

# Introduction to Bioinformatics

## Day 3: Genomics, Metagenomics and Transcriptomics

14<sup>th</sup> January 2026

# Outline Day 3

## Genomics, Metagenomics and Transcriptomics

- **Morning (9-12 pm):**
  - **Introduction to Genomics and Metagenomics:**
    - Introduction to Genomics
    - Genome sequencing (Sanger, NextGen, ThirdGen) technologies.
    - Introduction to Metagenomics
  - **Practical Session:**
    - Exploring NCBI Genome Data Viewer, Accessing Genomic and MGnify Metagenomic Data, and Analyzing with Biopython
- **Afternoon (2-5 pm):**
  - **Introduction to Transcriptomics:**
    - RNA sequencing and transcriptome analysis.
    - Expression profiling and differential expression analysis.
  - **Practical Session:**
    - Analyze RNA-seq data using python

# Introduction to Bioinformatics

## Day 3: Genomics, Metagenomics and Transcriptomics

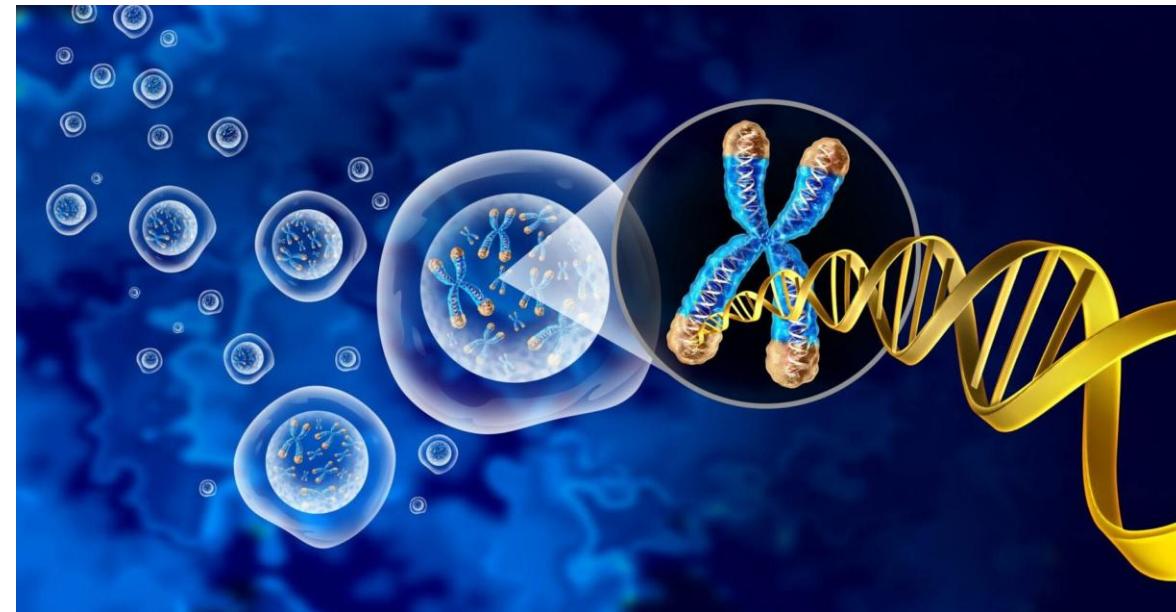
14<sup>th</sup> January 2026

Morning Session

# Introduction to Genomics

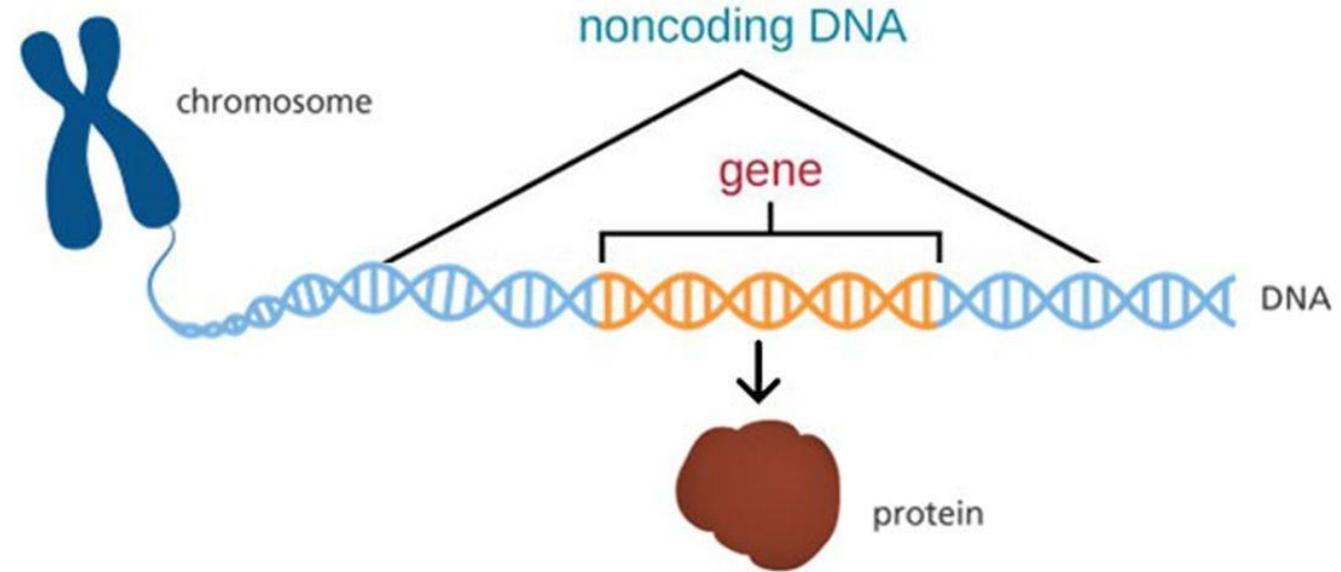
# What is a genome?

A genome is the complete set of DNA in an organism. From potatoes to puppies, every living organism has its own unique genome, containing all the instructions needed to build and sustain life.



# What is a genome?

- Genomics is the study of an organism's complete set of
- DNA, including all of its genes and non-coding regions.



# What is Genetics?

Genetics is the scientific study of genes, heredity, and variation in living organisms.

# Genetics vs Genomics

Aspect	Genetics	Genomics
<b>Definition</b>	The study of individual genes and their roles in inheritance.	The study of the entire genome, including all genes and non-coding regions.
<b>Focus</b>	Specific genes or groups of genes.	The entire set of an organism's DNA.
<b>Scope</b>	Narrow: How traits are inherited and passed on.	Broad: Interactions, functions, and organization of all genetic material.
<b>Methods</b>	Gene mapping, Mendelian inheritance, genetic crosses.	Genome sequencing, functional genomics, comparative genomics.
<b>Applications</b>	Studying inherited diseases, gene therapy.	Precision medicine, evolutionary studies, understanding complex traits.
<b>Example</b>	Investigating the BRCA1 gene in breast cancer.	Sequencing the human genome to identify all genes and variations.

# How Genetics is used in Genomics?

Genetics is the backbone of genomics, enabling a deeper understanding of the genome as a whole.

1. **Gene Function:** identify roles of individual genes, which genomics studies on a broader scale.
2. **DNA Variations:** understanding mutations and how they influence traits or diseases.
3. **Heredity Links:** Genomics uses genetic patterns to study inherited conditions.
4. **Genomic Tools:** Genetic data underpins resources like reference genomes and gene annotations.
5. **Data Interpretation:** provides the foundation for analyzing complex genomic interactions

# Genome sequencing technologies

# What is genome sequencing?

Genome sequencing is the process of determining the complete DNA sequence of an organism's genome, including all of its genes and non-coding regions.

# Why do we need to sequence genomes?

To understand the complete genetic blueprint of an organism, which helps in diagnosing diseases, advancing precision medicine, improving agriculture, studying evolution, and discovering new biological insights.

# Applications of Genomics

## ▪ Healthcare

- Personalized medicine: Tailoring treatments based on individual genetic profiles.
- Disease diagnosis: Identifying genetic mutations linked to specific diseases.
- Drug discovery: Identifying novel therapeutic targets.

## ▪ Agriculture

- Crop improvement: Developing disease-resistant and high-yield varieties.
- Livestock breeding: Enhancing traits like productivity and disease resistance.

## ▪ Environmental Science

- Biodiversity conservation: Understanding genetic diversity in populations.
- Bioremediation: Engineering microbes to clean pollutants.

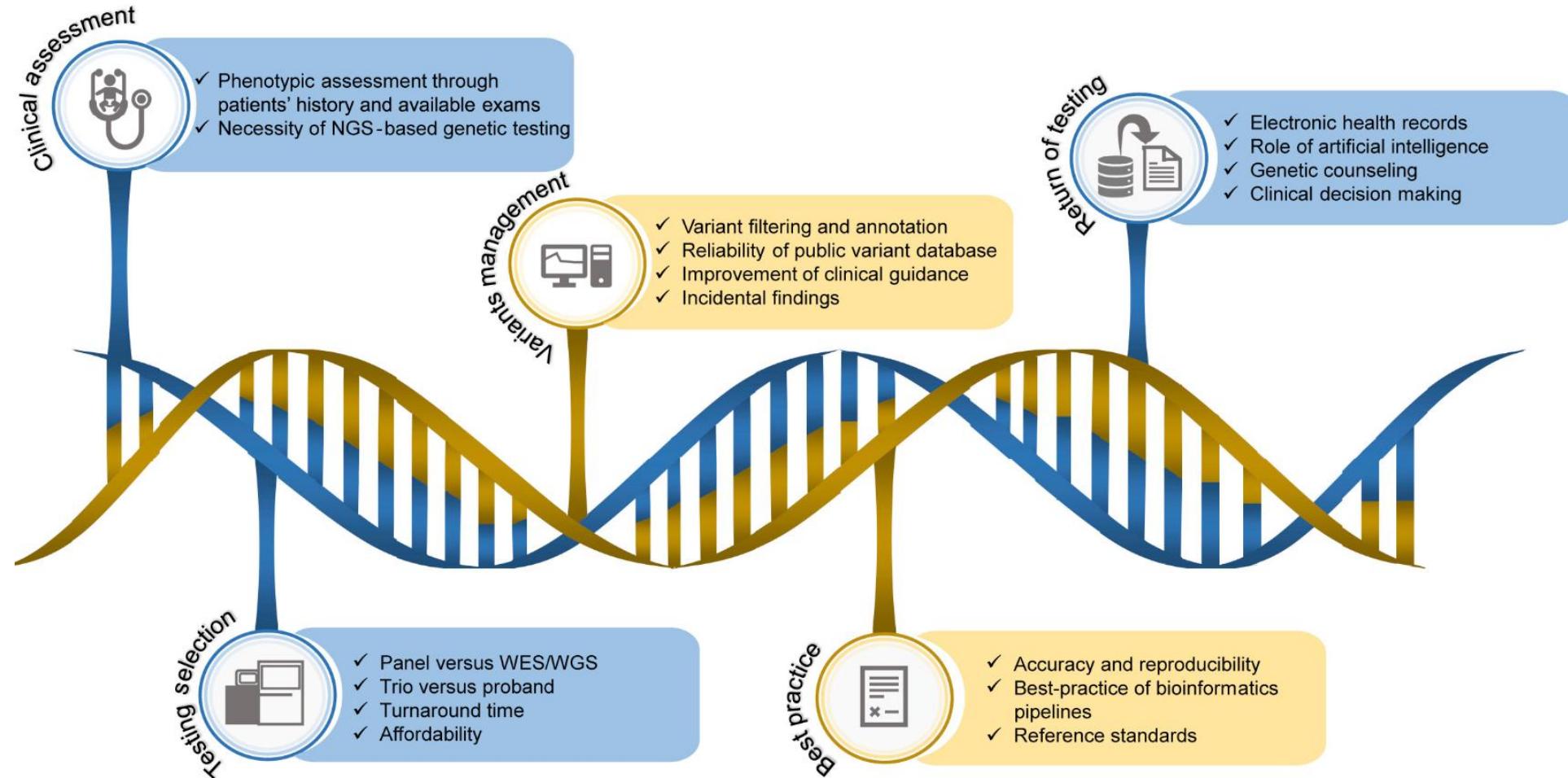
## ▪ Evolutionary Biology

Studying genetic variation and tracing species evolution.

## ▪ Forensic Science

DNA fingerprinting for crime scene investigation and identity verification.

# A Sample NGS Application: NGS is a powerful tool for rare disease diagnosis



source: <https://doi.org/10.1016/j.tig.2019.08.006>

# Types of genome sequencing technologies

## 1. First-Generation Sequencing

Sanger Sequencing: The original DNA sequencing method, accurate but limited in scale.

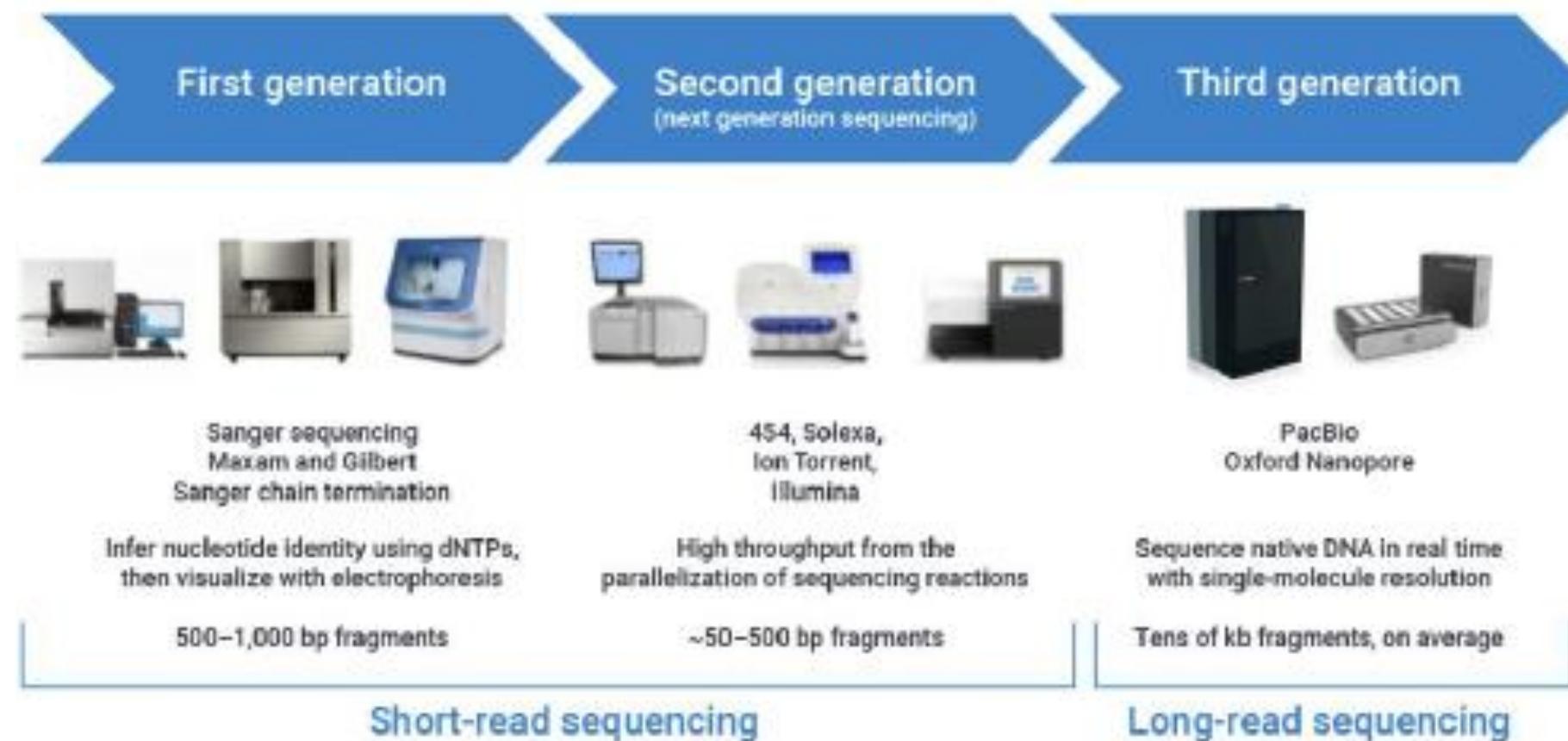
## 2. Next-Generation Sequencing (NGS)

- High-throughput sequencing methods, capable of sequencing millions of DNA fragments in parallel.
- Examples:
  - Illumina sequencing
  - Ion Torrent sequencing

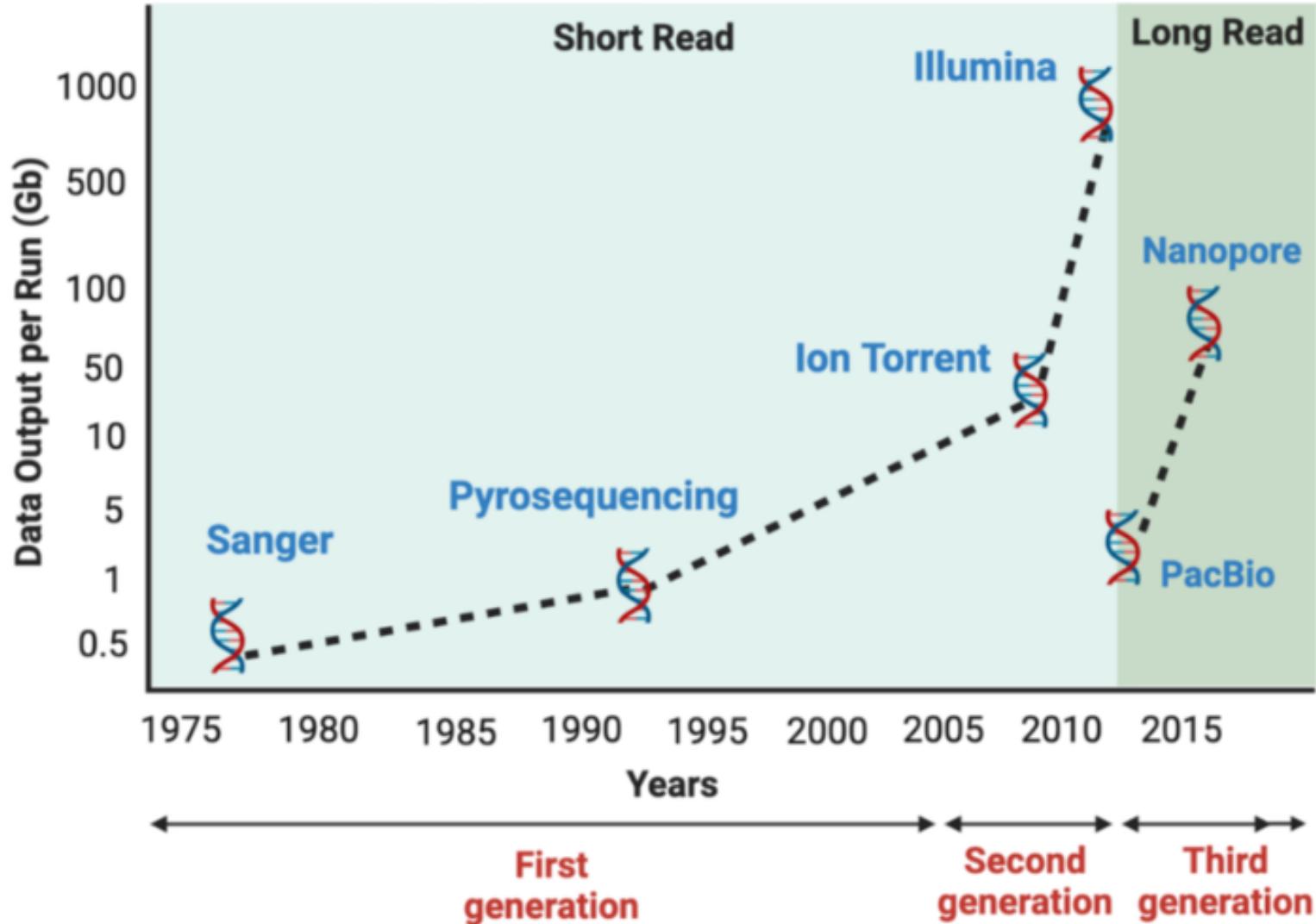
## 3. Third-Generation Sequencing

- Single-molecule sequencing, which reads longer DNA fragments.
- Examples:
  - PacBio (SMRT sequencing)
  - Oxford Nanopore sequencing

