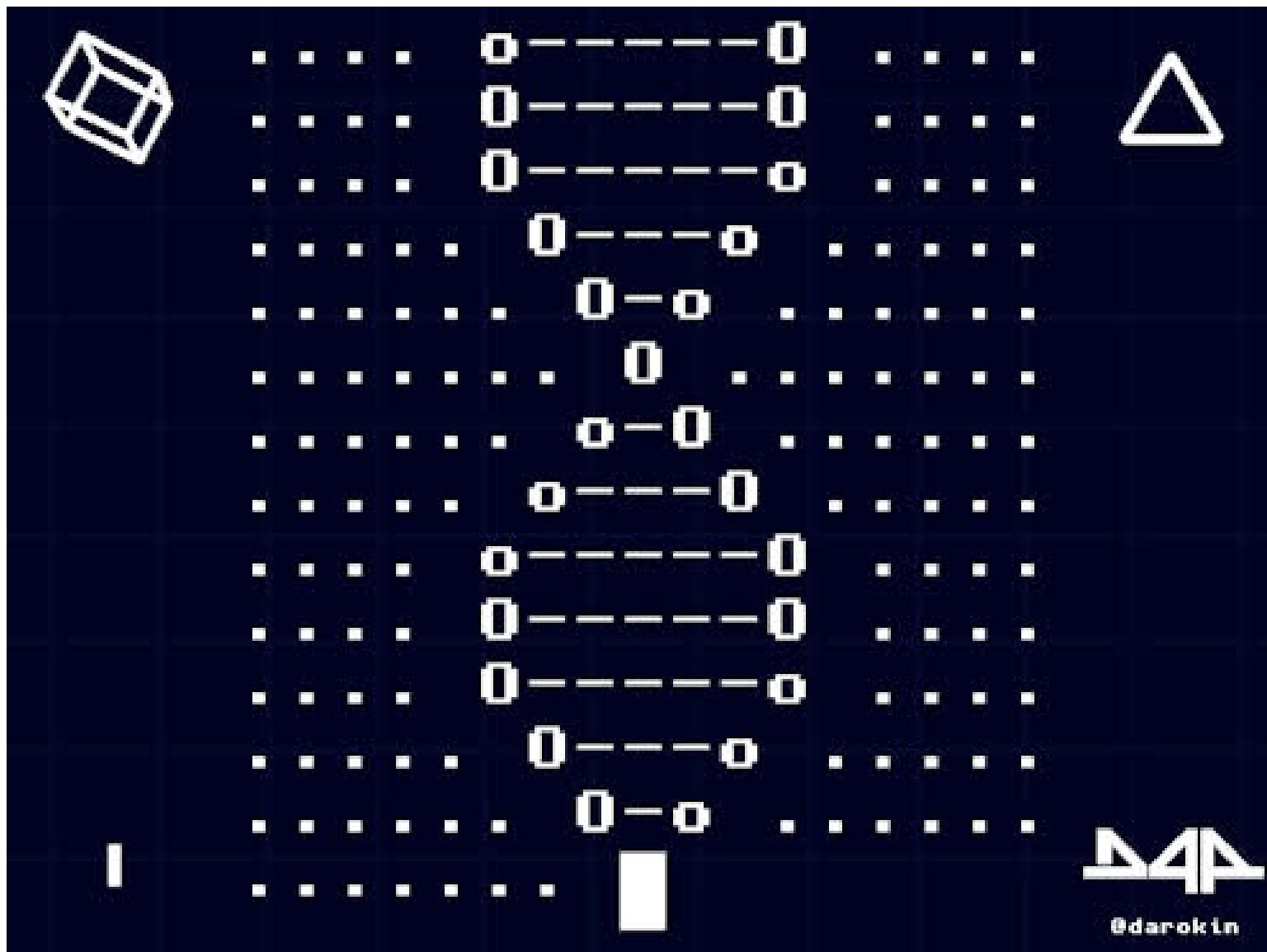


BY SO MANY
SIGNIFICANT
OTHERS

AND ME.

