liver	42	91	29	15		3	20		7		7	1				1	9		2				7
stomach	17	72	26	22		5	15		6		4	3				4	10		5				4
heart	5	79	15	21		3	16		7		9					1	10		2				8
spleen	18	58	23	6		8	11		1		3	4				4	6		4				1
lung	6	58	15	17		2	11		5		4	2				1	7		1				4
whole organisms	996		987			272	80	105															19
whole organism	996		987			268	72	105															15
carcass						4	8																4
primary cell	68	521	130	473	8	162	97		27	236	36	38				24	7		28	58	37	30	1
macrophage						78	14																
foreskin keratinocyte	3	25	9	2		3	5		13		3								13				
T-cell		11	3	60		3	1			6							1						
activated CD4-positive, alpha-beta T cell				83																			
CD4-positive, alpha-beta T cell		18	4	46			2			3													

РАЗНИЦА МЕЖДУ КЛЕТКАМИ



ОТ ОБРАЗЦА ДО РЕЗУЛТАТА



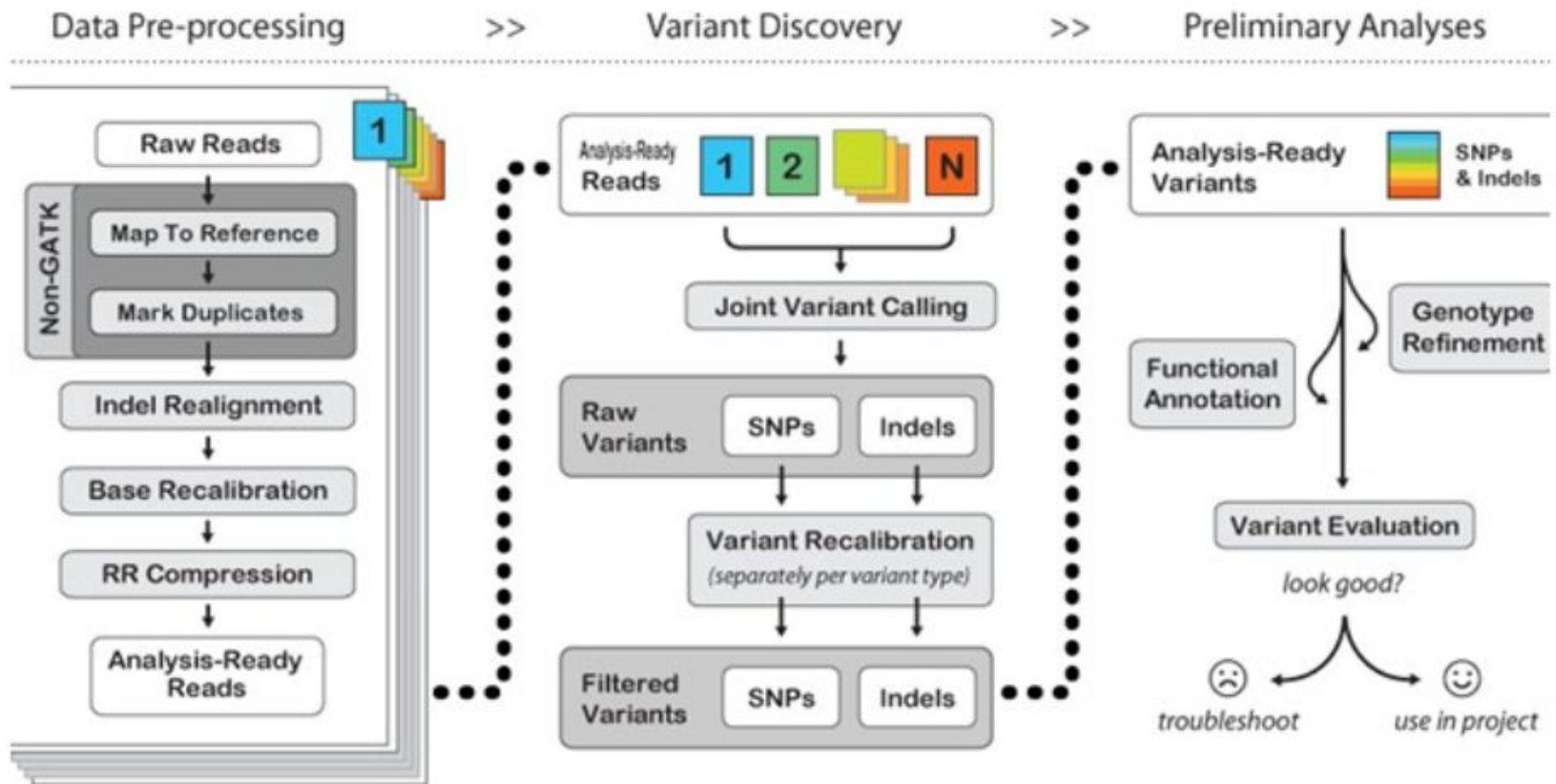
ОСНОВНЫЕ ЭТАПЫ - УЗНАТЬ ВСЁ

Постановка задачи:

- РНК / ДНК
- организм, тип клеток, клеточная линия
- платформа (Illumina, ...)
- метод (chip-seq, rna-seq, dna-seq, ...)
- биологическая задача (поиск мутаций, дифференциальная экспрессия, взаимодействия, ...)

