Bioinformatics
CS300
Introduction to Bioinformatics

Fall 2017

Oliver Bonham-Carter

Introduction to Bioinformatics

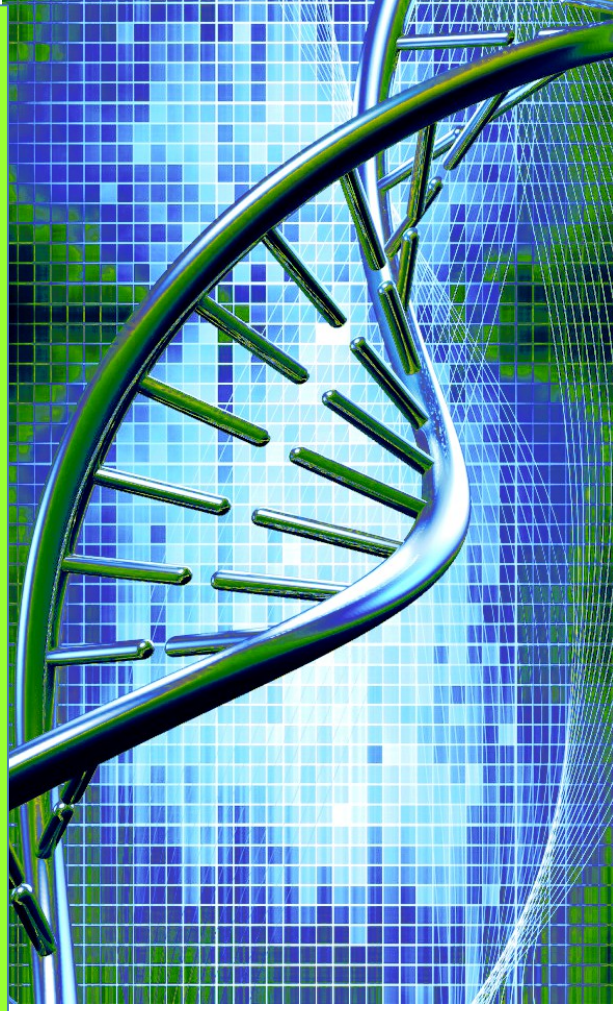
BIO*300/CMPSC*300

Fall 2017

T/Th 11:00-12:15pm

Wed 2:30-4:20pm

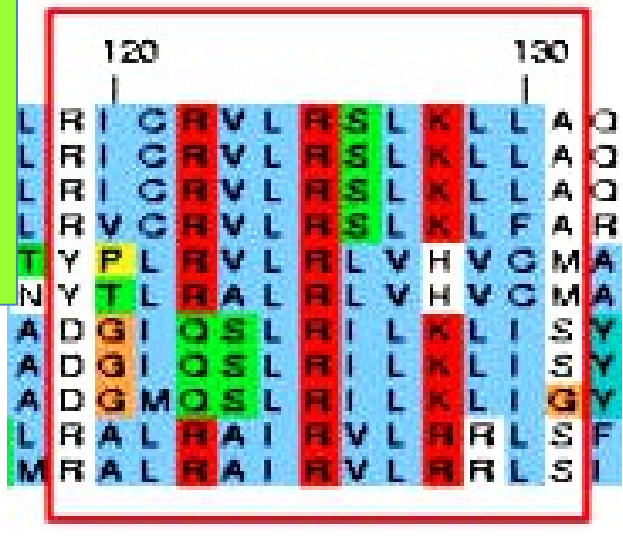
(Alden Hall 109)



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

Bioinformatics is many things...

