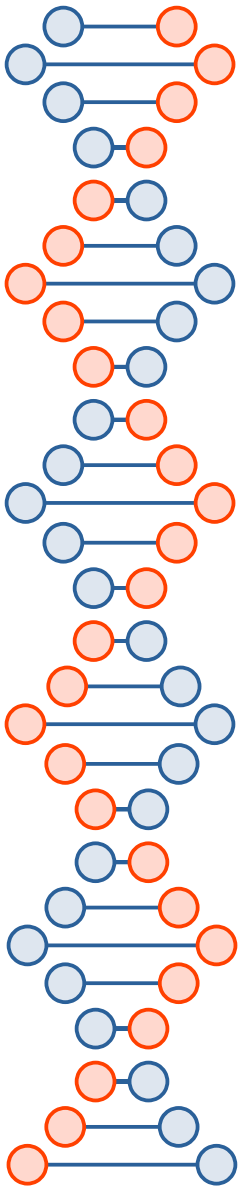


Computer Science and Medical Research

Session 1

Introduction to Bioinformatics
github.com/Yoyomanzoor/CSMR



Me

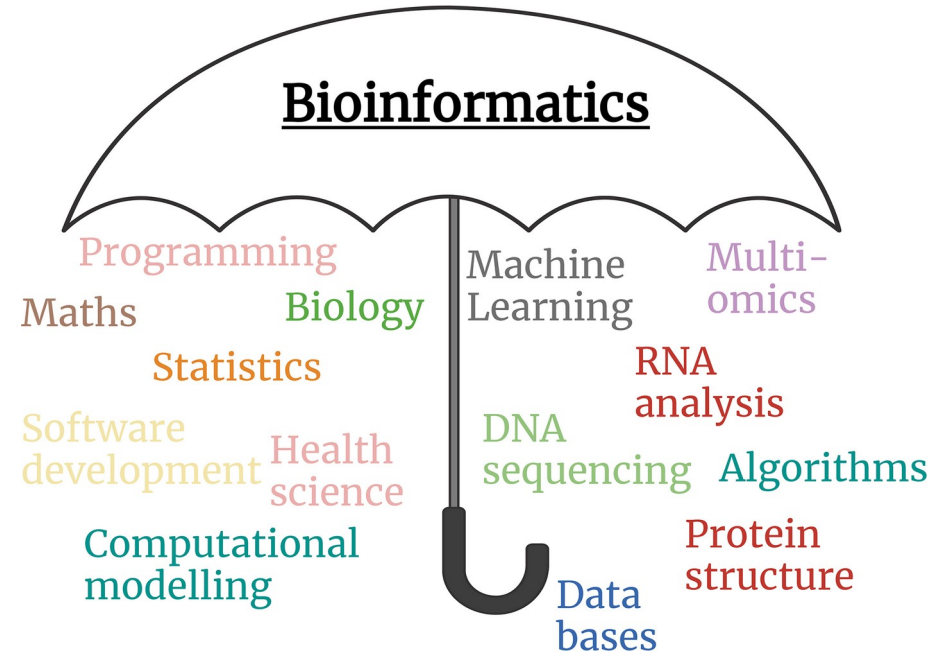
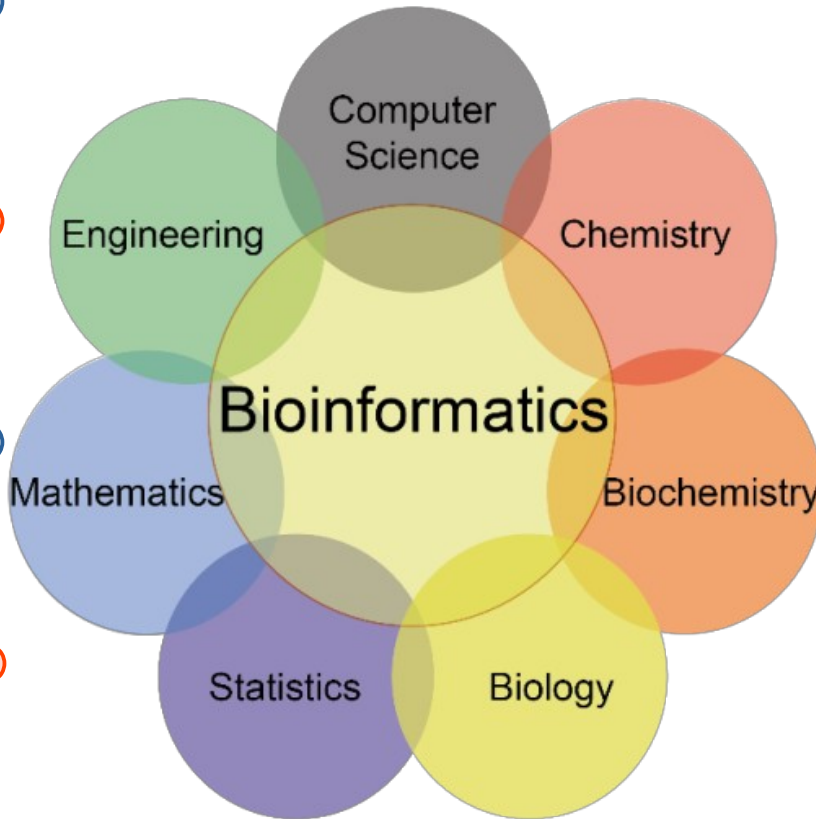
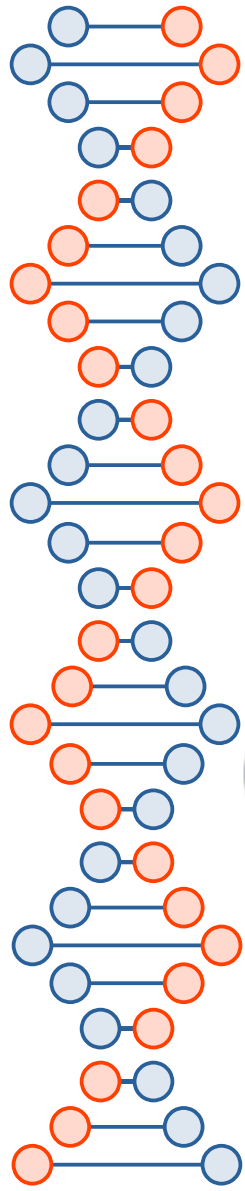
- Sumeed Manzoor
- smanzoor@umich.edu
- 2nd year grad student at the University of Michigan



