

## Sequence similarity- Basic concepts

**Sequence similarity is a measure of an empirical relationship between sequences.**

**A similarity score is therefore aimed to approximate the evolutionary distance between a pair of nucleotide or protein sequences.**

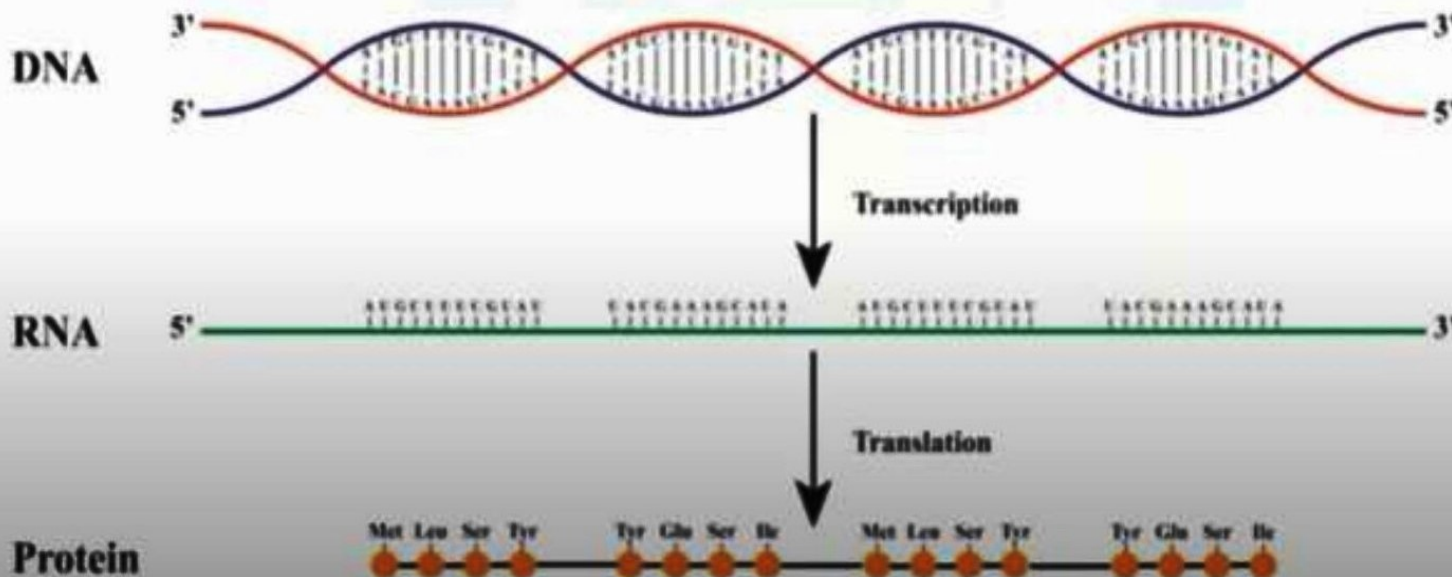
**Sequence similarity searching, typically with BLAST , is the most widely used, and most reliable, strategy for characterizing newly determined sequences.**

**Sequence similarity searches can identify "homologous" proteins or genes by detecting excess similarity – statistically significant similarity that reflects common ancestry.**

# What is sequence?

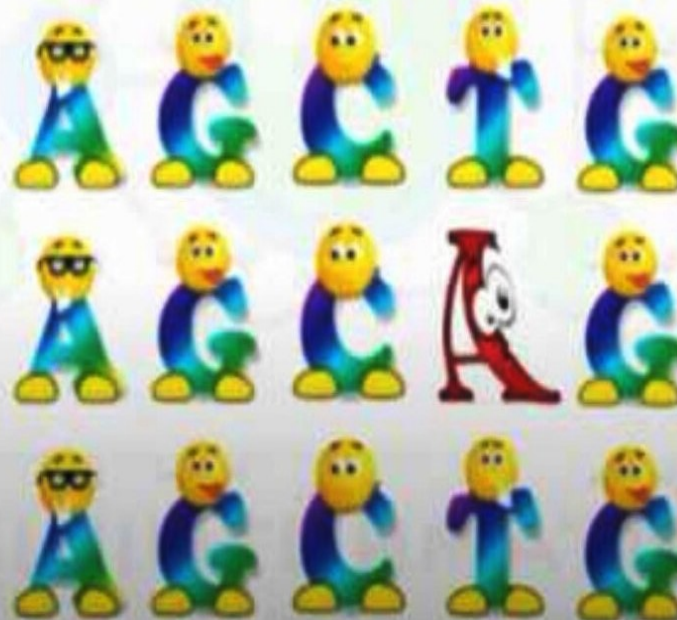
An arrangement of two or more **related things** in a successive order

**Related things:** Here these are DNA(base pairs as ACTG), RNA(base pairs as ACUG) or proteins(amino acids)



# Sequence Alignment?

The alignment of biological sequences is probably the most important and most accomplished in the field of bioinformatics. Their ultimate goal is to determine the similarity between different sequences.



**Search for Sequence similarity**

**One-One**

**One- Many**

**Many-Many**



**Compare two sequences  
in detail**

**In sequence data bases  
via BLAST/ FASTA**

**To compare multiple  
sequences at once**