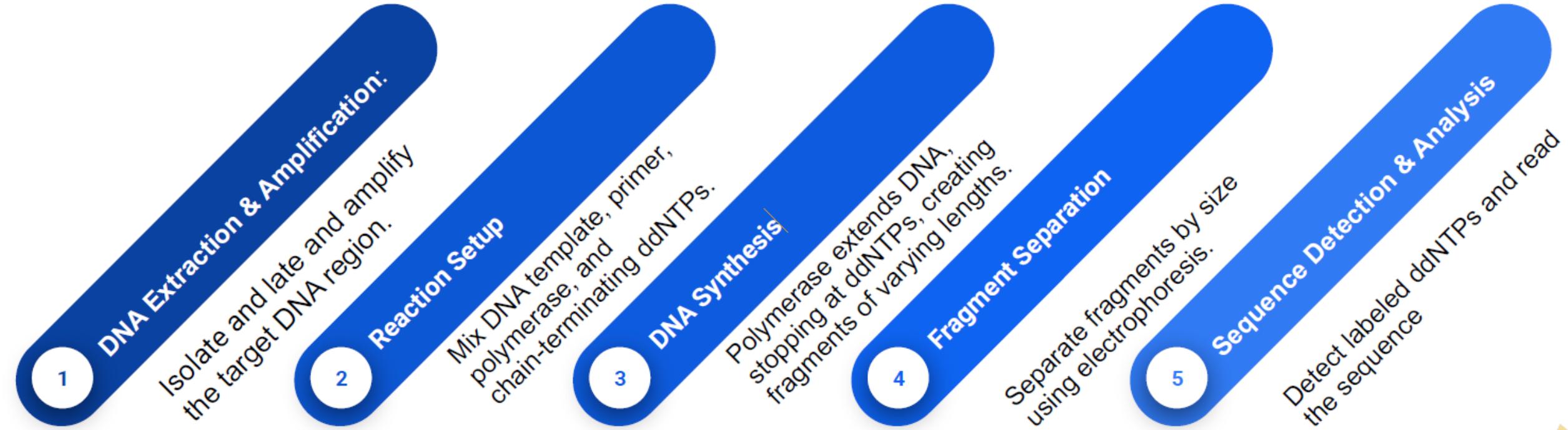
# Types of genome sequencing technologies



# Evolution of sequencing technologies



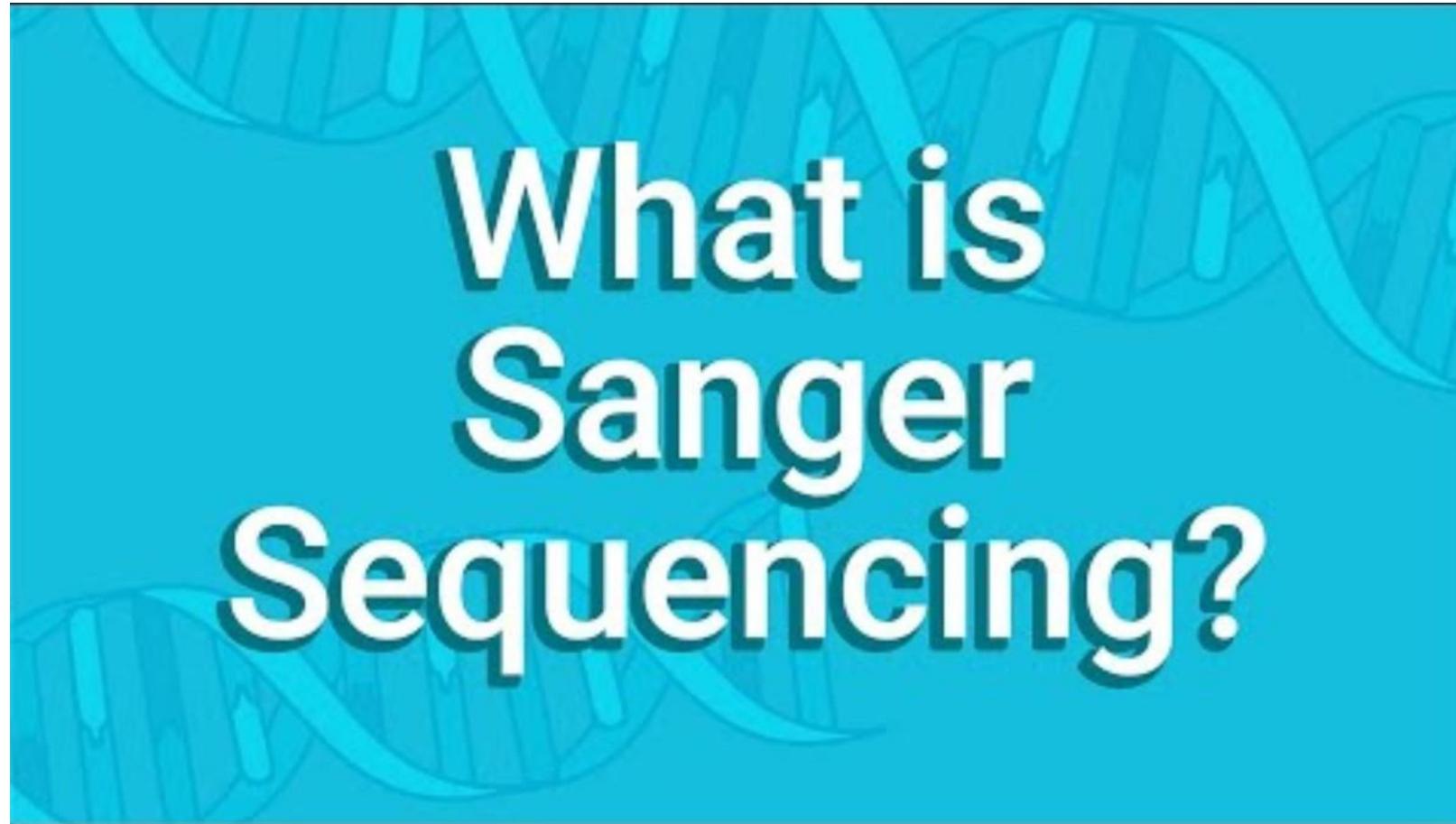
# Sanger Sequencing: Basic steps



ddNTPs (Dideoxynucleotide Triphosphates):

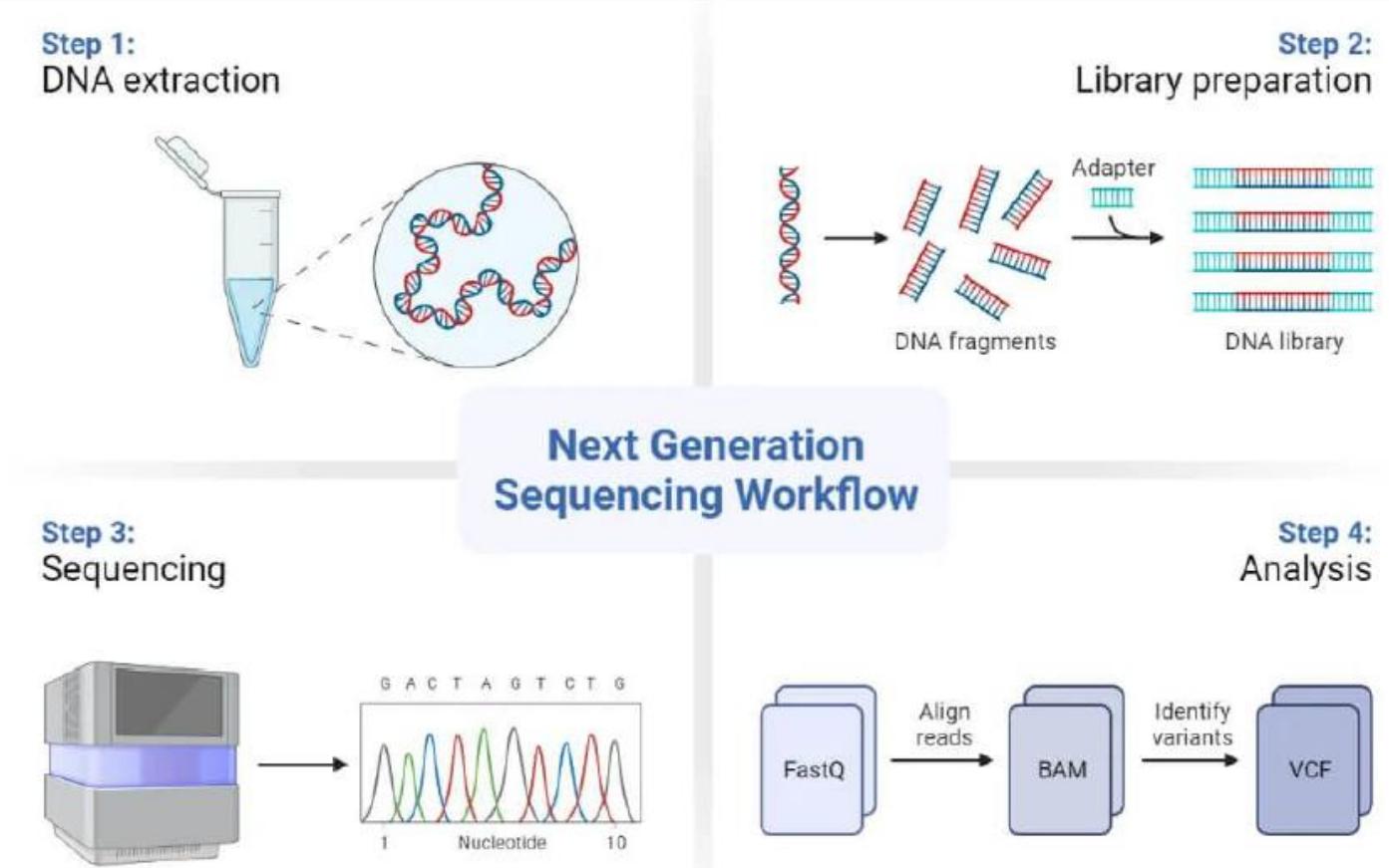
Modified nucleotides used in Sanger sequencing that lack a 3' hydroxyl group, causing DNA synthesis to stop when incorporated, resulting in fragment termination.

# Sanger Sequencing (demo)



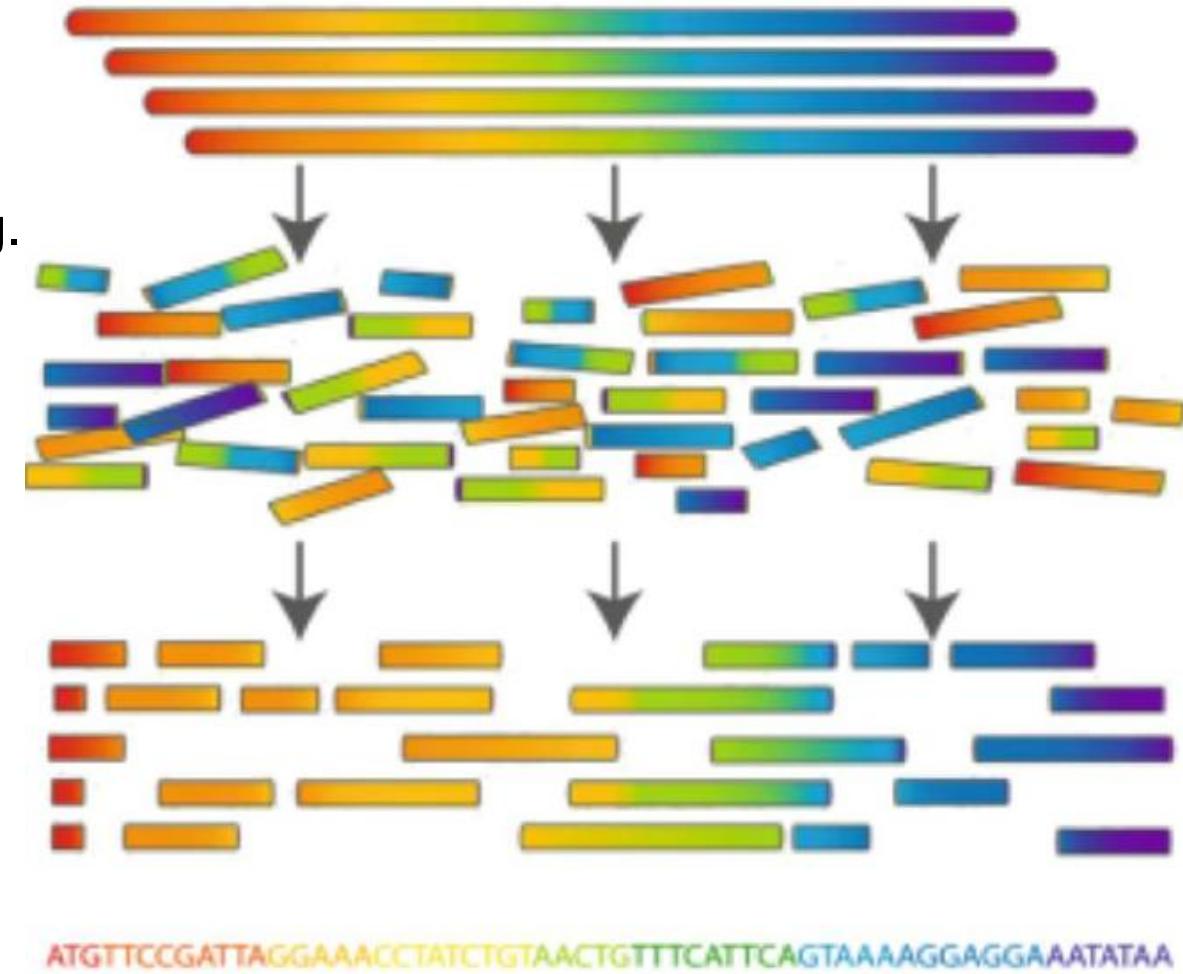
# How NGS Works?

- 1. DNA Extraction:** Isolate DNA from cells.
- 2. Fragmentation:** Break the DNA into smaller pieces.
- 3. Sequencing:** Use technologies e.g., Illumina) to read the DNA fragments.
- 4. Analysis:**
  - **Assembly:** Combine the fragments to reconstruct the entire genome.
  - **Annotation:** Identify genes and their functions.

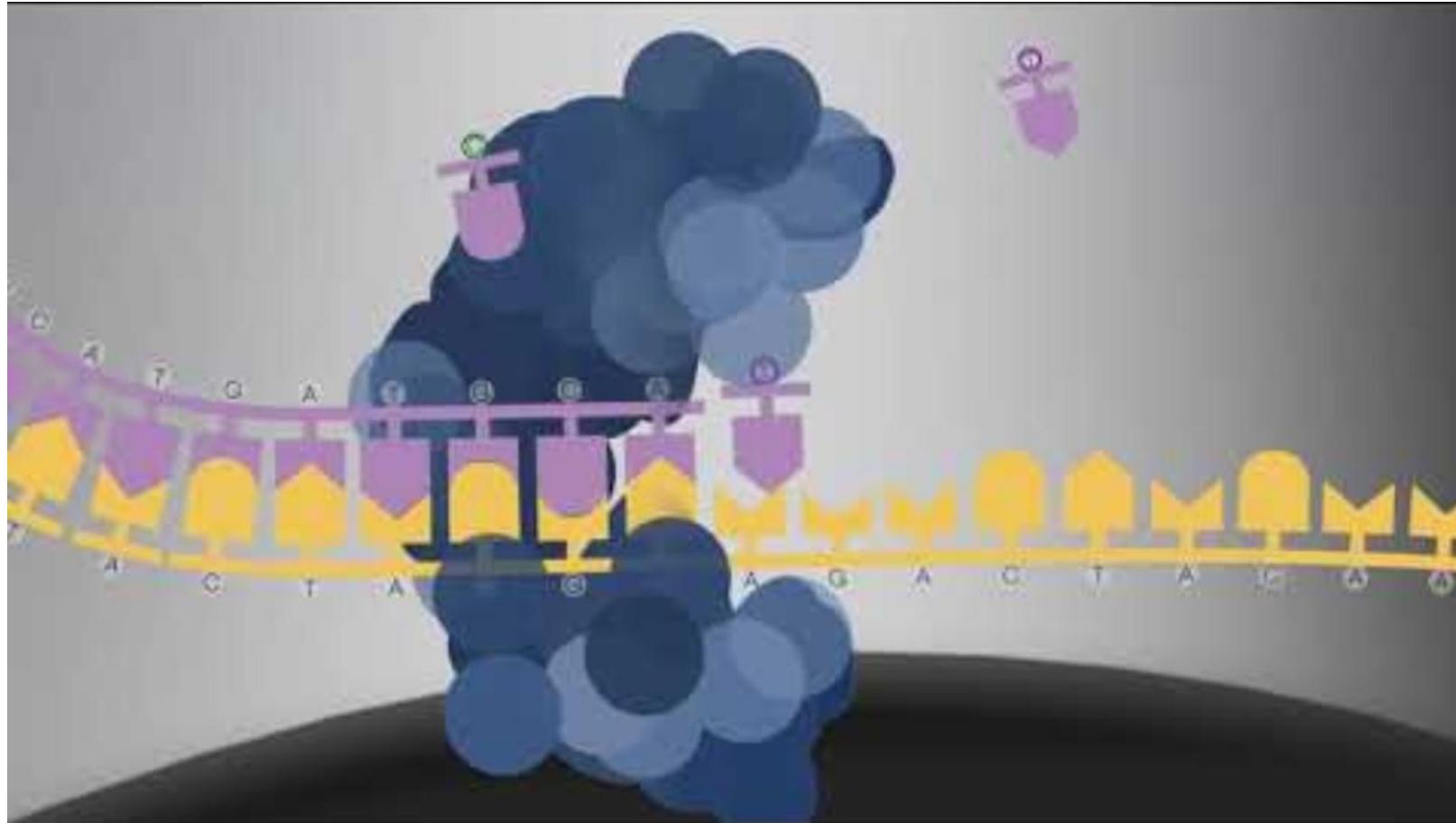


# Once a genome is sequenced, it needs to be Assembled

- Reconstructing a complete DNA sequence from short reads.
- **Steps:** Overlap → Contig Formation → Scaffolding.
  - **Fragment Overlap:** Identifying overlaps between short DNA sequences.
  - **Contig Formation:** Merging overlapping sequences into continuous segments.
  - **Scaffolding:** Ordering and orienting contigs using additional data (e.g., paired-end reads).
- **Types:**
  - De novo: Without a reference.
  - Reference-guided: Using a template.