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

A theoretical framework to detect of 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

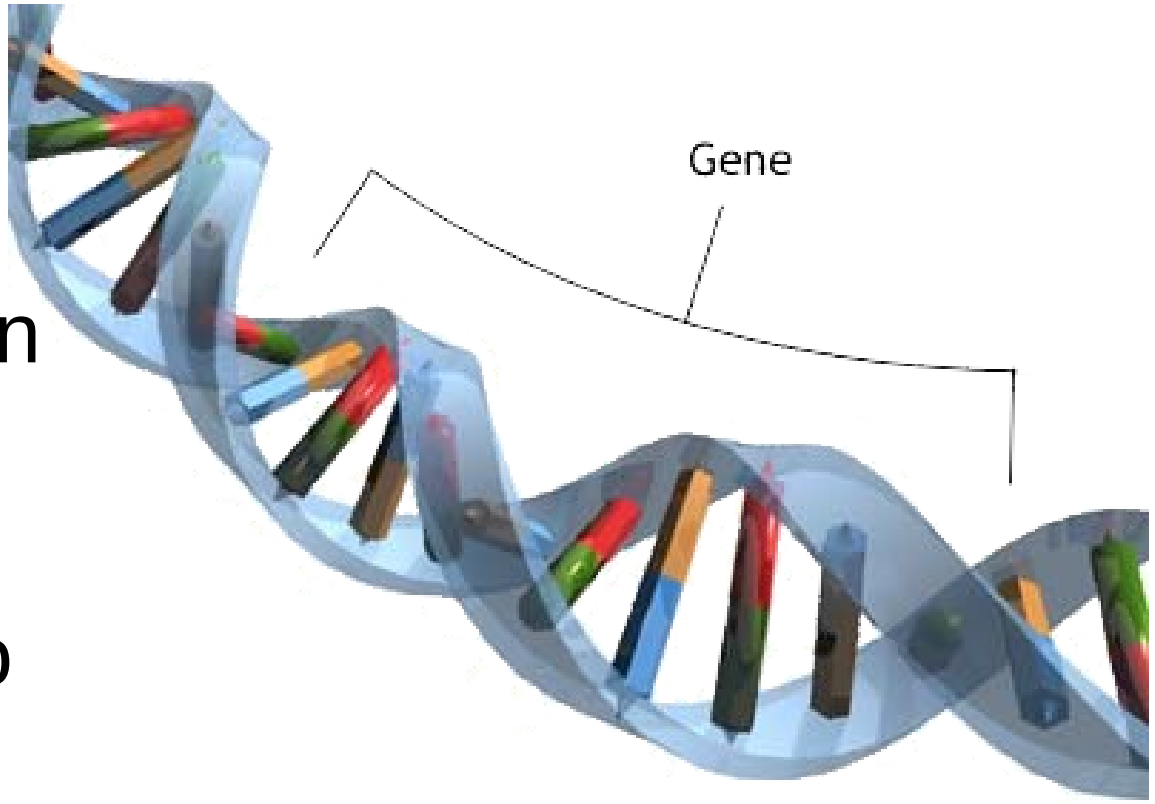
The study of data from living systems to determine patterns of life and health.

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!

