Cell Biology Education
Vol. 1, 101–104, Winter 2002

Meeting Report: Incorporating Genomics Research into Undergraduate Curricula

Betsey Dexter Dyer*,‡ and Mark D. LeBlanc†

*Department of Biology and †Department of Mathematics and Computer Science, Wheaton College, Norton, Massachusetts 02766

Submitted June 27, 2002; Revised August 18, 2002; Accepted August 26, 2002

In the first of two National Science Foundation (NSF)–funded workshops, 30 professors of biology and computer science from 18 institutions met at Wheaton College in Norton, Massachusetts, on June 6–7, 2002, to share ideas on how to incorporate genomics research into undergraduate curricula. The participants included nine pairs or trios of biologists and computer scientists, anticipating or already implementing collaborations. In a before-and-after format, the two workshops are intended to encourage experimentation in the classroom (June 2002) followed by reflection on and evaluation of ideas (June 2003).

Interdisciplinary work is the heart of any research in genomics. The magnitude of the data sets and the scale of the problems require the expertise of individuals from both computer science and biology or biochemistry (Brown, 1999; Colwell, 2002). Collaborations in the sciences are common in many research settings but can be difficult to implement in undergraduate classrooms, especially in a relatively new field such as genomics. Biology students need opportunities to ask original questions and participate in algorithm and software design to learn how to set up new projects with programmers. Likewise, programmers entering the field of genomics must have a richer facility with the types of analyses and hypotheses that are useful, and programmers could benefit from interactions with biologists who will be using their programs. Our goals and objectives for this initial workshop targeted these needs and are listed in Table 1.

OUTCOMES OF THE INITIAL WORKSHOP

Following is a discussion of the six main outcomes of the June 2002 workshop.

Confidence

From our perspective, one of the more significant outcomes of the first workshop was the overwhelming number of participants who left feeling more confident about their ideas and feeling justified that new teaching models are needed. The inherently interdisciplinary nature of bioinformatics and genomics can trigger unease in some departments with faculty who have taught only within a strict departmental paradigm. Viewing a healthy dose of examples of innovative models is especially helpful for junior faculty. The workshop discussion focused on two models, using a radar metaphor, for introducing new teaching models. One model was to introduce changes subtly by "flying under the radar." Specifically, several action items were recommended that do not necessarily need full departmental or institutional approval, including 1) handling some of your own recruitment and publicity by means of distributing flyers and creating professional-looking web sites, 2) linking your course with that of a professor in another department (see Linked Courses, which follows), 3) infusing genomics content into your syllabus, and 4) advising students to package courses in a less-than–official major concentration. A number of participants suggested the alternative approach of "flooding the radar"—that is, lobbying department members, administrators, alumni, and trustees with the advantages of and the need for courses and programs in bioinformatics. Participants also noted that most college publicity departments are eager to send out press releases about even the most modest research or classroom efforts in genomics.

Compare and Contrast Programs

Although a few participants would have preferred only a small liberal arts constituency, the workshop purposely featured a faculty with experiences in a wide range of models, including full programs at large universities (e.g., Rensselaer Polytechnic Institute [RPI] and Wright State University), new or planned majors or concentrations at small colleges (e.g., Colby, Ramapo, and Trinity Colleges), single dedicated courses in bioinformatics/genomics (e.g., Drake University and Williams College), and a series of courses with infused content (e.g., Dickinson College and Wheaton College). Participants noted that there are currently three major "flavors" of curriculum design for bioinformatics or genomics:

1. Biologists using preexisting software tools who have little need to interact with computer scientists
2. Biologists learning enough programming (such as in Perl) to write their own software who may undertake a reciprocal effort to encourage computer scientists to learn enough biology to formulate their own hypotheses

DOI: 10.1187/cbe.02-07-0016
‡Corresponding author. E-mail address: bdyer@wheatoncollege.edu.

Table 1. Goals and objectives for the first of two workshops on incorporating genomics into undergraduate curricula

1. That participants come away feeling energized, enlightened, confident, and justified in their current and future interdisciplinary endeavors in genomics with undergraduates
2. That participants have an opportunity to try out ideas with colleagues and obtain useful feedback
3. That participants compile some useful tips for facilitating classroom work on all levels—especially for interdisciplinary collaborations in genomics
4. That participants find (or renew acquaintances with) a core group of colleagues in the Northeast with whom to exchange ideas in the future
5. That participants begin or continue plans to infuse genomics into their curricula during 2002–2003 and report back in June 2003

3. Biologists and computer scientists collaborating to create new software tools to answer original questions not necessarily answerable with preexisting programs

A more detailed summary of work at each participating institution is shown in Table 2.