B
I
O
I
N
F
O
R
M
A
N
T
I
N
T
R
O
D
U
C
T
I
O
N
S



Presentation Overview

02

KEY TOPICS

- Biological data
- in vivo*, *in vitro*, and *in silico* processes
- What can bioinformatics do for you?
- Integration of various subjects in this realm
- Data stored in various databases and its types
- Some examples of databases

Biological data

03

```
cgactacgagaaccgtttcacgttgctggtggtggttctcgcttcaggatgtgc  
D Y E N R F Y V A G G G G F S L S S D A  
actggcggtctcgagtagcggtgctgcctcggtgccaccgatgccggctacgacgccttc  
T G G L E Y G A A S G A T D A G Y D A F  
tcctacagctacgacgaagtgcgtccatatggcaacggctccatcaactggatgccact  
S Y S Y D E V V L Y G N G S I N W D A T  
tacatgttggctaccaggctctgggtgaaatgaccaagatcgccaagcccctgaccggt  
Y M F G Y Q A L G E M T K I A K P L T R  
ggcttctacggcctccagcacaagaagatctacacctactacgagggctgtccgat  
G F Y G L S S D K K I Y T Y Y E G C S D  
ggtgtcggtatgagtcaggttcagcgctggggagatgaatatgacggtgttatac  
G G R E G M S Q V Q R W G D E Y D G V I  
gctggtgccctgcctccgcttgcagcagcagttcaccacgtttccctgccacc  
A G A P A F R F A Q Q Q V H H V F P A T
```