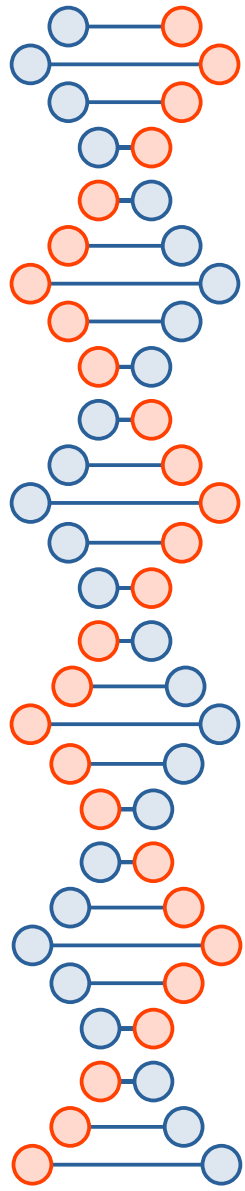
You?

- School?
- Year?
- What brought you here?

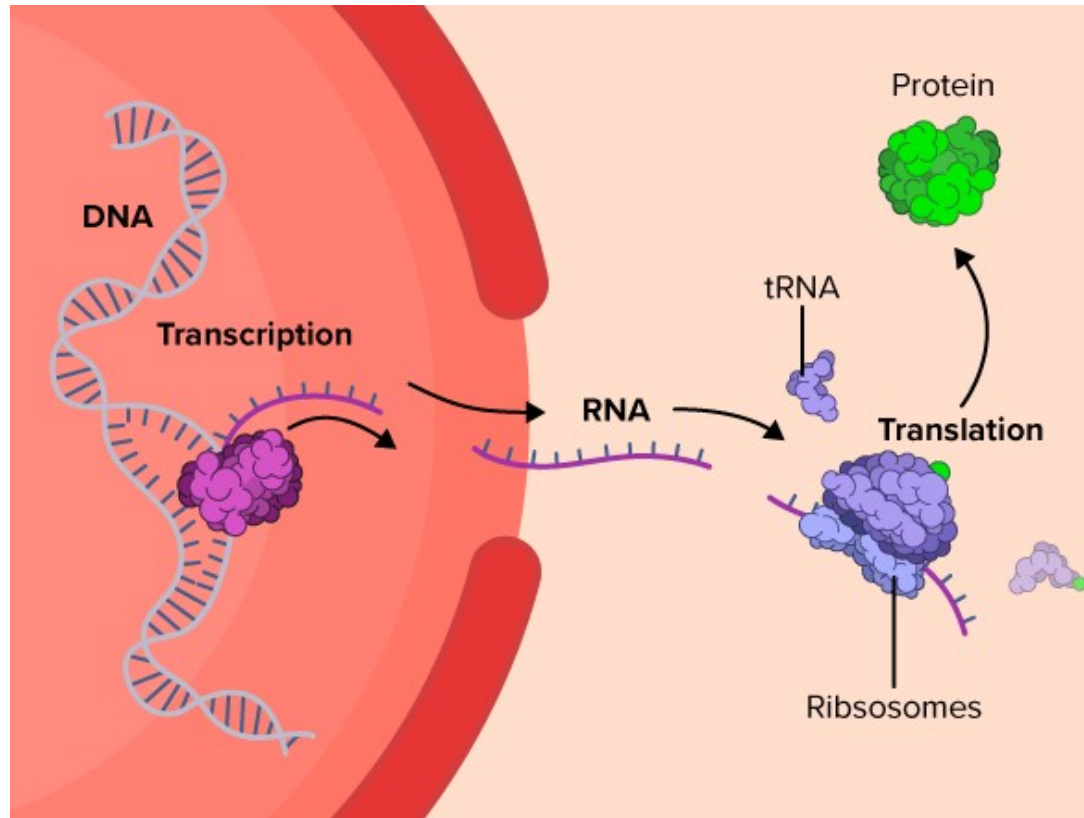
Bioinformatics



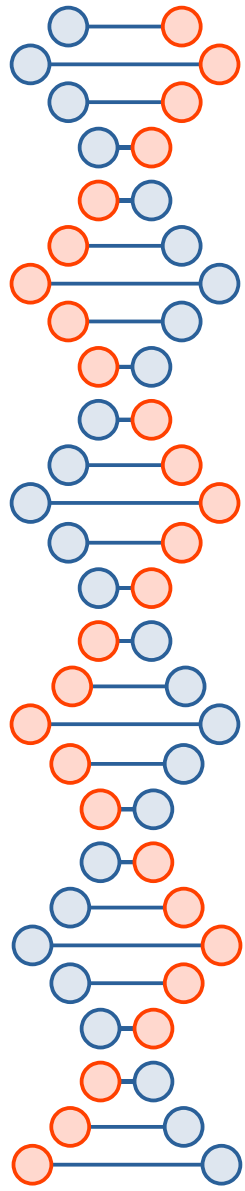
https://doi.org/10.1042/bio_2022_136



Central dogma

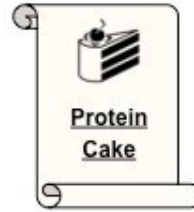


Central dogma



DNA
(cook book)

Transcription
(photocopying)



mRNA
(recipe)



Ribosome
(chef)

+

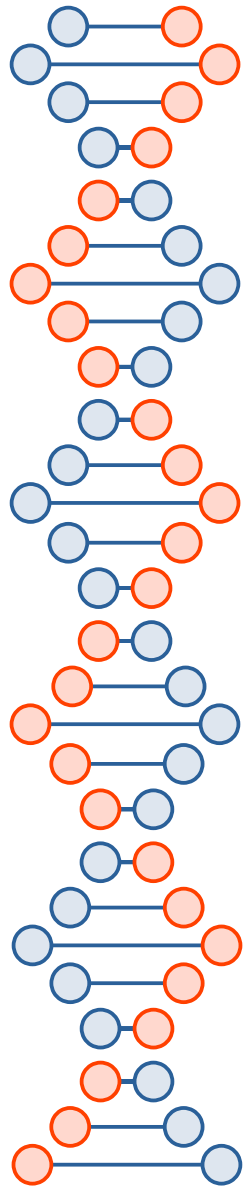


tRNA
(kitchen hand)