Infusing Genomics

The participants who are new to bioinformatics/genomics especially appreciated ideas for infusing genomics content into their existing courses. Regardless of the size of the institution, faculty in biology in particular teach the same or similar courses at the undergraduate level (e.g., genetics, molecular biology, evolution.) The uniform resource locators (URLs) in the References section point to some of the institutions that are sharing course materials. Eventually, we envision that some of these materials can help launch a genomics portal at the upcoming National Science, Mathematics, Engineering, and Technology Education Digital Library (NSDL).

Linked Courses

One original goal of the workshop was to find pairs of collaborators (biologists and computer scientists) from the same institution who may want to experiment with linking their courses. The idea for the workshops stems from our conviction that undergraduate institutions must do more to prepare students for a world of interdisciplinary work. For the past 3 yr at Wheaton College, we have infused genomics content into our respective courses within an interdisciplinary context by linking our biology and computer science courses. We define linked courses as two independent courses with certain shared elements; for example, the professors give reciprocal guest lectures, students from both classes come together for 4 or more of 12 lab periods, students collaborate on software specifications and designs, and teams of computer science and biology students work on a capstone research project. Our goal of infusion is intended to both complement and extend full programs in bioinformatics that are beginning to emerge at some colleges and universities. For most colleges, especially small liberal arts colleges such as ours, the infusion of genomics in linked courses will reach far more students than a new major would. The suite of courses for a bioinformatics major would attract an extremely small number of students, some of whom would be likely to switch out of already small and rigorous programs such as biochemistry. In the pilot iterations of our linking model at Wheaton (Algorithms and Genetics; Algorithms and Cell Evolution), we reached 100% of the computer science majors of a given class year (Algorithms is a required course) and approximately 70% of all majors in the biological sciences (biology, biochemistry, environmental science, and psychobiology).

Collection of Resources

Our genomics web site at Wheaton College (see References) will organize and collect links to sites providing pedagogical ideas on linked teaching and student collaborations, suggested homework and projects, and sources of bioinformatic and algorithmic content, including pointers to forthcoming textbooks (e.g., Campbell and Heyer, 2003; Krane and Raymer, 2003) and rich web sites of links (e.g., Bagga, 2002).

Communication Across the Disciplinary Boundaries

Forming interdisciplinary partnerships and teaching models is nontrivial, as evidenced in this suggestion to biologists who want to approach a computer scientist:

Wrong way: “Want to work on bioinformatics with me?”
Right way: “I hear you are an expert in pattern matching. Could you help me solve . . . .?”

Table 3, based on both formal discussions and informal discussions during the workshop, summarizes some of the differences between the two disciplines. These differences are often at the core of communication problems. Discussions began at Workshop I and are anticipated to continue into Workshop II as the participants experiment with ways to help both students and colleagues cross those boundaries.

Also, although many computer scientists could probably summarize what biologists do for research, the reverse is usually not true. Biologists often think of their computer science colleagues as resources for building web pages or troubleshooting commercial software packages. Computer science research might appear to biologists to be an odd subset of mathematics because so many computer scientists are still housed within mathematics departments, especially at small colleges. Therefore, biologists (both professors and their students) may underestimate the degree to which their computer science colleagues might enjoy the opportunity to apply their problem-solving and tool-building expertise to a large, meaningful data set such as a genome.

At Wheaton College, we have found it valuable to model and deconstruct for our students some of the working parts of our (LeBlanc and Dyer’s) research collaborations in genomics. Our class discussions on communication differences can facilitate interdisciplinary work and help to prepare students for the realities of working in teams at their future jobs. Our research has led to several publications with students (e.g., LeBlanc *et al.*, 2000), poster presentations (e.g., LeBlanc *et al.*, 2002), and a web site where we host our tools (see Wheaton College Genomics Group URL in References).

SUMMARY

In short, participants ranked the workshop a huge success toward helping them develop ideas for incorporating genomics

Table 2. List of participant institutions and various modes of integrating genomics

Participating institution	Currently have or have plans for dedicated course(s) in bioinformatics/ genomics	Currently have or have plans for interdisciplinary teaching	Currently have or have plans for infusing genomics into existing courses	Currently have or have plans for a major concentration or minor	Full undergraduate bioinformatics program in progress
Colby College Waterville, ME	X	X	X	X	
College of the Holy Cross Worcester, MA			X		
Dickinson College Carlisle, PA	X		X		
Drake University Des Moines, IA	X		X		
Ithaca College Ithaca, NY			X		
Molloy College Rockville Centre, NY			X	X	
Mount St. Mary College Newburgh, NY		X			
Quinnipiac University Hamden, CT		X	X		
Ramapo College Mahwah, NJ	X	X			X
Rensselaer Polytechnic Institute (RPI) Troy, NY	X	X	X	X	X
Springfield College Springfield, MA		X	X		
Saint Michael's College Colchester, VT	X		X		
State University of New York (SUNY) at Cortland, NY			X		
Trinity College Hartford, CT	X			X	
Union College Schenectady, NY	X		X		
Wellesley College Wellesley, MA			X		
Wheaton College Norton, MA	X	X	X	X	
Williams College Williamstown, MA	X		X		
Wright State University Dayton, OH	X	X	X	X	X

