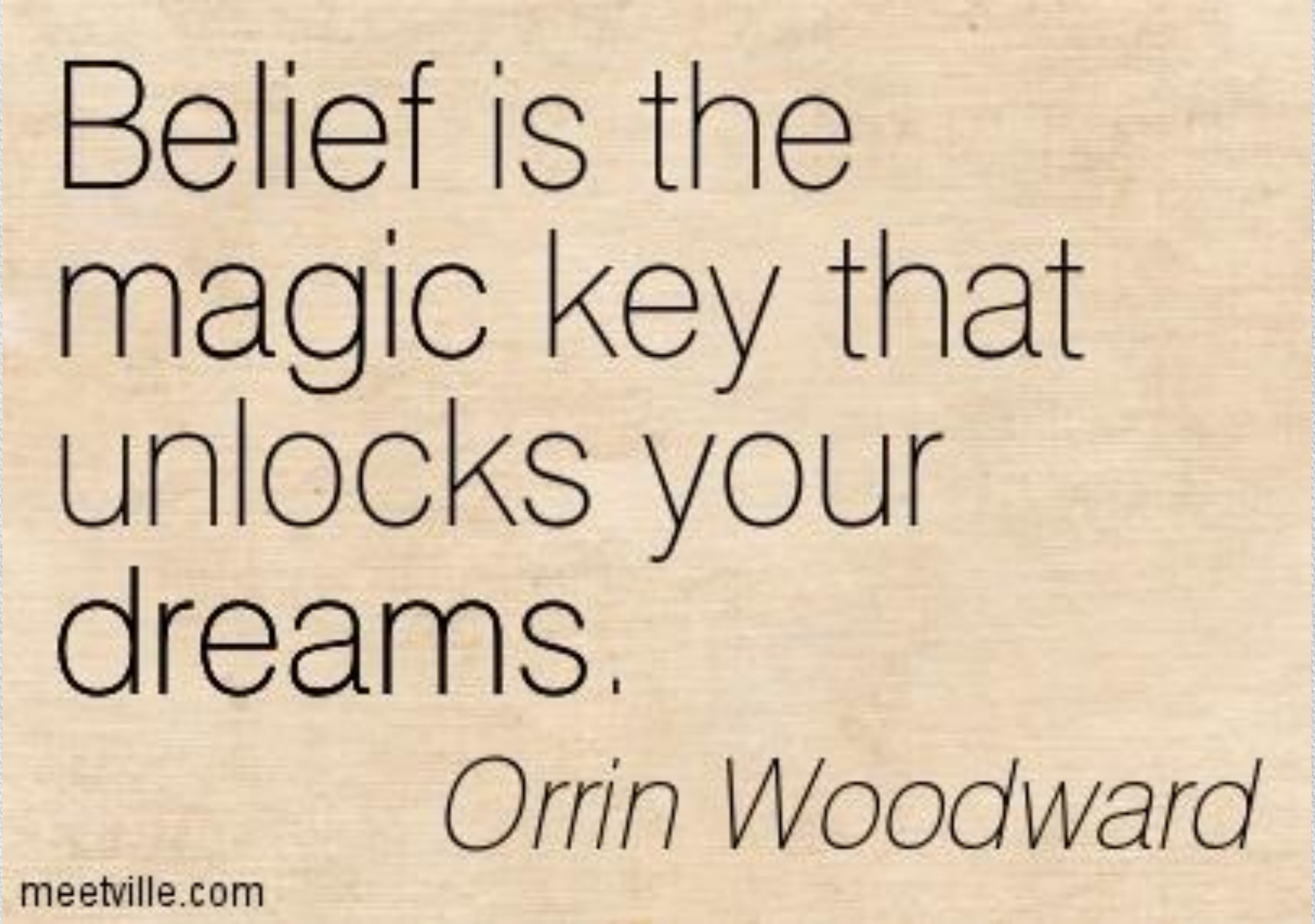


Belief is the magic key that unlocks your dreams.

