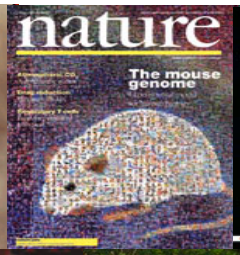
Structure Bioinformatics

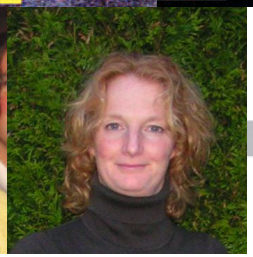
Xiaoman Shawn Li
HEC210
xiaomanshawnli@gmail.com
http://www.cs.ucf.edu/~xiaoman/spring

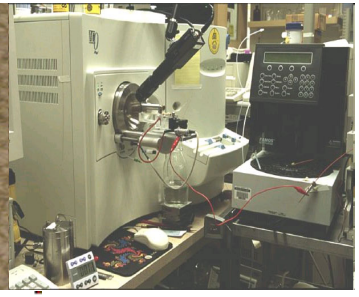
Office hours:
MW: 10:30am-11:30am

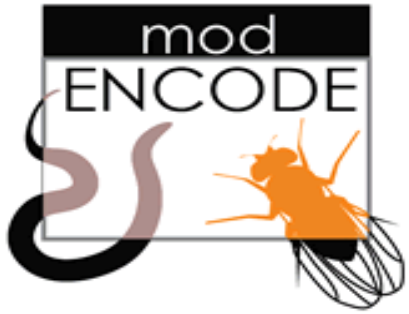
The big picture



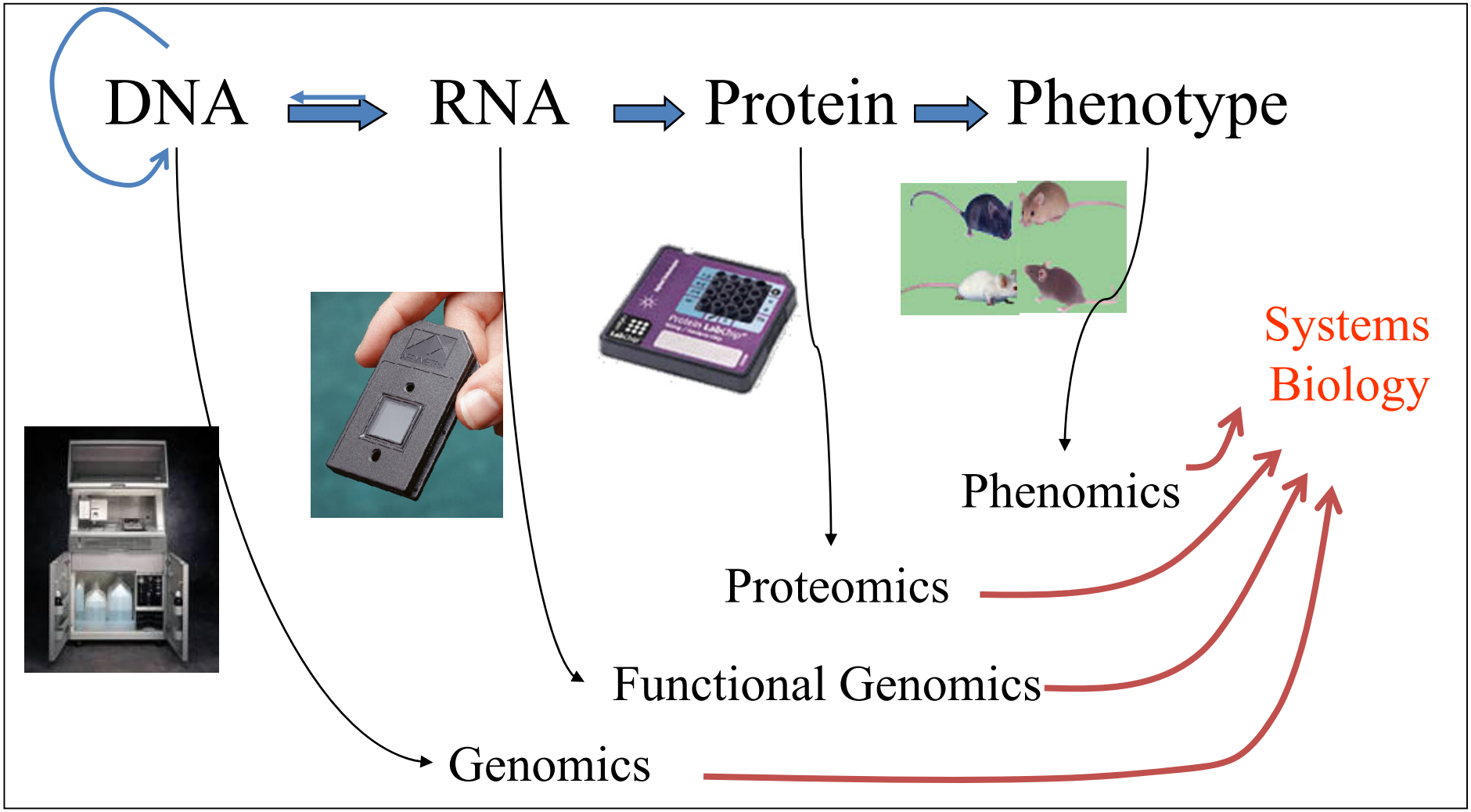




Computational Biology or Bioinformatics



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
- Machine Learning in Bioinformatics
- *Biomedical Informatics: Sequence Analysis
- *Biomedical Informatics: Structure Analysis

Topics will be covered

- **Next generation sequencing data analysis**
- **Basic tools covered in sequence analysis**
- **R programming**
- **miRNA**
- **RNA structure**
- **Protein linear motifs**
- **Protein structure**
- **Homology modeling and protein-DNA interaction**
- **Current trends in biology**

Graduate Assignments

- Read 2 related papers on the same topic. The papers must be from Science, Cell, or Nature, Nature Biotechnology, Nature Genetics, Nature Methods and must be published after 2021. Each paper should utilize high throughput data for its analyses, not for single gene or single pathway. No review, survey, or perspective paper is accepted. Email the instructor the links to the papers by Feb 3 and the slides for 30 minute presentation no later than March 5 (18%).
- Present your topics or tool and lead the discussion (40%).
- Write a 1 page of your own thoughts (not from the paper(s))on future directions of the topics or tool you presented (30%). You don't need to solve the problems. Due on April 23.
- Six in-class questions (6%), participation in class and discussion during others' presentations (6%).

Undergraduate Assignments

- Choose 1 paper published after 2021 from Science, Cell, or Nature, Nature Biotechnology, Nature Genetics, Nature Methods. The paper should utilize high throughput data for its analyses, not for single gene or single pathway. No review, survey, or perspective paper is accepted. Email the instructor the link to the paper by Feb 3 and the slides for 20 minute presentation no later than March 5 (18%).
- Alternatively, you can choose a bioinformatics tool that is approved by the teacher to present the tool, such as what the tool is about, and how to use the tool, and demonstrate the use on real examples.
- Present your topics or tool and lead the discussion (40%).
- Write a 1 page of your own thoughts (not from the paper(s))on future directions of the topics or tool you presented (30%). You don't need to solve the problems. Due on April 23.