**WHAT IS
BIOLOGICAL DATA?**

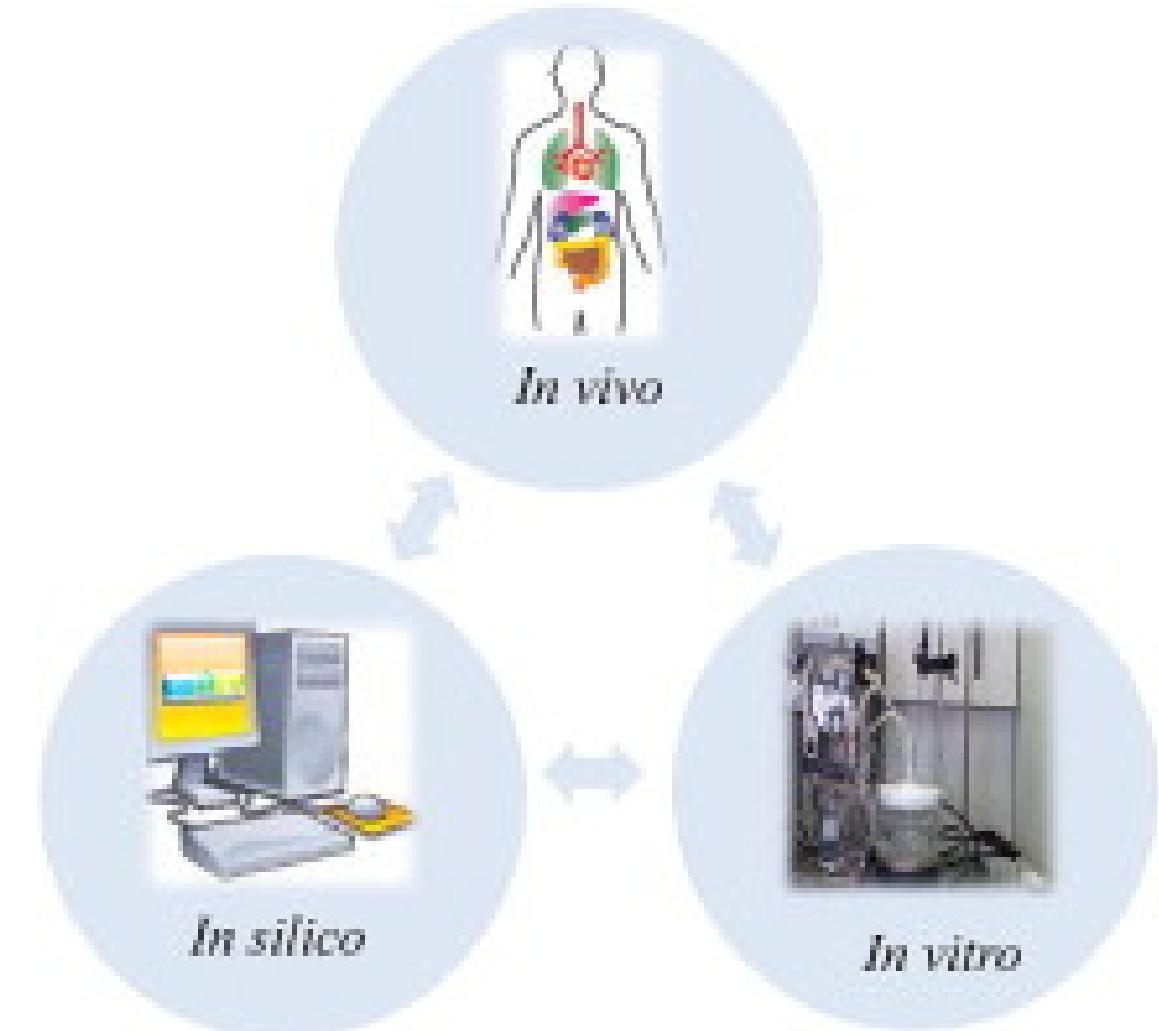
**HOW DO WE OBTAIN
BIOLOGICAL DATA?**

04

in vivo: inside the organism.

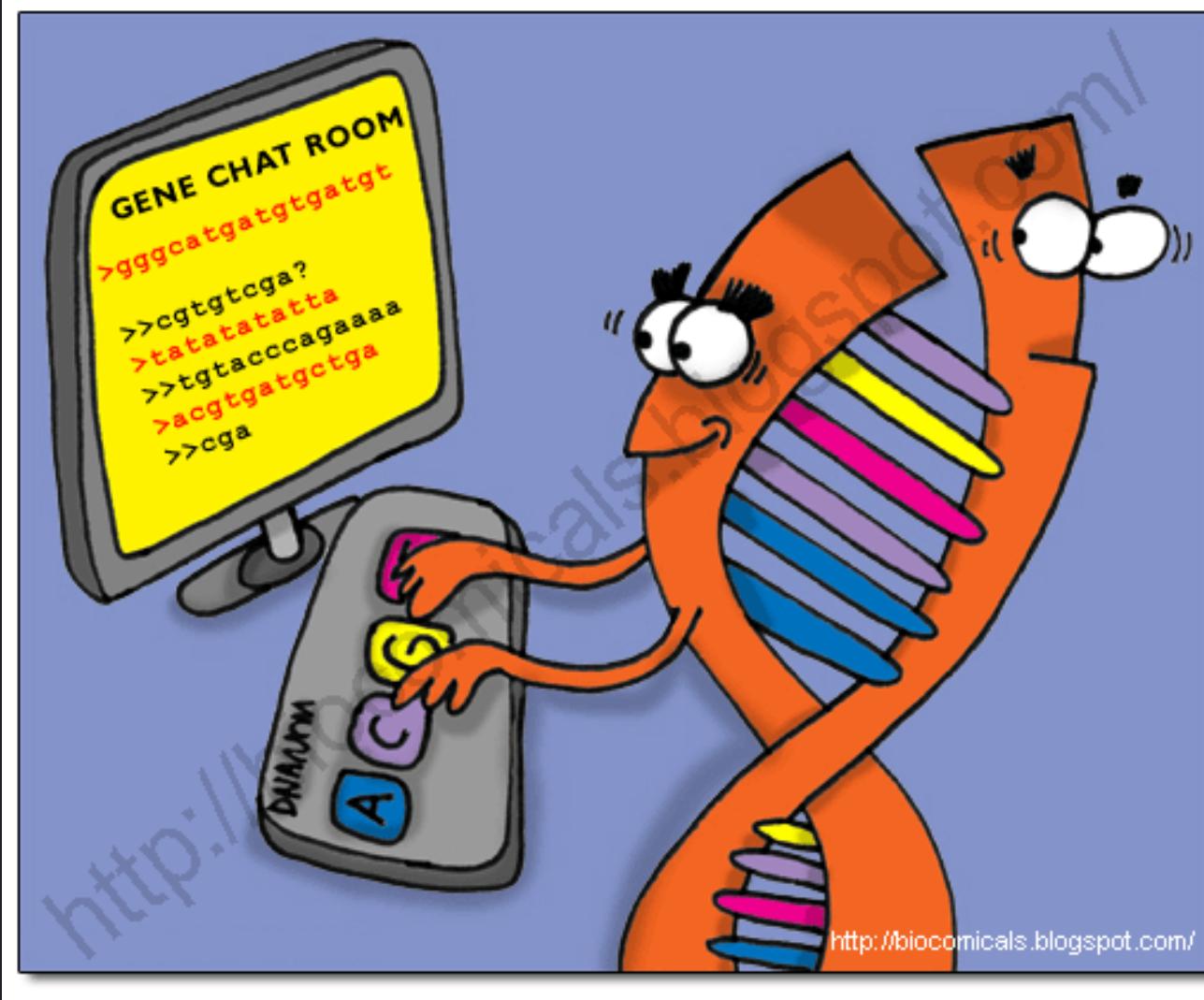
in vitro: in a false environment, mimicking the original.