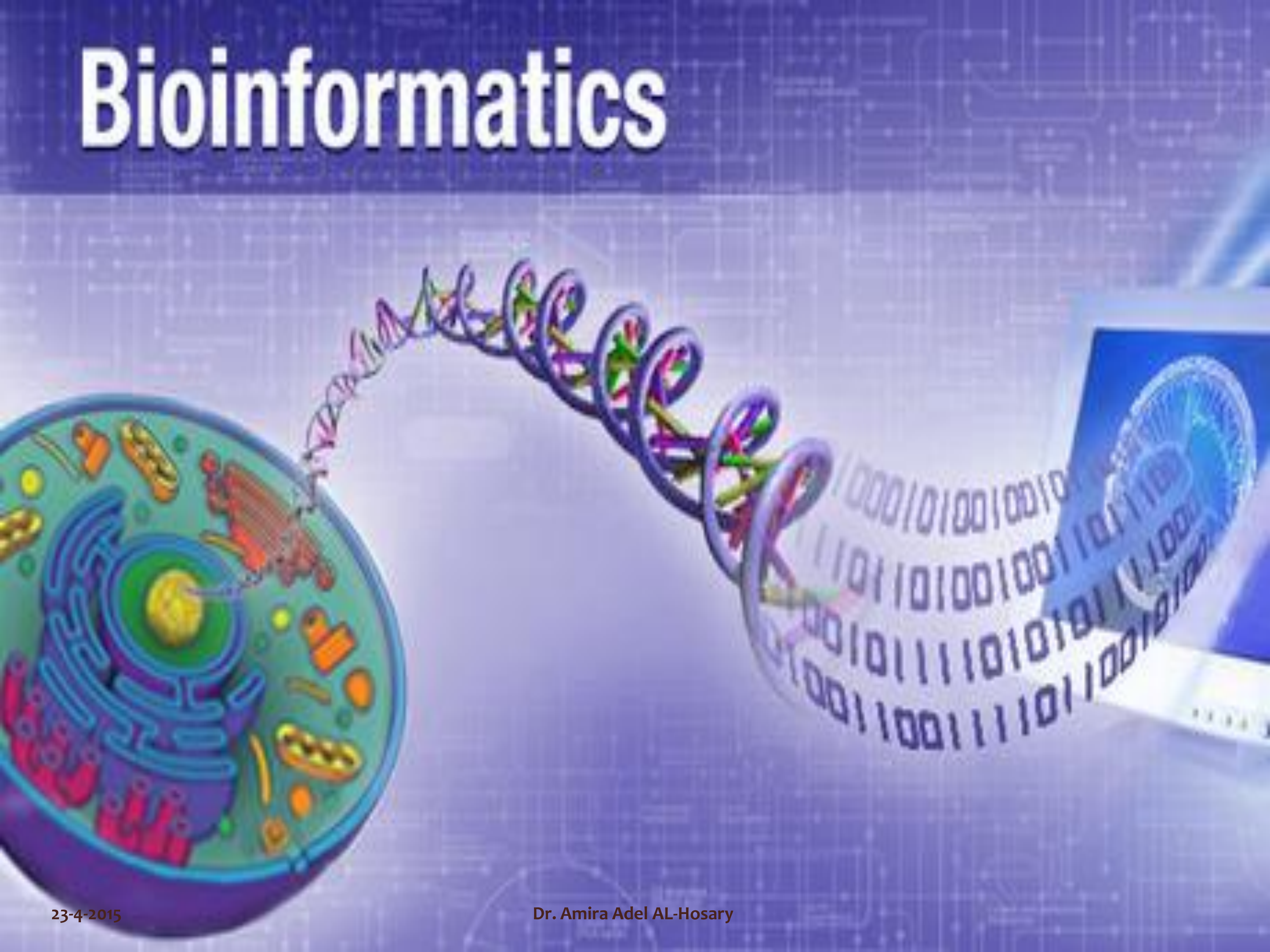
Orrin Woodward

meetville.com

Molecular Biology Research Unit

INTRODUCTION TO BIOINFORMATICS

Amira A. AL-Hosary
PhD of infectious diseases
Department of Animal Medicine
(Infectious Diseases)
Faculty of Veterinary Medicine
Assiut University-Egypt


Bioinformatics

Most biological research involves application of some type of mathematical, statistical, or computational tools to help synthesize recorded data and integrate various types of information in the process of answering a particular biological question.

