Bioinformatics
CS300
Introduction to Bioinformatics

Fall 2019
Oliver BONHAM-CARTER

Introduction to Bioinformatics
BIO*300/CMPSC*300

Fall 2019
Class: Tuesday / Thursday, 11:00am - 12:15pm
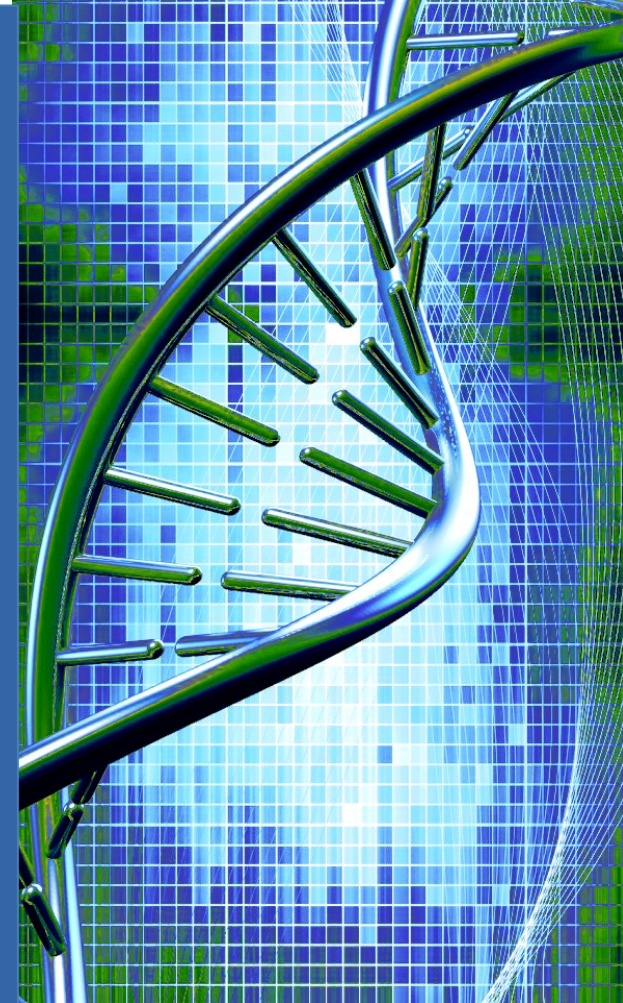**Lab: Monday 2:30-4:20pm**
(Alden Hall 101)

Bioinformatics is an emerging, rapidly expanding interdisciplinary discipline that studies how to effectively integrate concepts from computational sciences and biosciences. There is a high demand for scientists who are versed in both biology and computer science in the biomedical industry and research.

BIO 300/CMPSC 300 students will become familiar with the state-of-the-art bioinformatics software and the algorithms behind them. Through hands-on projects, students will explore current biological problems and will develop bioinformatics solutions to these issues.

