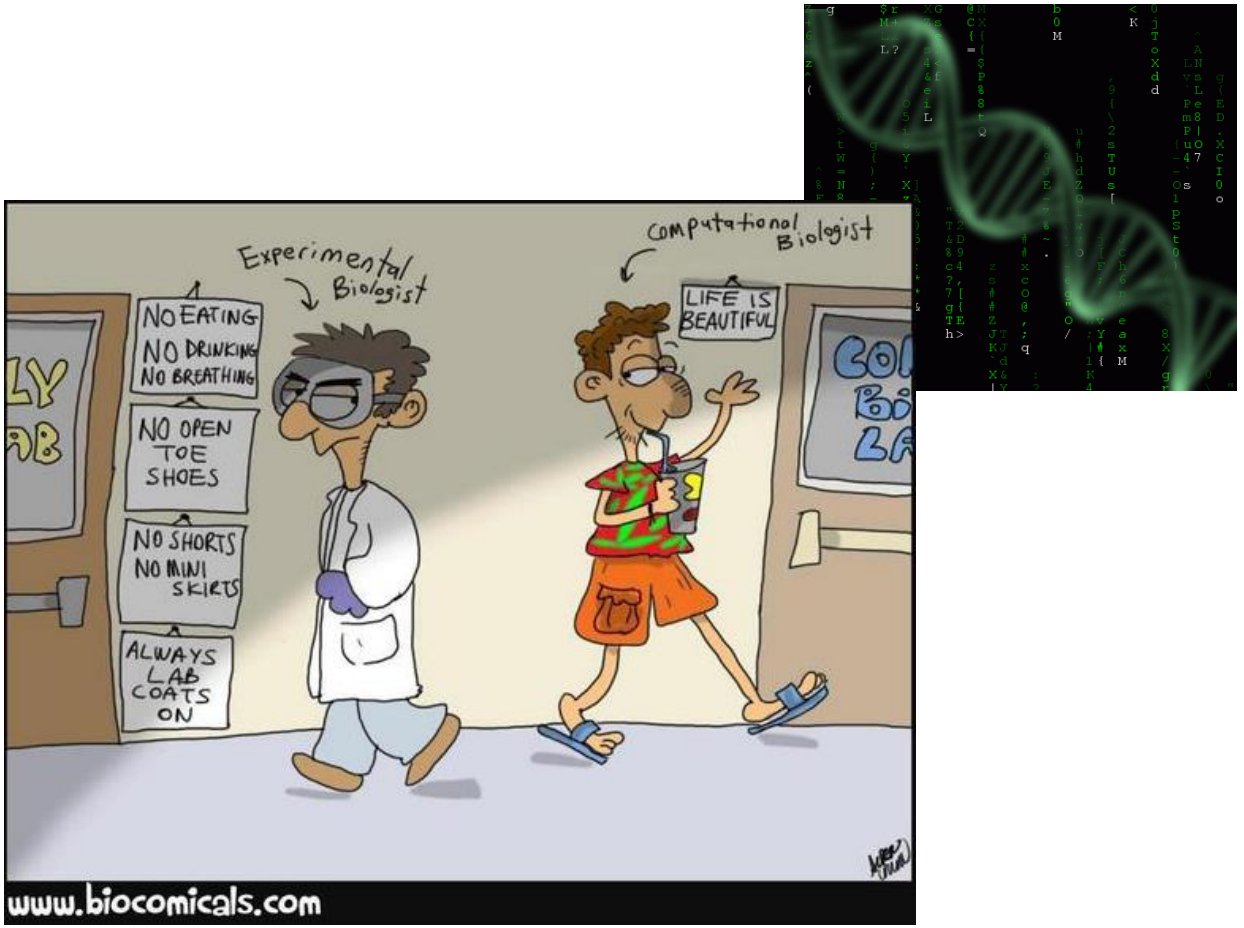
Welcome to BMI-330



Translational Bioinformatics

Li Liu, MD. MS. (BMI-330)

Get to Know Each Other

- Introduce yourself
  - Name, major, year of study
  - Where are you joining us from?
  - What is in your Zoom background? (real or virtual)