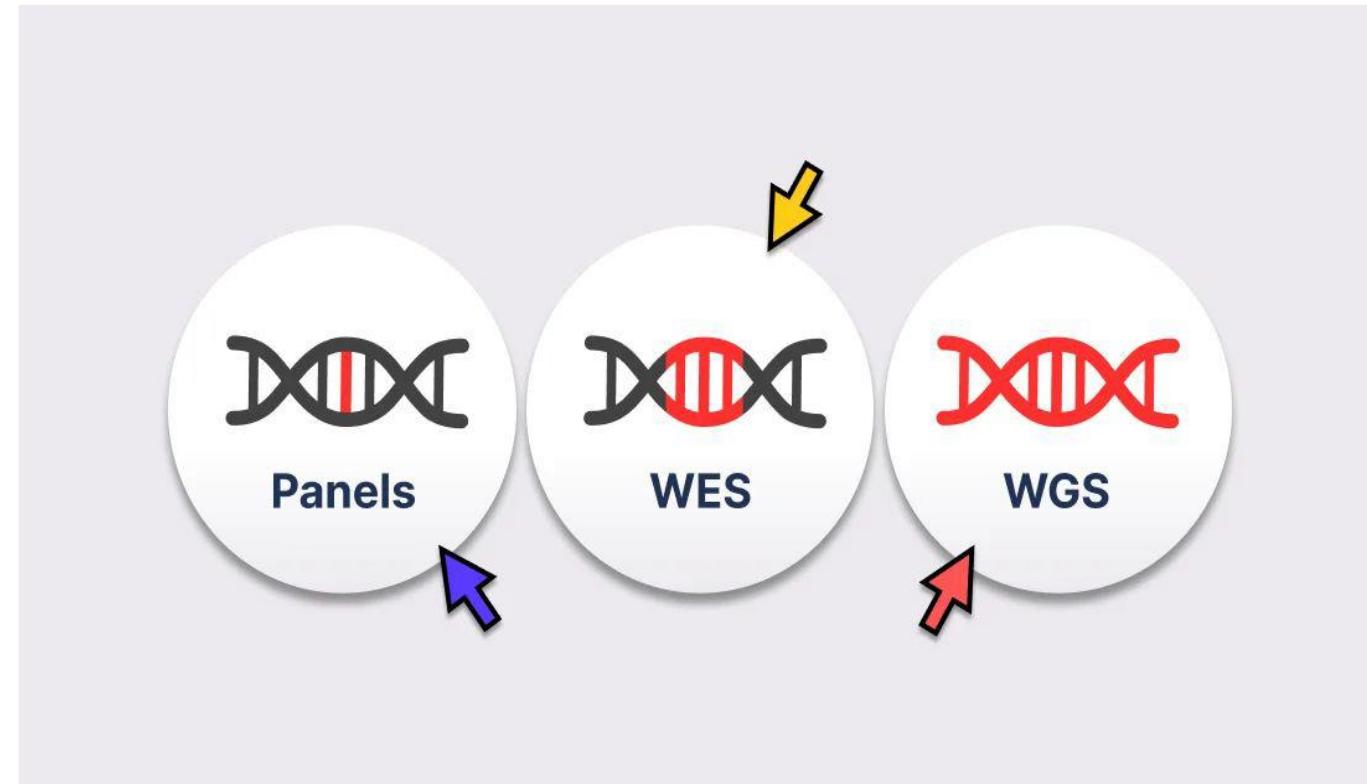


# Third Generation Sequencing



# Sequencing approaches

- **Whole Genome Sequencing (WGS)**
- Involves sequencing the entire genome, both coding and non-coding regions.
- **Whole Exome Sequencing (WES)**
- Focuses on sequencing only the exons (coding regions) of the genome.
- **Panel Sequencing – Targeted Sequencing**
- Involves sequencing a targeted set of genes or regions, typically relevant to specific diseases or conditions.

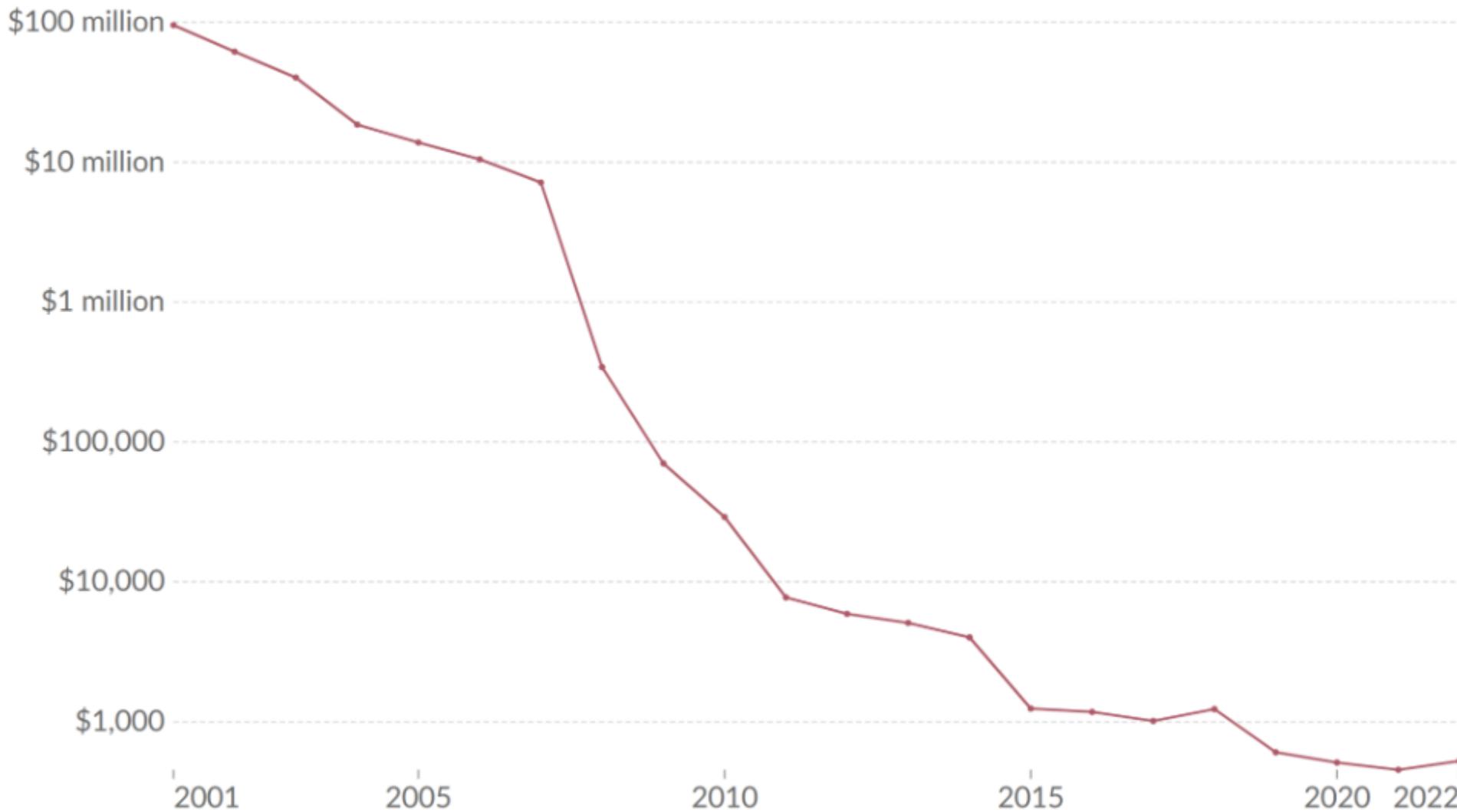


# First vs Second vs Third Generation Sequencing

Feature	First Generation (Sanger)	Second Generation (NGS)	Third Generation
Technology	Chain-termination method	Parallel sequencing (massive throughput)	Single-molecule sequencing
Read Length	500-1000 bp	150-300 bp	1000s to millions of bases
Throughput	Low	High (millions of reads)	High (real-time sequencing)
Accuracy	High	High but slightly lower than Sanger	Moderate (depends on platform)
Speed	Slow	Fast	Very fast (real-time)
Cost	High	Lower (compared to Sanger)	Varies (generally lower than NGS)
Applications	Small-scale projects, targeted sequencing	Whole genome/exome sequencing, RNA-seq	Long-read sequencing, complex genomes
Example Technologies	Sanger Sequencing	Illumina, Ion Torrent, SOLiD	PacBio, Oxford Nanopore

# Cost of sequencing a full human genome

The cost of sequencing the full genetic information of a human, measured in US\$. This data is not adjusted for inflation.



Data source: National Human Genome Research Institute (2022)

[OurWorldInData.org/technological-change](https://OurWorldInData.org/technological-change) | CC BY

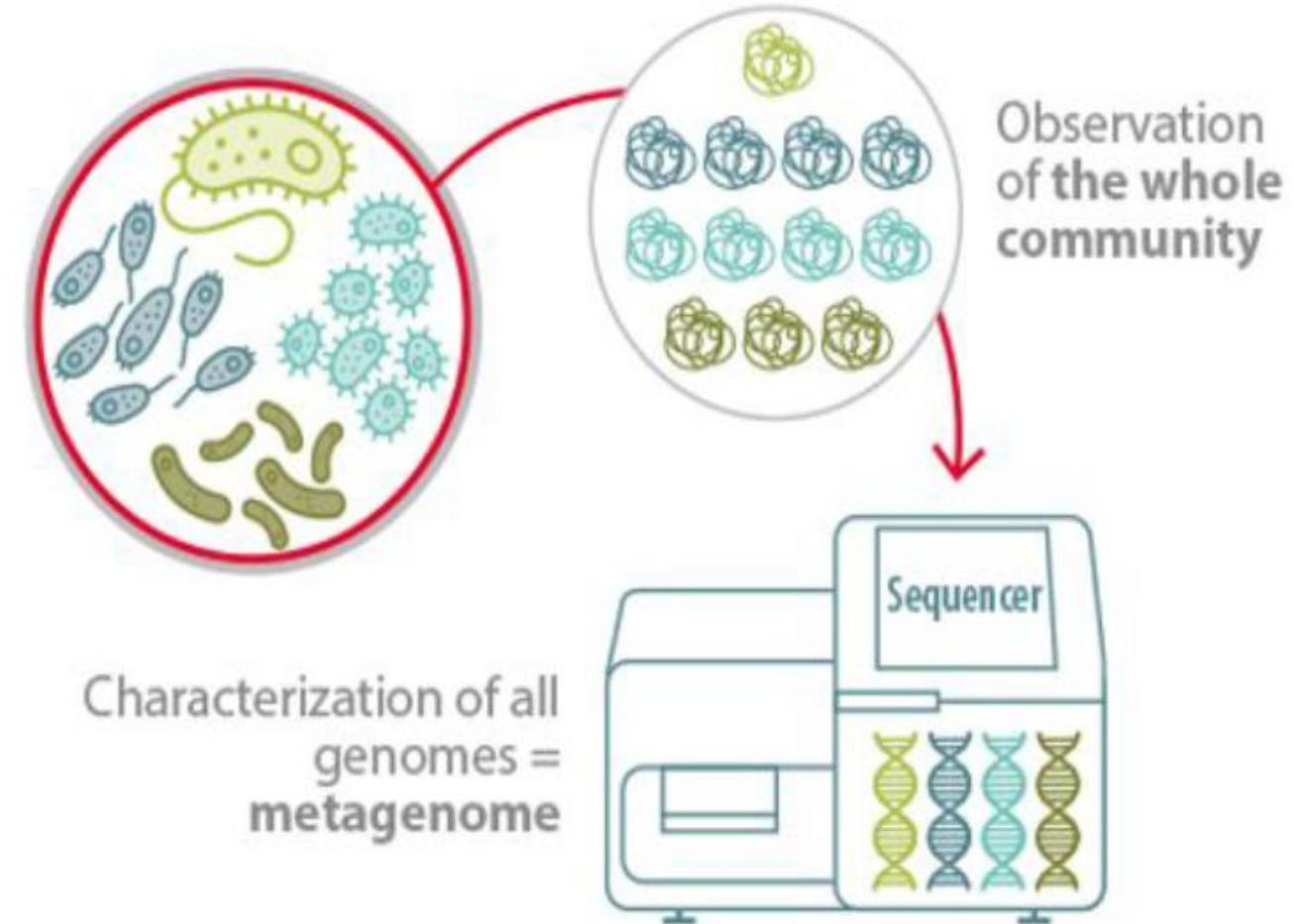
# Introduction to Metagenomics

# What is Metagenomics?

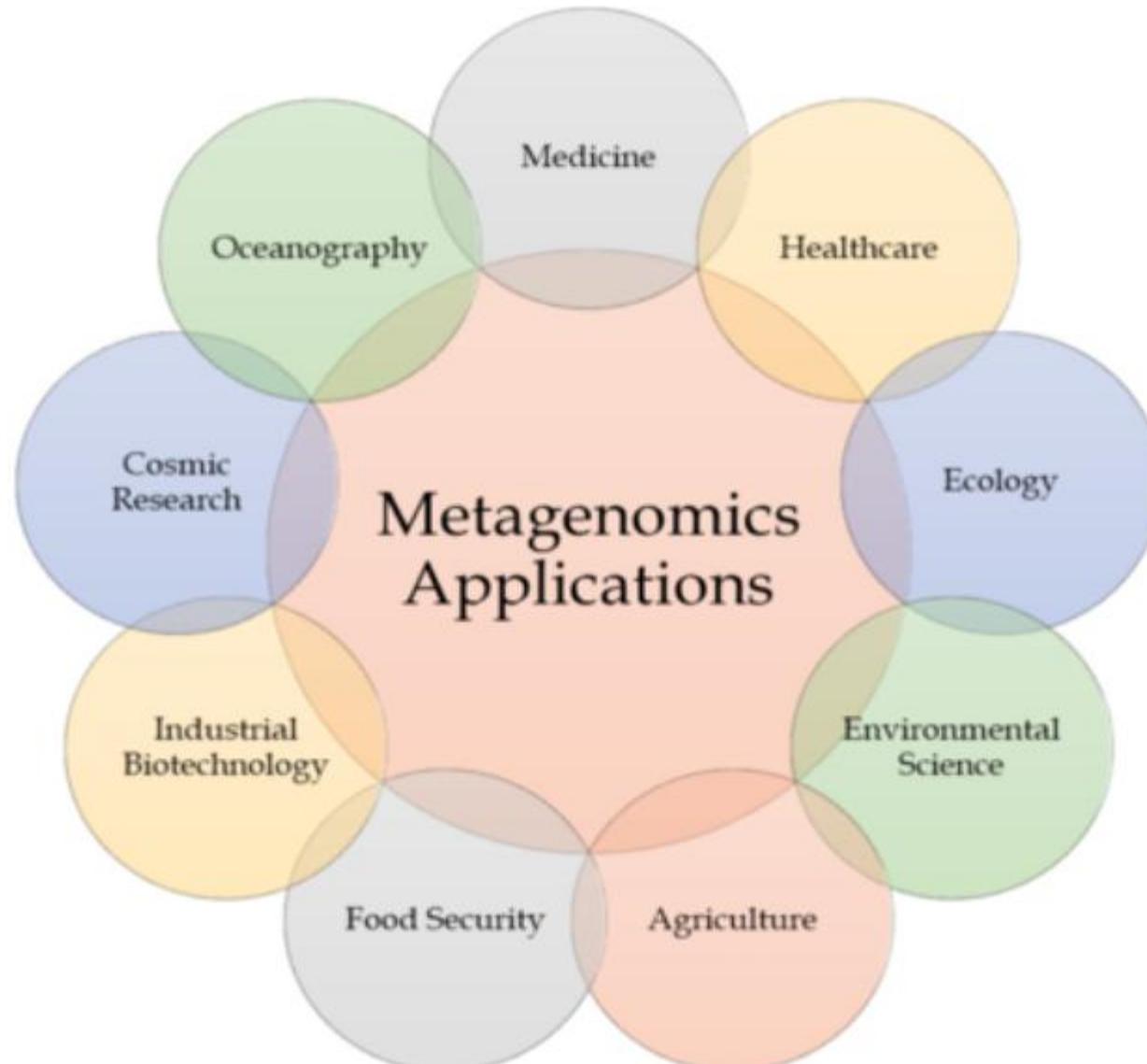
Metagenomics is the study of genetic material directly obtained from environmental samples, without the need for culturing organisms.

## Key Features:

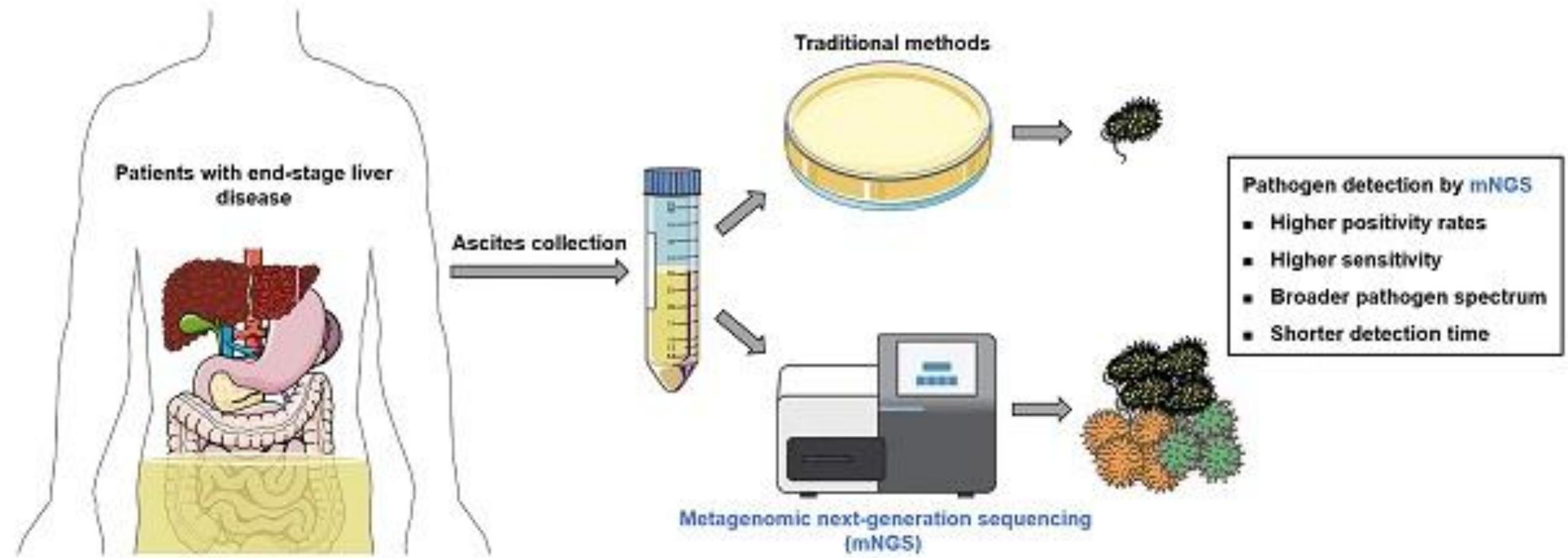
- Analyzes microbial diversity, functions, and interactions.
- Uses high-throughput sequencing (e.g., shotgun sequencing).
- Provides insights into microbiomes in environments like soil, water, and the human gut.