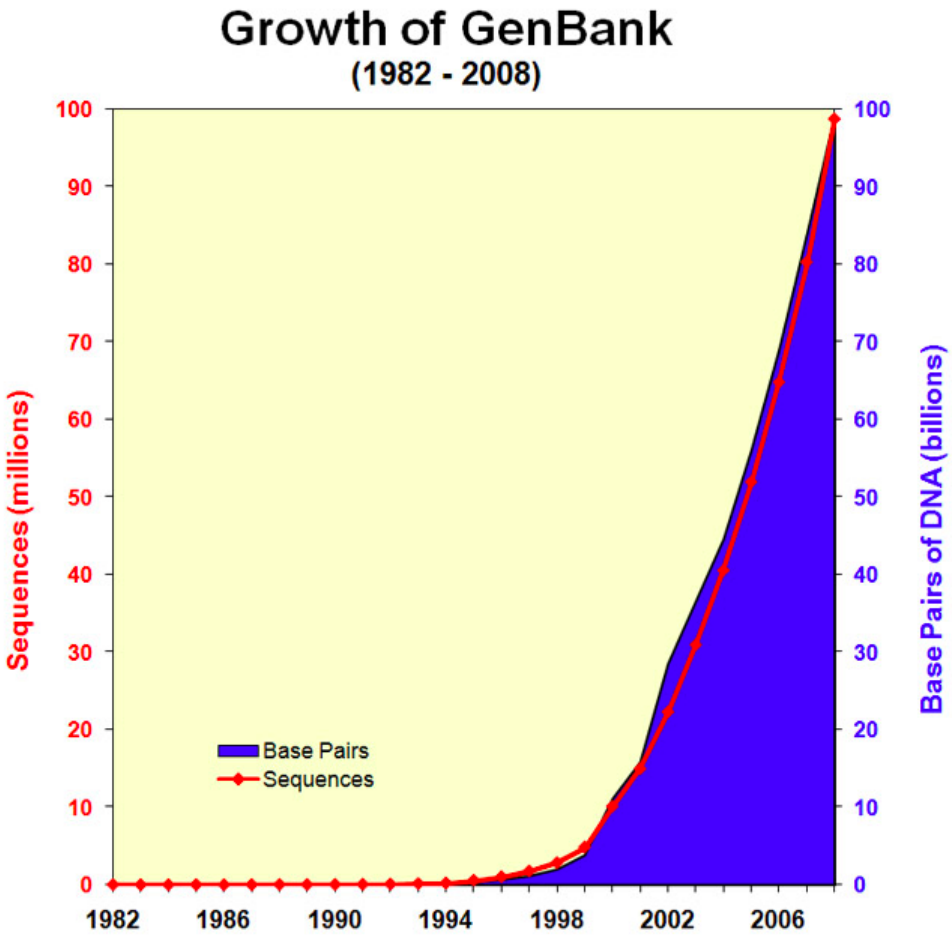
Bioinformatics: The study of DNA

- DNA is the genetic material that houses genetic information.
- Genes are written in this language
- Understanding DNA allows us to understand how genes work.



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



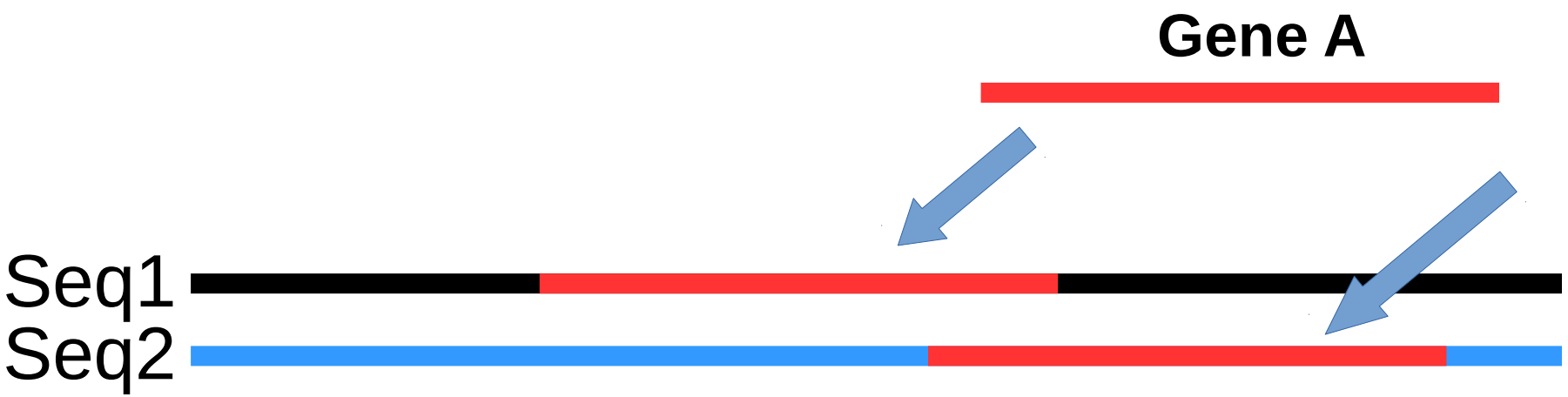
Use DNA to Compare...

- 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? Not the same?**

Comparing Regions?