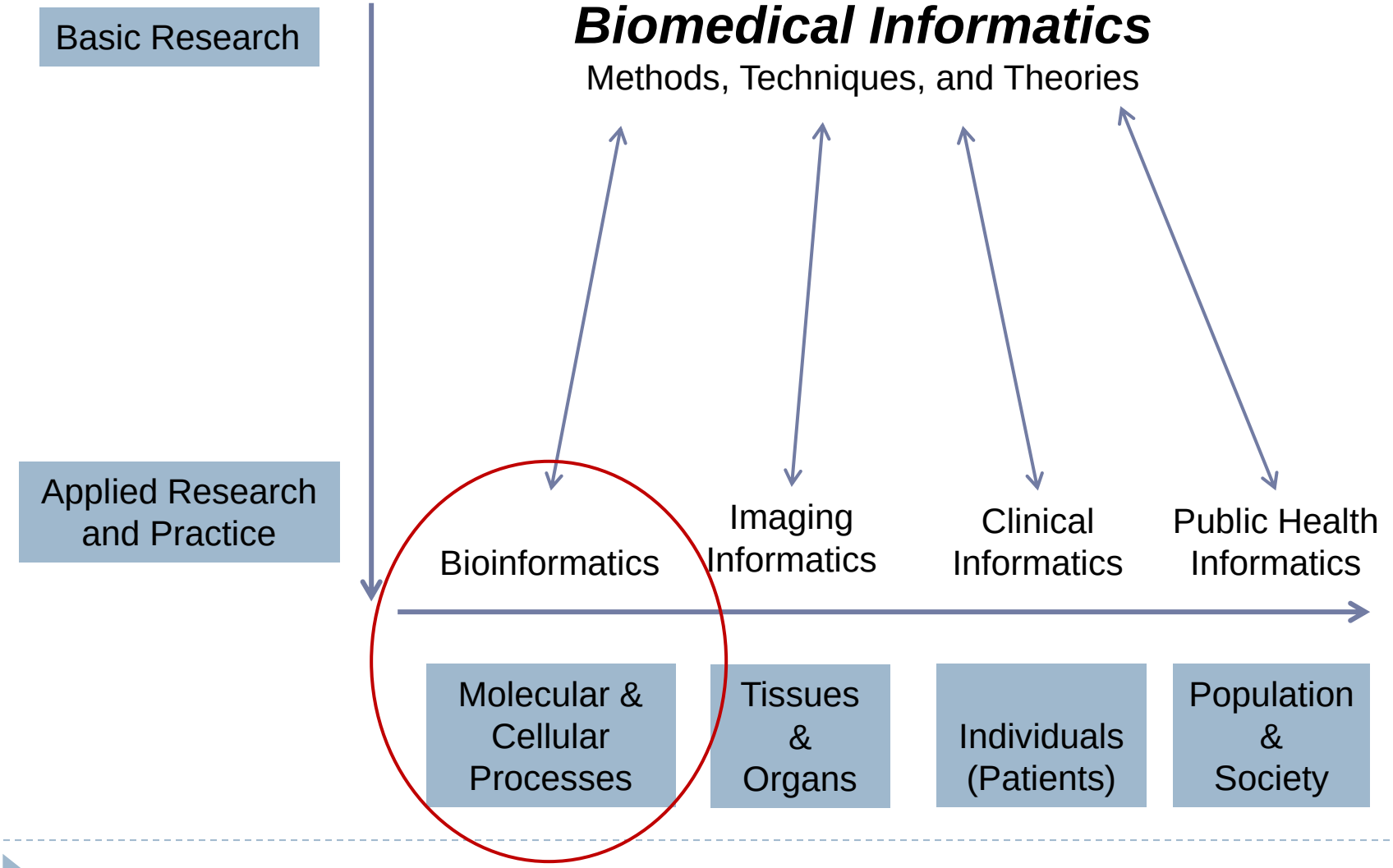
- Play a game
  - 3 students in a group
  - Tell a short story with random words

Bear	**Computer**	**Gene**	**?**	**?**

  - Share the story with the class

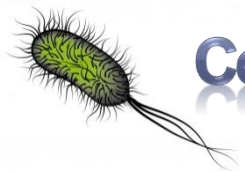
Start recording

Bioinformatics



Shortliffe, 2014

Why Cellular & Molecular Information?



Cells are fundamental units of life.

"*The key to every biological problem must finally be sought in the cell.*"

– E. B. Wilson

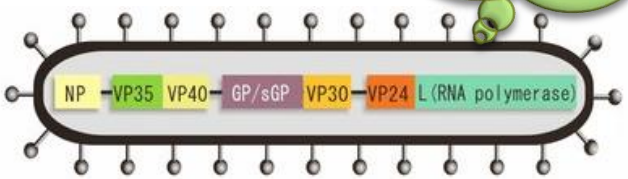
Motivating Questions in Molecular Biology

Bioinformatics

Use computational methods to help answer these questions?

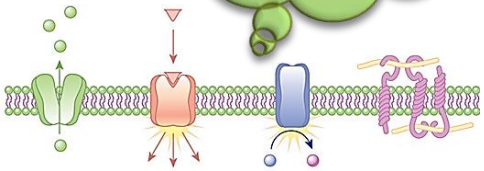
What instructions are encoded in a genome?

Ebola Virus

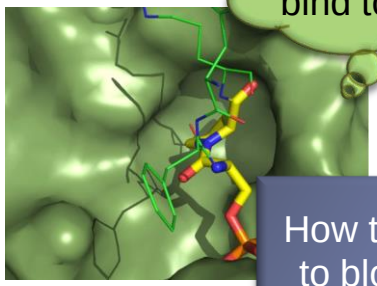


Which gene is associated with an epidemic outbreak, drug resistance, cancer progression, etc.?