# Metagenomic Application



# A sample metagenomics project: Pathogen detection



# Significance of Metagenomics

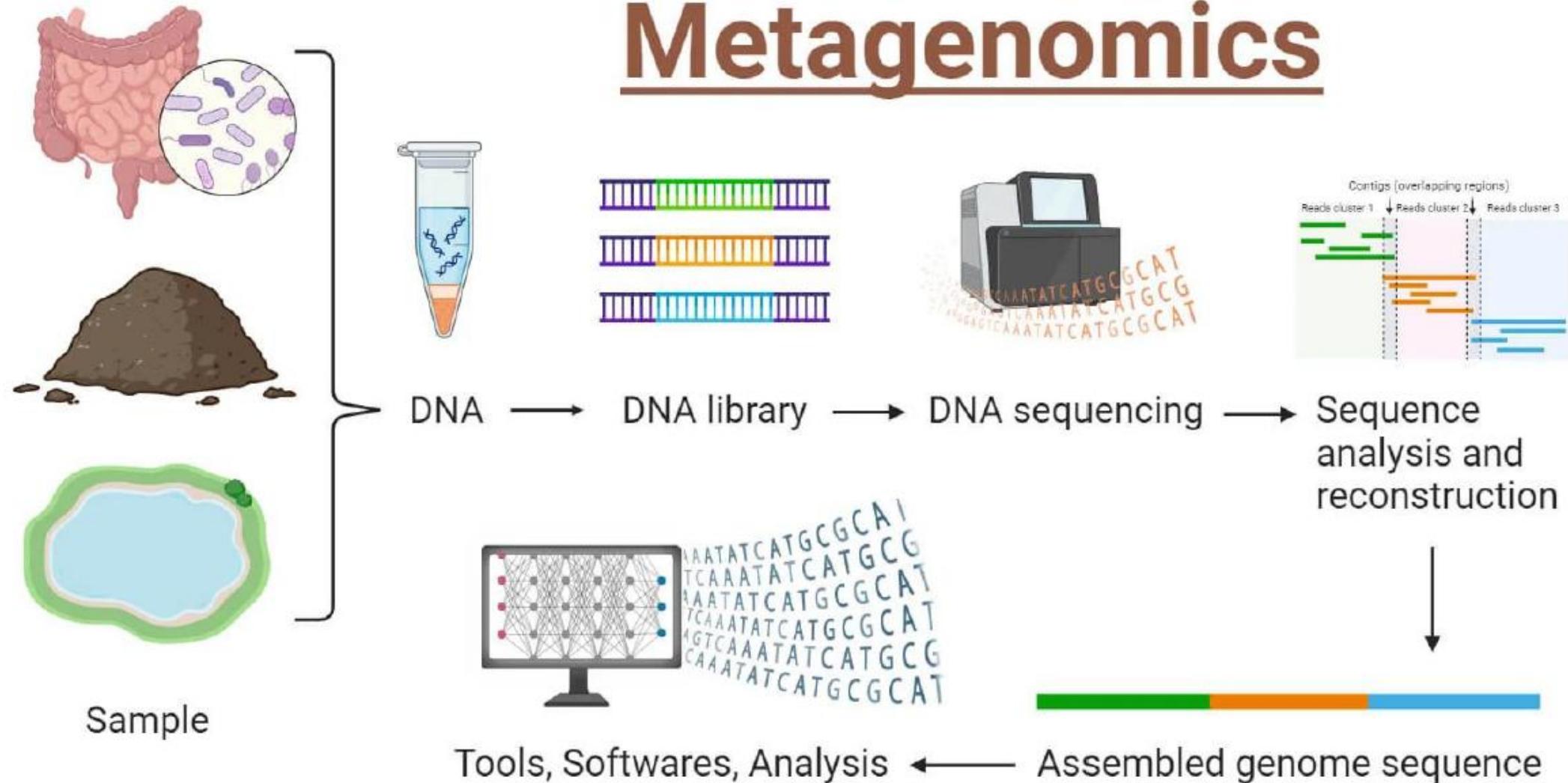
- Microbial communities are vital to ecosystem processes affecting plants, animals, and humans.
- These communities are diverse, complex, and dynamically influenced by the environment.
- Understanding their functional profiles is key to predicting ecosystem changes

# Significance of Metagenomics

- Metagenomics allows us to study microbes in their native environments, bypassing the need for isolation.
- Advances in high-throughput sequencing and computational biology can allow us to explore the functional diversity of metagenomes

# Steps involved in metagenomics

## Metagenomics



# Types of metagenomics methods

## 1. Targeted Metagenomics (Amplicon-based Sequencing):

- Focuses on specific genetic markers (e.g., 16S rRNA for bacteria, 18S rRNA or ITS for fungi).
- Identifies organisms at higher taxonomic levels.
- Does not provide strain-level resolution.
- Used for profiling microbial communities based on conserved regions.

## 2. Shotgun Metagenomics (WGS):

- Randomly sequences all genetic material in a sample.
- Provides detailed taxonomic and functional composition.
- Enables species-level identification.
- Offers greater depth and broader insights into microbial diversity.

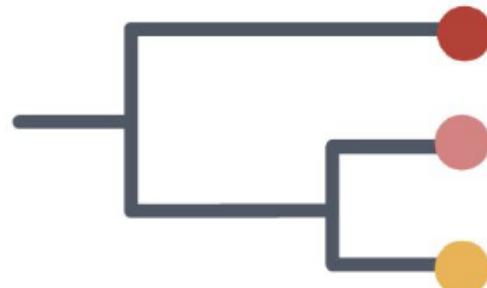
# Targeted vs Shotgun

## 16S Amplicon Sequencing



Microbe identification

Requires very little input DNA



## Whole Genome Sequencing



Microbe identification

Functional data

Metagenome-assembled genomes (MAGs)



# Day 3: Genomics, Metagenomics and Transcriptomics

## Morning Practical Session

# Hands-On: Exploring NCBI Genome Data Viewer, Accessing Genomic and MGnify Metagenomic Data, and Analyzing with Biopython

# Alternative to NCBI - GDV: UCSC and BLAT

- UCSC (University of California, Santa Cruz) Genome Browser is an online tool for visualizing and analyzing genomic data across species.

<https://genome.ucsc.edu/>

- BLAT (BLAST-Like Alignment Tool) is a fast alignment tool used to find similarities between sequences and reference genomes.

<https://genome.ucsc.edu/cgi-bin/hgBlat>



The screenshot shows the UCSC Genome Browser homepage. At the top, the logo for the University of California Santa Cruz Genomics Institute is displayed next to the "Genome Browser" title. Below the header is a navigation menu with links to Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Projects, Help, and About Us. The main content area features a genome track viewer with a blue background. On the left, a purple curve represents a signal across the genome. Several gene tracks are shown, including ATPB2, SHBG, TP53, and WRAPPI. To the right, a red pyramid-shaped visualization is overlaid on a yellow gradient background. At the bottom of the page, there are two sections: "Tools" and "News". The "Tools" section contains icons for hg38, hg19, and mm39, followed by a list of tools: Genome Browser, BLAT, In-Silico PCR, Table Browser, LiftOver, REST API, Variant Annotation Integrator, and More tools... The "News" section lists recent updates: New GENCODE gene tracks: Human V47 (hg19/hg38) - Mouse M36 (Nov 8, 2024), GIAB Problematic Regions tracks for human (hg38 and hs1) (Nov 4, 2024), New GENCODE gene tracks: V47 (hg38) - VM36 (mm39) (Oct 23, 2024), CADD v1.7 and ClinGen CSpec for hg19 and hg38 (Oct 9, 2024), New clinical tutorial (Oct 1, 2024), and New gnomAD v4.1 tracks for hg38 (Sep 30, 2024). There are also "More news..." and "Subscribe" buttons at the bottom.

# Lunch Break

# Introduction to Bioinformatics

## Day 3: Genomics, Metagenomics and Transcriptomics

14<sup>th</sup> January 2026

Afternoon Session

# Outline Day 3

- Afternoon (2-5 pm):
  - **Introduction to Transcriptomics:**
    - RNA sequencing and transcriptome analysis.
    - Expression profiling and differential expression analysis.
  - **Practical Session:**  
Analyze RNA-seq data using python

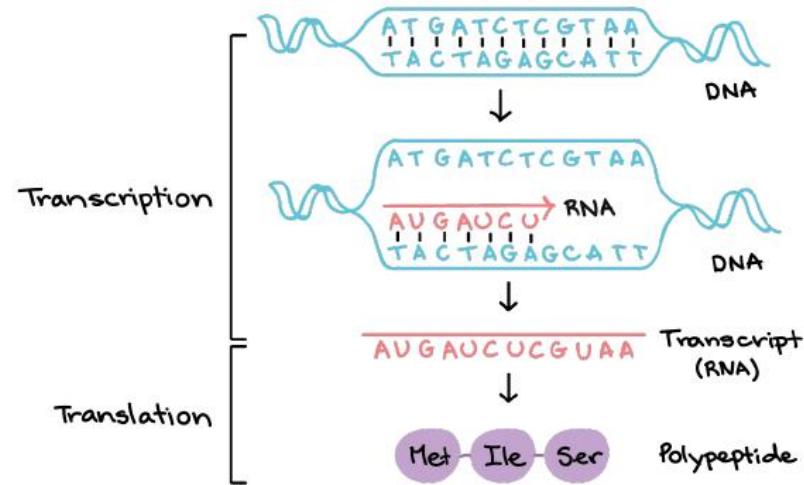
# Transcriptome

- The transcriptome is the complete set of transcripts in a cell and their quantity, for a specific developmental stage or physiological condition.
- Understanding the transcriptome is essential for interpreting the functional elements of the genome and revealing the molecular constituents of cells and tissues, and also for understanding development and disease.



# Transcriptomics

- Transcriptomics, the study of RNA in any of its forms.
- The transcriptome is the set of all RNA molecules, including mRNA, rRNA, tRNA, and other non-coding RNA produced in one or a population of cells.



# Transcriptomics scope

- The term can be applied to the total set of transcripts in a given organism, or to the specific subset of transcripts present in a particular cell type.
- Unlike the genome, which is roughly fixed for a given cell line (excluding mutations), the transcriptome can vary with external environmental conditions.

# Transcriptomics scope

- Because it includes all mRNA transcripts in the cell, the transcriptome reflects the genes that are being actively expressed at any given time, with the exception of mRNA degradation phenomena such as transcriptional attenuation.
- The study of transcriptomics, also referred to as expression profiling, examines the expression level of mRNAs in a given cell population, often using high-throughput techniques based on DNA microarray technology.

# Transcriptomics aims

- To catalogue all species of transcripts, including mRNAs, noncoding RNAs and small RNAs.
- To determine the transcriptional structure of genes, in terms of their start sites, 5' and 3' ends, splicing patterns and other post-transcriptional modifications.
- To quantify the changing expression levels of each transcript during development and under different conditions.

# Technologies

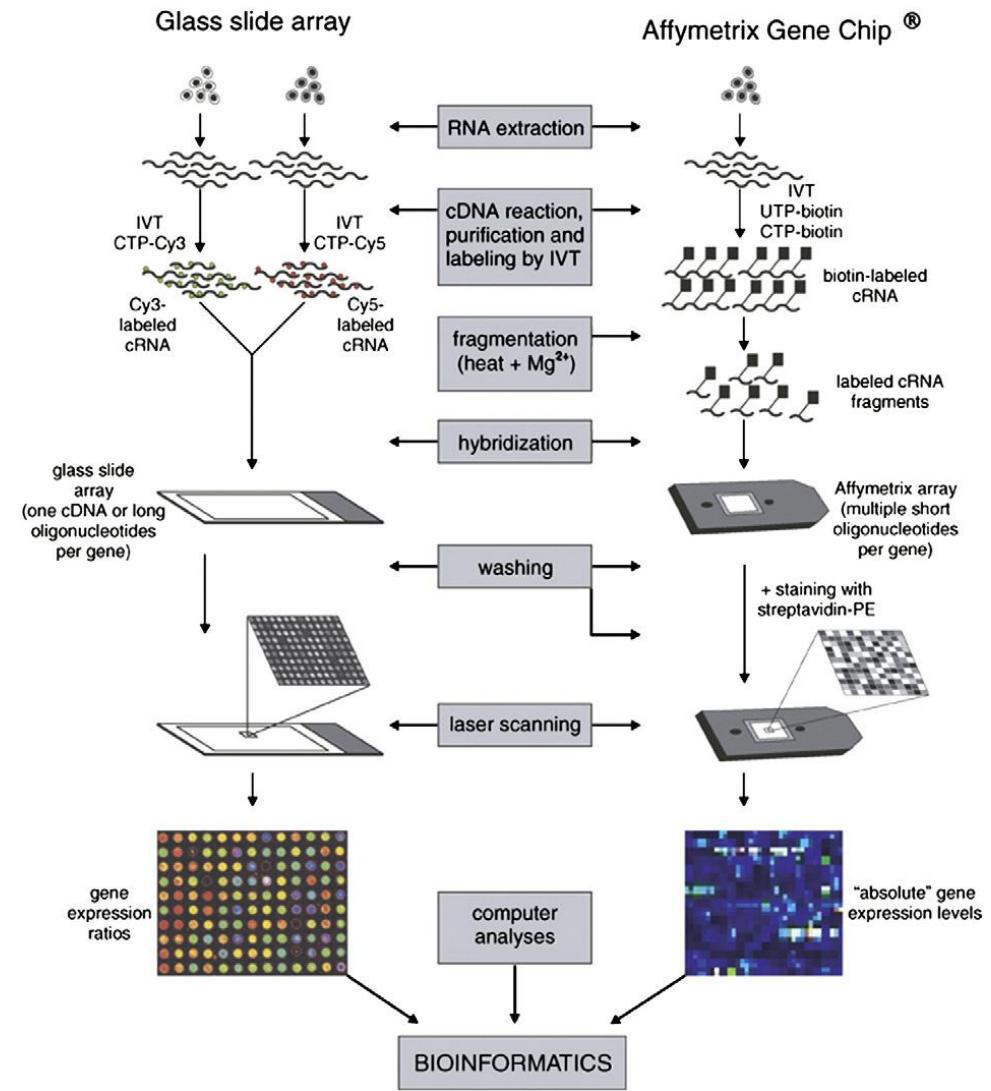
- Hybridization-based approaches
  - fluorescently labelled cDNA with custom-made microarrays
  - commercial high-density oligo microarrays
- Sequence-based approaches
  - Sanger sequencing of cDNA or EST libraries
  - serial analysis of gene expression (SAGE)
  - cap analysis of gene expression (CAGE)
  - massively parallel signature sequencing (MPSS)

# Hybridization approaches: microarrays and related techniques

- The technology has been developed in several variants but in the following we only discuss the two most popular:
  - “two colour” (or cDNA or two-channel) microarrays and
  - “one colour” (or oligonucleotides or one-channel) microarrays.
- Two colour microarrays are based on the competitive hybridization of two samples each of which has been labeled with a different fluorescent dye (e.g. red or green).
- After hybridization, the array is exposed to red and green laser light
- The array emits fluorescence proportional to the quantity of RNA
- The image produced is scanned yielding after some corrections a value which represents the expression of one sample relative to the other.

# Hybridization approaches: microarrays and related techniques

- One channel microarrays are based on RNA of one sample which has been labeled with a fluorescent dye and hybridized to a single array where millions of copies of short (around 24 base pairs) oligonucleotide probes representing all known genes (several probes for gene form a “probeset”) have been synthesized.
- After exposition to laser light and scanner the intensity of each location is measured yielding a value which represents an absolute measure of expression.



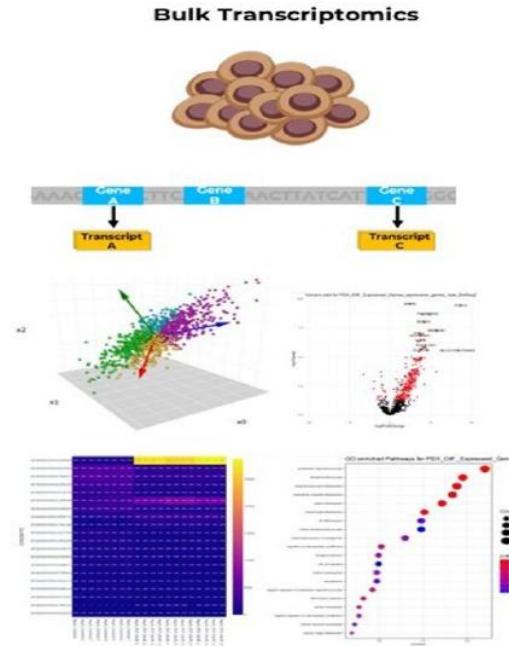
# Hybridization approaches: microarrays and related techniques

- Gene expression microarrays have been very useful to provide an overall view of how gene expression changes between two or more biological conditions.
- However, as the understanding of expression has evolved it has become apparent that more complex events than transcription and splicing actually occur within individual genes in a sample.

# RNA-seq: sequencing approaches to study the transcriptome

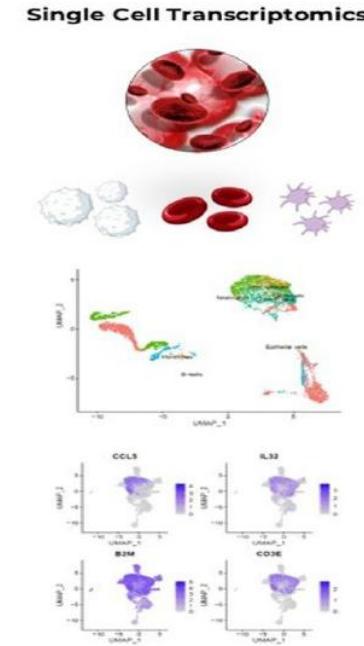
- RNA-Seq transcriptomics replaces the hybridization of nucleotide probes with sequencing individual cDNAs produced from the target RNA.
- Emerging methods for these fully quantitative transcriptomic analyses have the potential to overcome the limitations of microarray technology and there are ongoing discussions about whether sequencing approaches may replace microarrays in the middle or even short term.
- As a massively parallel process, next-generation sequencing (NGS) generates hundreds of megabases to gigabases of nucleotide sequence output in a single instrument run, depending on the platform.