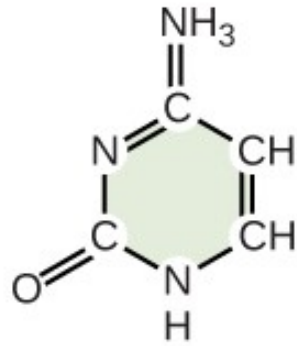
+



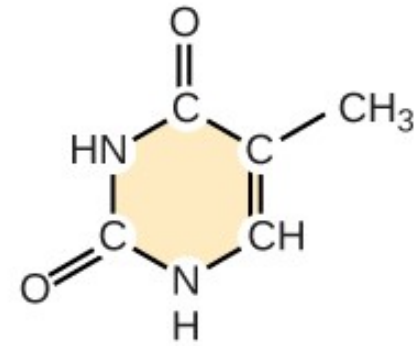
Amino acid
(ingredient)



Pyrimidines

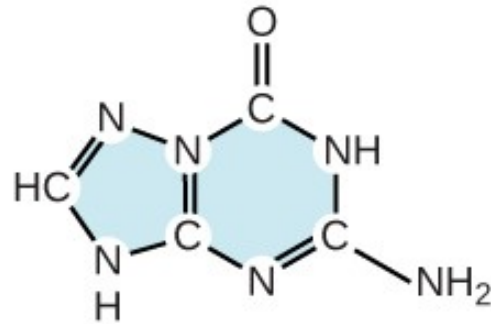