Bioinformatics

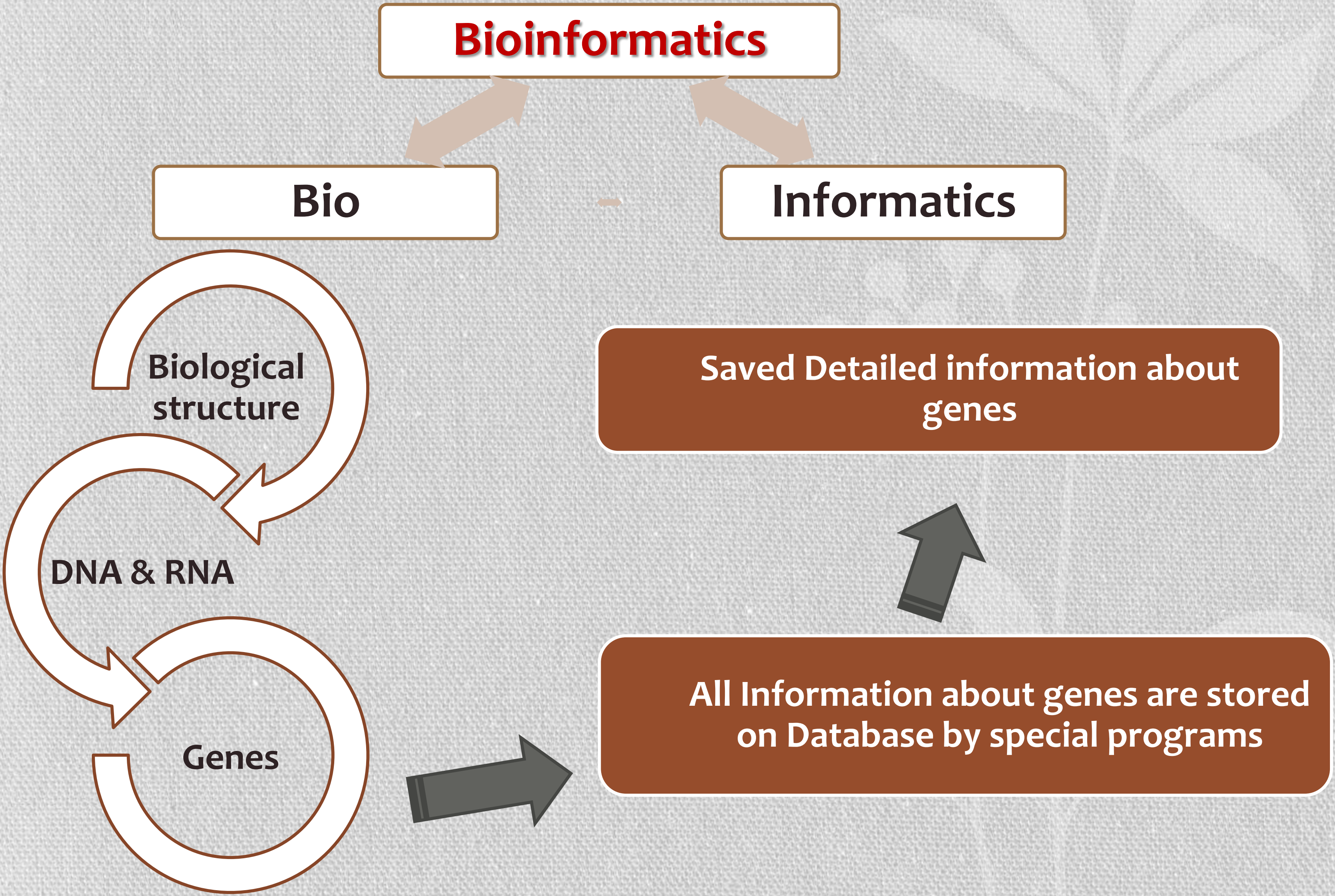
Bioinformatics *VS computational biology*

Bioinformatics differs from a related field known as *computational biology*.

Bioinformatics is limited to sequence, structural, and functional analysis of genes and genomes and their corresponding products and is often considered computational molecular biology.

However,

computational biology encompasses all biological areas that involve computation.



Bioinformatics

Bio

Informatics

Biological structure

DNA & RNA

Genes

All Information about genes are stored on Database by special programs

Saved Detailed information about genes

Bioinformatics

Bioinformatics:

Is the discipline of quantitative analysis of information relating to biological macromolecules with the aid of computers.

The development of bioinformatics as a field is the result of advances in both molecular biology and computer science over the past 40–50 years.

The earliest bioinformatics efforts can be traced back to the 1960s, although the word *bioinformatics* did not exist then.

