NGS (RNA-Seq)

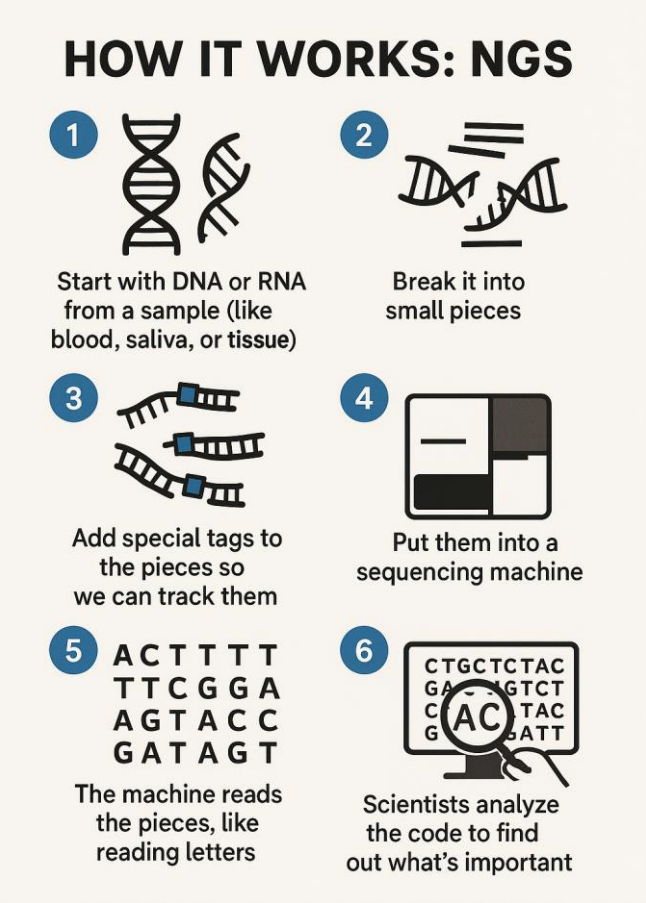
BASIC GUIDE FOR DATA ANALYSIS

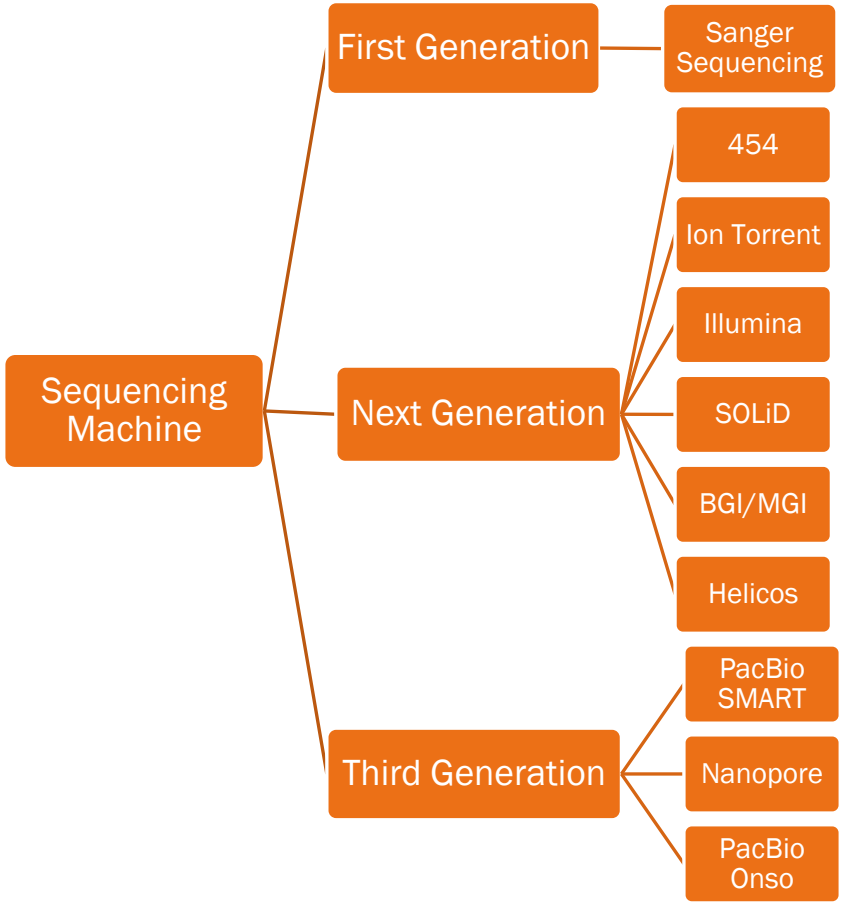
- SHOURYAN PATIL

What is NGS?