Cytosine
C

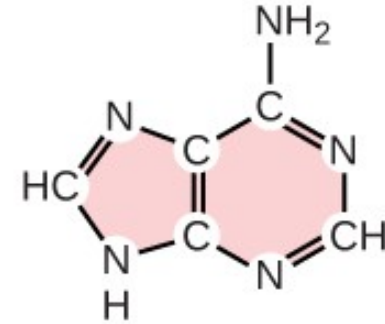


Thymine
T

Purines

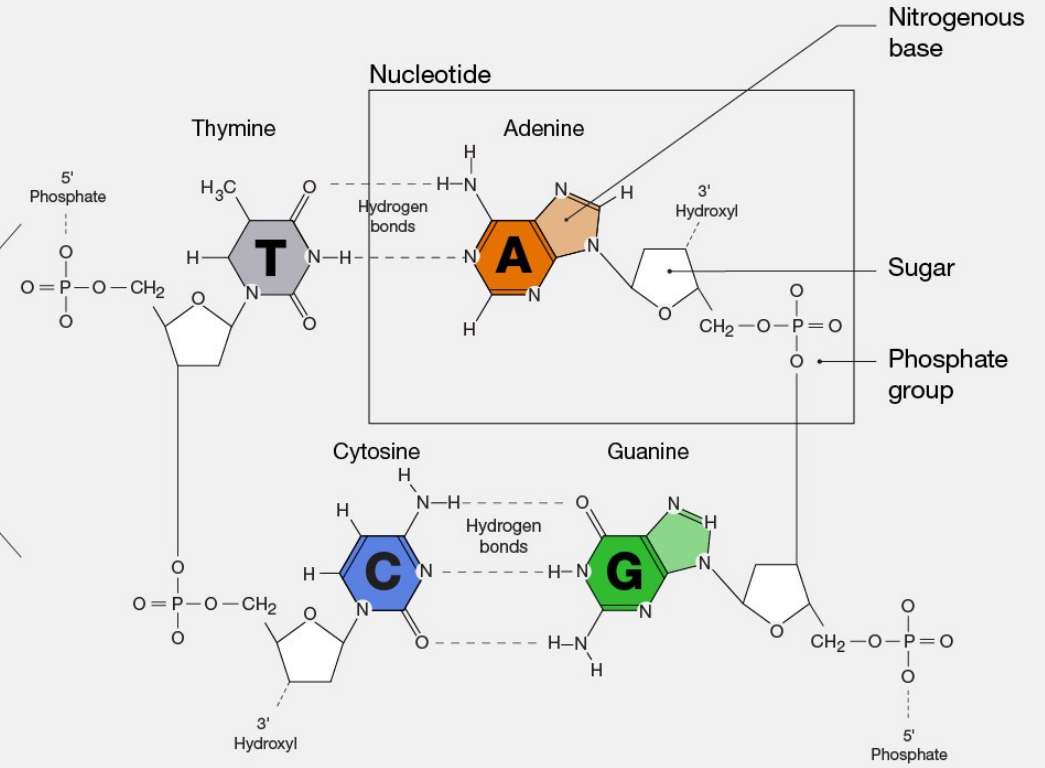
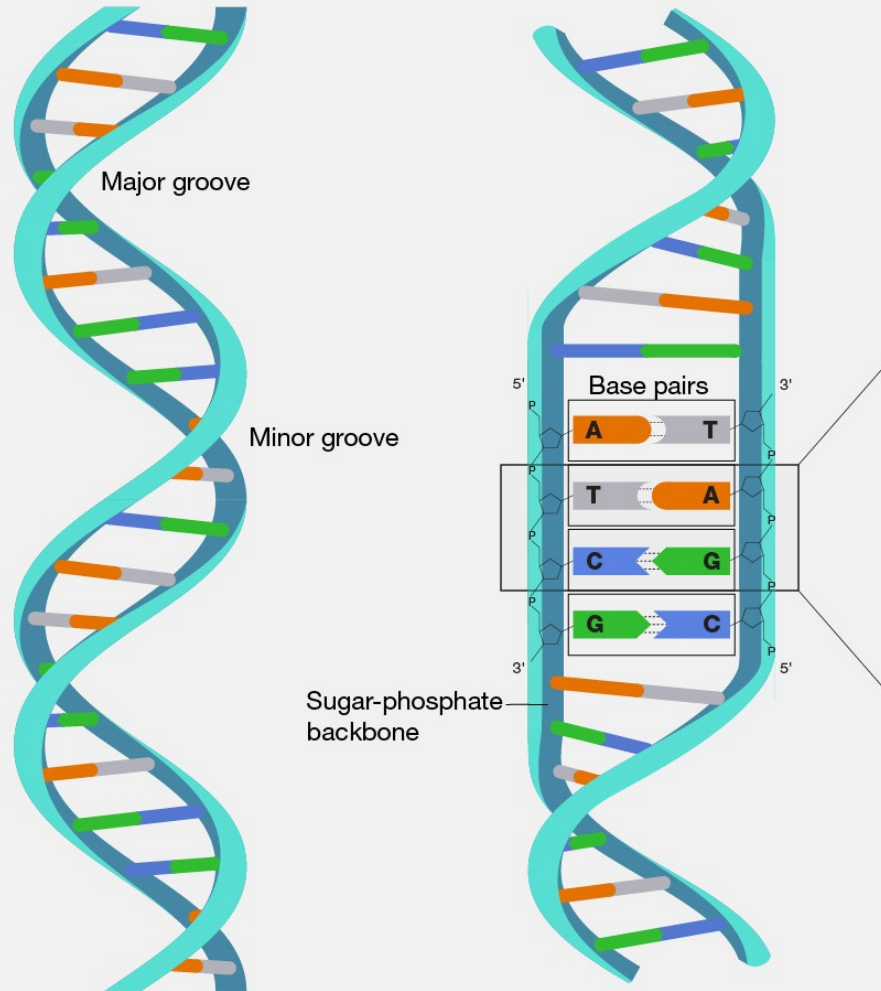