- Six in-class questions (6%), participation in class and discussion during others' presentations (6%).

Grade criteria

- A: 95-100; A-:90-94; B+:87-89; B: 84-86; B-:80-83; C+:77-79; C:74-76; C-:70-73; D+:67-69; D:64-66; D-: 61-63; F: <=60
- Review, opinion and perspective papers cannot be served as the main paper(s) for presentation.
- In-class question needs to be emailed to xiaomanshawnli@gmail.com by 5pm on the same date the question is asked. If you miss half of the six in-class questions, you will automatically lose five points. For doctor visit and other special reasons, you should have doctor notes or other documents submitted immediately after each occurrence to excuse from the in-class question.
- For the first version of the slides submitted, for undergraduate students, you will need to have at least 18 slides. For graduate students, you will need to have at least 26 such slides. The title slide and the "thank you" or reference slides are not counted. Fewer slides will result in the loss of 1 to 2 points.
- For the presentation, for graduate students, you are expected to present 28 to 30 minutes; for undergraduate students, you are expected to present 18 to 20 minutes. Longer than the expected time will be stopped; shorter than the expected time will result in the loss of 1 to 2 points.
- If you provide the topic after Feb 3, or provide your slides later than March 5, you will lost 1 to 3 points. If you provide your topic after March 1st, or submit your slides after April 1st, you lose 5 points.
- You need to make your own slides. If you use other person's slides, you will lose 3 to 5 points.
- You are expected to explain the slides you made instead of reading the slides. If you read more than 50% of time, you will lose 3 to 5 points.
- In the final essay, you are expected to point out one or more directions of your own for further research based on the paper you read, assume you actually the author of the paper. An essay with only the points from the paper or others will lose 10 points.

How to find papers yourself

- Search pubmed http://www.ncbi.nlm.nih.gov/pubmed to find something interesting to read and present.
- You go over the titles of the papers from 2022 in the following journals: Science, Nature, cell, Nature Genetics, Nature Biotechnology, Nature methods. If you present computational methods, you can also choose from bioinformatics, Nucleic Acids Research, Genome Biology, Genome Research, PNAS.