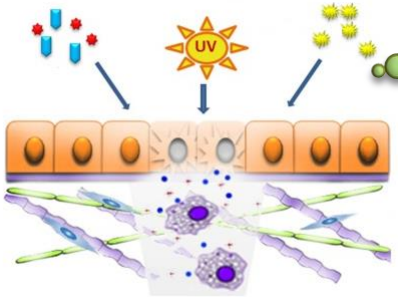
What are the functions of a protein?



Where does a small molecule bind to a protein?



How to design a drug to block the binding site in a protein?



How does a cell interact with its environment?

What are the genetic and environmental factors of diabetes?

…

Translational Bioinformatics (TBI)

Bench

Basic Bioinformatics Research

Bedside

Improved Human Health

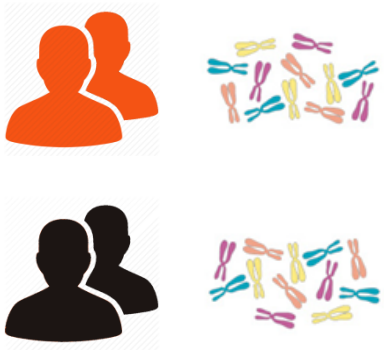
Genomic Medicine

Pharmacogenomics

Precision Medicine

Bioinformatics + Genomic Medicine

Miller's syndrome



~8,000 mutations
in ~4,600 genes

filter agst. dbSNP

~650 mutations

predict deleteriousness

10 mutations
in one gene
(*DHODH*)

Site	19	135	152	199	202	202	244	284	346	392
Variant	E	C	R	C	A	D	W	I	W	G
Human	G	R	G	R	G	G	R	T	R	D
Chimp	G	R	G	R	G	G	R	T	R	D
Gorilla	G	R	G	R	G	G	Q	T	R	D
Orangutan	G	R	G	R	G	G	R	T	R	D
Macaque	G	R	G	R	G	G	Q	T	R	D
Baboon	G	R	G	R	G	G	Q	T	R	D

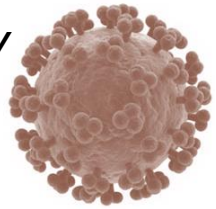
30 20 10 0
Millions of Years Ago

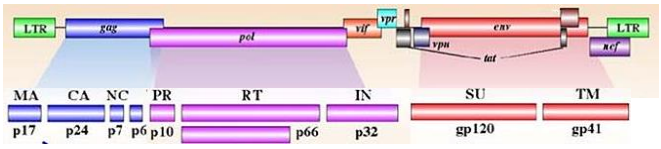
PMID: 19915526

Bioinformatics + Pharmacogenomics

- Drug target identification

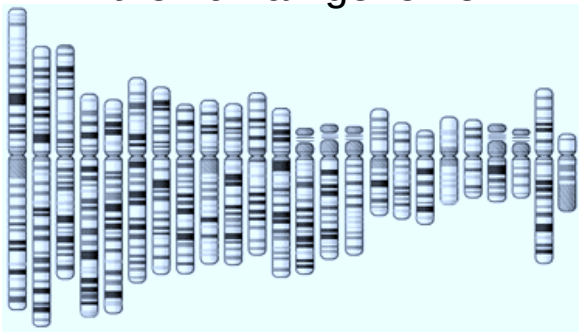
HIV



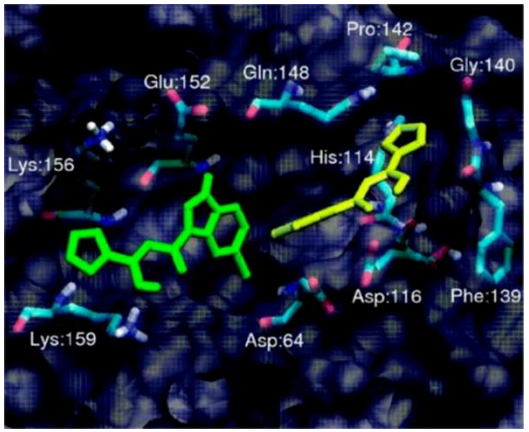


integrase

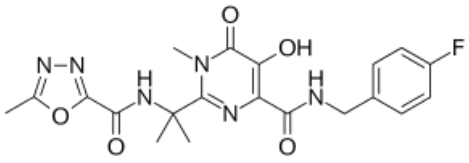
No functional analogue in the human genome



- Drug Design



IN inhibitor

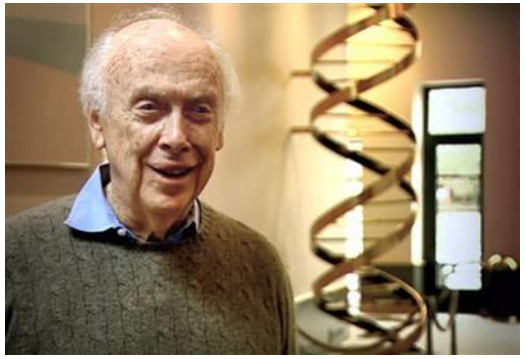


Raltegravir

PMID: 17366700

PMID: 15055986

Bioinformatics + Precision Medicine



- Dr. Watson
  - Genome sequenced in 2007
  - Hypertension, treated with a β-blocker that makes him inappropriately sleepy.