This course counts as an Area A Biology Course and an Applications Course for Computer Science.
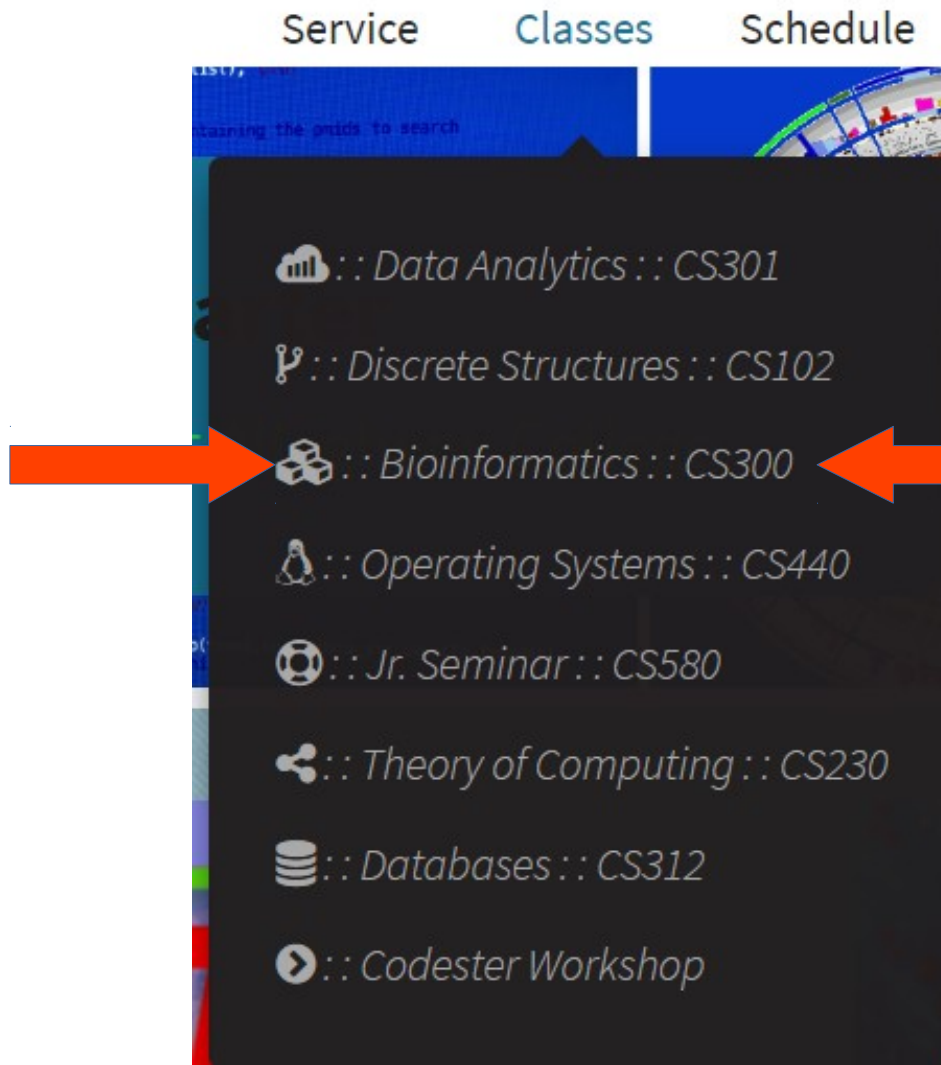
Questions? Contact Dr. BONHAM-CARTER at
obonhamcarter@allegheny.edu

Class WebSite

Find:
- Syllabus
- Meeting place and times
    - Class
    - Lab
- Calander
- First slides
- Etc.

https://www.cs.allegheny.edu/sites/obonhamcarter/

How To Be Successful

- Come to classes!!!

- Come to lab!!!

- Read the book and handouts!

- Keep up with your lab homework!

An Ethical Approach

- We will spend time to learn about responsibly in the bioinformatics field and in its inquiry

- Ethical considerations for research, building tools, implementing algorithms, working with life, etc ...

Bioinformatics is Many Things...

The science of collecting and analyzing complex biological data such as genetic codes.

A theoretical framework to detect genes which contribute to the onset of unhealthy development.

The field of exploration into data to describe the onset of Illness, disease and medical disorder

Is an interdisciplinary field that develops methods and software tools for understanding biological data.

Bioinformatics is both an umbrella term for the body of biological studies that use computer programming as part of their methodology

The study of data from living systems to determine patterns of life and health.

The development of tools to aid in the comparison of genetic and genomic data and more generally in the understanding of evolutionary aspects of molecular biology.

A framework used to determine the relatedness between people, dogs, cats, mice, rats, rabbits, or any living thing!

ALLEGHENY COLLEGE

Bioinformatics:
The Study of DNA

- DNA is the genetic material that houses genetic information.