- ✓ **Margaret Dayhoff** in 1965 developed a first protein sequence database called *Atlas of Protein Sequence and Structure.*
- ✓ **Brookhaven National Laboratory** in the early 1970s, established the Protein Data Bank for archiving three-dimensional protein structures.
- ✓ **Needleman** and **Wunsch** in 1970 developed the first sequence alignment algorithm.

Sequence alignment:

This was a fundamental step in the development of the field of bioinformatics, which paved the way for the routine sequence comparisons and database searching practiced by modern biologists.

Term Bioinformatics was invented by Paulien Hogeweg and Ben Hesper in 1970

Paulien Hogeweg is a Dutch theoretical biologist and complex systems researcher studying biological systems as dynamic information processing systems at many interconnected levels.

- ✓ **Stephen Altschul** and **coworkers in** The 1980s establishment of GenBank and the development of fast database searching algorithms such as **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool)
- ✓ **The start of the human genome project in the late 1980s** provided a major boost for the development of bioinformatics.
- ✓ **The development and the increasingly widespread use of the Internet in the 1990s** made instant access to, and exchange and dissemination of, biological **data** possible.

Bioinformatics:

Uses many areas of computer science, statistics, mathematics and engineering to process biological data.

1- Complex machines are used to read in biological data at a much faster rate than before.

2- Databases and information systems are used to store and organize biological data.

3- Analyzing biological data may involve algorithms in artificial intelligence, soft computing, data mining, image processing, and simulation.

GOALS

The ultimate goal of bioinformatics is to better understand a living cell and how it functions at the molecular level.

By analyzing raw molecular sequence and structural data, bioinformatics research can generate new insights and provide a “global” perspective of the cell.

SCOPE

Bioinformatics consists of two subfields:

1-The development of computational tools and databases.

2-The application of these tools and databases in generating biological knowledge to better understand living systems.

These two subfields are complementary to each other.

APPLICATIONS:

Bioinformatics has not only become essential for basic genomic and molecular biology research, but is having a major impact on many areas of biotechnology and biomedical sciences.

1- Medical uses:

The high speed genomic sequencing coupled with sophisticated informatics technology will allow a doctor in a clinic to quickly sequence a patient’s genome and easily detect potential harmful mutations and to engage in early diagnosis and effective treatment of diseases.

Molecular Medicine:

Every disease has a genetic component. This may be inherited as hereditary disease or occurs as a result of the body's response to an environmental stress which causes alterations in the genome.

Bioinformatics help us to search for the genes directly associated with those diseases and understand their molecular basis more clearly. This will enable better treatments, cures and even preventative tests to be developed.

Personalized medicine: Clinical medicine will become more personalized with the development of the field of **pharma-co-genomics**. This is the study of how an individual's genetic inheritance affects the body's response to drugs.

Preventative medicine: With the specific details of the genetic mechanisms of diseases occurrence and development of diagnostic tests will help us to measure a person's susceptibility to different diseases.

Preventative actions such as change of lifestyle or having treatment at the earliest possible stages when they are more likely to be successful, could result in huge advances in our struggle to conquer disease

2-Forensic Medicine and DNA analysis.

In forensics, results from molecular phylogenetic analysis have been accepted as evidence in criminal courts.

3-Agricultural biotechnology:

Plant genome databases and gene expression profile analyses have played an important role in the development of new crop varieties that have higher productivity and more resistance to disease.

4-Pharmaceutical and Drug design:

Computational studies of protein–ligand interactions provide a rational basis for the rapid identification of novel leads for synthetic drugs.

5-Agricultural biotechnology:

Plant genome databases and gene expression profile analyses have played an important role in the development of new crop varieties that have higher productivity and more resistance to disease.

6-Pharmaceutical and Drug design:

Computational studies of protein–ligand interactions provide a rational basis for the rapid identification of novel leads for synthetic drugs.

7-Knowledge of the three-dimensional structures of proteins:

Allows molecules to be designed that are capable of binding to the receptor site of a target protein with great affinity and specificity.

This informatics-based approach significantly reduces:

- **The time and cost necessary to develop drugs with higher potency.**
- **Reduces the side effects and toxicity when compared with traditional trial-and-error approach.**

8-Gene therapy: The potential for using genes themselves to treat disease may become a reality. Gene therapy is the approach used to treat, cure or even prevent disease by changing the expression of a persons genes.