Homozygous variants

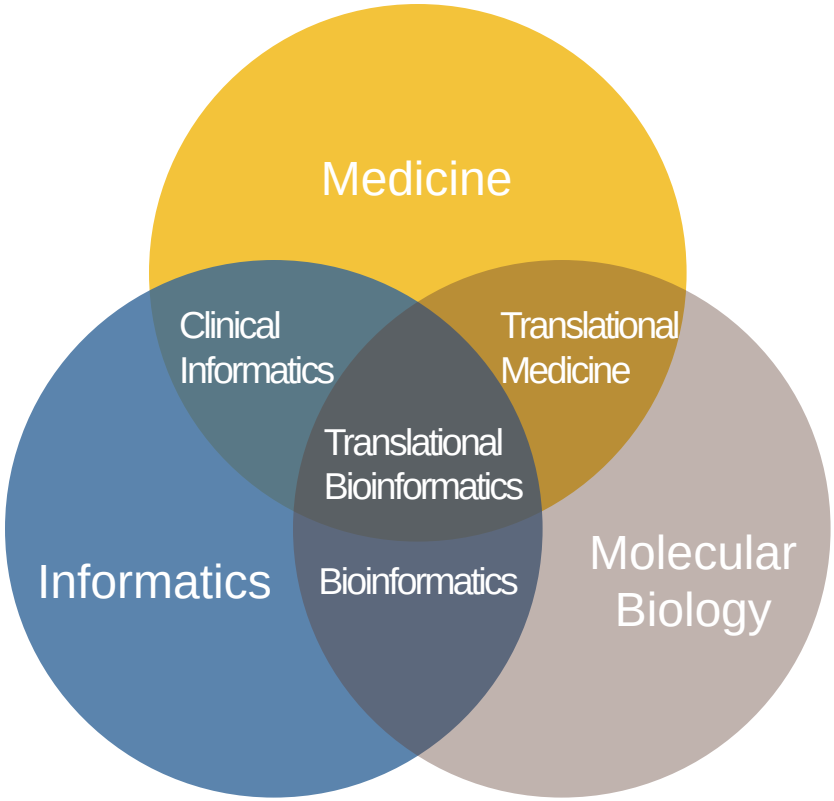
```
          C
GCACGCTACTCACCAGGCCCC
A   R   Y   S   P   G   P
          P
```

CYP2D6: involved with metabolism of 25% of commonly prescribed drugs, including β-blockers. Carriers of homozygous variants are poor metabolizers.

The health provider reduced the dosage of Dr. Watson's medicine to avoid unwanted side effects and keep his blood pressure under good control.