- Genes are written in this language

- Understanding DNA allows us to understand how genes work.
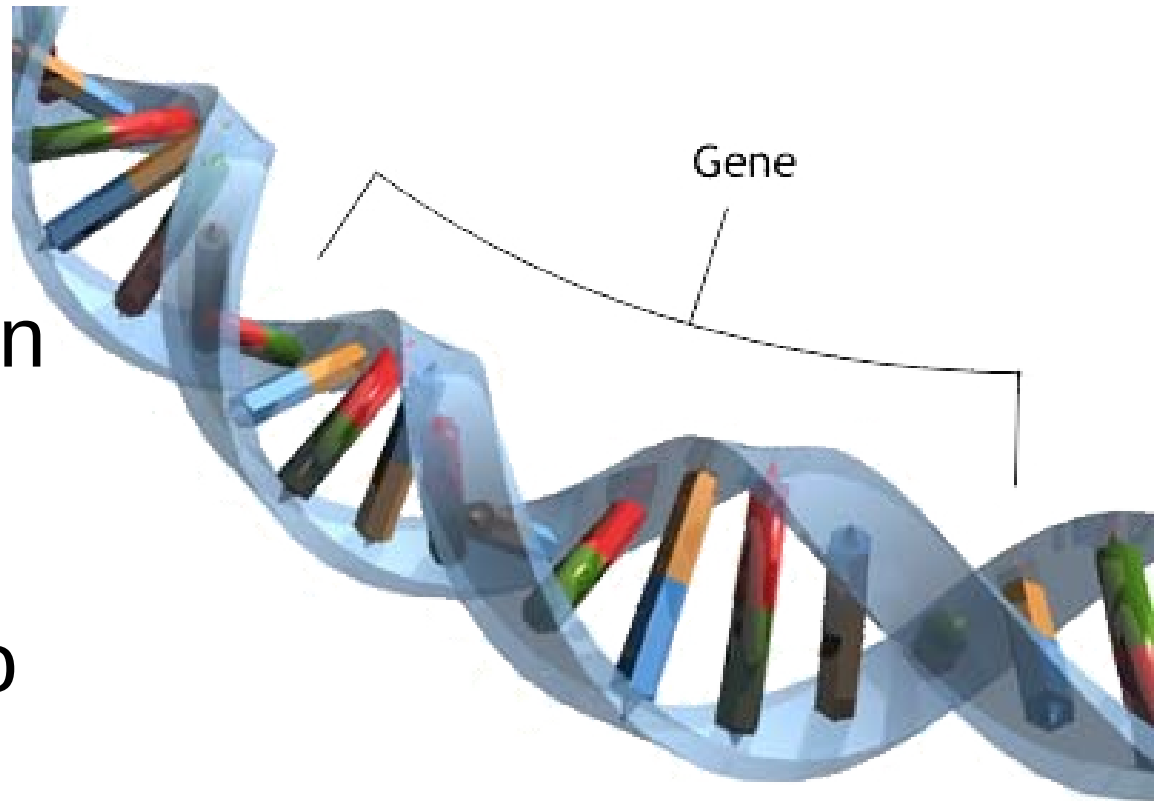
Gene

How Much DNA to Study?

- Gen Bank: Public repository of DNA sequences and related data.