- Next-generation sequencing (NGS) is a way to read the order of letters (bases) in DNA or RNA much faster and cheaper than older methods((Sanger Sequencing).
- Instead of reading one piece at a time, NGS can read millions of pieces at once, all in a single experiment
- Think of DNA like a book full of letters (A, T, G, C). NGS is like a fast, automatic machine that can read millions of books at once, instead of reading one page at a time by hand.

Here’s how it works ?



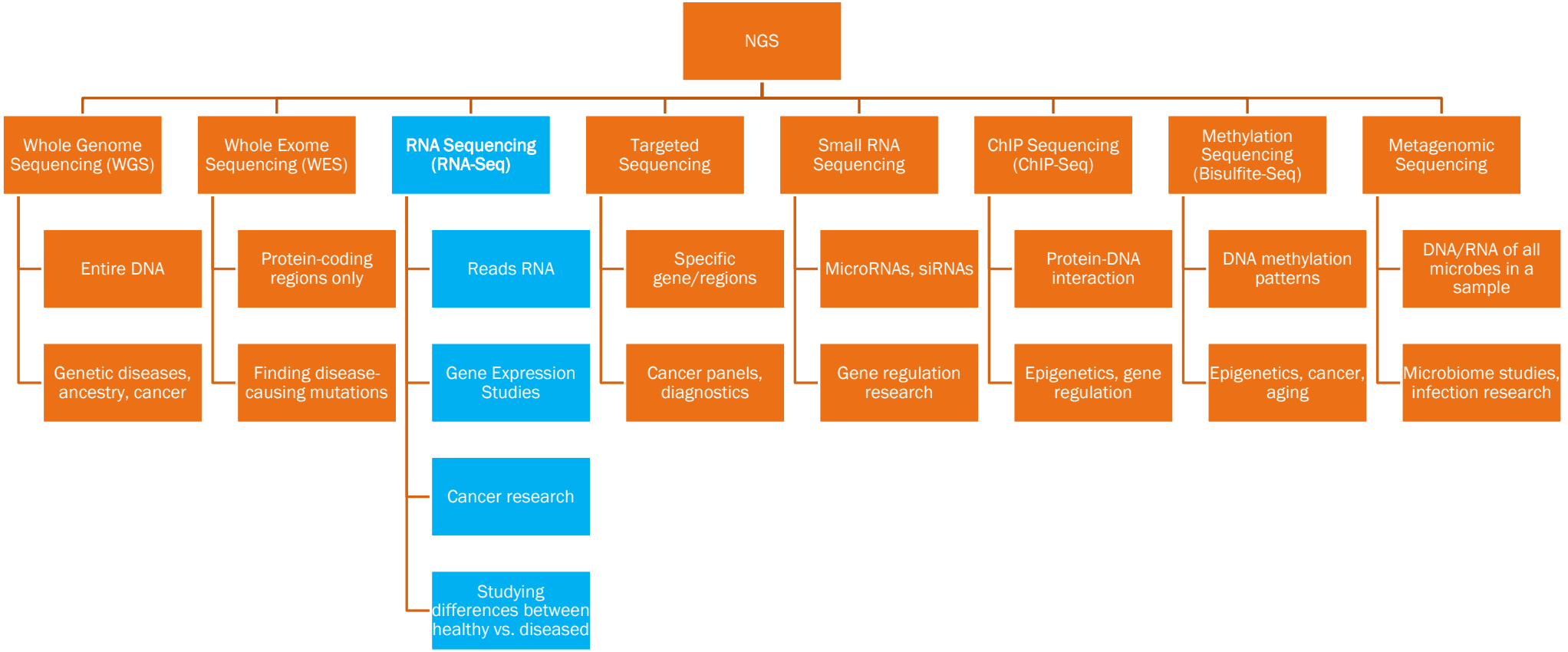


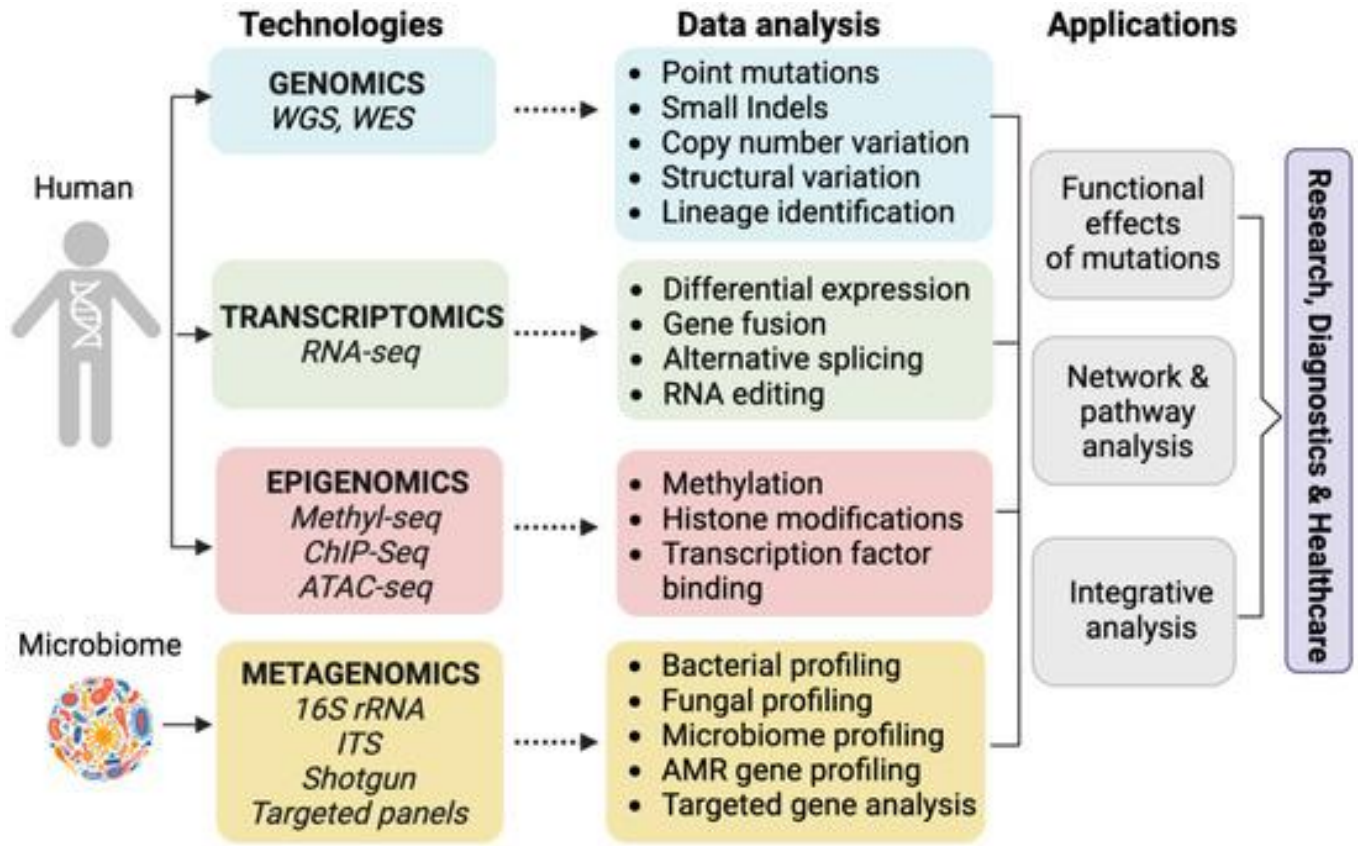
Why do we use NGS?

To find out

- Which genes are present in a person, plant, or animal.
- Which genes are active (turned on) or inactive (turned off).
- If there are mistakes or changes (mutations) in the DNA.
- To study diseases like cancer or genetic disorders.
- To study how bacteria or viruses spread.