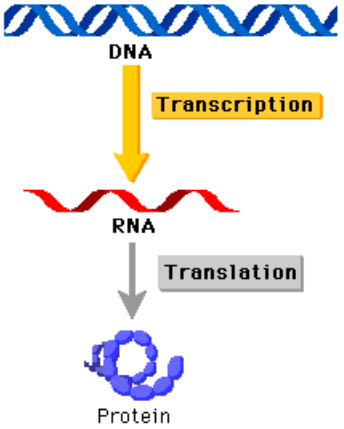
PMID: 18421352

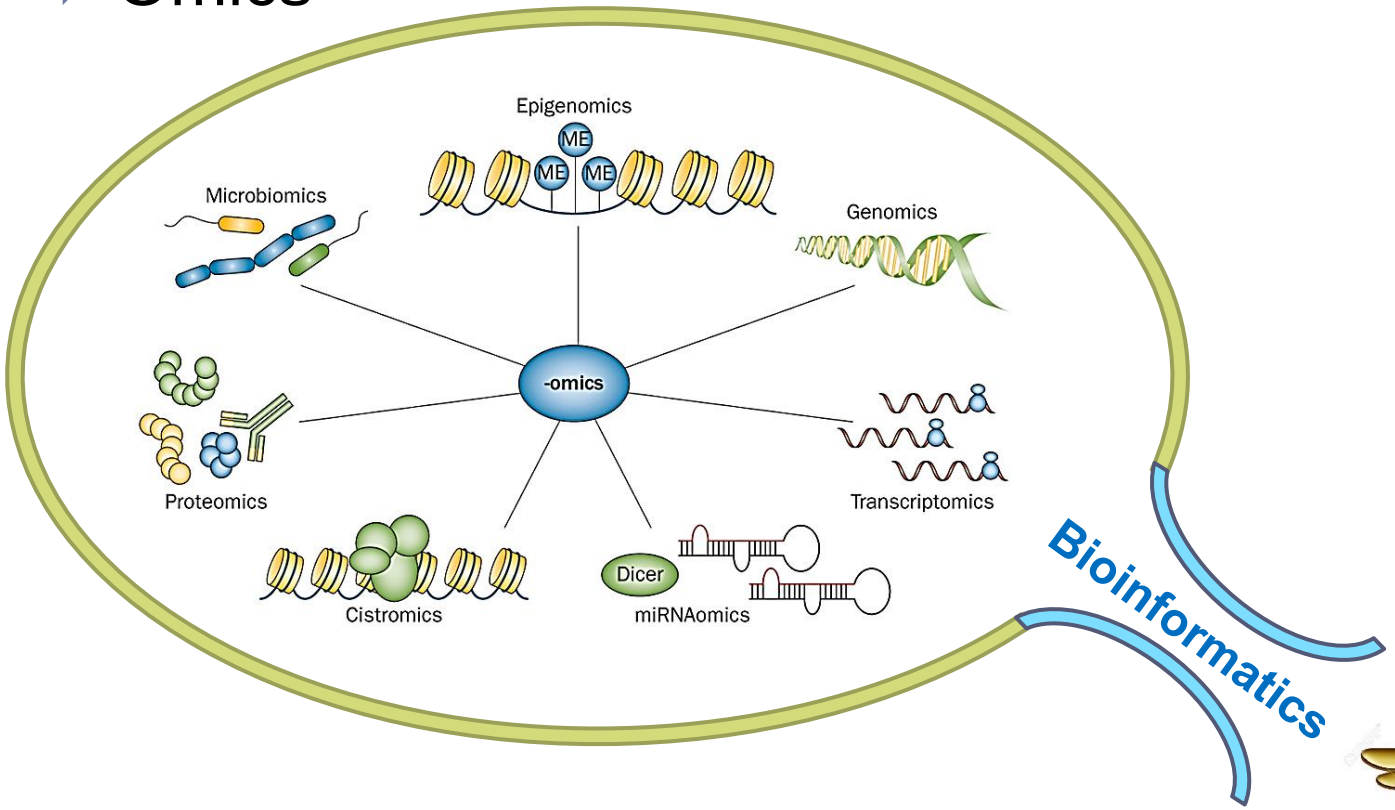
Interdisciplinary



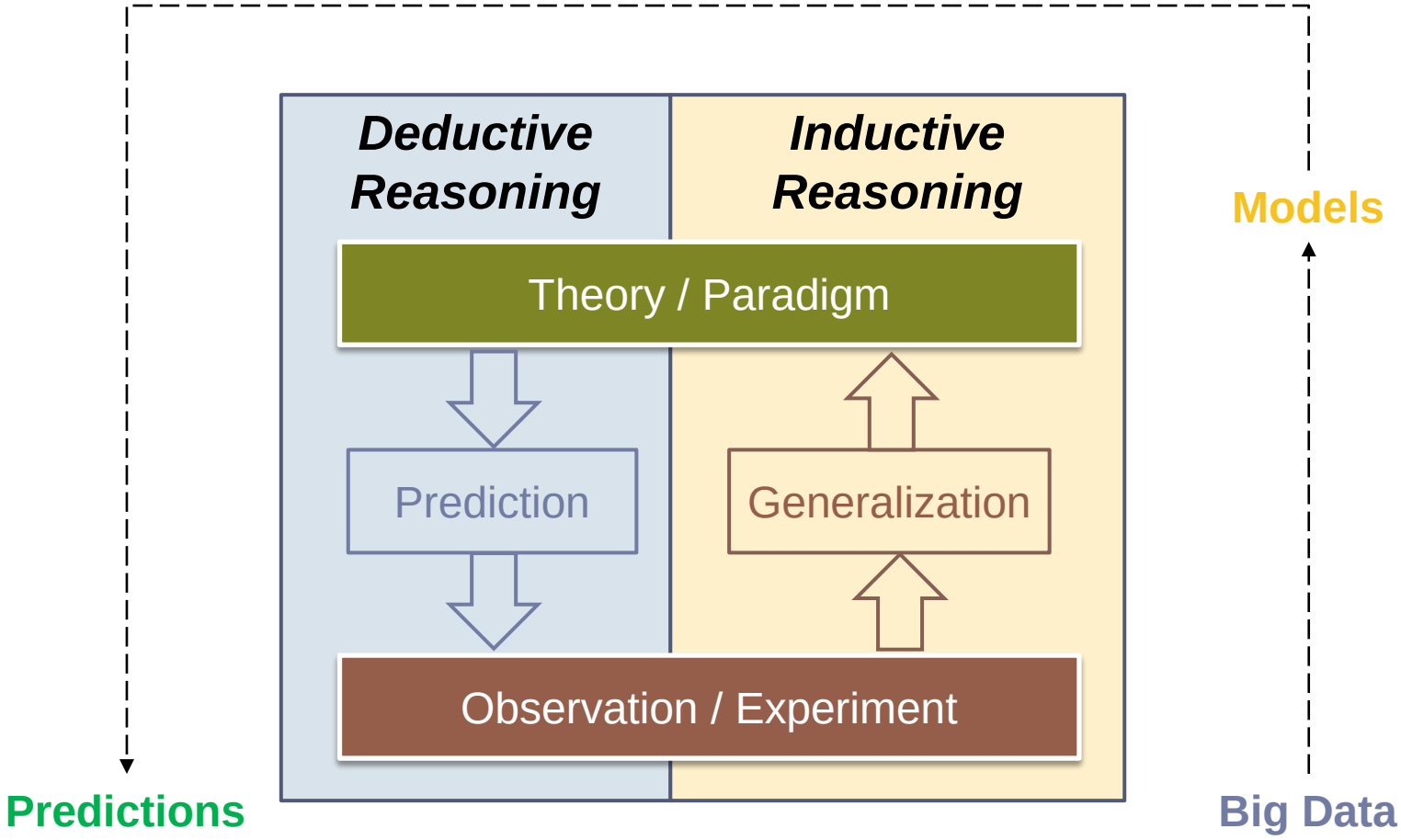
Big Data

- Omics