- Seemingly exponential growth in amount of sequence data available for study

Growth of GenBank (1982 - 2008)
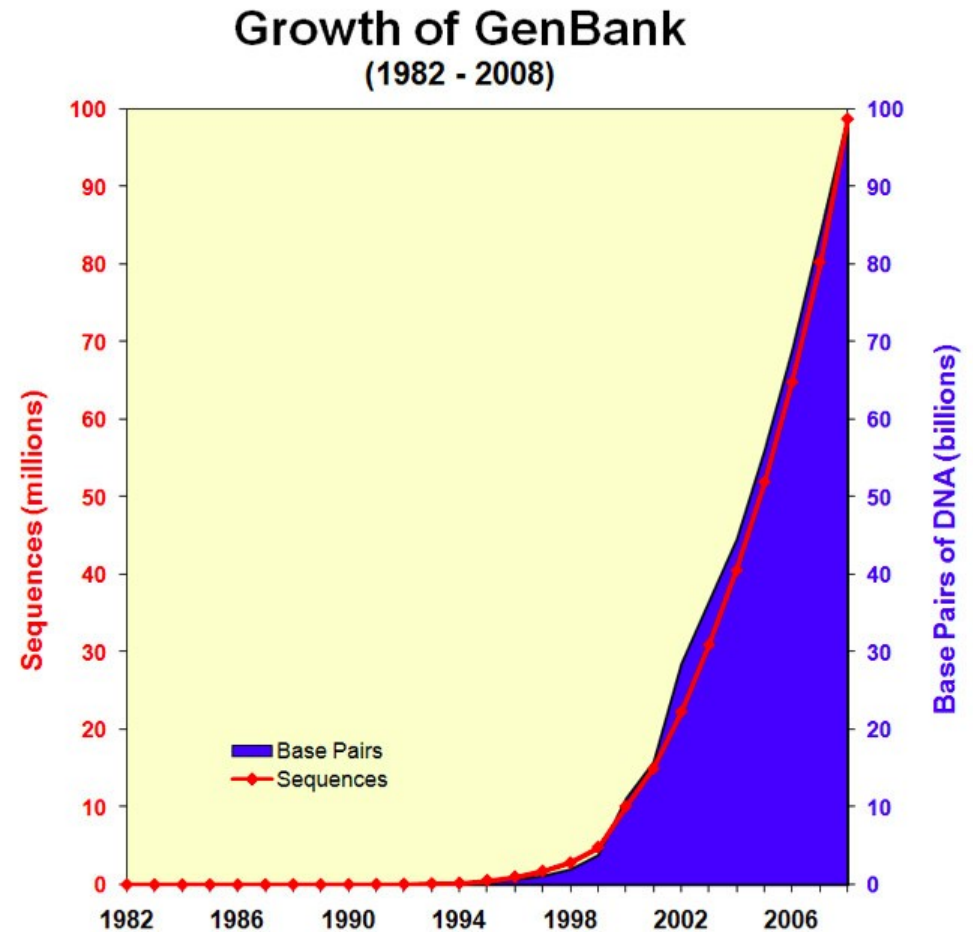
Use DNA to Compare...

- DNA sequences

- Genes

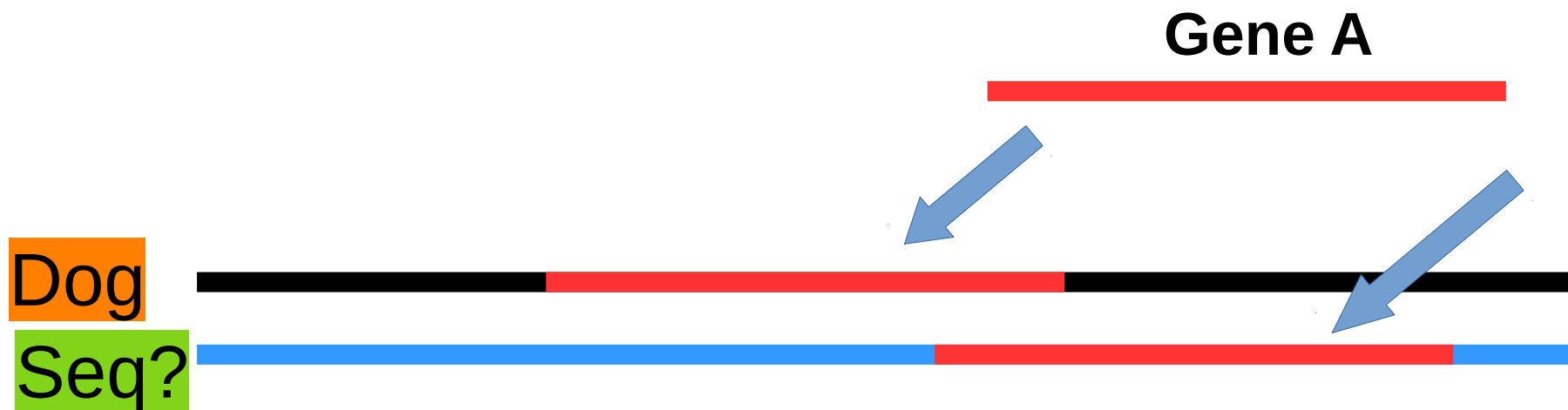- Proteins

- Organisms

- **Why do we compare these things?**

- **What is there to learn when we find that two *things* are the same? Or are <u>not</u> the same?**

Comparing Regions?

- We learn to scan millions of unknown DNA sequences to find familiar genes.