9-Microbial genome applications: Microorganisms are found every where The arrival of the complete genome sequences and their potential to provide a greater insight into the microbial world and its capacities.

10. Bio-weapon creation:

11. Waste cleanup: *Deinococcus radiodurans* is known as the world's toughest bacteria and it is the most radiation resistant organism known. It is an extremophilic bacterium, one of the most radiation-resistant organisms known. It can survive cold, dehydration, vacuum, and acid, and is therefore known as a polyextremophile. So, it will be usefulness in cleaning up waste sites that contain radiation and toxic chemicals.

12. Climate change Studies: Increasing levels of carbon dioxide emission, mainly through the expanding use of fossil fuels for energy, are thought to contribute to global climate change. One method of doing so is to study the genomes of microbes that use carbon dioxide as their sole carbon source.

13. Alternative energy sources

Scientists are studying the genome of the microbe *Chlorobium tepidum* which has an unusual capacity for generating energy from light.

14. Bio-Technology

Some M.Os. have potential for practical applications in industry and environmental remediation. Those M.Os. thrive in water temperatures above the boiling point therefore it may provide knowledge of heat-stable enzymes suitable for use in industrial processes. ***Corynebacterium glutamicum*** which is of high industrial interest as a research object because it is used by the chemical industry for the biotechnological production of the amino acid lysine.

15. Antibiotic resistance:

Enterococcus faecalis causes infection among hospital patients. Studying of its genome lead to discover of a virulence region made up of a number of antibiotic-resistant genes which known as a **pathogenicity island**

16. Evolutionary studies.

17. Veterinary Medicine: Sequencing projects of many farm animals under work for improving the production and health of livestock and ultimately have benefits for human nutrition.

18- Comparative Studies Analyzing and comparing the genetic material of different species is an important method for studying the functions of genes.

Limitation

After recognized the power of bioinformatics, it is also important to realize its limitations and avoid over-reliance on and over-expectation of bioinformatics output.

In many ways, the role of bioinformatics in genomics and molecular biology research can be likened to the role of intelligence gathering in battlefields.

- Intelligence is clearly very important in leading to victory in a battlefield.
- Fighting a battle without intelligence is inefficient and dangerous.
- Having superior information and correct intelligence helps to identify the enemy's weaknesses and reveal the enemy's strategy and intentions.

- The gathered information can then be used in directing the forces to engage the enemy and win the battle.
- However, completely relying on intelligence can also be dangerous if the intelligence is of limited accuracy.
- Overreliance on poor-quality intelligence can yield costly mistakes if not complete failures.

- It is no stretch in analogy that fighting diseases or other biological problems using bioinformatics is like fighting battles.
- Bioinformatics and experimental biology are not independent, but complementary, activities.
- Bioinformatics depends on experimental science to produce raw data for analysis.

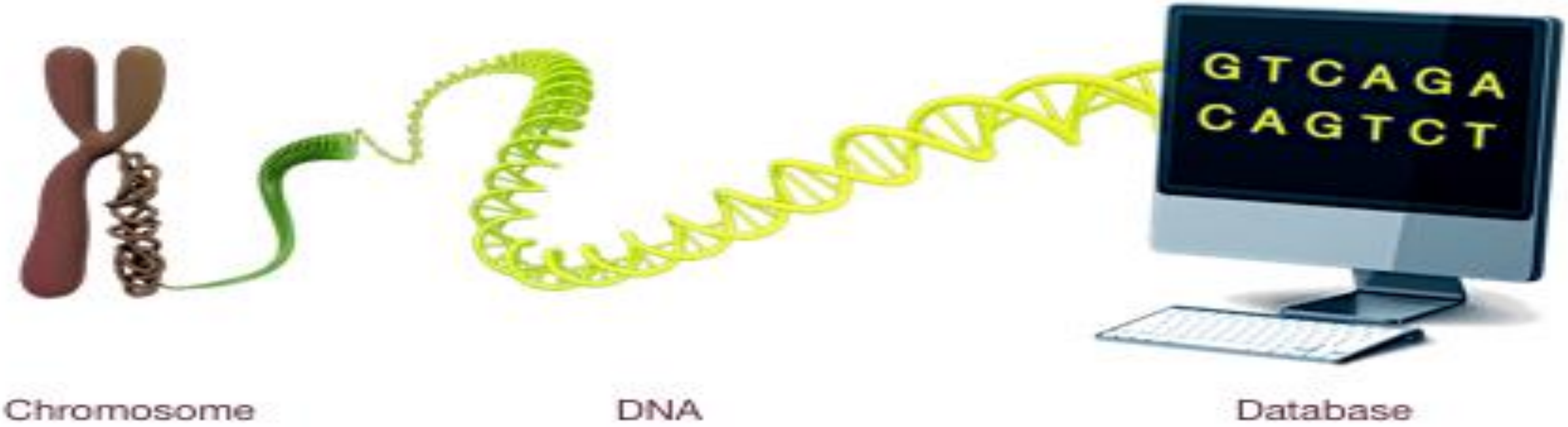
- Bioinformatics predictions are not formal proofs of any concepts.
- They do not replace the traditional experimental research methods of actually testing hypotheses.
- In addition, the quality of bioinformatics predictions depends on the quality of data and the sophistication of the algorithms being used (Sequence data often contain errors).
- If the sequences are wrong or annotations incorrect, the results from the downstream analysis are misleading as well.