Планирование анализа

АНАЛИЗ ДАННЫХ NGS



Taken from: <http://www.broadinstitute.org/gatk/guide/best-practices>

ОСНОВНЫЕ ЭТАПЫ

Оценка качества и количества чтений

Картирование на геном / Сборка

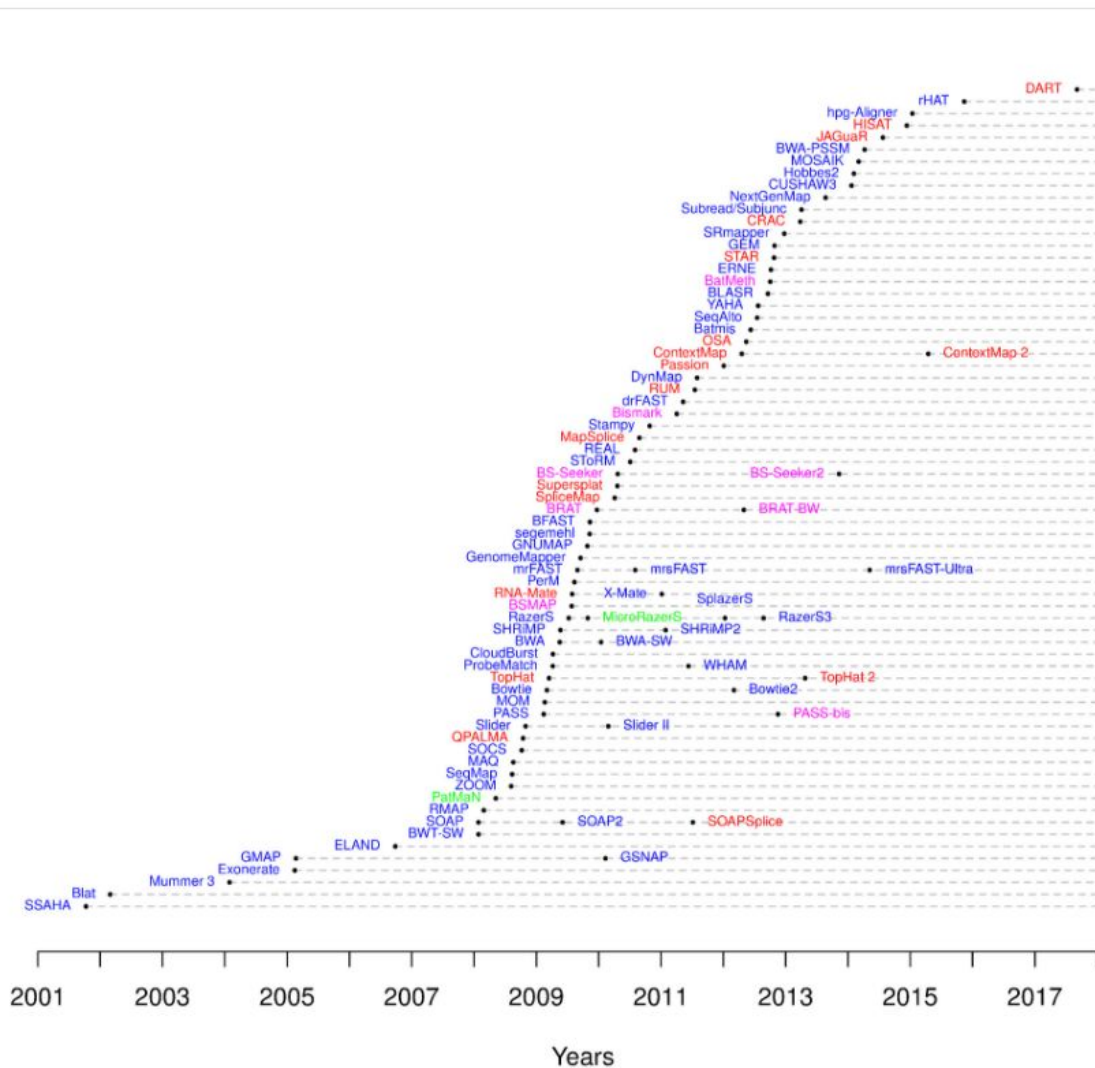
ВЕРСИИ ГЕНОМОВ

Release name	Date of release	Equivalent UCSC version
GRCh38	Dec 2013	hg38
GRCh37	Feb 2009	hg19
NCBI Build 36.1	Mar 2006	hg18
NCBI Build 35	May 2004	hg17
NCBI Build 34	Jul 2003	hg16

Recent human genome assemblies chart, Wikipedia.

МНОЖЕСТВО ПРОГРАММ

Read mappers timeline



ФОРМАТЫ ФАЙЛОВ

fasta

fastq

bed

bedGraph

bigBed

sam

bam

gtf

gff

vcf

gvcf

wig

bigWig

Для КАЖДОГО МЕТОДА СВОЙ НАБОР ПРОГРАММ

Необходимо учитывать:

- специфику метода
- артефакты метода
- “шум”
- биологическую задачу
- детекция результата
- визуализация
- ...

НАВЫКИ БИОИНФОРМАТИКА В НАУКЕ

Генетика / молекулярная биология / ботаника / зоология

Программирование / администрирование

Знание английского

Дизайн

Спикер

Писатель

Мыслитель

Философ



ДОМАШНЕЕ ЗАДАНИЕ

2 варианта:

- описание одного метода NGS
- исследование хроматограммы

Для зачета нужно выбрать и сделать ОДИН вариант задания

Можно выполнить ДВА варианта - будут бонусные баллы