BUILDING THE NEXT GENERATION OF QUANTITATIVE BIOLOGISTS

KRISTINE A. PATTIN

Institute for Quantitative Biomedical Sciences, Dartmouth College
Hanover, NH 03755, USA
Email: Kristine.A.Pattin@Dartmouth.edu

ANNA C. GREENE

Institute for Quantitative Biomedical Sciences, Dartmouth College
Hanover, NH 03755, USA
Email: Anna.C.Greene@Dartmouth.edu

RUSS B. ALTMAN

Department of Genetics, Stanford University
Stanford, CA 94305, USA
Email: russ.altman@stanford.edu

LAWRENCE E. HUNTER

Computational Bioscience Program, University of Colorado School of Medicine
Aurora, CO 80045, USA
Email: larry.hunter@ucdenver.edu

DAVID A. ROSS

Celera
Alameda, CA 94502, USA
Email: David.Ross@celera.com

JAMES A. FOSTER

Department of Biological Sciences, University of Idaho
Moscow, ID 83844, USA
Email: foster@uidaho.edu

JASON H. MOORE

Institute for Quantitative Biomedical Sciences, Dartmouth College
Hanover, NH 03755, USA
Email: Jason.H.Moore@Dartmouth.edu

Many colleges and universities across the globe now offer bachelors, masters, and doctoral degrees, along with certificate programs in bioinformatics. While there is some consensus surrounding curricula competencies, programs vary greatly in their core foci, with some leaning heavily toward the biological sciences and others toward quantitative areas. This allows prospective students to choose a program that best fits their interests and career goals. In the digital age, most scientific fields are facing an enormous growth of data, and as a consequence, the goals and challenges of bioinformatics are rapidly changing; this requires that bioinformatics education also change. In this workshop, we seek to ascertain current trends in bioinformatics education by asking the question, “What are the core competencies all bioinformaticians should have at the end of their training, and how successful have programs been in placing students in desired careers?”

1. Background

Bioinformatics is an intrinsically multidisciplinary field, which makes educating students across this educational continuum difficult. Therefore bioinformatics education has been dubbed an NP-hard problem.[1] In 1998, Professor Russ Altman led the charge to formalize bioinformatics education at the graduate level. He specified five competency areas for bioinformatics training: biology, computer science, statistics, ethics, and core bioinformatics yet cautioned against defining the curriculum too narrowly.[2] Previous workshops, RECOMB-BE and ISMB-WEB, have discussed bioinformatics education at the graduate and undergraduate levels[1], and the AMIA has suggested a list of biomedical informatics core competencies.[3] Defining bioinformatics curricula is a global effort.[4–8]

1.1. *Challenges*

There are many challenges associated with bioinformatics education, including faculty/instructor knowledge, computing resources, and breadth of knowledge required for bioinformatics training.[9] One of the major challenges graduate programs face is a lack of widespread adoption of these courses at the high school and undergraduate levels. Professors Ned Wingreen and David Botstein state, "The problem begins early in undergraduate education, and by the doctoral level there are severe interdisciplinary communication difficulties that are encountered by even the most motivated of collaborators."[10] Many have examined tactics to bring bioinformatics into the high school classroom in order to make the connection between biology and computation.[11–13] Others have incorporated bioinformatics into biology, chemistry, and computer science undergraduate courses.[14–23] Summer programs have been shown to be efficacious in helping students step into bioinformatics careers.[24] Strategies to mitigate problems surrounding this educational divide are paramount to having an efficacious program that trains successful students.

Another challenge for graduate training programs is the depth of multidisciplinary understanding needed to produce bioinformatics scientists instead of technicians.[9, 25] Ranganathan et al. have proposed a "minimum skillset" for bioinformaticians[26], and we would like to further address this to ask, "What are the key concepts and skills that one must graduate with in order to be successful today?" The modern biomedical researcher must be able to speak more than one language to successfully collaborate in a highly interdisciplinary environment; therefore it is beneficial to train a new generation of researchers who are well versed in the quantitative biomedical sciences and thus crucial to understand how such training programs succeed or fail. It would be valuable to faculty and directors of current programs or for those interested in bringing multidisciplinary programs to their institution to learn approaches and strategies for program development from existing and nascent programs. A few of the presentation and discussion topics include: curriculum design, effectively integrating multiple disciplines into a training program, getting first year students up to speed, the impact of a bioinformatics training program on faculty research, effectiveness (what works & what does not?), online curricula, student success, and the skills and tools students should have at program completion.

2. Workshop Contributions

Russ B. Altman is the Director of the Biomedical Informatics Training Program (BMI) at Stanford University. BMI is an interdisciplinary program that focuses on learning to develop and apply quantitative and computational methods to various biological and medical problems. BMI students represent a spectrum of educational backgrounds such as biology, research and clinical medicine, computer science, statistics, engineering, and a number of other fields. The faculty members come from a broad range of departments, including Bioengineering, Computer Science, Genetics, Medicine, Pediatrics, Radiology, and Statistics to provide research training and coursework in a number of related fields.