- Does a gene have the same function each time we find it in a sequence? Same origins? Do the same disorders affect the copies of the gene?
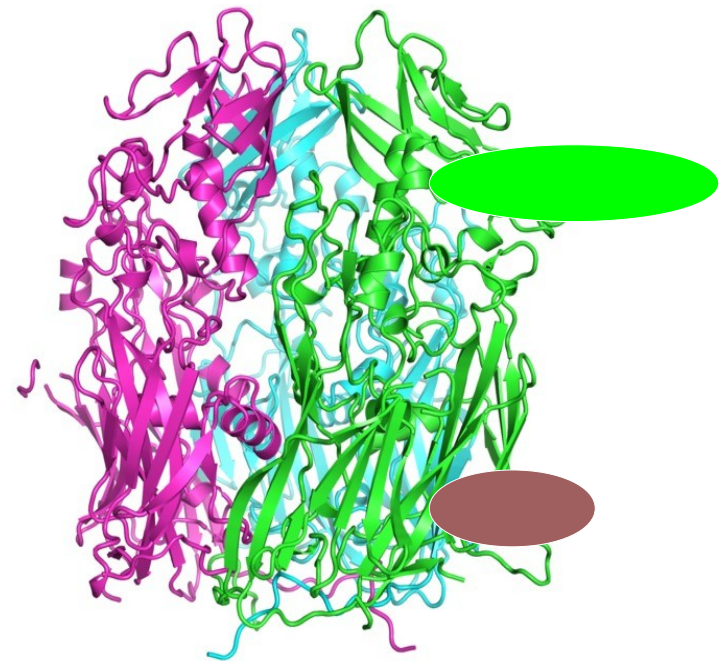
Gene A

Dog

Seq?