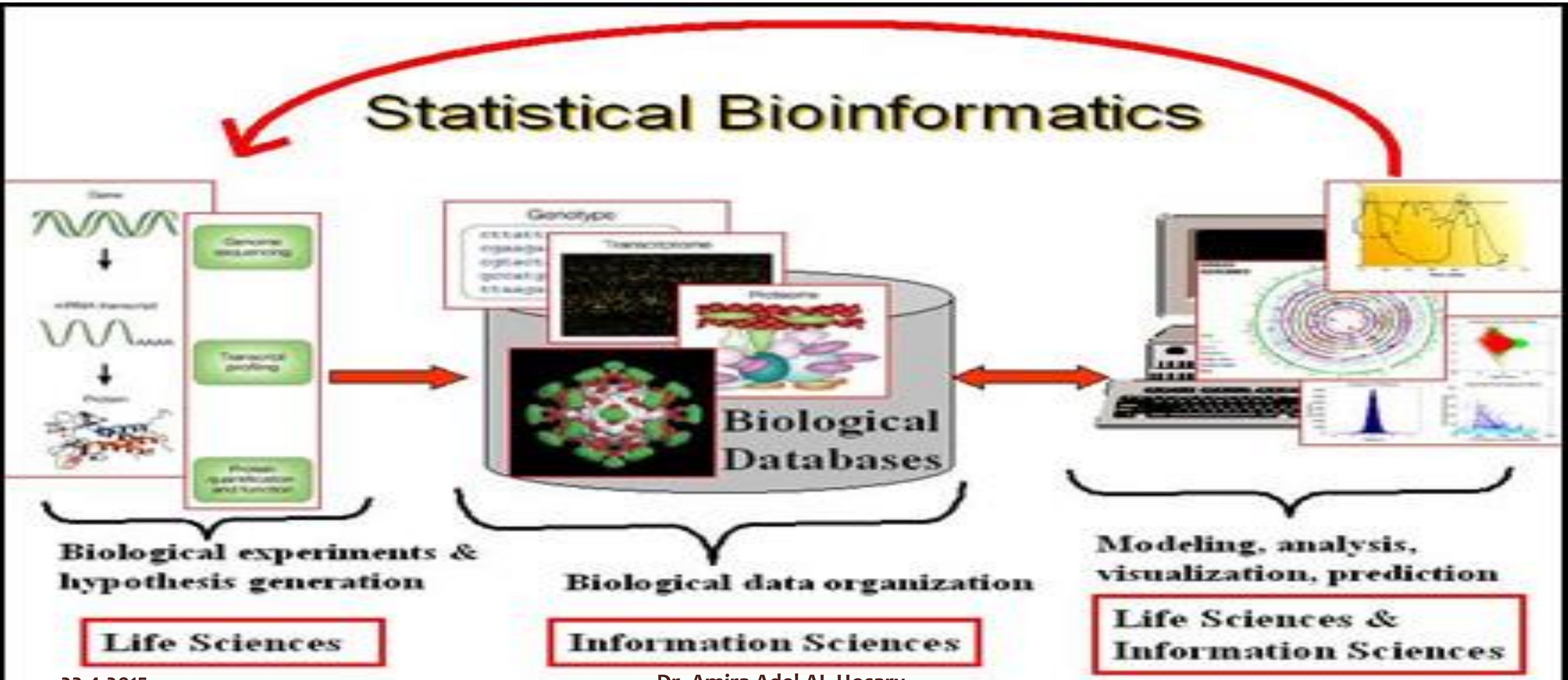
- The outcome of computation also depends on the computing power available.
- Many accurate but exhaustive algorithms cannot be used because of the slow rate of computation.