in silico: inside a computer



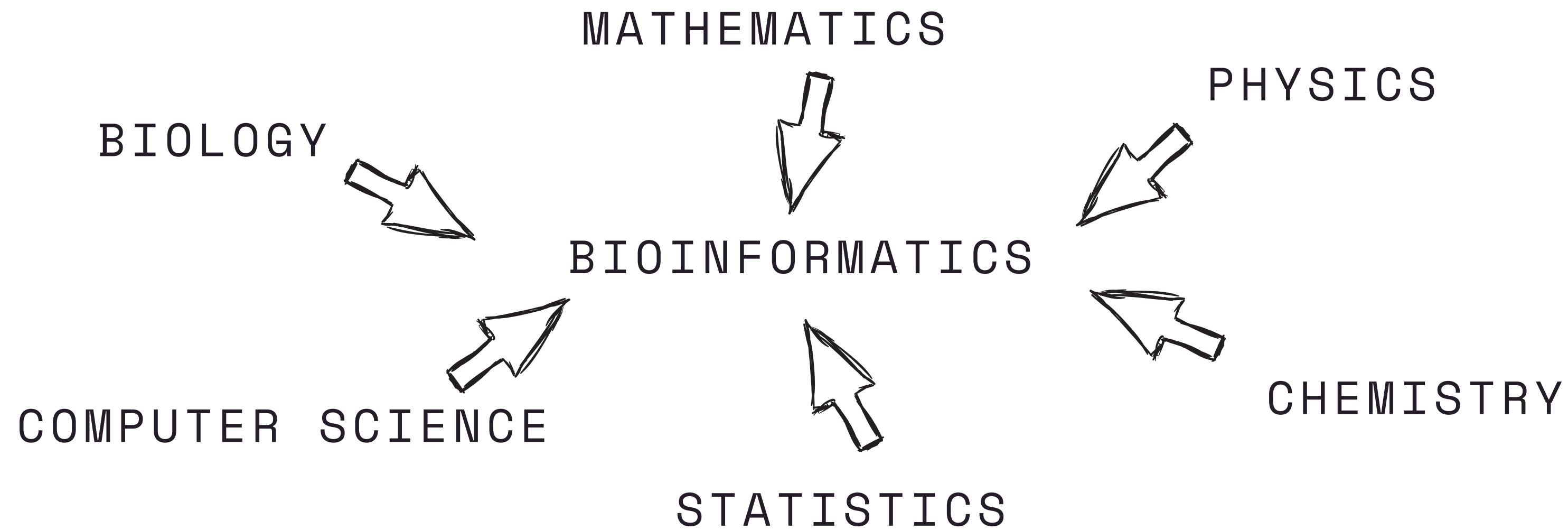
What can bioinformatics do for you?

05



1. Sequence analysis
2. Sequence comparison
3. Structural analysis
4. Finding the family and domain of a protein
5. Showing 3D structures of a protein
6. Nucleotide and protein sequencing

When the witch stirs the cauldron!



TYPES OF BIOLOGICAL DATABASES

1. Primary Databases
2. Composite Databases
3. Literature Databases
4. Structural Databases
5. Secondary/Derived
Databases
6. Protein Sequence
Databases
7. Genome Databases

1. Different kinds of data are stored in a database, eg nucleotide sequences, protein sequences, organisms they are retrieved from, etc.
2. Data can be visualized, annotated, and worked upon from a database.
3. Databases are updated regularly by the founder organization.

- Primary Databases - Nucleotide sequence databases, eg: GenBank, DDBJ, EMBL
- Composite Databases - Composite databases have been created to simplify the sequence search for a query in a single compilation in the context of the many different primary database searches, by merging a variety of different primary resources. These databases are non-redundant and render sequence searching, eg: OWL
- Literature Databases - Various publications can be found here, eg: PLoS, Biomed Central, PubMed
- Structural Databases - Protein structure databases, eg: PDB
- Secondary/Derived Databases - Contains information from the primary databases along with some specific data, eg: SWISS-PROT, PIR
- Protein Sequence Databases - Contains protein sequences, eg: Uni-Prot
- Genome Databases - Contains genomic sequences of various organisms, eg: ICTV, MBGD