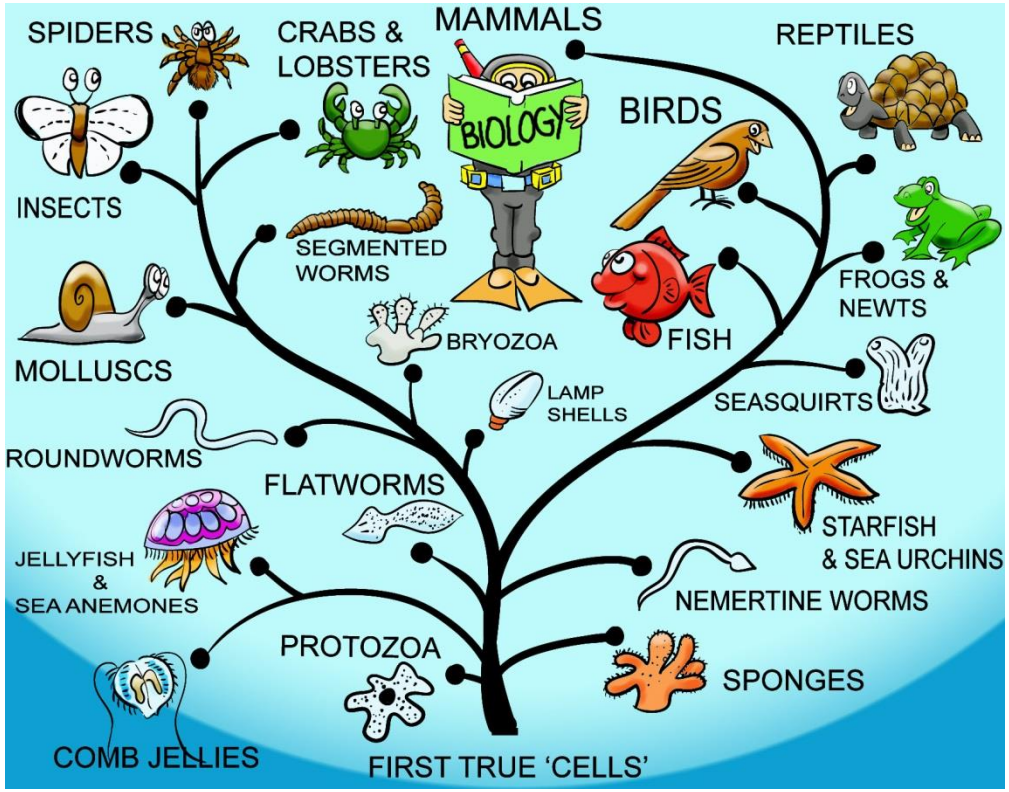


Scientific Approaches



Attention!

- Bioinformatics is not only about human.

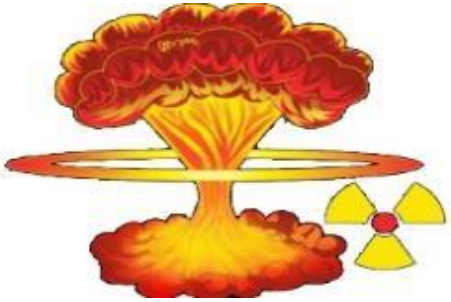


Attention!