Guanine
G

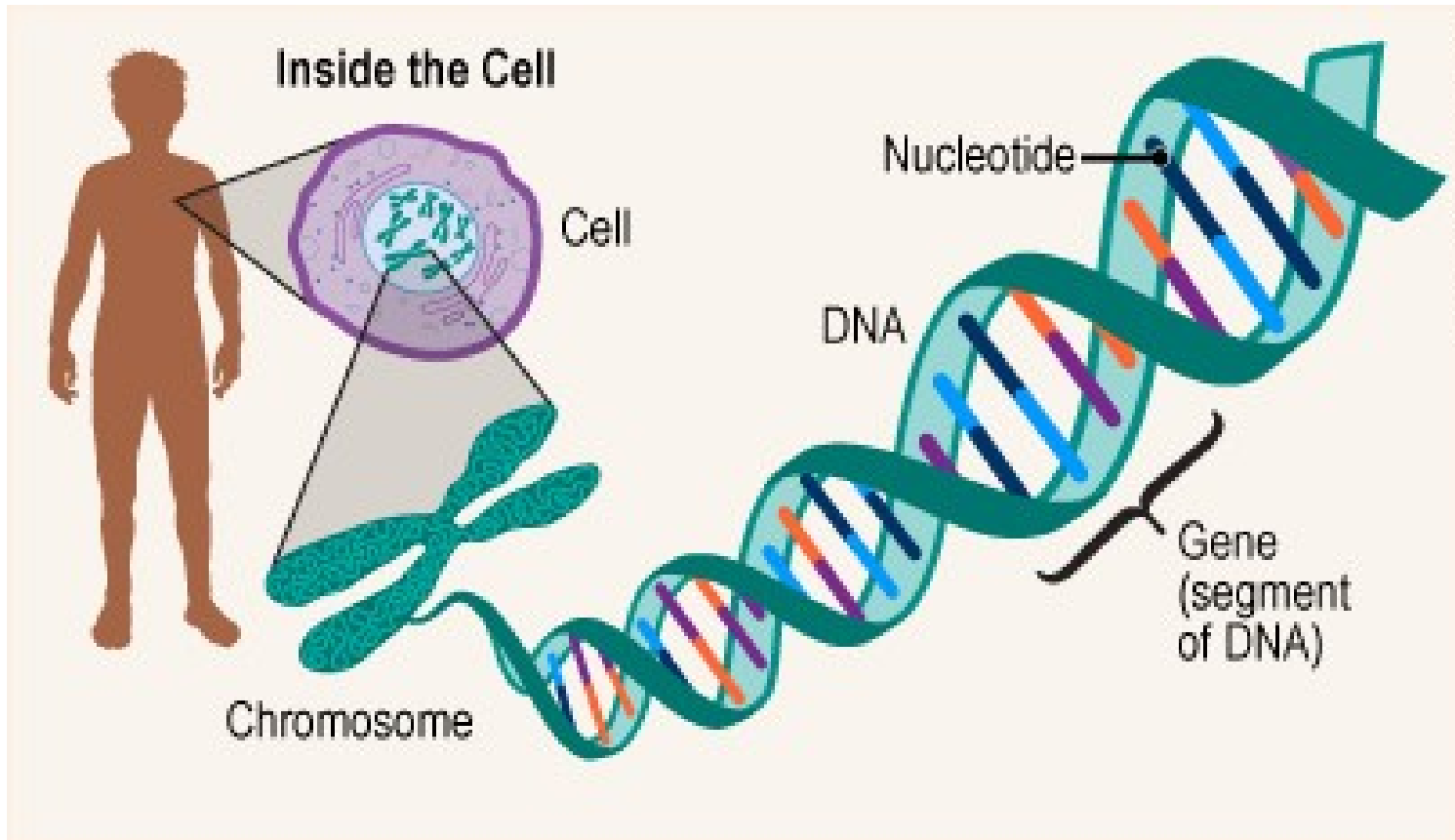


Adenine
A

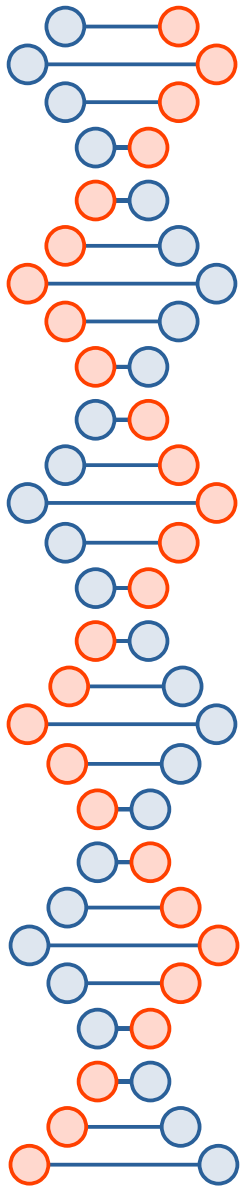
Deoxyribonucleic acid (DNA)



Genes are made of DNA



KidsHealr® All rights reserved.