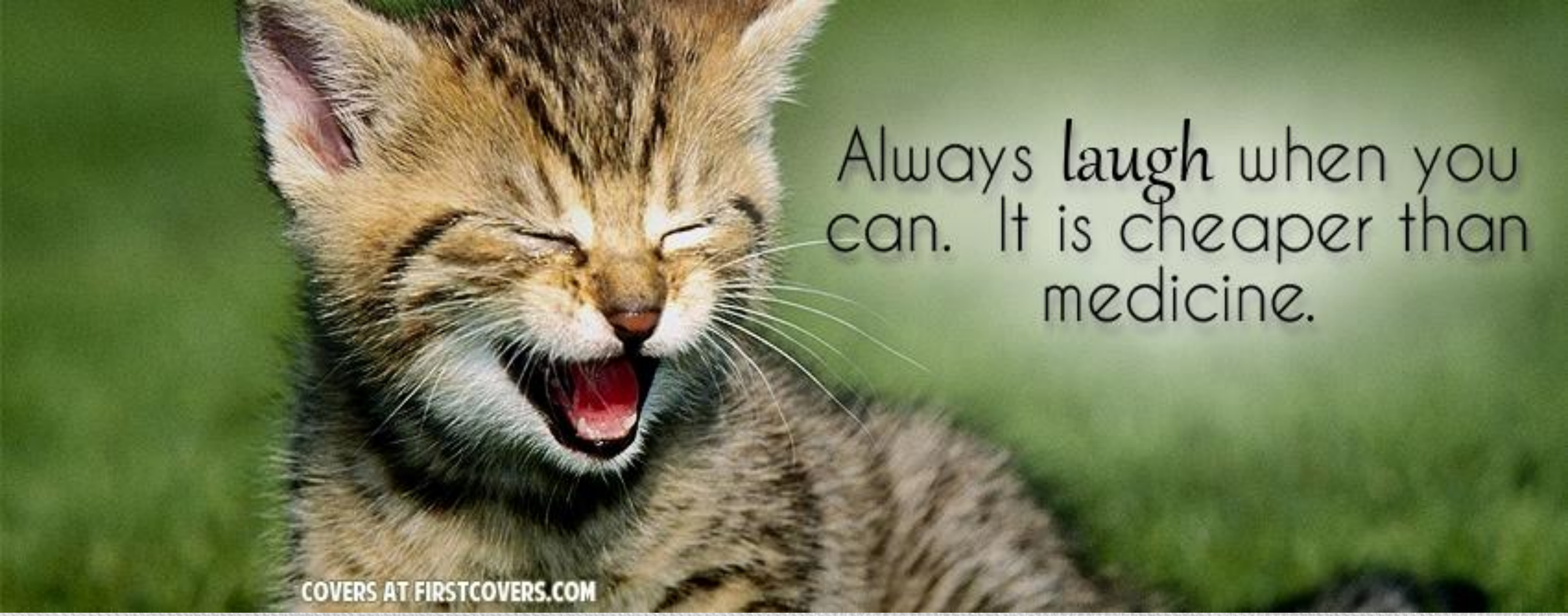
SCOPE

Caution

- You should always be exercised when interpreting prediction results.
- It is a good practice to use multiple programs, if they are available, and perform multiple evaluations.
- More accurate prediction can often be obtained if one draws a consensus by comparing results from different algorithms.


GTCAGA
CAGTCT
Chromosome
DNA
Database


Statistical Bioinformatics
Gene
Genome sequencing
Transcript profiling
Genotype
Transcriptome
Proteome
Biological Databases
Biological experiments & hypothesis generation
Biological data organization
Modeling, analysis, visualization, prediction
Life Sciences
Information Sciences
Life Sciences & Information Sciences



Thanks a lot

with my Best Regards and My Best wishes

Amira A. AL-Hosary
E-mail: Amiraelhosary @yahoo.com
Mob. (002) 01004477501