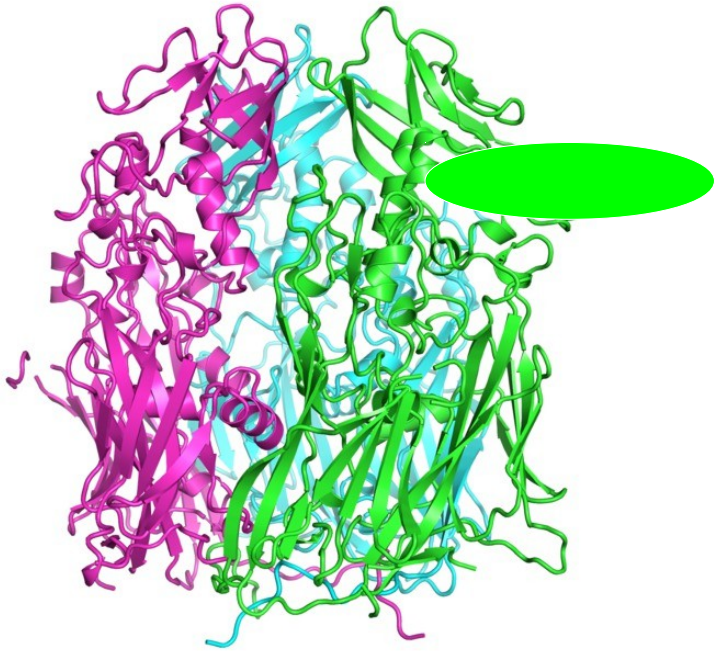
- We learn to scan millions of unknown DNA sequences to find familiar genes.
- Does gene have same function in each sequence? Same origin?



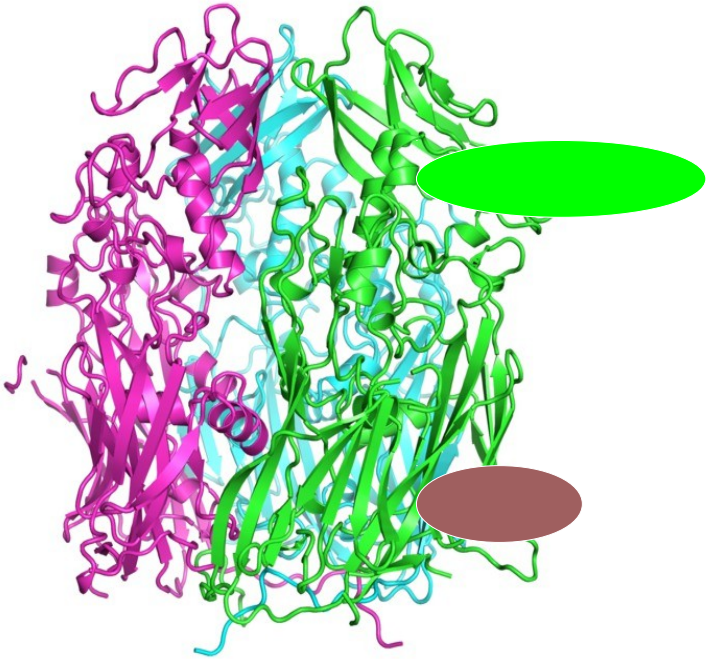
Gene A is found in this unknown sequence

