Quantitative Training in the Life Sciences: Designing an Undergraduate Curriculum in Computational Biology

Louis J. Gross
Ecology and Evolutionary Biology and Mathematics
University of Tennessee, Knoxville

Overview:

1. Curriculum development and the structure of interdisciplinary reform of quantitative training in the life sciences - the CPA approach

2. Background from past and ongoing projects for quantitative training of undergraduates

- **a. Workshop summaries**
- **b. Dual approach to quantitative training**
- **c. Key quantitative concepts for undergraduate biology**
- **d. Content of an entry-level math course**
- **e. Quantitative enhancement of life science courses**

3. Computational biology training

- **a. Dual approach**
- **b. Flexible approaches for different audiences**
- **c. Short course content for advanced undergraduates**
- **d. Designing a full curriculum - limitations and options**

Computational Biology:

The U.S. National Institutes of Health is being urged to create biocomputing centers and enhance the development of computer-savvy biologists (Science 284 June 11, 1999)
David Botstein says "You can count on the fingers of one hand" the number of researchers with top-flight training in both biology and computer science.

How do we go about meeting the needs for researchers and educators in computational biology? How do we modify our courses and curricula so students have the opportunity to obtain jobs in the field?

Building on experience in quantitative training of undergraduates, I suggest a *Dual Approach is Needed*:

1. Enhance the opportunities for biology undergraduates and graduate students to obtain training in mathematics/computer science that is directly applicable to their biological interests and opens up the possibility of positions in the field.
2. Provide opportunities for computer science/math undergraduates and graduate students to develop interests in biological areas requiring extensive computation.

How do we go about doing these? I believe flexibility is important in a developing interdisciplinary field. This means multiple approaches are important.

Some Possible Approaches:

1. Short courses
(a) for undergraduate life science students
(b) for undergraduate computer science/math students
(c) for graduate students and faculty in life sciences
2. Formal year-long courses in computational biology
(a) for undergraduate life science students
(b) for undergraduate computer science/math students
3. Summer institutes for graduate students from both fields to work collaboratively on problems in computational biology.
4. Development of interdisciplinary curricula with degrees in computational biology.

What are the key concepts/skills life science students need to be successful in computational biology?
In an focus article in Science on Careers in Bioinformatics, John Greene (Senior staff scientist at Gene Logic, Inc.) suggests the needed skills include:
knowledge of UNIX, ; a good grasp of the concept of relational databases and skill with Structured Query Language (SQL); knowledge of object-oriented databases; programming skills in C, Perl, Java; knowledge of sequence-analysis programs like BLAST and FASTA; Web skills.

The above focuses on the CS skills, but I would argue that just as important are the theoretical underpinnings of genomics, mainly combinatorics and other areas of discrete math.

Attention Biology Majors

Do you want a high paying job ($40,000 - $80,000 typical to start) without completing another degree?

Consider a career in Bioinformatics! [1]

Use your biology skills to:

**** analyze fascinating basic biological questions**

**** help derive new drugs**

**** help find new cures for genetic diseases.**

This can be in your future if you sign up for Computational Biology 401- a course designed to teach you about genomic databases, UNIX, C and Java programming, and combinatorics. All the skills you will need for a new career in bioinformatics!

[1] **Warning: This course will be the most difficult you have ever taken, requiring enormous hours of effort, great concentration, dealing with completely new concepts, and utilizing all those mathematics skills that you thought you could forget. THIS IS NOT FOR WIMPS.**

Curricula in Computational Biology:

Many institutions are now developing new graduate programs in computational biology. At the undergraduate level this has really not started yet. What might be included in such a curriculum and how do we do it?

To overcome compartmentalization of fields of study - do this within a current program i.e. have a degree in Biology with a concentration in computational biology.

Rough curriculum:
2 years of math with one year calculus/differential equations and one year of discrete math/probability/statistics
2 years of computer science with intro courses on data structures, compilers, operating systems, networks and databases.
2 years of chemistry including biochemistry
3 years of biology including molecular biology, genetics, and special courses on genomics.

Is the above possible? Only if we work collaboratively to entice our best science students into this, which means the pre-medicine students.

What about Computer Scientists? Collaboration is the key - they will learn adequate biological skills if they carry out team projects with biologists.