Genomics

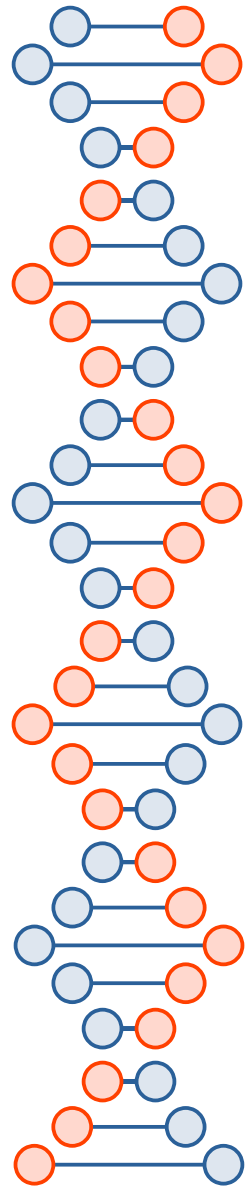
- The study of an organism's complete set of genetic information.
- The genome includes both genes (coding) and non-coding DNA.
- 'Genome': the complete genetic information of an organism.

VS



Genetics

- The study of heredity
- The study of the function and composition of single genes.
- 'Gene': specific sequence of DNA that codes for a functional molecule.