# Types of RNA-Seq

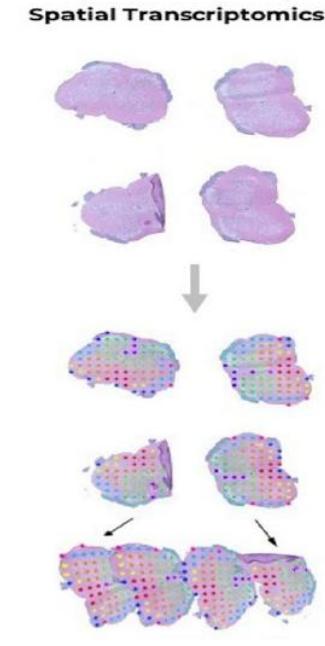


- provide a global view of the gene expression

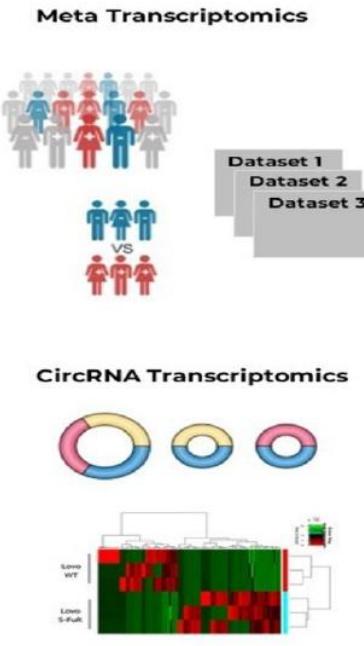
Ray, 2023



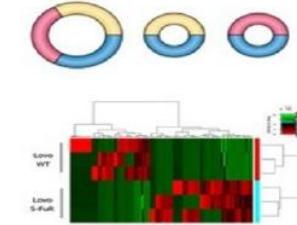
- Uncovers cellular diversity



- map gene expression patterns within tissues /organisms

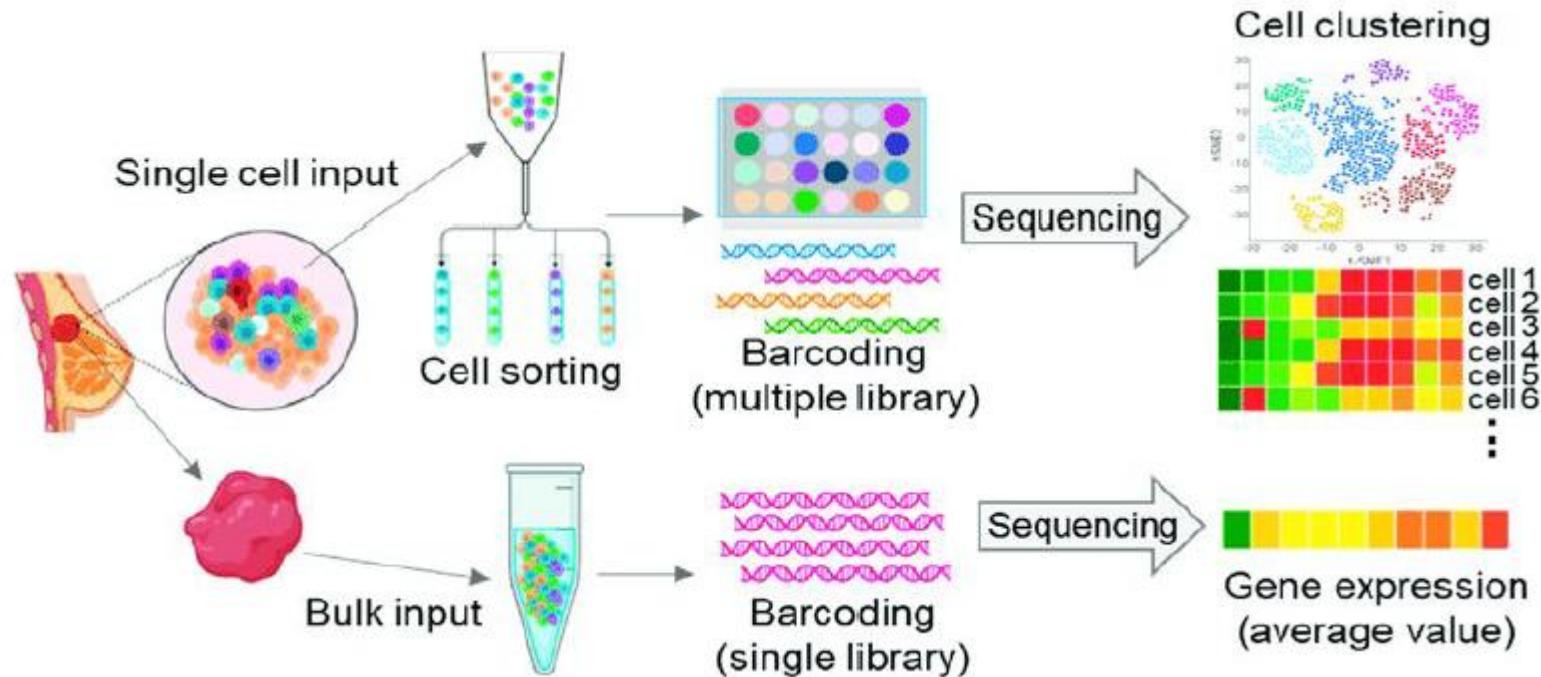


## CircRNA Transcriptomics

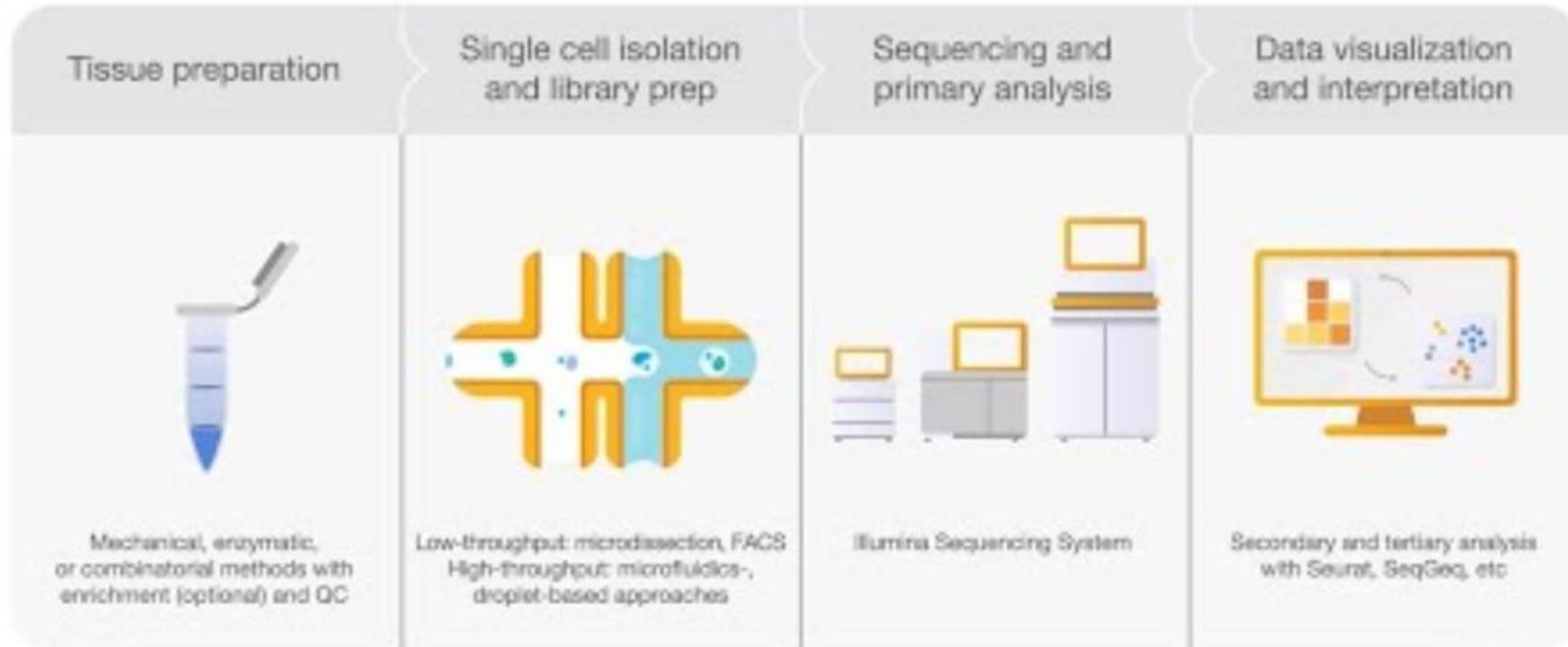


- explore transcriptional expression in microbial communities
- Circular RNAs have been associated with various diseases including neurological

# Bulk vs. Single-Cell RNA-Seq



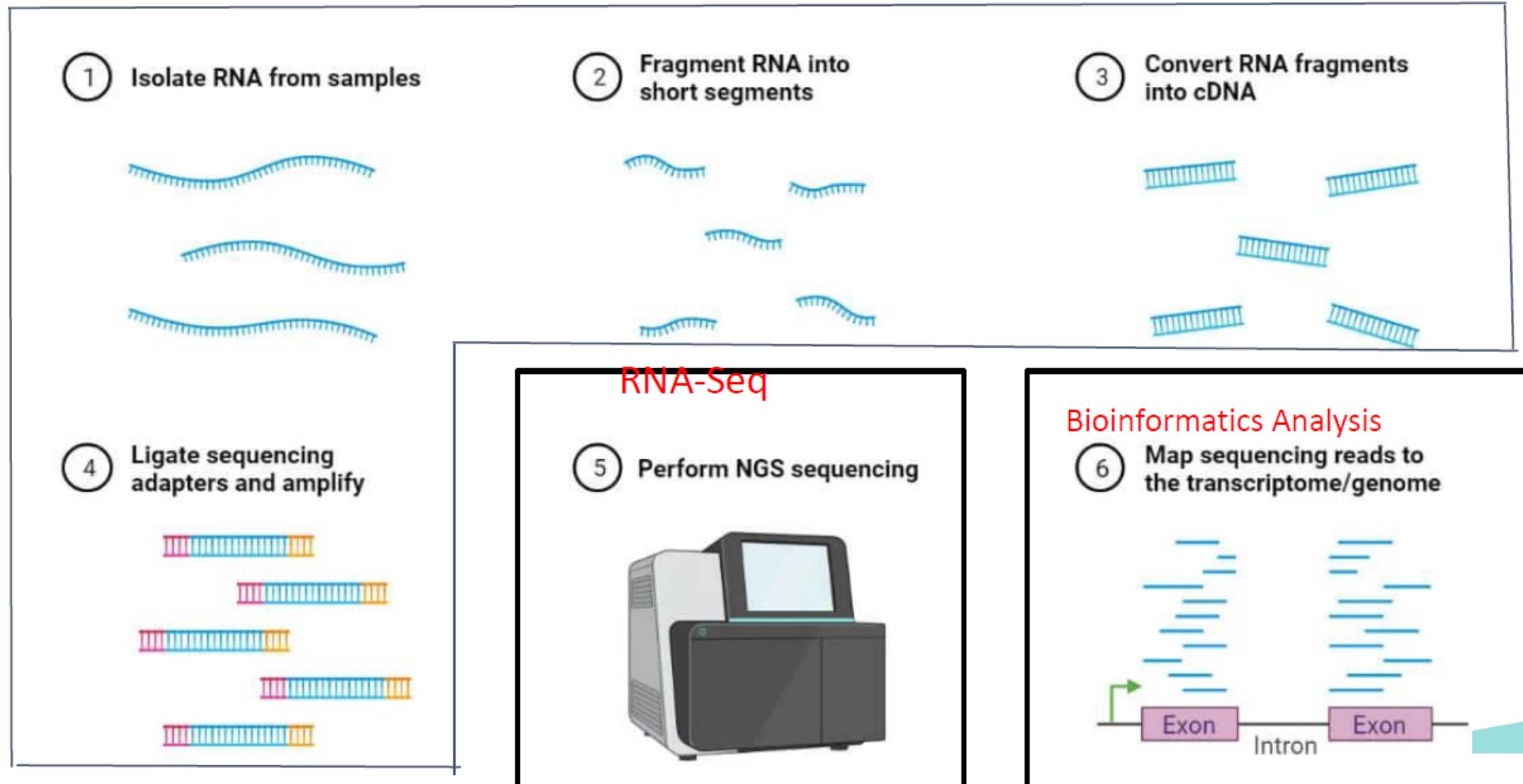
# Single-cell Sequencing



The workflow begins with initial tissue preparation, which involves isolating the cells from their native

# Overview on the processes

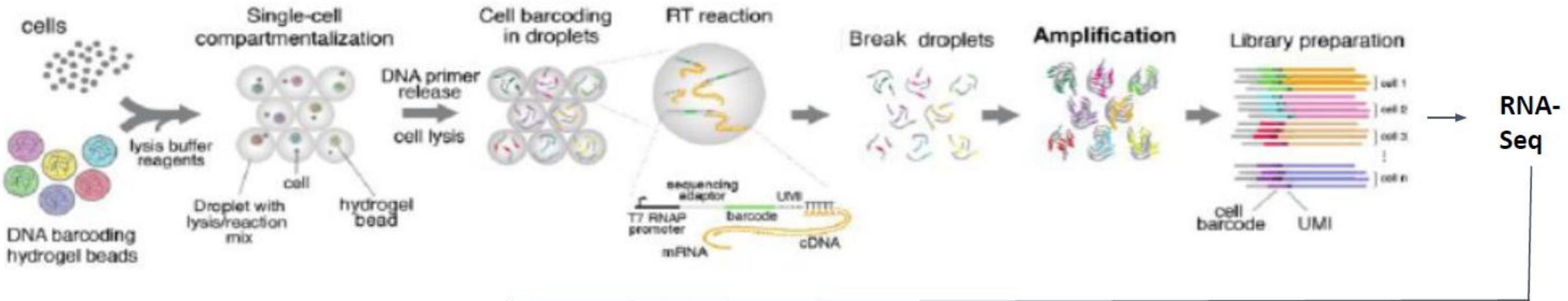
Wet Lab



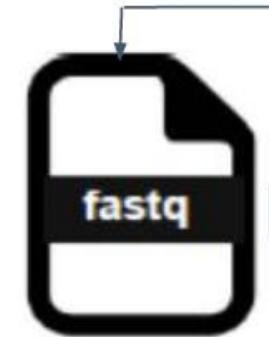
# Main bioinformatics analysis

- Quality Control and Preprocessing
- Read Alignment
- Gene Expression Quantification
- Normalization of Counts
- Differential Gene Expression Analysis platform.

# Step 1: Raw reads:

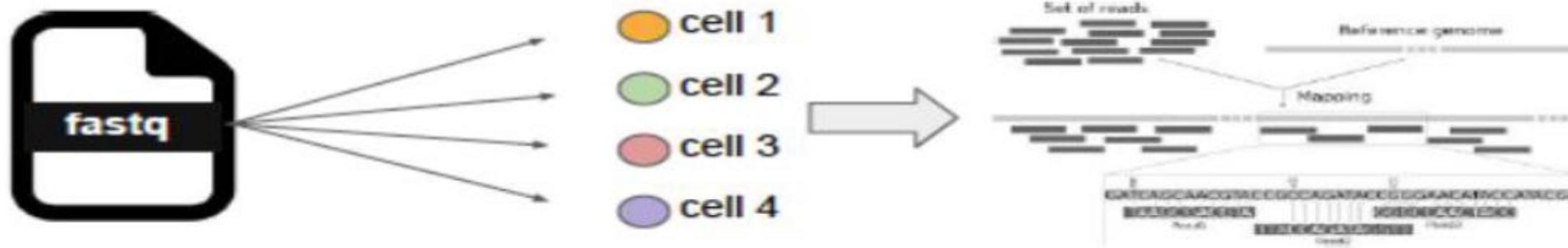


[Raw reads >](#)



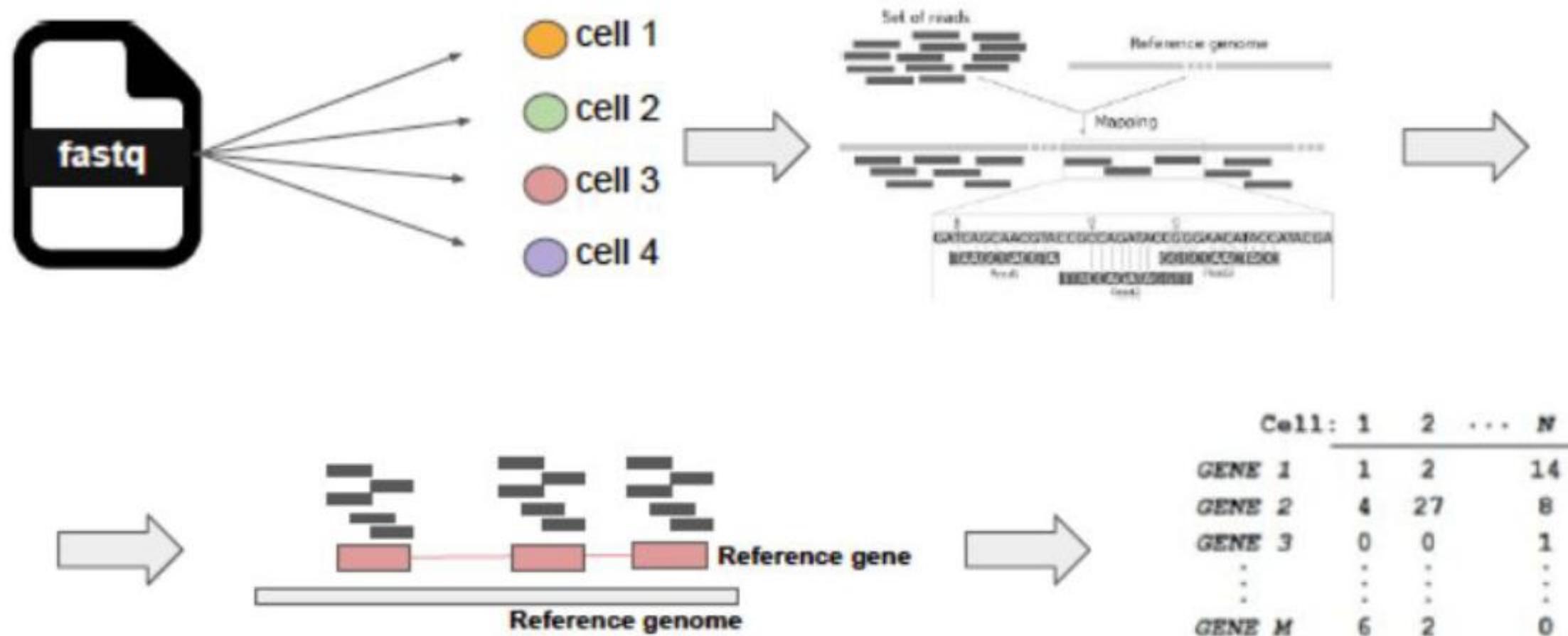
Identifier	• @SRR566546.970 HWUSI-EAS1673_11067_FC7070M:4:1:2299:1109 length=50
Sequence	• TTGCCTGCCTATCATTAGTCCTGTGAGGTGGAGATGTGAGGATCACT
'+' sign	• +
Quality scores	• hhhhhhhhhhhhhhhhhhhhhhhhhfffffe'ee['X]b[d[ed'[Y[^Y
Identifier	• @SRR566546.971 HWUSI-EAS1673_11067_FC7070M:4:1:2374:1108 length=50
Sequence	• GATTGTATGAAAGTATAACTAAAACGTGGATCAGAGTAAGTC
'+' sign	• +
Quality scores	• hhggfhhcghghggfcffdhfehhhhcehdchhdhaehffffde'bVd

# Step 2: Read mapping:



Raw reads > Barcoding inspection > Reads mapping >

# Step 3: Quantification and Count tables



Raw reads > Barcoding inspection > Reads mapping > Quantification >

# What does the count data actually represent?

samples: want to see if differences across condition are significant  
(w.r.t. biological and technical variation)