Gene A is found in this unknown sequence

Comparing Protein

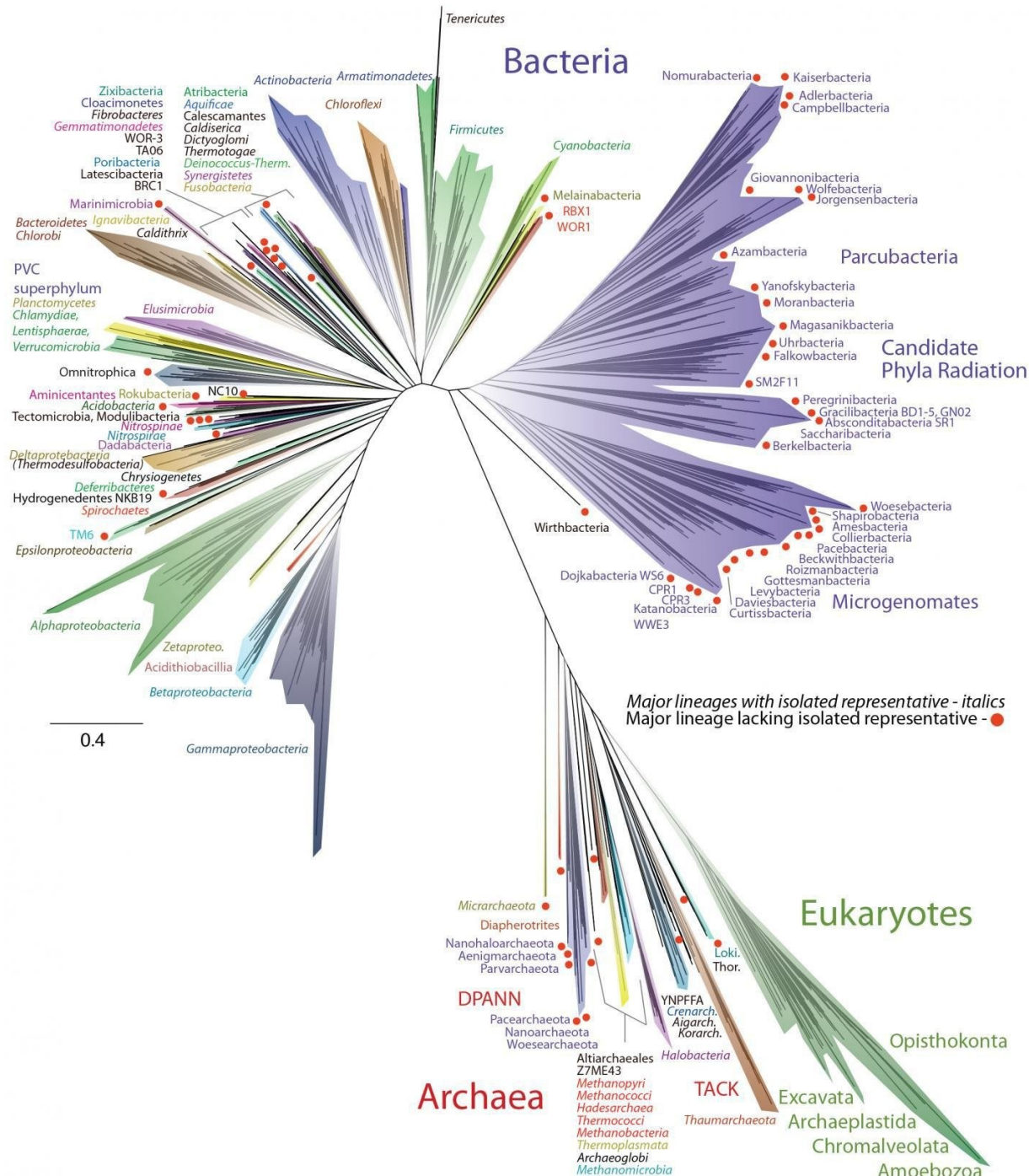- Two proteins (wildtype, non-wildtype) are compared to find causes of disorder.
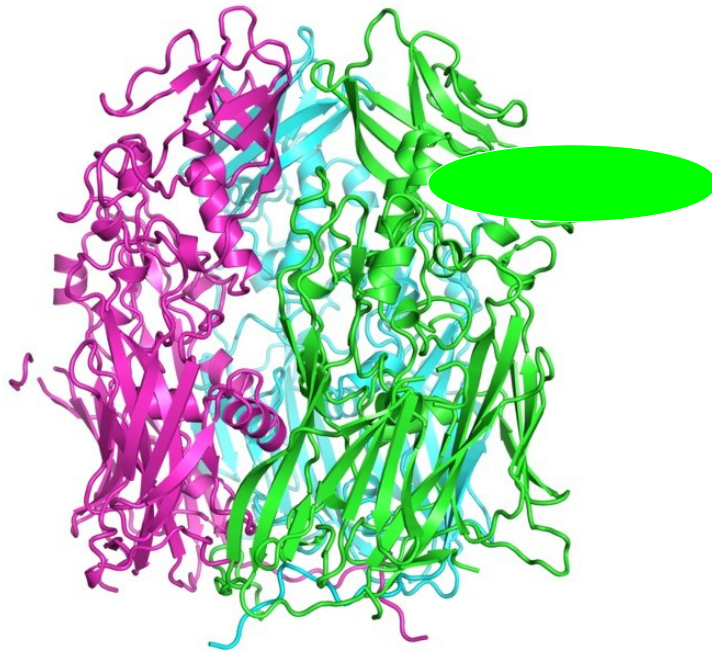
Healthy

Unhealthy

Tree of Life:
Determining Relatedness

So, Why Computer Science?

- Comparisons of *lots* of sequence data
- Ex: Compare seq A and seq B
  - Find any differences that exist

Seq A: *This it s goat!*

Seq B: *This is a boat!*

Seq C: *actcgaattt ctcgcattta cttttgtttt gaattcgcgc*

Seq D: *actcgaactt ctcgcattta ctttagttg gatttagcgc*

**What if these sequences get *really* long!**

Objectives in Bioinformatics

- Sequence analysis
- Make/ use software tools to perform analyses
- Add (publicly) available bio data
- Identification of sequence similarities
- Learn the function of sequences
- Use DNA to answer questions
- And more!

As a Discipline

- Discipline: a branch of knowledge, typically one studied in higher education.

- Bioinformatics is formed out of three disciplines.

