Biomedical Informatics: Sequence Analysis
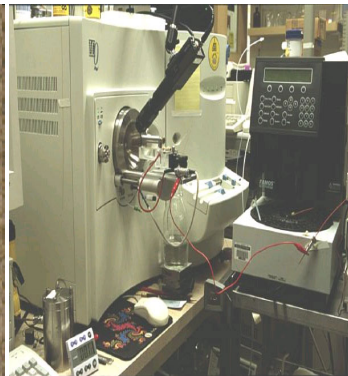
Xiaoman Shawn Li

xiaomanshawnli@gmail.com
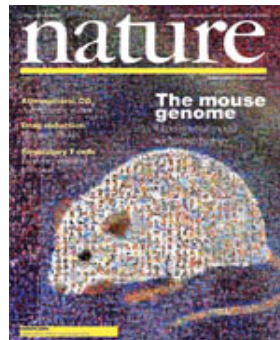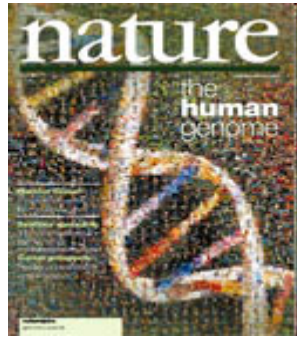
Office: HEC210
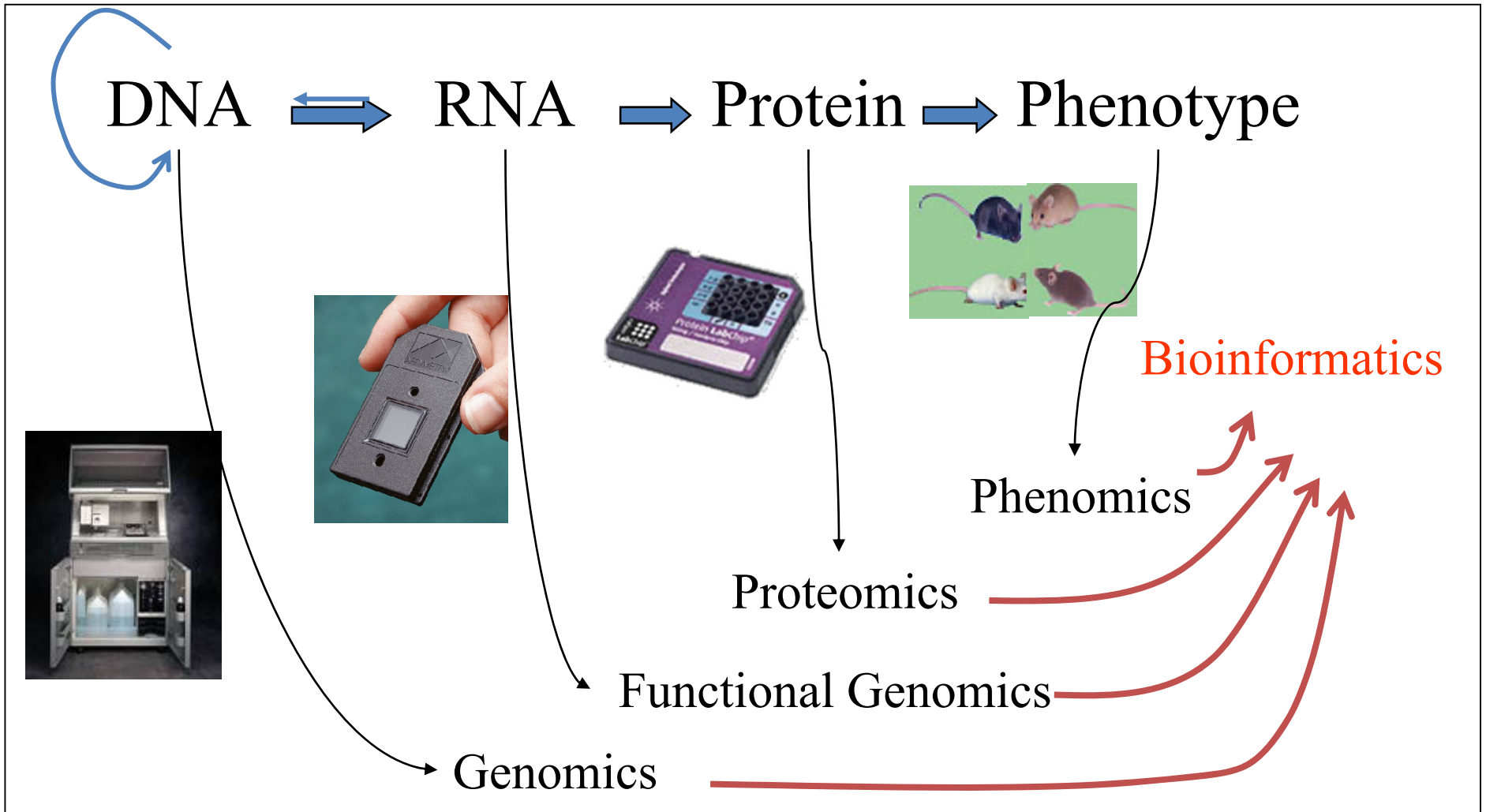
Phone: 823-4811