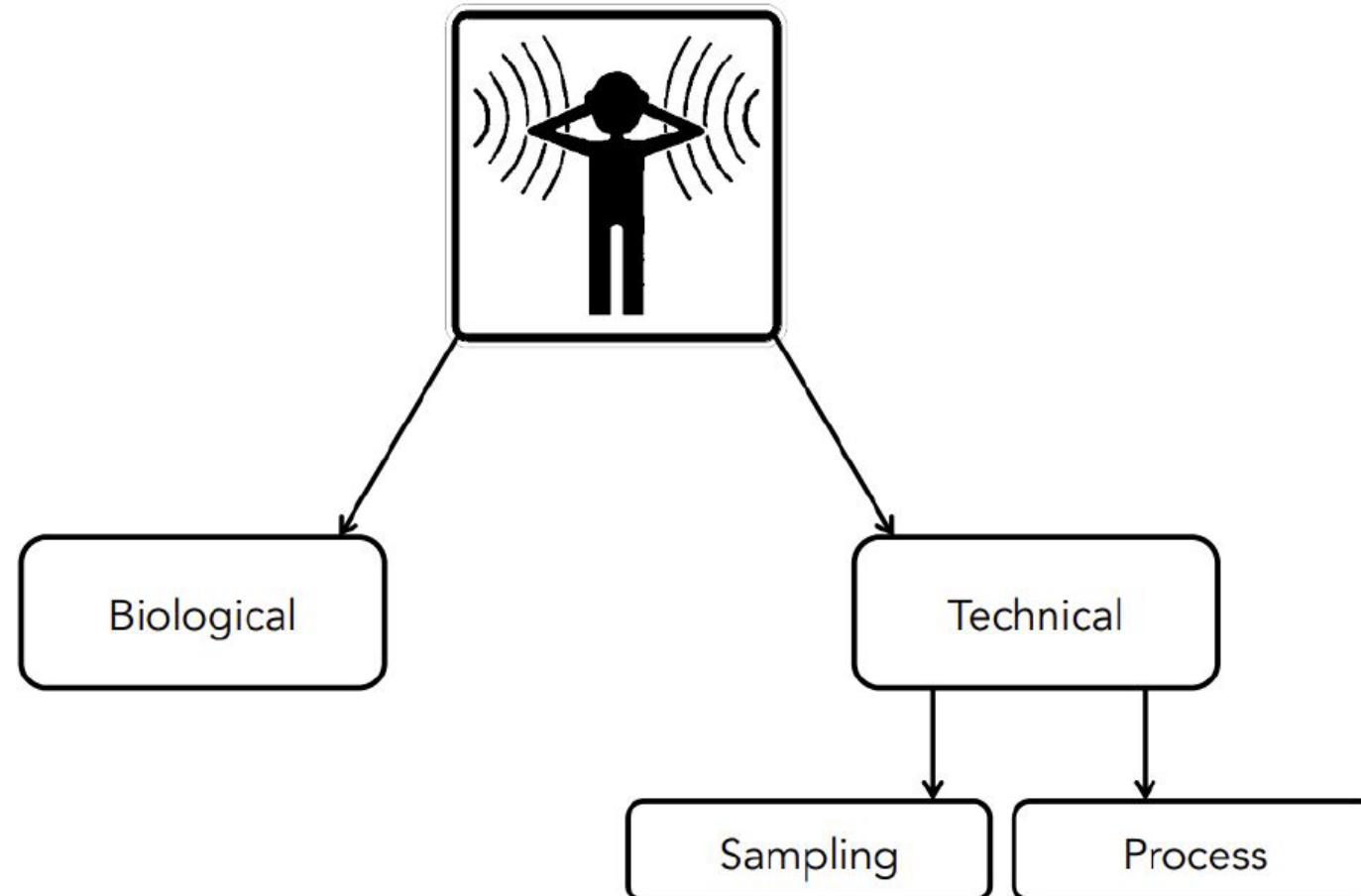
features (e.g. genes)



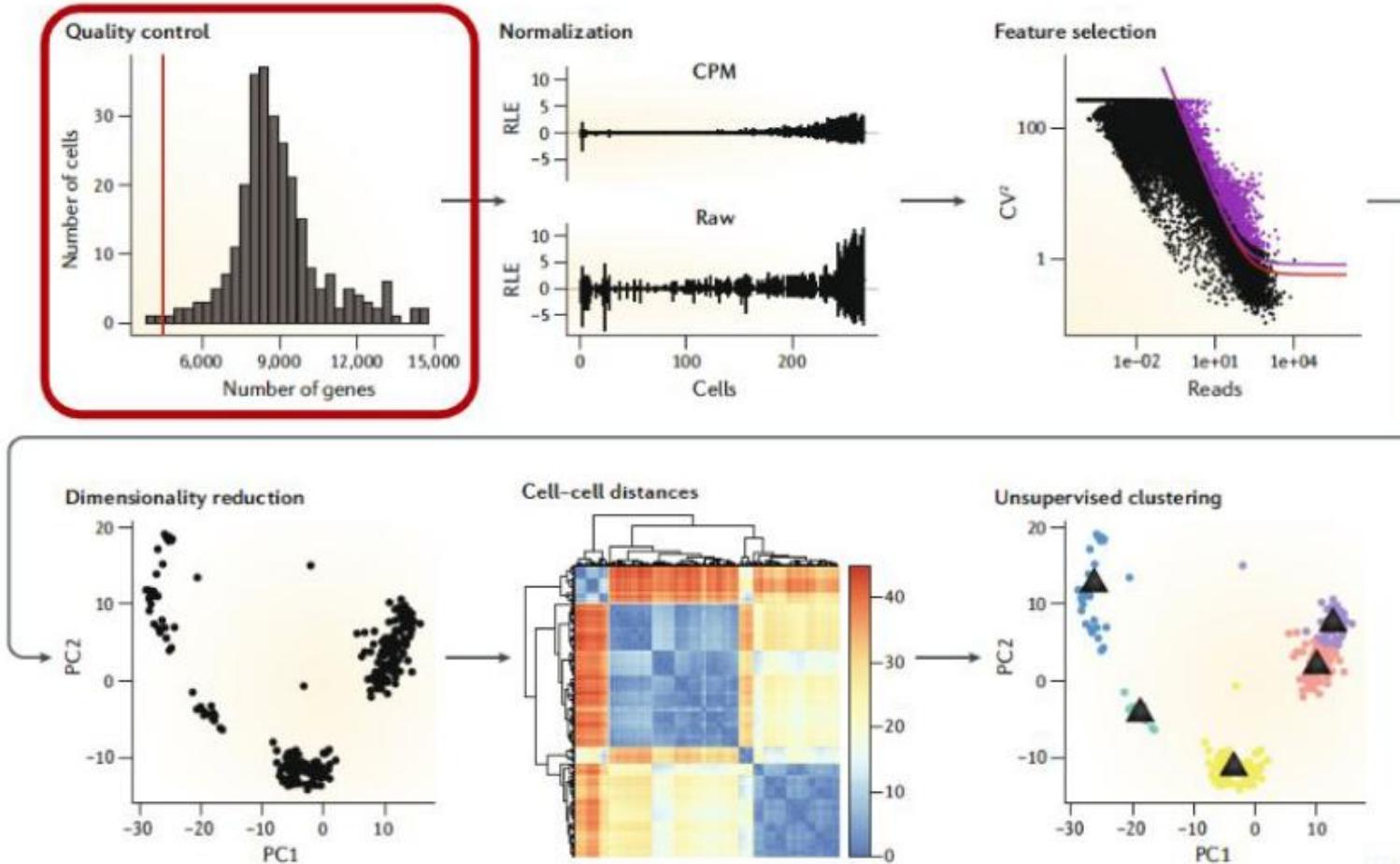
	SRR1039508	SRR1039509	SRR1039512	SRR1039513	SRR1039516
ENSG000000000003	679	448	873	408	1138
ENSG000000000005	0	0	0	0	0
ENSG00000000419	467	515	621	365	587
ENSG00000000457	260	211	263	164	245
ENSG00000000460	60	55	40	35	78

raw counts that indicate the number of sequencing reads mapped to each gene

# Source of noise:



# Step 4: Quality Control - removing low counts:



Motivation:

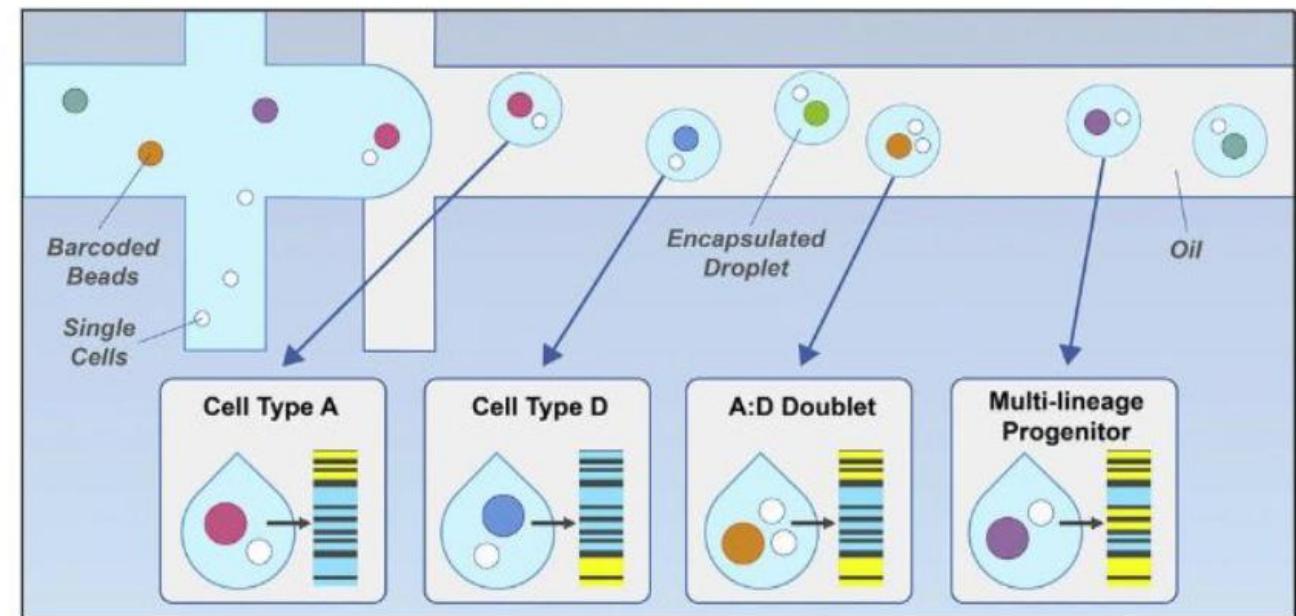
'Cells' featuring one or more of:

- ⚠ Low total counts
- ⚠ Few expressed genes
- ⚠ High proportion of reads coming from mitochondria

# Step 4: Quality Control - removing Doublets:

Doublet = when two cells are lysed and sequenced within the same droplet

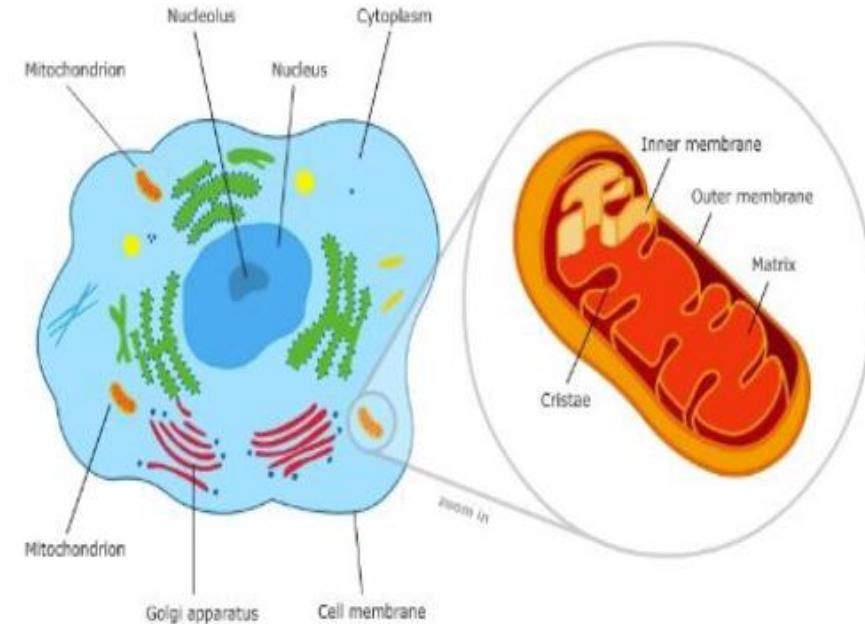
- Number of unique genes detected in each cell.
- Low-quality cells or empty droplets will often have very few genes.
- Cell doublets or multiplets may exhibit an aberrantly high gene count.
- Technical terms:
  - Feature count = number of genes
  - RNA count = number of UMIs



DePasquale et al 2019

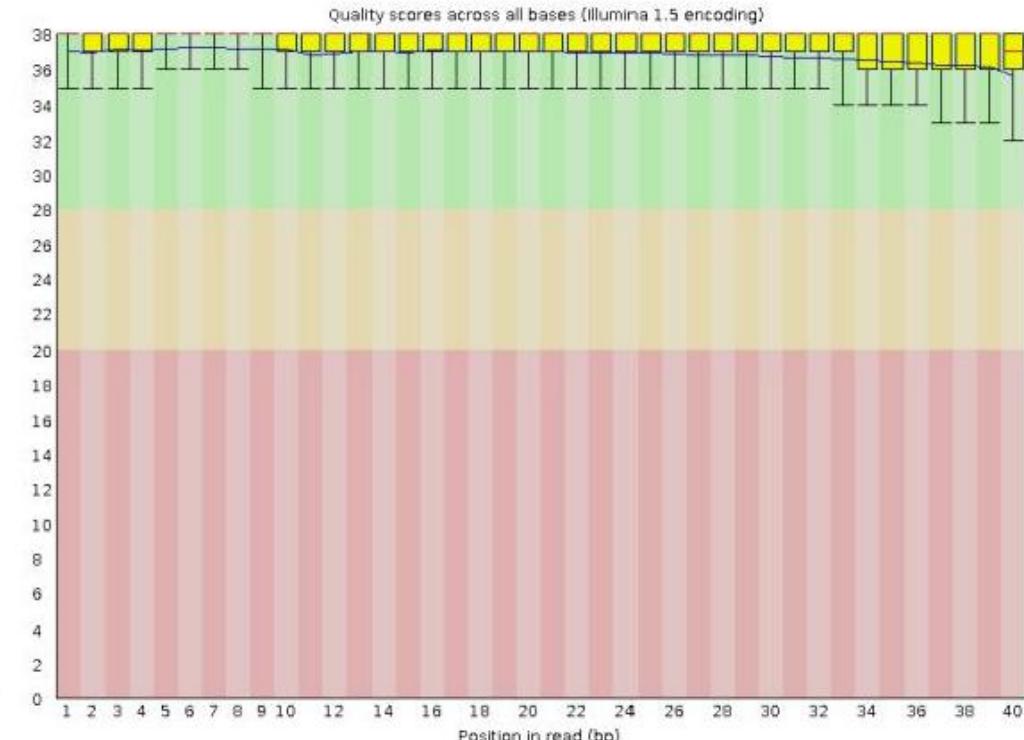
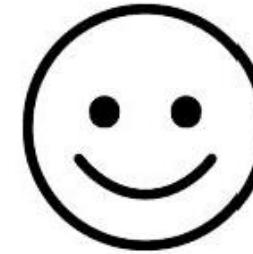
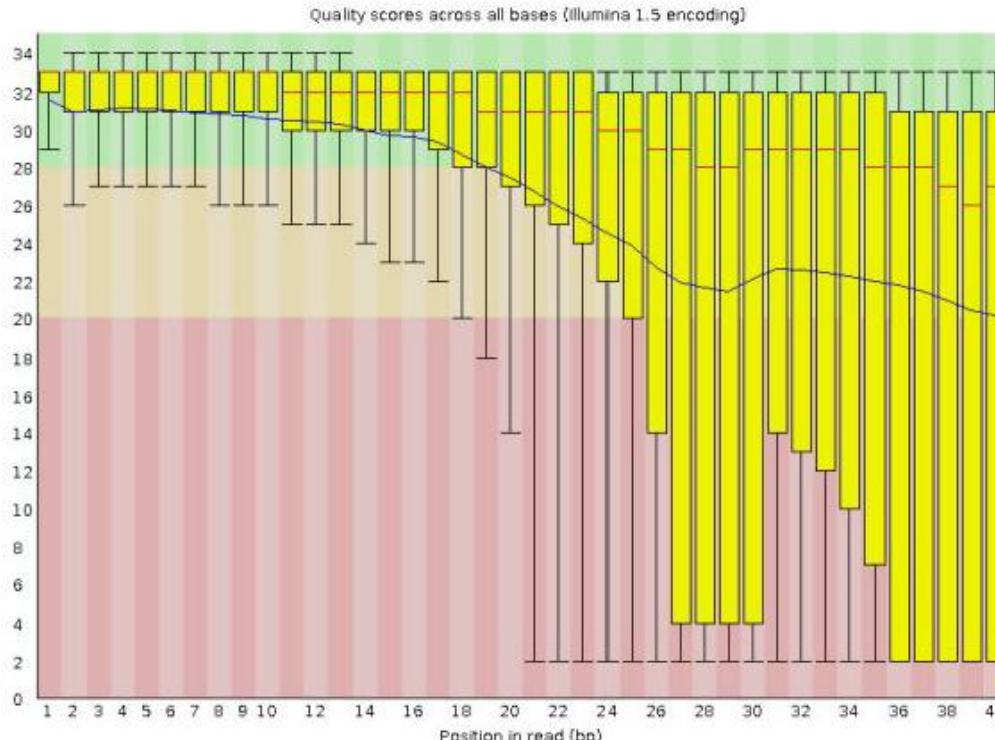
# Step 4: Quality Control - remove Mitochondrial genes:

- Mitochondrial RNA
  - Due to very harsh conditions in tissue dissociation step.
  - Dying cells release their cytoplasmic contents.
  - The percentage of reads that map to the mitochondrial genome.
  - + Low-quality / dying cells often exhibit extensive mitochondrial contamination.



© 2007-2011 The University of Waikato | www.sciencelearn.org.nz

# Step 4: Quality Control: Sample outcomes



# Step 5: Normalization of scRNA-Seq Data

Normalization **adjusts for technical variability and ensures accurate, biologically meaningful comparisons** across cells for reliable downstream analysis.

