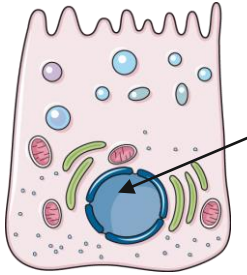


# The biology data dive



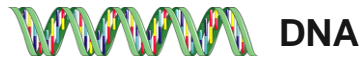
## Our body has vast amounts of information

There are ~ 30 trillion cells in the human body.



There are different types of cells but each one has a **NUCLEUS**. This stores genetic material called **DNA**.

Each cell has millions of protein molecules. Proteins are needed to generate cells so:



- DNA is a recipe book for making proteins (20,000 protein-coding genes)



- Messenger RNA (mRNA) carries a copy of the gene sequence to protein-making machines (~360,000 mRNA molecules in a single cell)

## What are the different types of biological data?

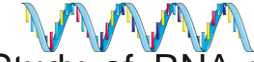
On the DNA level:



**Genetics:** Study of genes to understand inheritance and evolution

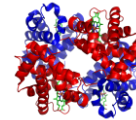
**Epigenetics:** Study of DNA chemical markers that switch genes on or off

On the RNA level:



**Transcriptomics:** Study of RNA molecules. Focus on mRNA, captures genes that are on.

On the protein level:



**Proteomics:** Study of protein molecules

**Phosphoproteomics:** Study of phosphate groups on proteins which controls its function

## Why do we need to dive into the biological data?

There are various applications:



- To understand how disease occurs and progresses
- To develop novel diagnostics and drugs
- To predict disease outcome and treatment response

Example of research translating to clinic:

- Screening for changes in *BRCA1/2* genes for risk of breast/ovarian cancer

Who takes the biology data dive?  
Bioinformaticians (check out the careers in perspective poster)