- Bioinformatics is not only about Medicine.
  - Agriculture

  - Environment

  - Homeland security



  - Outer space

Class Objectives

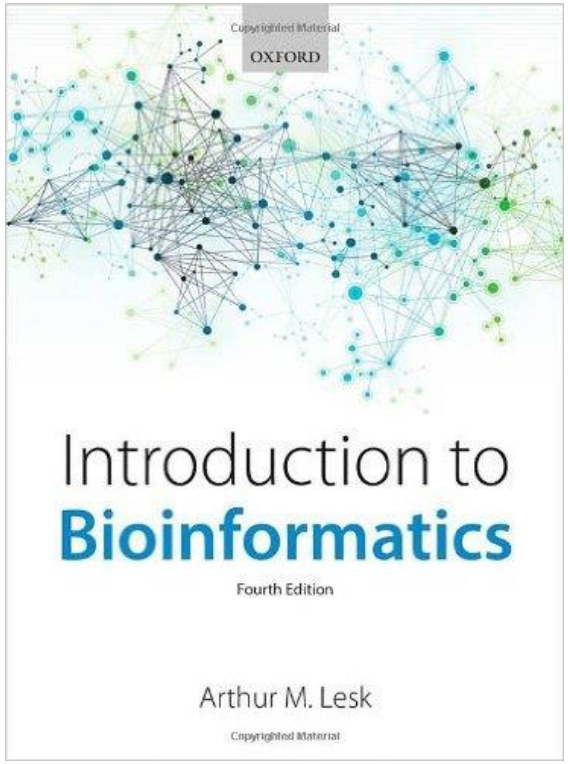
- Identify and describe the major areas of TBI
- Critically appraise existing applications and methods in these areas
- Hands-on practice to analyze example datasets
- Describe likely future applications and probable growth areas

Prerequisites

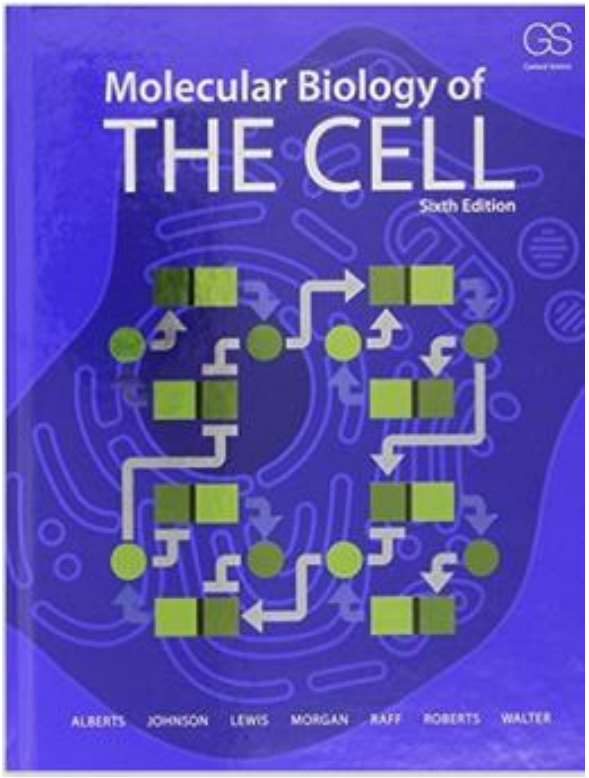
- Biomedical Knowledge
- Molecular Biology / Genetics
- Programming skill (optional)
  - Which language to learn?
- Operating system
  - LINUX (second half of the semester)
  - ASU HPC Cluster

Textbook

- Required



- Recommended



ISBN 978-0199651566

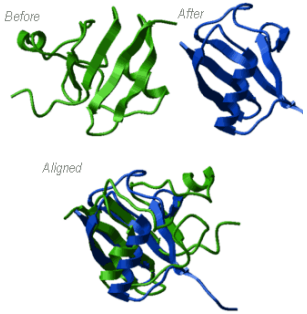
http://www.ncbi.nlm.nih.gov/books/NBK21054/

Areas in Bioinformatics & Example Topics

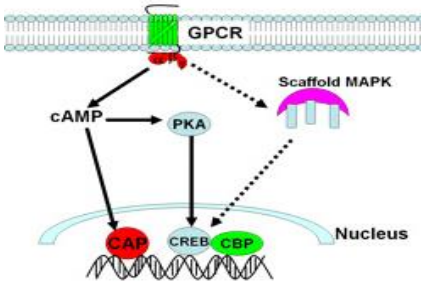
- **Sequence Analysis**
- **Structure Analysis**
- **Functional Analysis**