Major types (tools) of NGS





DOI: HTTPS://DOI.ORG/10.3390/BIOLOGY12070997

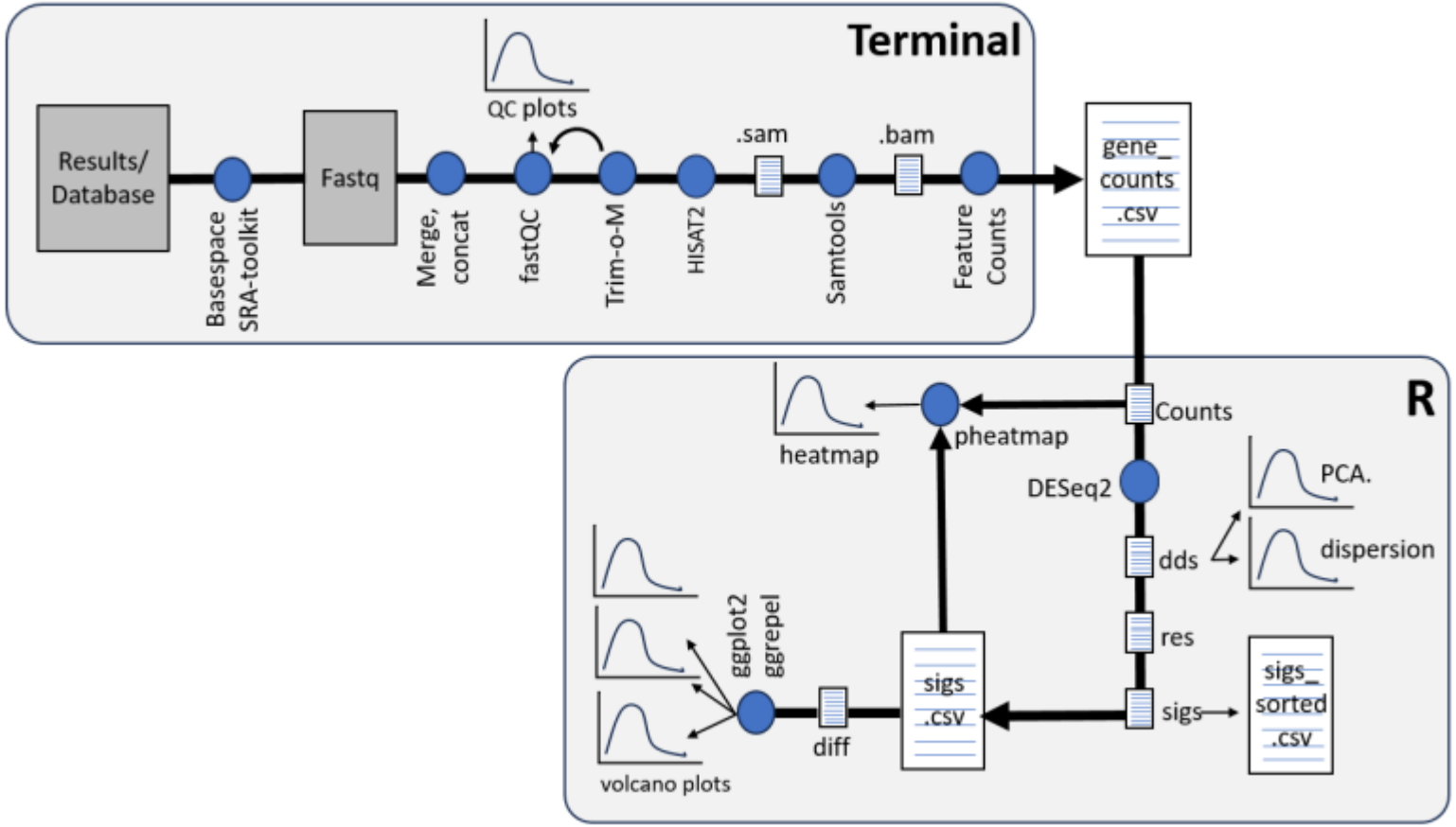
RNA Sequencing

- **RNA-Seq (RNA sequencing)** is **one specific method using NGS** to **study RNA**.
- It's used when you want to know **which genes are turned on or off** in cells.

1. **Gene Expression** — How much each gene is active in the sample.
2. **Differential Expression** — Find genes that are different between healthy and disease.
3. **Alternative Splicing** — Detect different versions of the same gene.
4. **New Genes/Transcripts** — Discover unknown genes or RNAs.
5. **Gene Fusions** — Find mixed/fused genes, often in cancer.
6. **Non-coding RNAs** — Study RNAs that don't make proteins but control other genes.
7. **Mutations/RNA Editing** — Spot changes in RNA or editing events.
8. **Pathway Analysis** — See which biological pathways are active or affected.
9. **Cell Types (Single-cell RNA-Seq)** — Identify different cell types in a tissue.
10. **Transcript Counts** — Get accurate counts of RNA molecules.