Comparing Protein

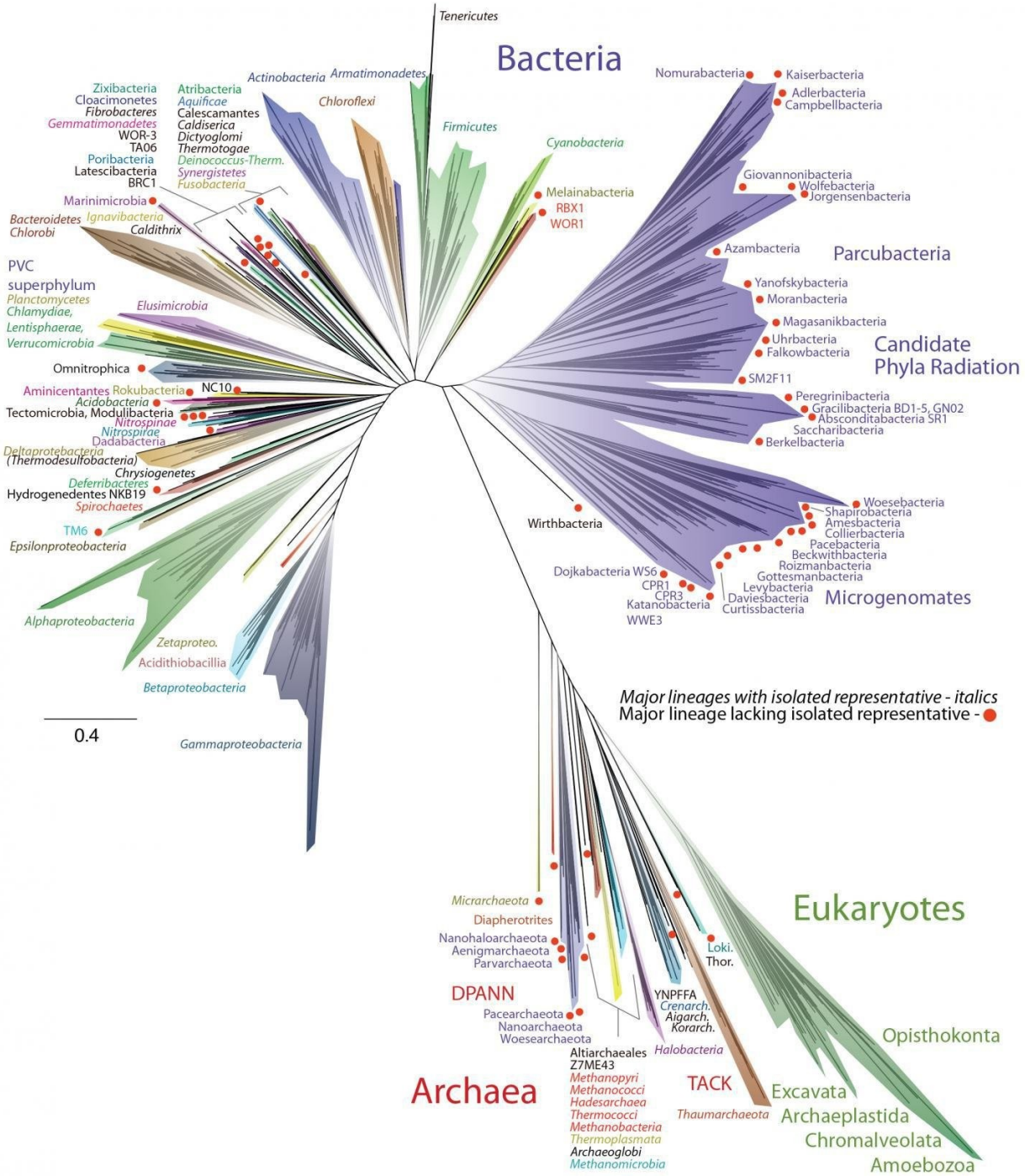
- Two proteins (wildtype, non-wildtype) are compared to find causes of disorder.



Healthy



Unhealthy



Determining Relatedness

Computer Science?

- To perform comparisons of sequence data
- Compare seq A and seq B.
  - Find any differences