Biology, Medicine **Bioinformatics** **Computer Science**

Mathematics, Statistics
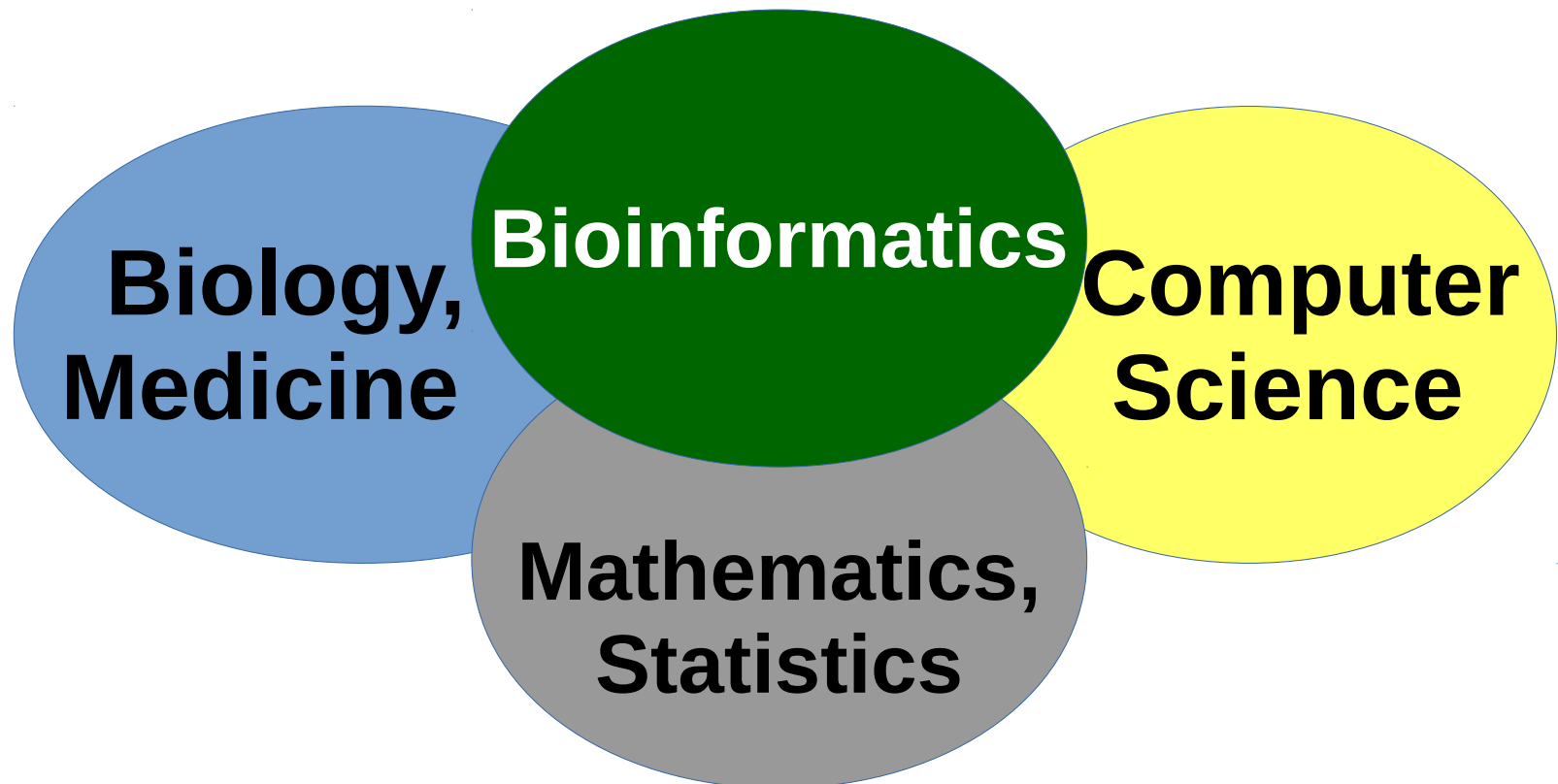
Consider This!

- Get into groups to discuss the following. Keep some notes.

- List some potential applications of bioinformatics (inferring from today's class)

- Why else would you want to compare genetic material? How could this knowledge be meaningful?

- How could knowing the relationships between organisms improve science and medicine?