Overview

(Image DOI: 10.21769/BioProtoc.5295)



Git Hub Repo - https://github.com/Shouryanpatil/A-Guide-to-Basic-RNA-Seq

Where to Download data

FASTQ

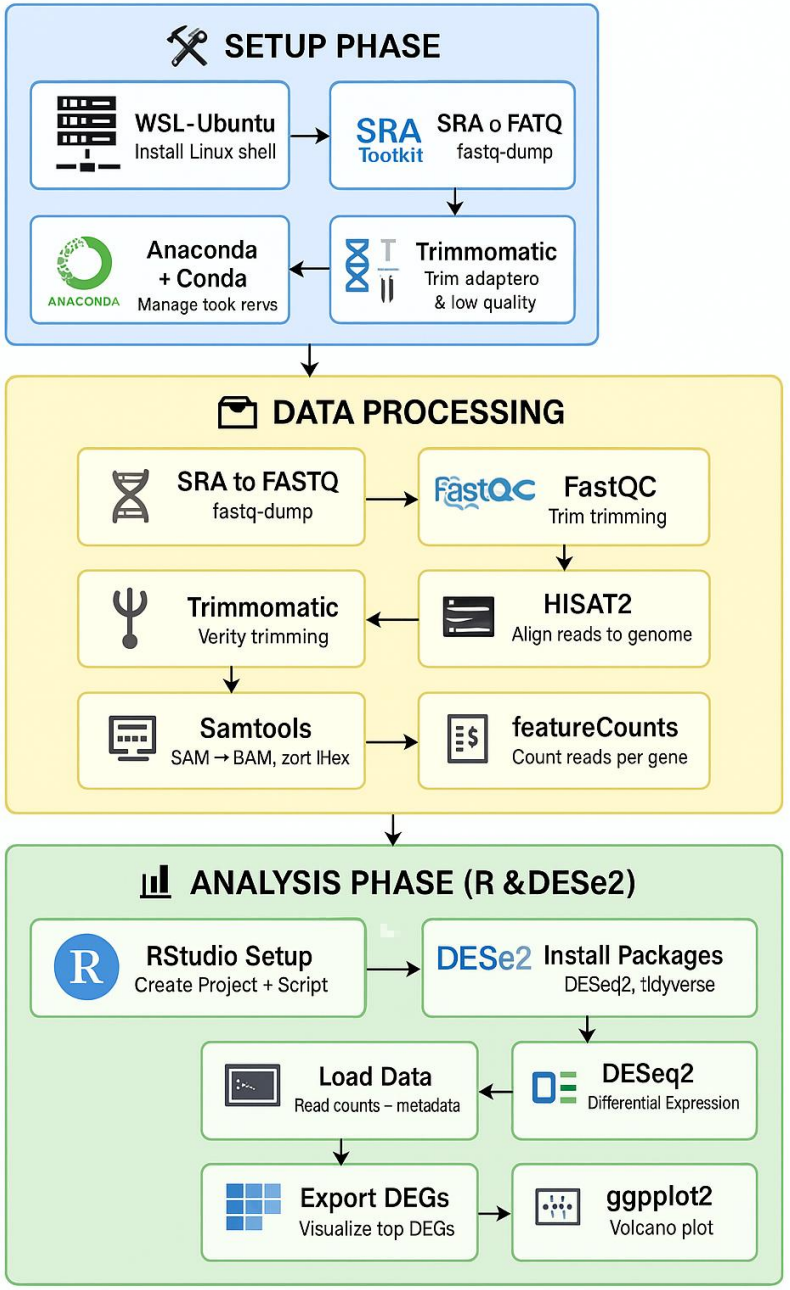
- NCBI SRA (Sequence Read Archive)
- ENA (European Nucleotide Archive)
- DDBJ (DNA Data Bank of Japan)

FASTA

- NCBI
- GenBank
- UniProt
- Ensembl

Workflow

- Set UP
- Data Processing (Linux)
- Analysis Phase (R Studio)



NeXT Step

Next step

Read Linux_RStudio_Guide.docx

Upload on git hub repo - https://github.com/Shouryanpatil/A-Guide-to-Basic-RNA-Seq

For Linux command refer – CODE.txt

For R Studio – RNA_SEQ.R file