Seq A: *This it s goat!*

Seq B: *This is a boat!*

Seq C: *actcgaattt ctcgcattta cttttgtttt gaattcgcgc*

Seq D: *actcgaactt ctcgcattta ctttagttg gatttagcgc*

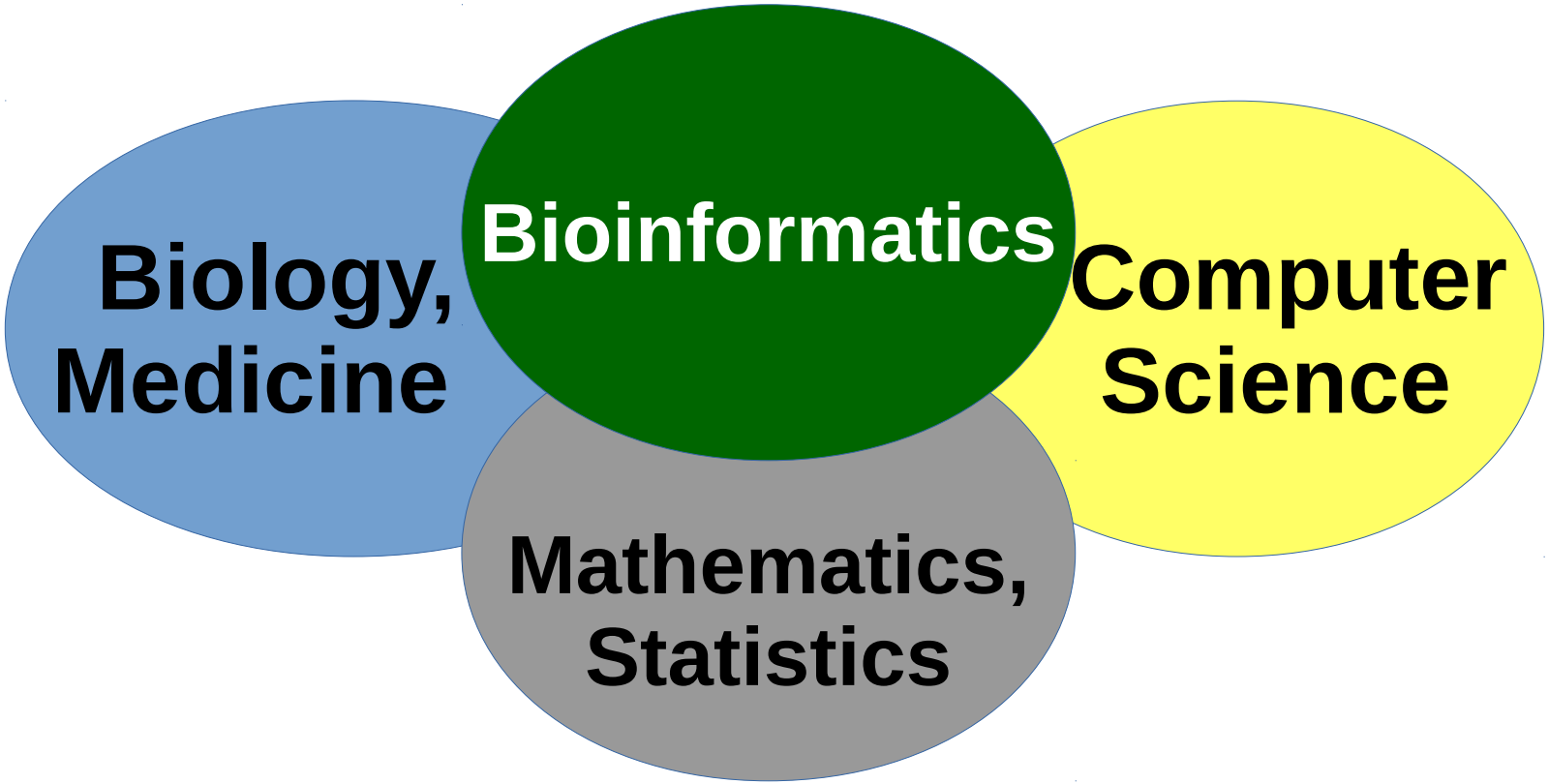
What if these sequences get *really* long!

Objectives to Bioinformatics

- Sequence analysis
- Use software tools to perform analyses
- Use publicly available bio-data
- Identification of sequence similarity
- Defining the function of sequences
- Use sequence information to hypothesize function
- And more!

As a Discipline

- Discipline: a branch of knowledge, typically one studied in higher education.
- Bioinformatics is formed out of three disciplines.



What is Bioinformatics?

- Clinical informatics
  - Systems used to deal with patient health
  - Clinical trial management systems, electronic health records, etc.
- Laboratory information
  - Systems to deal with scientific instrumentation and data management
  - Connecting instruments together, managing laboratory flow, etc.
- Bioinformatics
  - Systems to deal with basic research data
  - DNA, proteins, *molecular* things

Why do we need Bioinformatics?

- Able to fight disease with medicine
- Avoid types of disorders by helping people make better health decisions before it's too late
- To process the massive amounts of health data that already exists.
  - What can we learn from our experiments?
  - Can we incorporate data into computer models to be run instead of using animal models?
  - Other themes...