Raw counts

Normalization

Log-transformation

Regressing out confounders

Scaling

# Normalization of scRNA-Seq Data

Raw counts

Normalization

Log-transformation

Regressing out confounders

Scaling

	GGCACCA	CGTCGG	ACTCAT
LRRN3	0	0	0
CCR7	0	1	0
SELL	0	0	0
PASK	0	0	0
IL7R	0	4	0
KLRB1	0	0	3
TNFSF13B	0	0	0
CCL4	0	0	1
GZMH	0	0	3
GZMA	0	4	4
GNLY	0	0	2
NKG7	0	0	12
CST7	0	1	2

# Normalization of scRNA-Seq Data

Raw counts

Normalization

Log-transformation

Regressing out confounders

Scaling

	GGCACCA	CGTCGG	ACTCAT
LRRN3	0	0	0
CCR7	0	2.56937	0
SELL	0	0	0
PASK	0	0	0
IL7R	0	10.2775	0
KLRB1	0	0	8.9021
TNFSF13B	0	0	0
CCL4	0	0	2.9674
GZMH	0	0	8.9021
GZMA	0	10.2775	11.869
GNLY	0	0	5.9347
NKG7	0	0	35.608
CST7	0	2.56937	5.9347

$$size_i = \sum_j counts_{ij}$$

$$counts_{ij} = \frac{counts_{ij}}{size_i} \cdot 10^7$$

- $size(i)$  is the total count for cell  $i$ ,
- $counts(i,j)$  is the read count for gene  $j$  in cell  $i$ .
- $10^7$  is a scaling factor to bring the values into a comparable range, often used to express counts per million or similar units.

# Log transformation:

Raw counts

Normalization

Log-transformation

Regressing out confounders

Scaling

	GGCACAG	CGTCGG	ACTCAT
LRRN3	0	0	0
CCR7	0	1.27239	0
SELL	0	0	0
PASK	0	0	0
IL7R	0	2.42281	0
KLRB1	0	0	2.2927
TNFSF13B	0	0	0
CCL4	0	0	1.3781
GZMH	0	0	2.2927
GZMA	0	2.42281	2.5549
GNLY	0	0	1.9365
NKG7	0	0	3.6003
CST7	0	1.27239	1.9365

$$counts_{ij} = \log(counts_{ij} + 1)$$

Log transformation stabilizes variance, reduces the impact of extreme values, and enhances the detection of relative changes in scRNA-seq data.

# Confounding:

Raw counts

Normalization

Log-transformation

Regressing out confounders

Scaling

	GGCACA	CGTCGG	ACTCAT
LRRN3	-0.1444	-0.3176	-0.2511
CCR7	-0.7783	0.61731	-0.7795
SELL	-1	-0.7431	-0.9459
PASK	-0.4364	-0.5668	-0.5561
IL7R	-0.8651	0.8313	-1.2465
KLRB1	0.06623	-0.3682	2.1679
TNFSF13B	0.14242	-0.174	-0.0181
CCL4	-0.0371	-0.0673	1.3528
GZMH	-0.0573	-0.0612	2.2605
GZMA	-0.0785	2.19866	2.4734
GNLY	-0.0548	-0.1525	1.8888
NKG7	-0.1141	-0.1876	3.5241
CST7	-0.1069	0.96818	1.7942

Removes unwanted variation, isolating true biological signals in scRNA-seq data for more accurate analysis.

$$counts_{ij} = expression_{ij} + c_1 + c_2 + \dots + c_n$$

# Scaling scRNA-Seq Data

Raw counts

	GGCACCA	CGTCGG	ACTCAT
LRRN3	-0.23308909	-0.5128038	-0.40535638
CCR7	-0.9702567	0.76956093	-0.9717808
SELL	-1.2036221	-0.8944871	-1.1384836
PASK	-0.52249885	-0.6786562	-0.66580796
IL7R	-0.958673	0.92126876	-1.3813727
KLRB1	0.09821216	-0.5460074	3.2148056
TNFSF13B	0.3473894	-0.424432	-0.044154156
CCL4	-0.14821634	-0.26928374	5.411061
GZMH	-0.21939094	-0.23415962	8.654181
GZMA	-0.1635857	4.583083	5.155674
GNLY	-0.120534256	-0.33541664	4.152941
NKG7	-0.22848022	-0.375761	7.05939
CST7	-0.19288397	1.7462314	3.236064

$$counts_{ij} = \frac{counts_{ij} - mean_j(counts_{ij})}{std_j(counts_{ij})}$$

Log-transformation

Regressing out confounders

Scaling

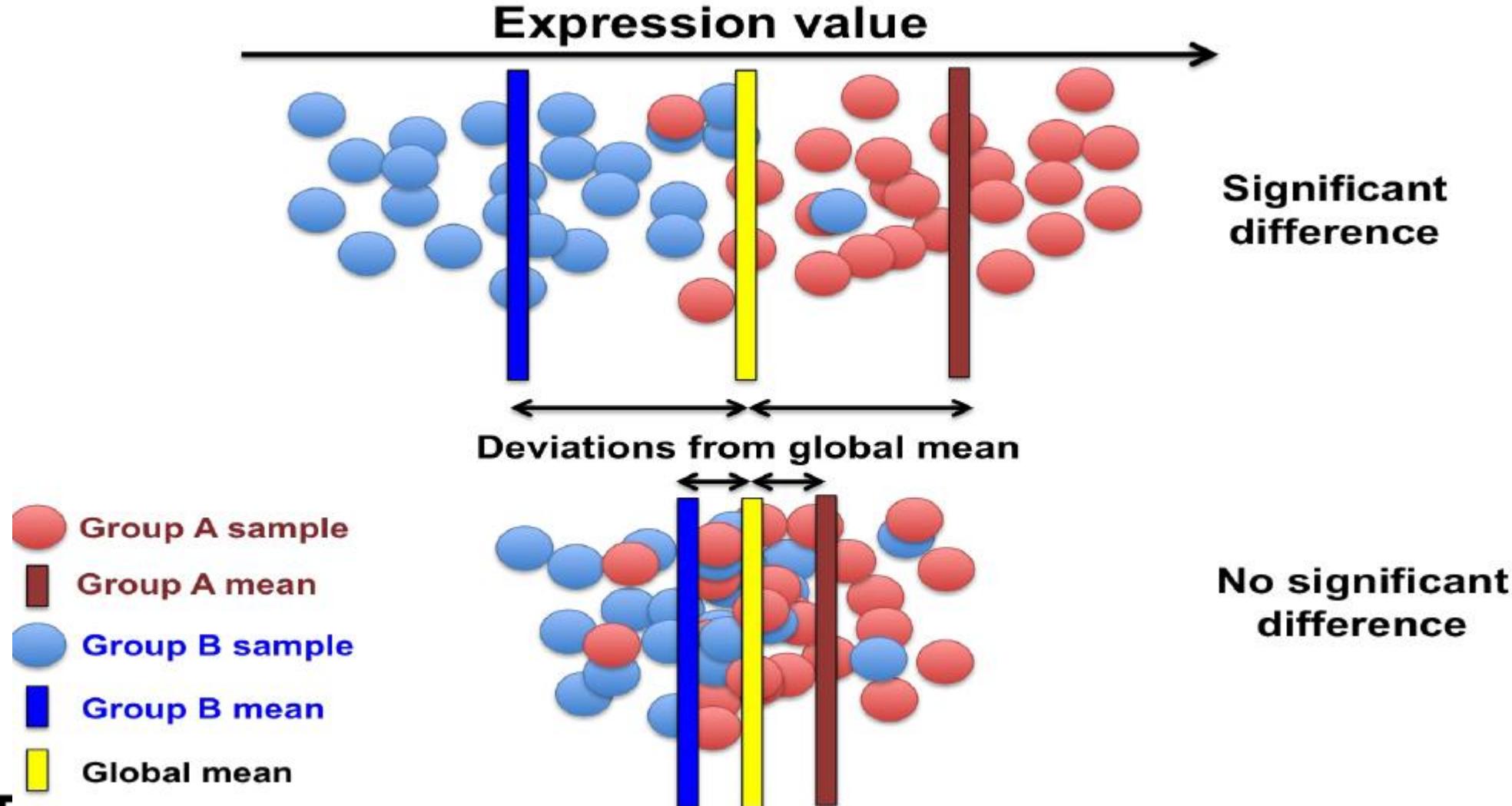
Scaling standardizes gene expression values across cells, ensuring comparability and optimizing the performance of downstream analyses like clustering and dimensionality reduction.

# Step 6: Differential Expression Analysis and Cell Annotation tools:

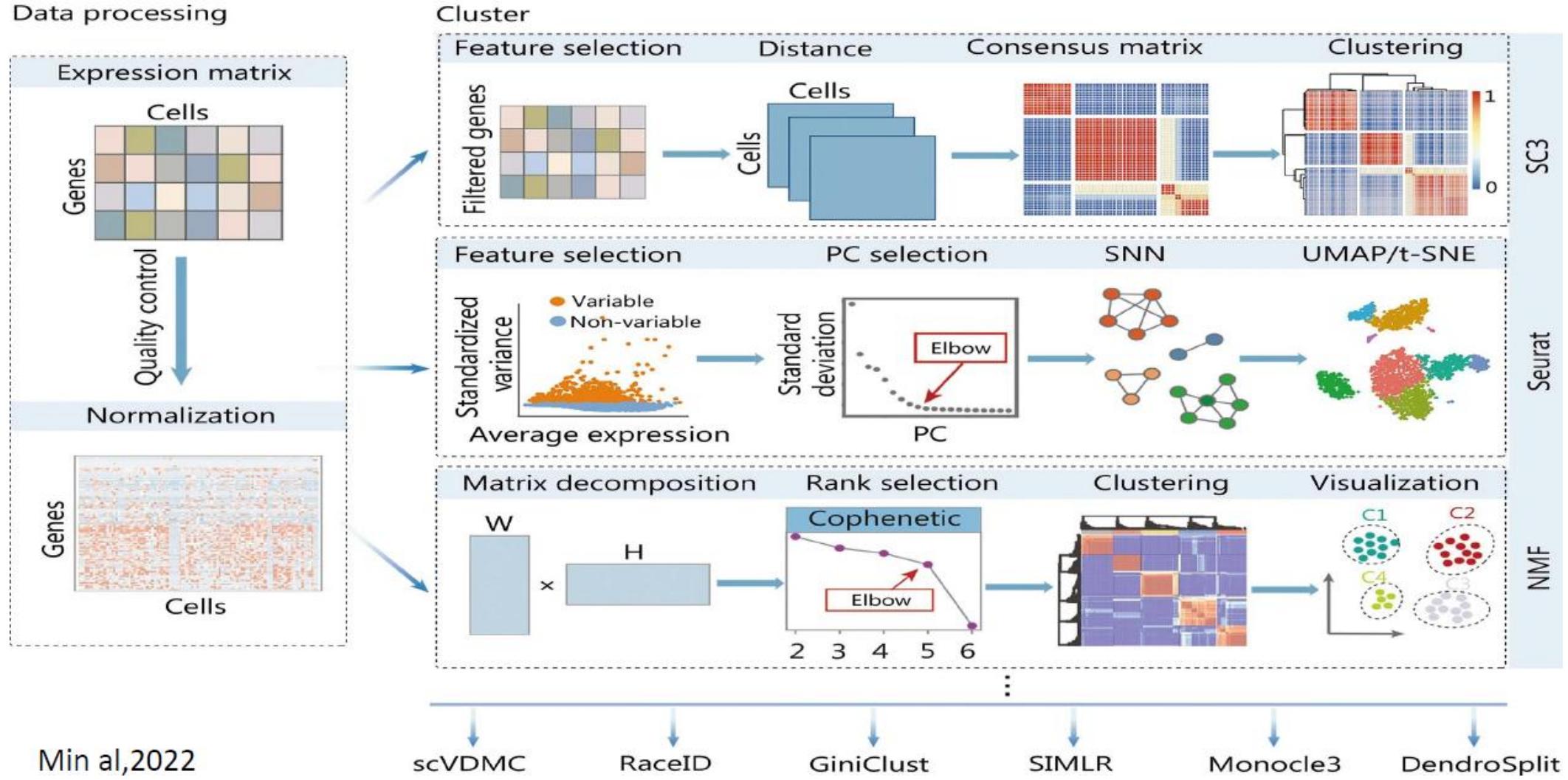
Attempts to infer genes that are statistically significantly over- or underexpressed between any compared groups (commonly between healthy and condition per cell type).



# Differential Expression Analysis: Division based on the gene attitude

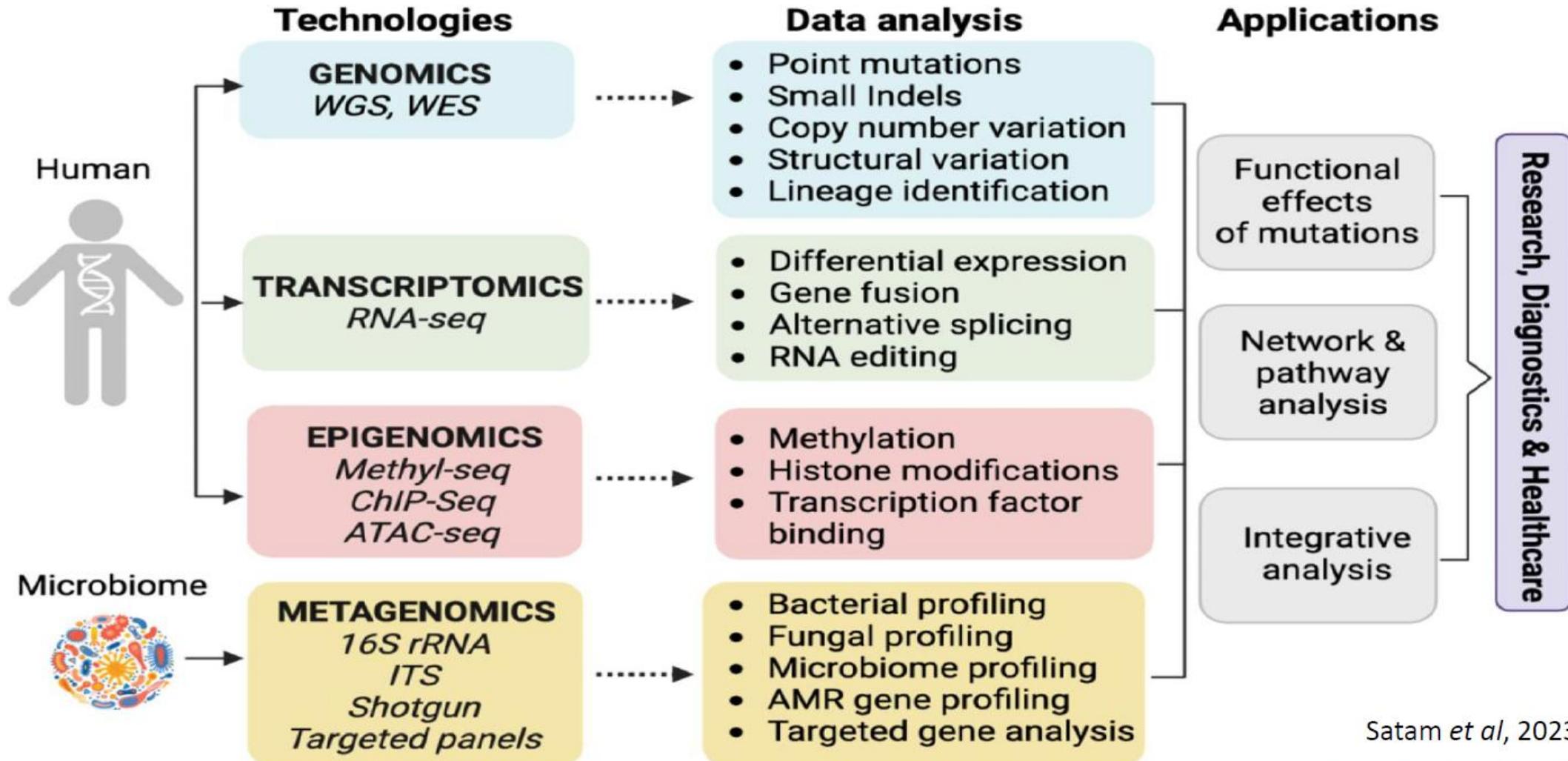


# Expression profiling and differential expression analysis



Min al,2022

# Various approaches used for genome analysis and applications of NGS, including technological platforms, data analysis, and applications



