What is Computational Biology Research?

David D. Pollock, PhD, August 31, 2009

An introduction to the BIOL 7711 course and the aims of the Computational Bioscience Program. A discussion of what it means to be a graduate student in computational biology.

Welcome and Introduction to BIOL 7711

Get exact names, contact information for all participants. Get information on their backgrounds (general as well as computation and biology). Go over the syllabus. Keep it short.

What is Computational Biology Research?

Format: loose lecture and discussion

Goals: to discuss what it means to be a graduate student, how to be a graduate student, especially in computational biology.

Discuss the difference, if any, between **computational biology** and **bioinformatics**. Informatics is more about processing information, building databases, building pipelines, making information accessible. Computational Biology is more about using computation as a tool to solve biological problems. Some time ago, this would have mostly been population biology, evolutionary biology, and epidemiology, with some morphometrics thrown in. Today, there is

- high-throughput production of data from DNA (**genomics**),
- protein (**structure**, prediction, **proteomics**), and
- RNA (**expression** data, hybridization, functional genomics).
- Also **metabolomics**, etc.

New experiments and high-throughput data are coming down the pike all the time. There is also a huge field of **medical informatics**, including a lot of basic statistics on doing studies right, processing medical data, and even hospital management and design.

What is our mission?

To **train** computational biologists who aspire to **achieve excellence** in **research**, **education**, and **service**, and who will **apply the skills** they learn toward **improving human health** and **deepening our understanding** of the living world.

Our training integrates training and computation and biomedical sciences with student research and teaching activities that grow increasingly independent. Graduates can do **independent research**, **collaborate** effectively, and **communicate** their knowledge to both students and the broader scientific community. There are **four major goals**: Knowledge, communication, professional behavior, self-directed learning skills.

Knowledge: core concepts and principles, ability to apply computation to gain insight into biological problems. This includes mastery and integration of biology, statistics, and computer science.

- Scientific principles of molecular biology, statistics, and computer science
- Ability to productively integrate knowledge to solve problems
- Know types and sources of data
- Know classes of common algorithms and how to apply them
- Principles and practices of scientific method, or how to discover new truths about the world that are correct (good predictors of future observations)

Communication: interpersonal, oral, and written skills. Clearly communicate results of their work and teach others. Bridge the gap between molecular biology and computational cultures.

- Communicate effectively
- Be familiar with biomed and computational modes of expression
- Commitment and skill in teaching to and learning from students, colleagues and other scientists

Professional behavior: highest standards of integrity, commitment to ethical conduct of research, professional development, and thoughtfulness regarding broader implications of your work.

- Act in an ethically responsible manner (integrity, honesty, appropriate conduct) at all times
- Recognize limits of your knowledge, skills, and behavior through self-reflection and seek to overcome these limits
- Consider the broad implications of professional actions, including implications for society and the living world

Self Direction: develop good habits and skill for self-directed and life-long learning. Recognize that computational biology is a rapidly evolving discipline. We focus on developing adaptive, flexible and curious scientists, able to assimilate new ideas and technologies. [able to develop new directions of research as conditions warrant].

- Recognize need to engage in lifelong learning, stay abreast of the new
- Locate, evaluate, and assimilate relevant new knowledge and techniques from a wide variety of sources

Being a Graduate Student

Types of papers:

- analysis of new data and demonstration of new understanding of biological reality,
- introduction of a novel (useful) method,
- new understanding of a problem, clarification
- mathematical proof (e.g., proof of hardness),
- demonstration of substantial speed/efficiency improvement,
- collation of a useful database.

The point here is not so much the division into categories, as it is to start thinking about these things. Write papers and drafts as early and as professionally (e.g., clean, legible figures) as possible.

What makes up a dissertation:

- a hardbound stack of paper?
- Years of toil (regardless of the product)?
- Your new-found abilities to research, write and communicate?
- Publication of manuscripts?
- Your ability to define and carry out a scientifically relevant research project?
- To pluck the wheat from the chaff, and to know when to finish or discard something to not get bogged down?
- Meeting people in the field and joining a research community?

Who do you want to be when you finish?

- What is going to be your central area of expertise, where nobody knows more than you do?
- What is your more broad area of expertise, in which you are well educated and easily able to hold up your end of a discussion?
- How good is your understanding of biology and computation in the broadest sense?

What else should you be doing?

1. Seek out money wherever you can, now and when you're heading for a postdoc. Get an NSF DIGG. Write for NIH, NSF, Burroughs-Wellcome, Ford, Gates, UNCF/Merck, Sigma Xi, Etc. **Search the web**.
2. Begin **research** early. Often. Start projects. Don't be afraid to begin.
3. **Seek out training**, in and out of class
    a. Build relationships with **multiple professors**
    b. Learn **multiple techniques**, **approaches** to research (rotations)
4. Set **publishing goals** (early and often)
5. Consider your **long-term research goals**: what **kind of scientist**? What will you **know**? What will you be **capable** of?

What will you do **beyond graduate school**?

1. Consider your **options**. Postdoctoral **fellowships**. **Overseas**?
2. **Faculty** positions. Teaching load, graduate students, setup and facilities, type of college or university, type of department (who will your colleagues be?)
3. **Research institutes**. **Industry**.

Carpe Diem!

1. Changing **demographics** affect job opportunities (as do recessions)
2. Get to know **funding programs** and opportunities, and **program officers**
    a. NSF, NIH, USDA, NASA, DARPA, DOE, Foundations
    b. NLM, NIGMS, NHGRI, etc, etc, etc.
3. Study how to **write a grant**
4. **Seek advice**
    a. In your field and relatively far from it
    b. Leave **time to get feedback** and incorporate it
    c. Be careful with **perfectionism and pride** (timing, amount)
5. **MentorNet** (www.Mentor.Net) The e-mentoring network for diversity in engineering and science