Table 3. Communication differences between biologists and computer scientists (based on formal and informal discussions)

Biologist	Computer scientist
Uses tools and equipment	Designs and makes tools
Focuses on some aspect of biology; uses trial and error methodically; uses scientific method	Focuses on (is motivated by) tool user (sometimes muse-like); thinks trial and error is not the best model for designing or debugging (although is surprisingly resilient when trying trial and error)
Usually has little experience with large data sets	Regardless of experience is sometimes overconfident about ability to scale up to any size
Is likely to think in small subsets of larger problems; thinks of effect of X on Y	Has a "macho" attitude toward huge problems; thinks of the effect of all possible Xs on all possible Ys
Has a large "jargony" vocabulary full of Greek and Latin etymology	Has a large "jargony" vocabulary of acronyms and fanciful etymologies
May write up a table of communication differences between biologists and computer scienctists	Probably would not write up a table of communication differences between biologists and computer scientists

into undergraduate curricula. Bioinformatics and genomics are new fields of research, and teaching in these areas is newer still. In many ways, this initial workshop began a dialogue between colleagues and uncovered issues and questions that we must address together so that we can design labs, collaborations, and programs to meet the needs of the next generation of students. These open questions include the following five:

1. What are the right courses or tracks for students in bioinformatics/genomics?
2. What concepts in computing and in particular what types of algorithms do biologists really need to know?
3. What are the biological concepts that computer science students need to learn for them to be effective players?
4. What are the differences in language between biology and computer science and what can we do pedagogically to educate our students to work with these differences?
5. What are the different ways biologists and computer scientists solve problems (e.g., the scientific method vs. decomposition/encapsulation)?

The pursuit of answers to these and other emerging questions is part of our relentless effort to bring genomics to undergraduate biology and computer science students.

ACKNOWLEDGMENTS

This material is based on work supported by the NSF under Grant DUE No. 0126643.

REFERENCES

Brown, S.M. (1999). Dealing with genome project data. Biotechniques *26*, 266–268.

Campbell, A.M., and Heyer, L.J. (2003). Discovering Genomics, Proteomics, and Bioinformatics, San Francisco: Benjamin Cummings.

Colwell, R. (2002, April). NSF director: "New bioscientist" will be future of genomics [Interview]. Genome Technol. (20), 72.

Krane, D.E., and Raymer, M.L. (2003). Fundamental Concepts of Bioinformatics, San Francisco: Benjamin Cummings.

LeBlanc, M., Aspeslagh, G., Buggia, N., and Dyer, B. (2000). An annotated catalogue of inverted repeats of *Caenorhabditis elegans* Chromosome III with observations concerning odd/even biases and conserved motifs. Genome Res. *10*(9), 1381–1392.

LeBlanc, M., Baron, M., Christoforou, A., Doolittle, N., Kimball, M., Villa, A., Williams, G., and Dyer, B. (2002). The DNA motif lexicon—cataloguing and annotating genomes. Presented at the 14th International Genome Sequencing and Analysis Conference (TIGR), Boston.

Relevant Web Sites

Bagga, P.S. (2002). Bioinformatics resources on the Internet. http://orion.ramapo.edu/~pbagga/binf/binf_int_res.htm

Bioinformatics: A Faculty Development Course—Computer Applications in the Biosciences. http://www.haverford.edu/biology/genomicscourse/genomics.htm

Dolan DNA Learning Center, Cold Spring Harbor Laboratory. DNA from the Beginning. http://www.dnaftb.org/dnaftb/

Genome Consortium for Active Teaching (GCAT). http://www.bio.davidson.edu/Biology/GCAT/GCAT.html

Honts, J. (2002). Introducing bioinformatics within the undergraduate cell biology curriculum. http://www.cellbio.drake.edu/ASCB/present.html

National Science, Mathematics, Engineering, and Technology Education Digital Library (NSDL). http://www.nsdl.nsf.gov

Raymer, M.L., Doom, T., Krane, D. (2002). Introduction to bioinformatics. http://birg.cs.wright.edu

Wheaton College Genomics Group.
http://genomics.wheatoncollege.edu