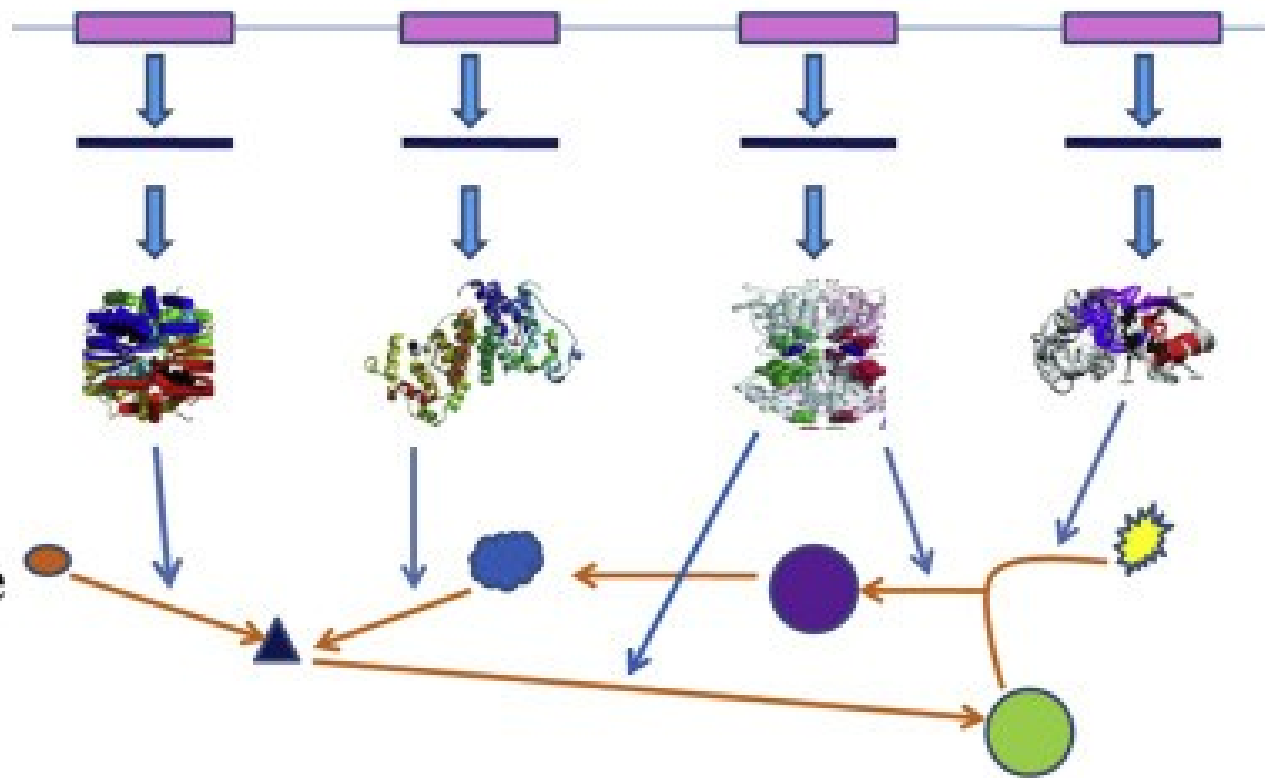


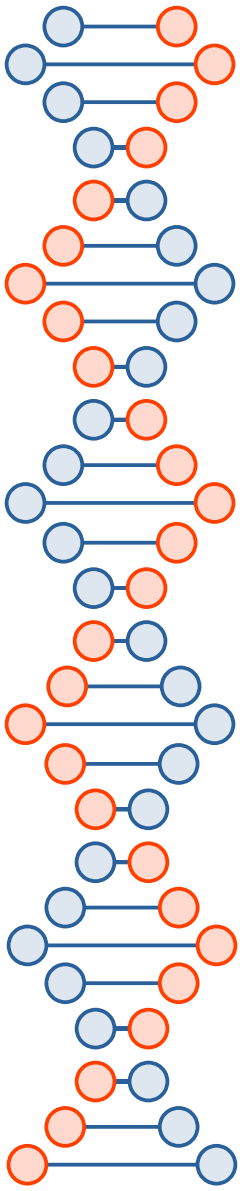
Genome

Expression

Proteome

Metabolome





Big Data!

- The genome consists of ~3 billion base pairs! That's a 3.5 gigabyte text file.
- There are >45k genes, and ~20k protein coding genes
- Enter bioinformatics! We can use computational methods to analyze this data and develop hypotheses, form predictions, and make discoveries



- We will use Rosalind to learn the basics of Python

ROSA LIND | Variables and ...

https://rosalind.info/problems/ini2/

ROSA LIND About Problems Statistics Glossary search f t My Classes yoyom Log out

Variables and Some Arithmetic solved by 42055

Dec. 7, 2012, 12:42 p.m. by Rosalind Team Topics: Introductory Exercises Programming

Variables and Some Arithmetic [click to expand](#)

Problem

Given: Two positive integers a and b , each less than 1000.

Return: The integer corresponding to the square of the hypotenuse of the right triangle whose legs have lengths a and b .

Notes:

1. The dataset changes every time you click "Download dataset".