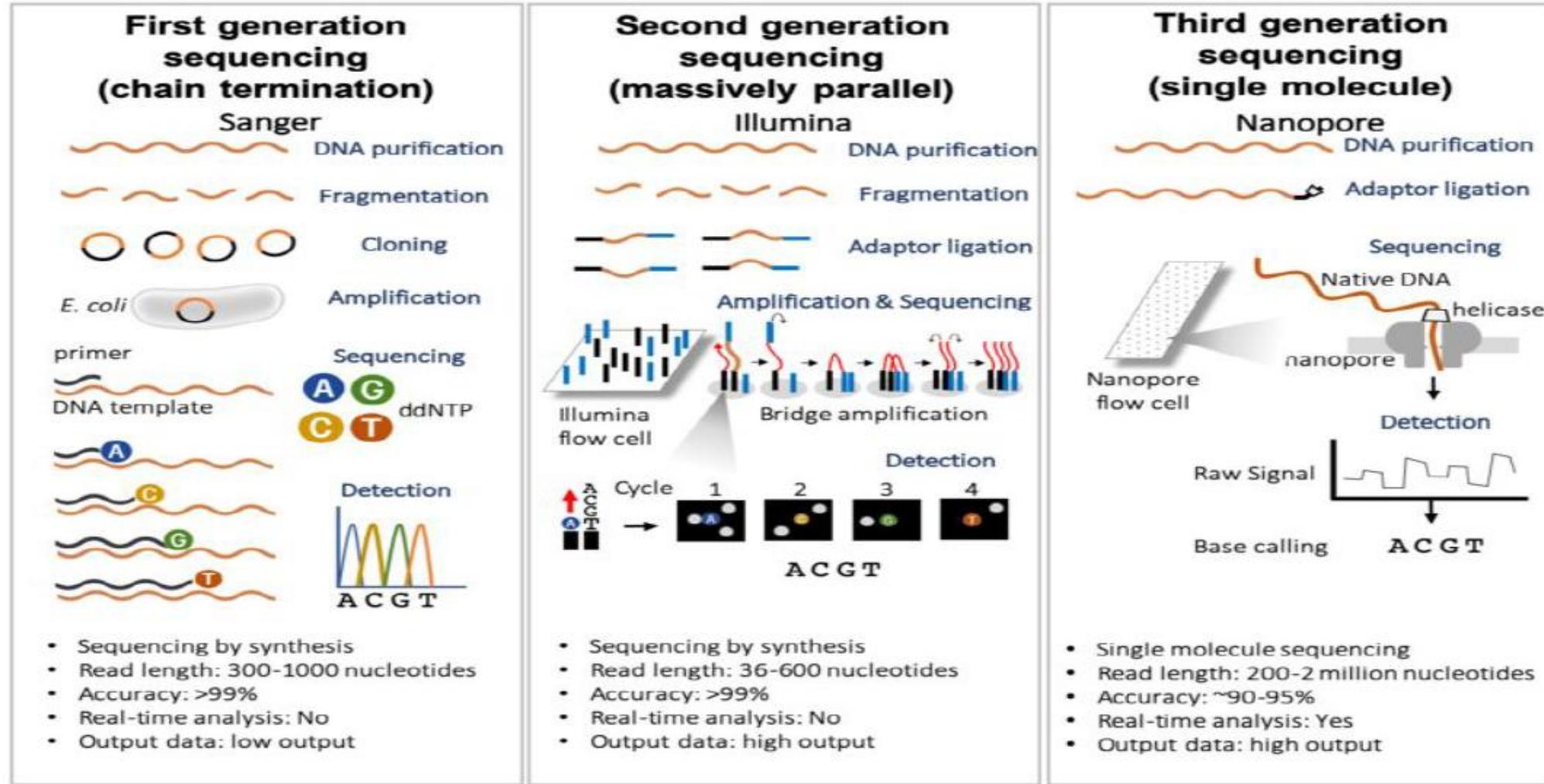
Satam et al, 2023

# Day 3: Genomics, Metagenomics and Transcriptomics

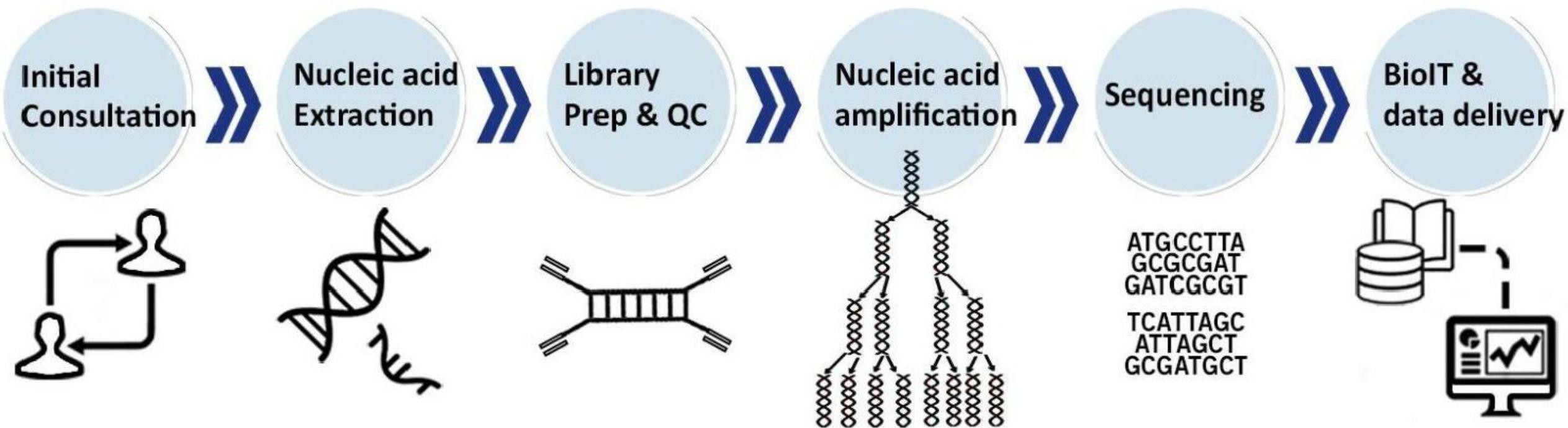
## Afternoon Practical Session

# Hands on: Transcriptomics Data Analysis Using Python

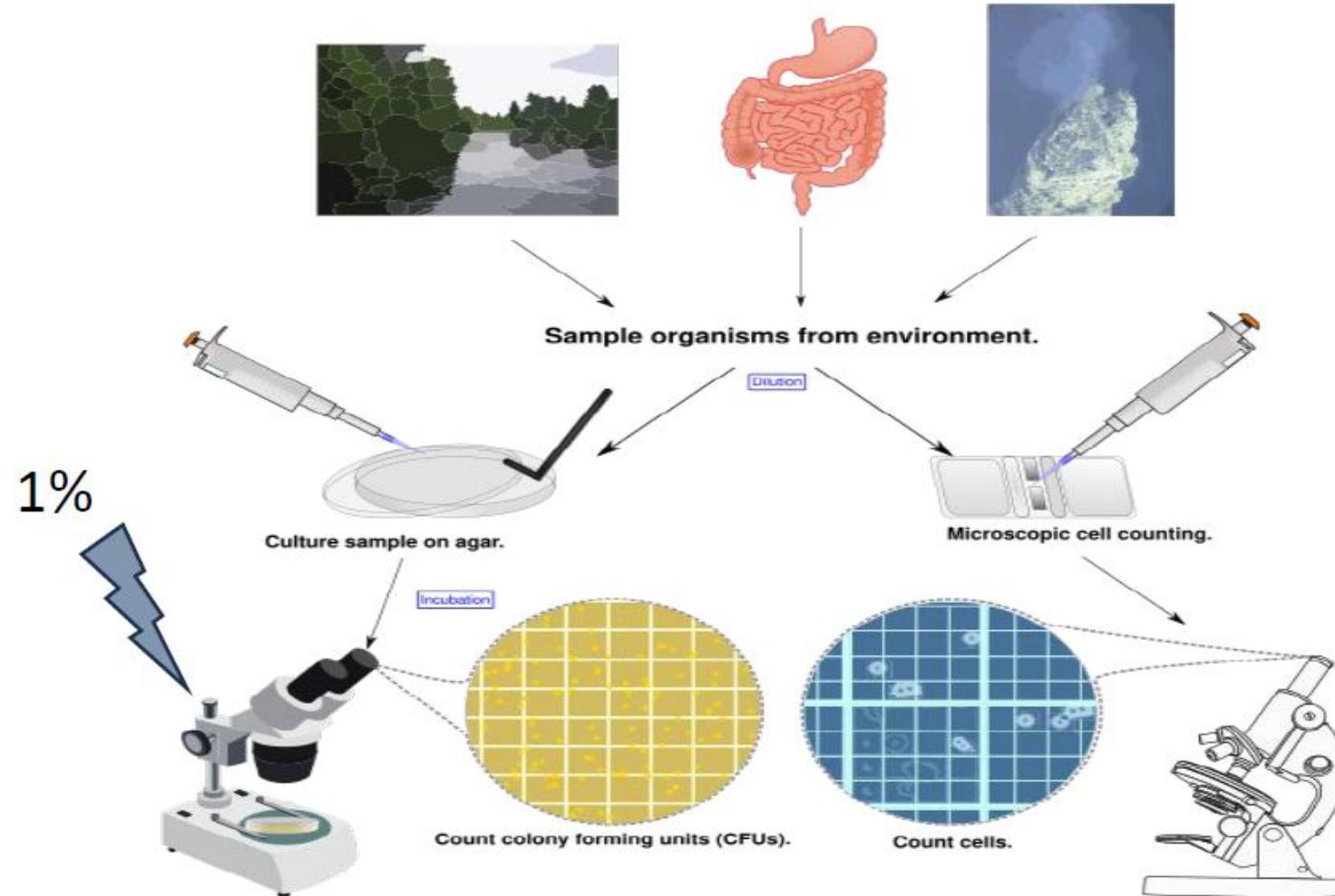
# Schematic Representation of first, second and third generation sequencing



# NGS Pipeline

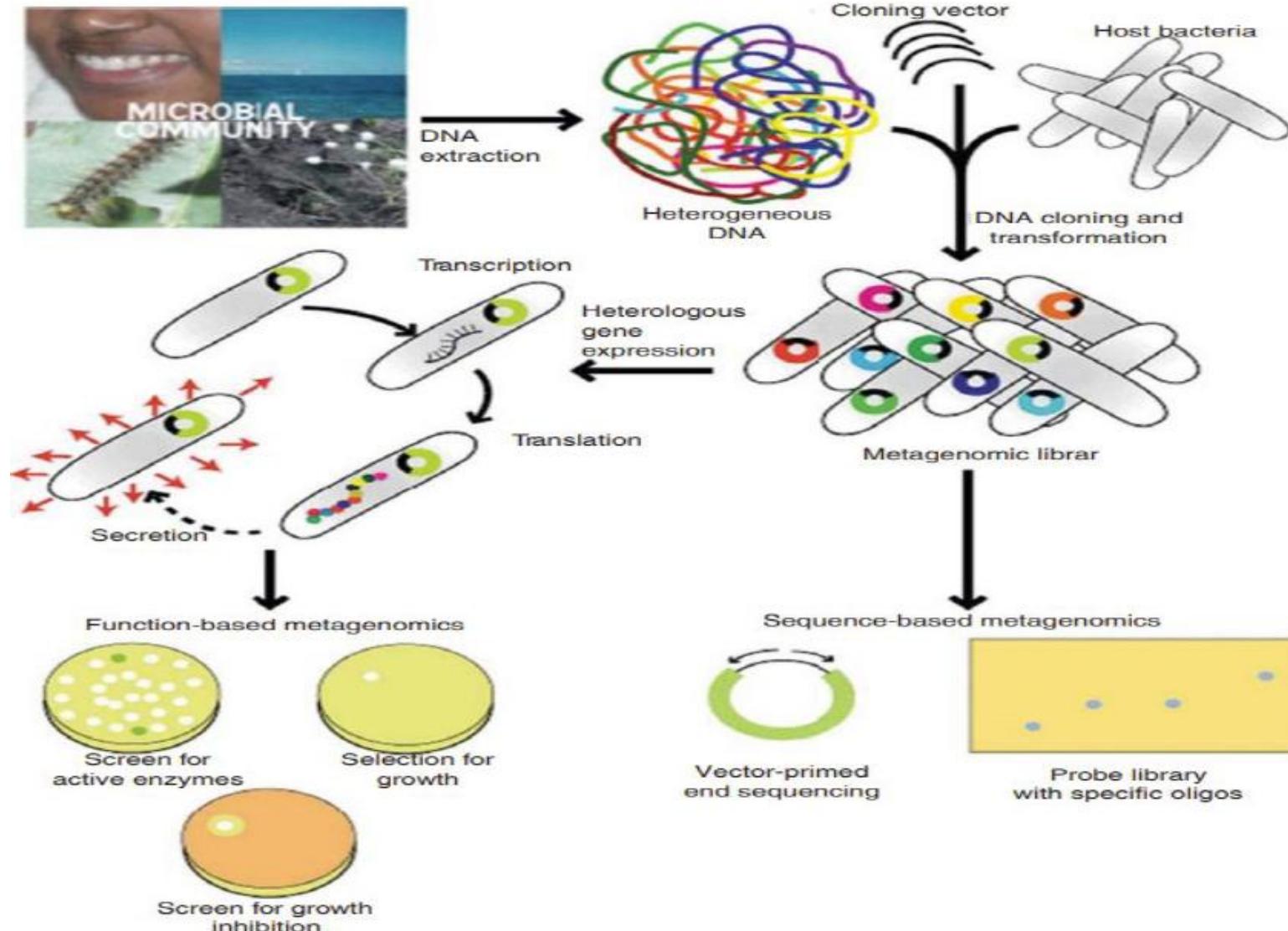


# Why metagenomics?



"The Great Plate Count Anomaly"

# Metagenomics pipeline analysis:

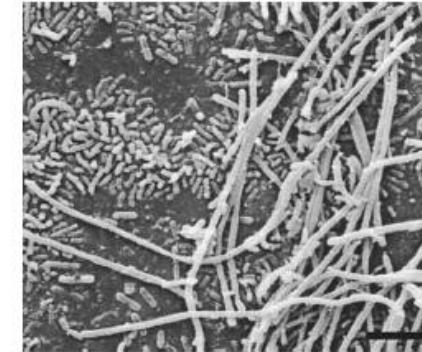


# Genomics and Metagenomics

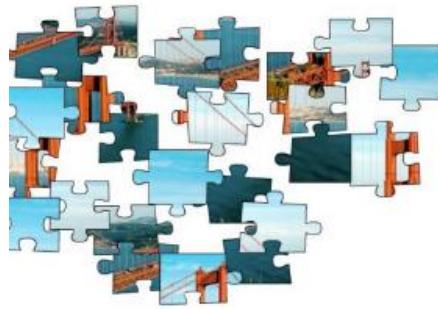
Genomics/ DNA  
Isolated



Sequences



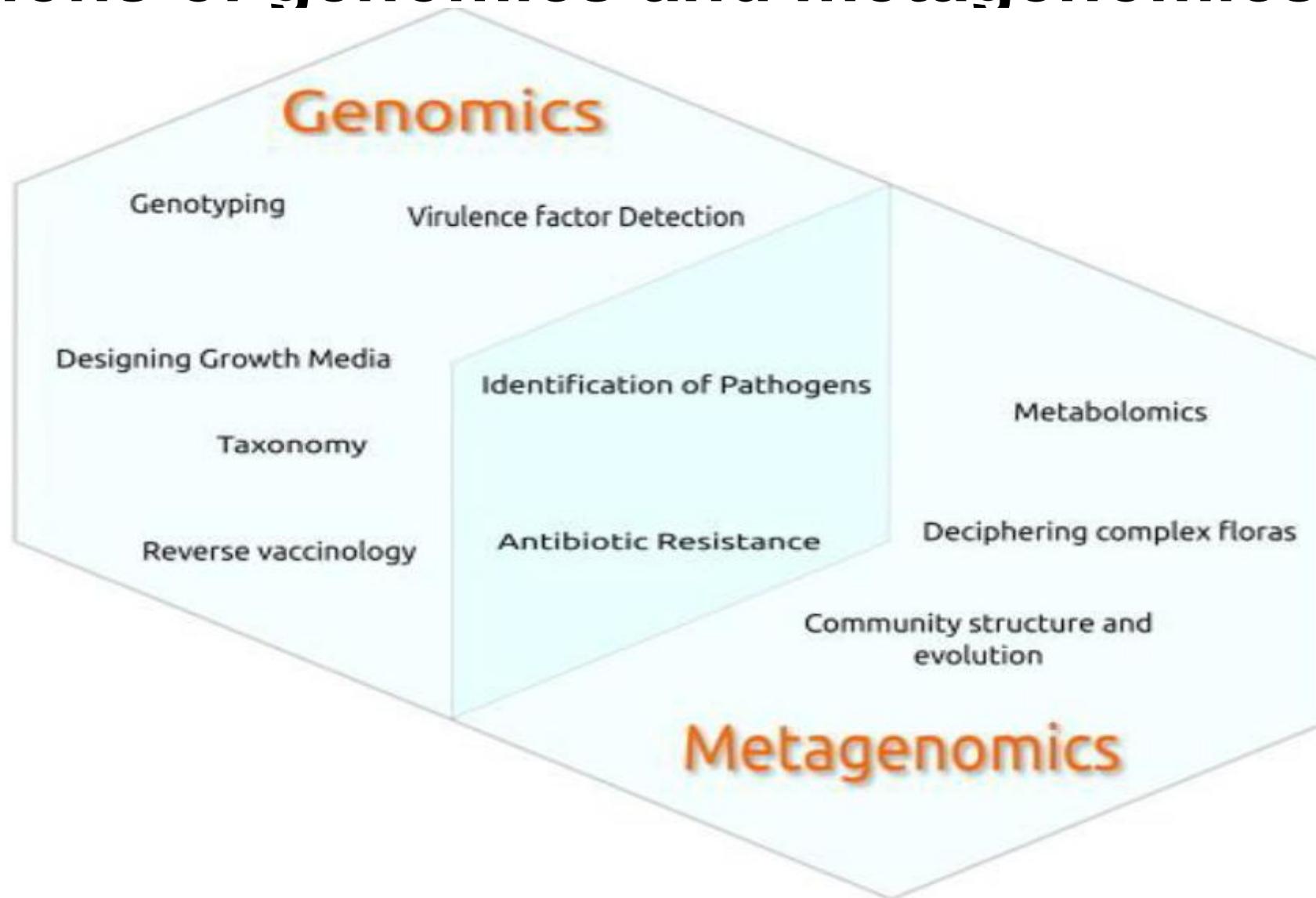
Metagenomics/  
Community



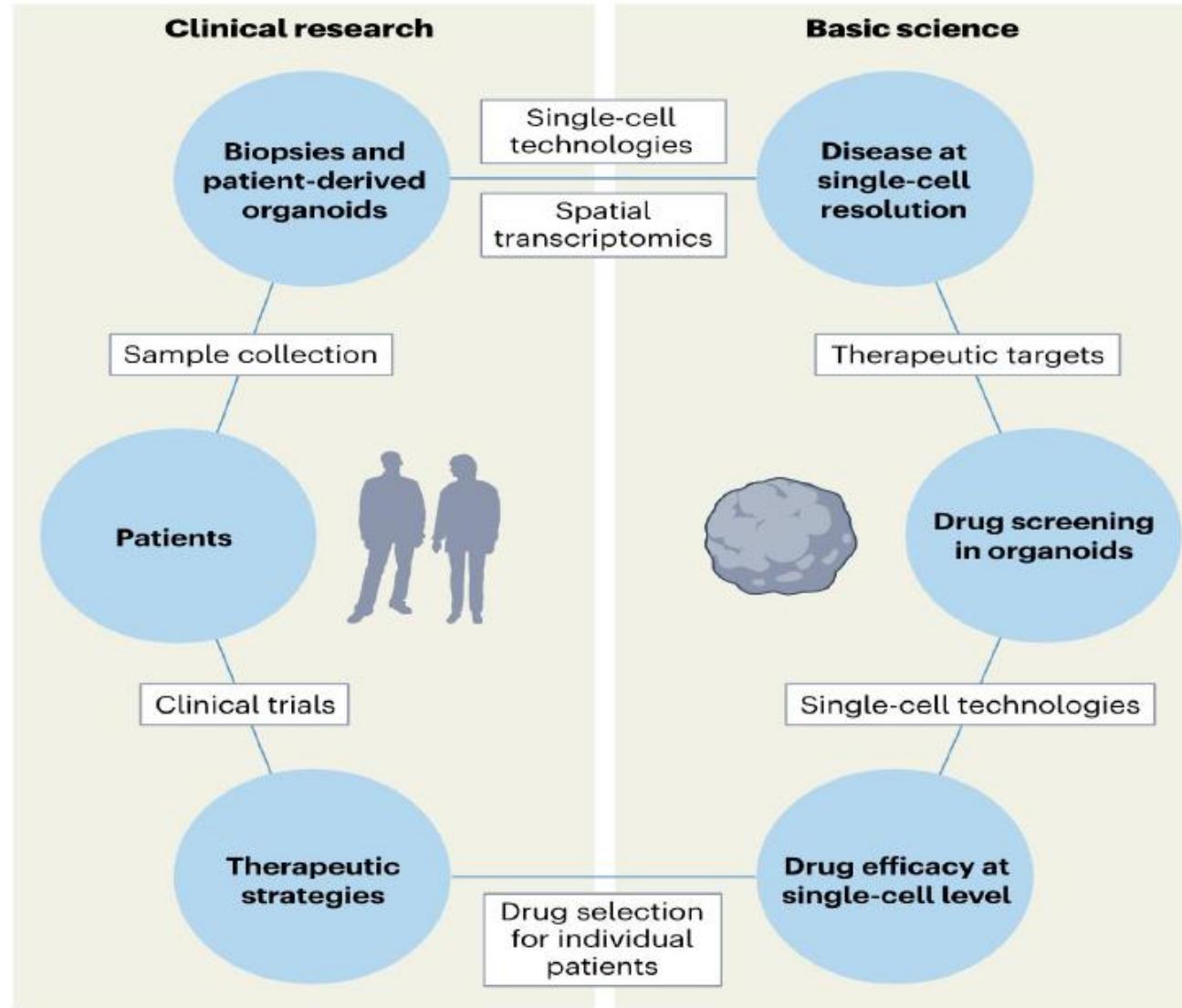
Assembly



# Applications of genomics and metagenomics



# Future directions for the application of single cell and spatial transcriptomics in clinical use.



# Future directions: Integrative omics approach

- Sequencing DNA or RNA is not enough to understand a disease or define a cell type!
- Integrate analysis using bioinformatics, AI, machine learning platforms.