Office hour: MW9:30am-10:30am

http://www.cs.ucf.edu/~xiaoman/fall/

The big picture

Bioinformatics

DNA ⇔ RNA → Protein → Phenotype

Phenomics

Proteomics

Functional Genomics

Genomics

Bioinformatics

One of the best Bioinformatics work

Phoebus Levene identified the nucleotide unit: base, sugar, and phosphate.

Erwin Chargaff 's rule: #A=#T, #G=#C

Rosalind Franklin: X-ray structure of DNA

Linus Pauling: Triple structure model of DNA

Why you want to know bioinformatics

- Bioinformatics to you is like the computer to the people in the 70th.
- Bioinformatics training will be one of the routine training for biologists.
- You are looking for answers to some questions, which may be addressed from the data people already generated.
- What you have taught in the past and what you take for granted may be completely wrong.
- You will be paid much higher
- You want to be the next Richard Karp, Wing Wong, Walter Gilbert……

Bioinformatics courses offered at UCF

- Introduction to Bioinformatics
- *Algorithms in Bioinformatics
- Machine Learning in Bioinformatics
- *Algorithms in Comparative Genomics
- Biomedical Informatics: Sequence Analysis
- *Biomedical Informatics: Structure Analysis

Topics will be covered

- **Next generation sequencing**
- **Programming with Python**
- **Basic tools and resources**
- **Sequence analysis**
- ***Gene regulation analysis at the transcriptional level**
- **Current trends in bioinformatics**

Graduate Assignments

- choose one topics you are interested in and read two papers. The preferred topics include but not limited to miRNA, single cell, and metagenomics, drug resistance, enhancers. The papers must be from Science, Nature, Nature Genetics, Nature biotech, Nature methods, or cell after 2021. All papers should utilize next generation sequencing data for their analyses. The instructor should be informed of the selected topic by Sept 11 and the slides should be submitted by email no later than Oct 9 (20%).

- If you present methodology papers applied on biological problems, you can choose papers from other journals with impact factor larger than 5 and published after 2018.

- Present your topics in 30 minutes and lead the discussion (40%).

- Write a 1 page of your thoughts on future directions of the topics you presented (30%). The essay must include at least two points about what can be done for future research by yourself, not from the papers. You don't need to solve the problems. Due on Nov 27.

- The attendance (0%), class writing questions (6%), participation in class and discussion during others' presentations (4%).

Undergraduate Assignments

- Choose one topic and read one paper on the selected topic. The preferred topics include but not limited to: miRNA, single cell, metagenomics, drug resistance, and enhancers. The main paper selected must be from Science, Cell, or Nature, Nature Biotechnology, Nature Genetics, Nature Methods and must be published after 2021. The main paper should use the next generation sequencing data. Make slides for 20 minute presentation. The topic choice is due by <u>Sep 11</u> . Your presentation ppt file is due on <u>Oct 9</u> (20%).

- Present your topics and lead the discussion (40%).

- Write a 1 page of your own thoughts (not from the paper) on future directions of the topics you presented (30%). You don't need to solve the problems. Due on <u>Nov 27</u>.

- The attendance (0%), class writing questions (6%), participation in class and discussion during others' presentations (4%).

Grade criteria

· A: 95-100; A-:90-94; B+:87-89; B: 83-86; B-:80-82; C+:77-79; C:73-76; C-:70-72; D+:67-69; D:63-66; D-: 60-62; F: <60

· Review, opinion and perspective papers cannot be served as the main paper(s) for presentation.

· For the slides submitted at the first time, you will need to have at least 18 (undergraduate) or 25 (graduate) slides (not counting the title slide and the "thank you" or reference slides). Fewer slides will result in the loss of 1 to 2 points.

· For the presentation, you are expected to present 18 to 25 minutes (undergraduate) or 28 to 35 minutes (graduate). Longer presentations may be stopped; shorter ones will result in the loss of 1 to 2 points.

· Late submission penalty: the topic after Sept 11, or slides after Oct 9, 1 to 3 points penalty; topic after Oct 1st, or slides after Nov 1st, 5 points deduction.

· Make your own slides. Using other person's slides, lose 5 to 10 points.

· Explain the slides instead of reading. Reading > 50% of time lose 5 points.

· In the final essay, you are expected to point out one or more directions for further research. An essay with only the points from the paper will lose at least 3 points.

· Submit your in-class question answer by 5pm of the same date. Submitted later will not be counted. Only those submitted answers to more than half of the in-class questions will be considered for A.