Skills for Careers

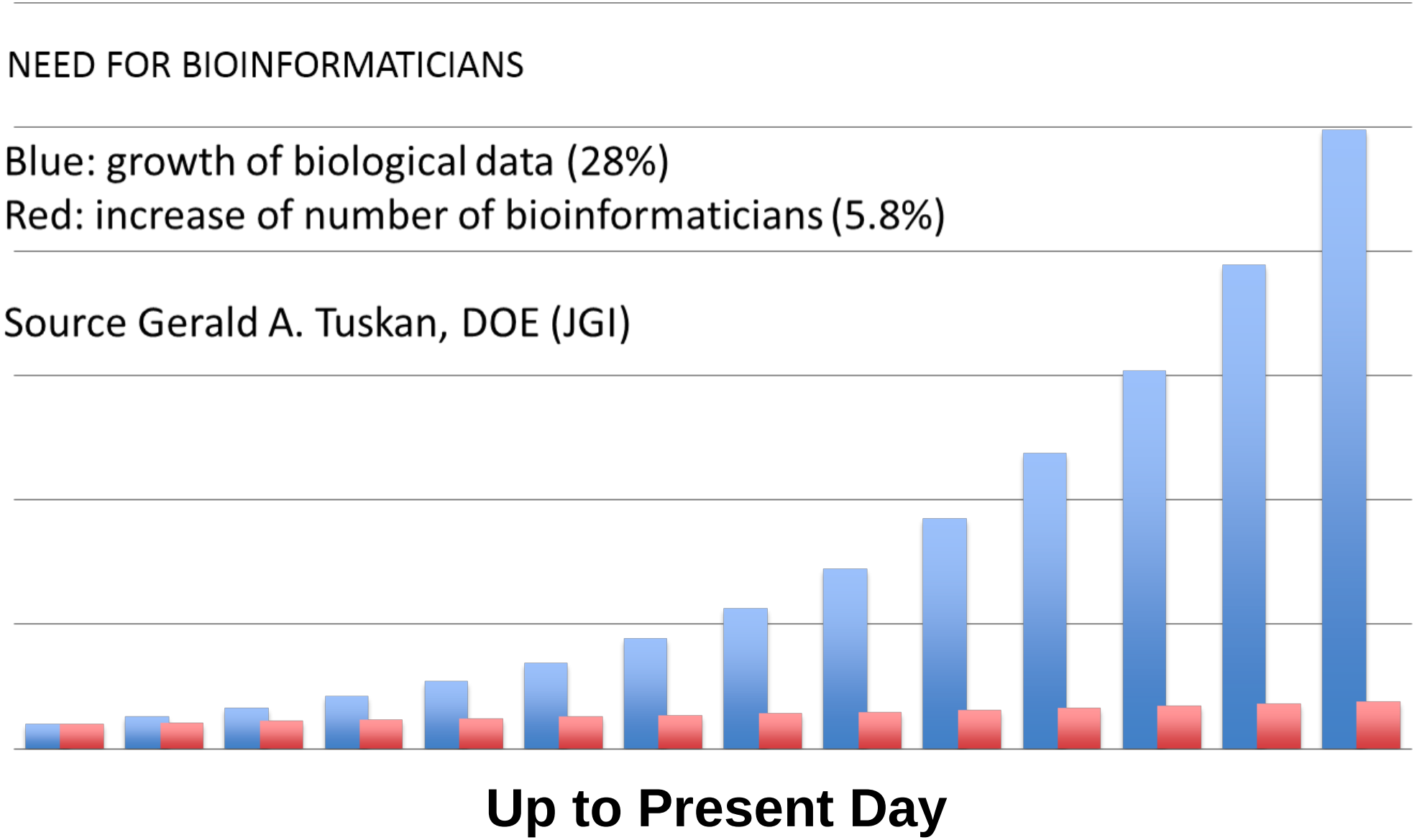
- Biologists:
  - Computational skills
  - Mathematical /statistical
  - Programming for Automation

- Computer scientists
  - BioMedical skills
  - Understanding of biological systems and mechanisms

Career Titles

- Software (bioinformatics) engineer
- Research scientist in biotechnology
- Data scientist
- Project manager (pharmaceuticals, medical, etc)
- Computational immunologist
- Medical doctor (in clinical and research applications)

They Need You Out There!



Rising Amounts of Data to Study

- Exponential Growth of NIH base pairs through December 2013
- Data is to be assembled into sequences (jigsaw puzzle pieces).