2. We check only your final answer to the *downloaded* dataset in the box below, not your code itself. If you would like to provide your code as well, you may use the upload tool. Please also note that the correct answer to this problem will not in general be 34; it is simply an example of what you should return in the specific case that the legs of the triangle have length 3 and 5.

Sample Dataset

3 5

Sample Output

34

Time limit You'll have 5 minutes to upload the answer.

Congratulations You solved this problem (attempt #1). Now you may like to try problem "Strings and Lists".

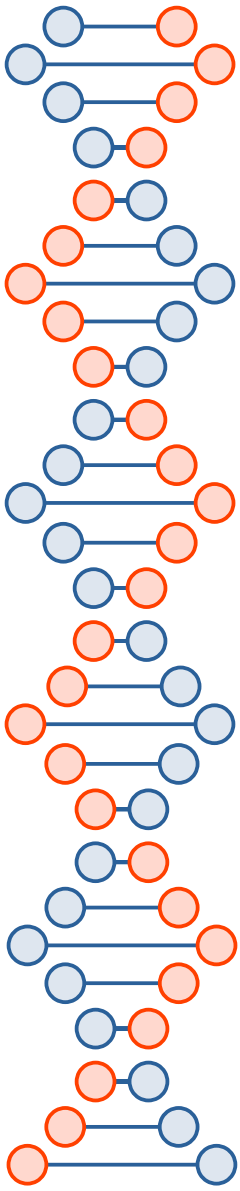
Download dataset You may make an unlimited number of attempts without being penalized.

Questions Solutions Explanation

Attempts

#	Begin Time	Submit Time	Result	Dataset	Submission	Code
1	July 14, 2023, 7 p.m.	July 14, 2023, 7:01 p.m.	Correct	Ⓢ	Ⓢ	—

Found a typo? Suggest a new problem Take a tour



Expectations

- Be excited! We will first start with learning the tools of bioinformatics and analyzing data, but soon you will be working with real-world data and making scientific discoveries!
- I will challenge you to think deeply and broadly to learn this topic. Programming is learned through practice, not lectures!
- By the end of this course, I hope that you will be exposed to scientific literature, medical research, and computer science in the context of biology. You will get out what you put in.