BIOLOGICAL DATABASE

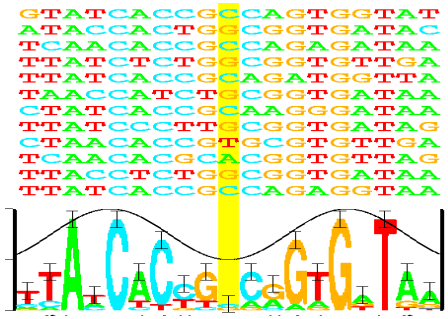
STRUCTURE ALIGNMENT



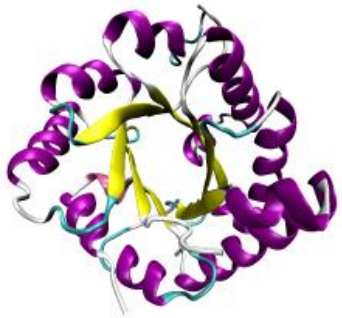
PATHWAY ANALYSIS



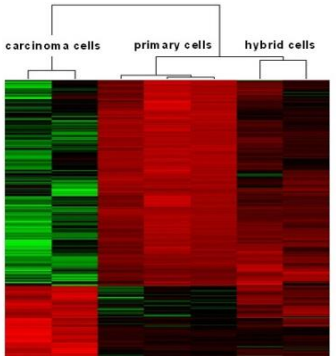
ALIGNMENT & ASSEMBLY



STRUCTURE PREDICTION



PATTERN ANALYSIS



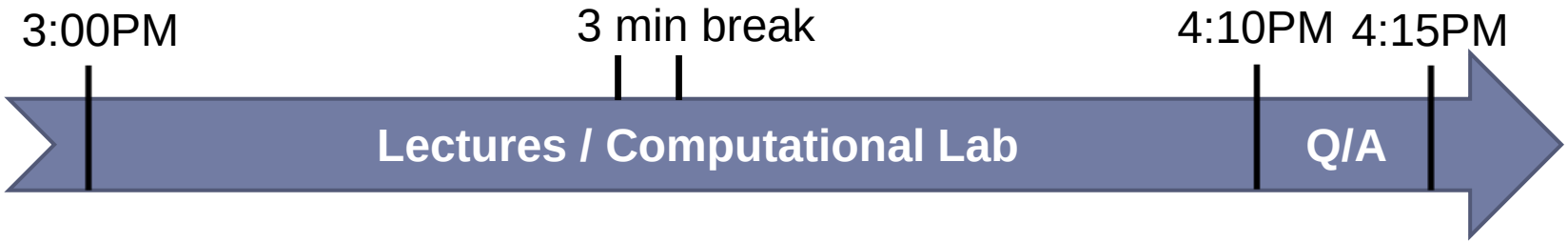
Topics & Schedules

Date	Topic	Homework	Chapter
Jan. 12	Course Overview		Ch. 1
Jan. 14	Primer on Genetics and Genomics		NCBI eBook
Jan. 19 – Jan. 26	Biological Databases	HW-1	Ch. 3 & 4
Jan. 28 – Feb. 9	Sequence Alignment, Motifs & Domains	HW-2, 3	Ch. 5
Feb. 11	Phylogenetic Analysis		Ch. 5
Feb. 16	Gene Calling	HW-4	Ch. 2
Feb. 18	Structure Analysis		Ch. 6
Feb. 23	Functional Analysis of Genetic Variants		
Feb. 25	Case Study – Ebola Virus Genome		
Mar. 2	Study Time		
Mar. 4	Midterm Exam		
Mar. 9 – Mar. 16	Genomics: Technologies and Analysis		Ch. 2
Mar. 18	Linux & ASU High-performance Computing		
Mar. 23 – Mar. 25	Computer Lab – Analyze NGS data		
Mar. 30	Disease Association Analysis	HW-5	Ch. 2
Apr. 1 – Apr. 6	Transcriptomics: Technologies and Analysis	HW-6	Ch. 9
Apr. 8	Proteomics, Metabolomics & Other Omics		Ch. 9
Apr. 13	Pathway and Network Analysis		Ch. 8
Apr. 15	Project Final Presentation		
Apr. 20	Project Final Presentation		
Apr. 22	Study Time		
Apr. 27	Final Exam (9:50 - 11:40 AM)		

Class Activities

- Before class:
  - Read assigned materials: book chapters and papers
  - Install necessary computational packages
  - Download example datasets

- During class:



- After class (homework)
  - Unfinished computational lab exercises
  - Submit results on Blackboard.

Term Project – Disease Genes

- Bioinformatics analysis of coronaviruses
- Team
  - 4 persons in each team
  - Decide by March 9th (the first class after midterm)
- Oral presentation

Grading

- Homework: 20%
  - Reading: textbook chapters
  - Written assignments:
    - Usually due in one week after assignment
    - Submit through the Blackboard.
    - Late submission: 10% off for each late day, no assignment accepted after answers are posted.
- Class discussion participation: 10%
- Midterm Exam: 25%
- Final Exam: 25%
- Final project: 20%
  - 10% from presentation
  - 10% from peer evaluations

Academic Integrity

- University Policies (http://provost.asu.edu/academicintegrity/students).

- All assignments and projects submitted must be the original work of the student(s) submitting it.
  - **cite references**.

Cite References

- **Examples:**
  - Books:
    - Lesk AM (2014) Introduction to Bioinformatics. 4th edition. *Oxford University Press, UK*.
  - Journal Articles:
    - Venter JC, et al. (2010) The sequence of the human genome. Science 291(5507):1304-1351.
  - Websites:
    - Alzheimer's Foundation of America (2016) Alzheimer's statistics www.alzfdn.org/AboutAlzheimers/statistics.html

Office Hours

- Time:

  2:00PM – 3:00PM (Tues. and Thur.)

- Phone:

  (480)727-9813

- Email:

  liliu@asu.edu

  hchen294@asu.edu (TA)

- Office:

  BDB220B (Biodesign Institute Bldg. B 220B – Tempe)

Homework

- Reading
  - Chapter 1: pages 1 – 18
  - Chapter 2: pages 59 – 64, 88 – 95

Online resources on the website of the textbook:
http://global.oup.com/uk/orc/biosciences/bioinf/leskbioinf4e/

QUESTIONS?