Ph.D. candidates are required to take core BMI classes, electives in computer science, statistics, mathematics, engineering, and allied informatics-related disciplines, coursework in social, legal, and ethical issues, and have access to unrestricted electives. BMI also aims to ensure students develop the ability to communicate their ideas and research effectively by having requirements and opportunities to present in journal clubs, colloquia, and lab meetings. BMI has successfully produced over 100 graduates who have pursued a number of different careers and can be found as distinguished faculty at top universities and medical schools as well as industry leaders at major corporations and startups.

James A. Foster is a Professor of Biological Sciences and Founding Member of the Institute for Bioinformatics and Evolutionary Studies (IBEST) and the Bioinformatics and Computational Biology (BCB) graduate program at the University of Idaho. Students in the BCB graduate program participate in research that encompasses the disciplines of computer science, biology, mathematics, and statistics and are affiliated with IBEST. BCB Ph.D. students may choose to focus their training in either the computer/mathematical sciences or the biological sciences.

Faculty members include physicists, chemists, molecular biologists, organismal biologists, ecologists, behavioral biologists, mathematicians, statisticians, and computer scientists. With access to such resources and technology in numerous scientific disciplines, BCB graduate students can no doubt become versatile members of the scientific community in whatever career path they decide to follow.

Lawrence E. Hunter is the Director of the Center for Computational Pharmacology and the Computational Bioscience Program at the University of Colorado Denver. The Computational Bioscience Ph.D. program curriculum is designed to integrate training in both computation and biomedical sciences with a student's research and teaching activities. Therefore, graduate students are expected to emerge from this program as independent researchers with a solid foundation in computational methods and molecular biomedicine, the science and technology comprising the two, as well as the skills to collaborate, communicate, and develop computational approaches that can be applied to a wide variety of biological problems.

Faculty represent a breadth of scientific research with multiple appointments in the departments of Medicine, Pharmacology, Biometrics, Biochemistry & Molecular Genetics, Computer Science, as well as from National Jewish Health. Four key competencies: knowledge, communication, professionalism, and life-long learning skills structure the core of the program’s teaching philosophy. The Computational Bioscience Program recognizes that bioinformatics is a rapidly evolving field, and while these core objectives will remain steadfast, they are committed to continually reviewing, revising, and improving their curriculum to keep pace with this evolution.

Jason H. Moore is the Director of the Institute for Quantitative Biomedical Sciences (iQBS) at Dartmouth College. IQBS is based on the idea that biomedical research studies move through a series of related activities that require a specific skillset and a specific scientific language. In any given study, there are specialists who design the experiment, who collect and organize the data, who analyze the data and who interpret the data. The major areas of expertise are bioinformatics, biostatistics and epidemiology. While there is some overlap, there is no one discipline that incorporates understanding across the entire process.

To cover these three disciplines both in student research and coursework, the program is interdepartmental with faculty from the Departments of Biological Sciences, Community and Family Medicine, Computer Science, Genetics and Medicine at Dartmouth College and the Geisel School of Medicine. Numerous collaborations exist between QBS members and those in other Ph.D. programs at Dartmouth including the Molecular and Cellular Biology Program, the Program in Experimental and Molecular Medicine and the Graduate Programs at The Dartmouth Institute for Health Policy and Clinical Practice.

The overarching goal of QBS is to prepare students for productive careers as quantitative scientists in the biomedical sciences by cross-training Ph.D. students and providing in-depth collaborative experiences in specific applications. Successful students will be able to effectively lead biomedical research studies from start to finish or participate as interdisciplinary members of large collaborative groups. QBS is an innovative approach to graduate training that combines multiple disciplines to train the next generation of collaborative scientists.

David A. Ross is the Director of Computational Biology at Celera. Celera is well known for its original mission to sequence the human genome and subsequently provide clients with early access to this data. Since then Celera has made important contributions to scientific research by developing “shotgun” sequencing and commencing the Applera Genomics Initiative, an effort that identified over 40,000 novel SNPs which became the foundation for new genetic tests that Celera is developing. Today, they provide a number of services including diagnostic products used for personalized disease management. Celera represents the interdisciplinary nature of biotechnology, where a student with interdisciplinary training would be an asset in many ways. Dr. Ross’ perspective on the bioinformatics training needed to be successful in an industry-based career will be welcomed.

References

1. S. Ranganathan, PLoS Computational Biology **1**, 6, 447–448 (2005).

2. R.B. Altman, Bioinformatics **14**, 7, 549–550 (1998).

3. C.A. Kulikowski, E.H. Shortliffe, L.M. Currie, P.L. Elkin, L.E. Hunter, T.R. Johnson, I.J. Kalet, L.A. Lenert, M.A. Musen, J.G. Ozbolt, J.W. Smith, P.Z. Tarczy-Hornoch, and J.J. Williamson, Journal of the American Medical Informatics Association : JAMIA **19**, 6, 931–8 (2012).

4. I. Koch and G. Fuellen, Briefings in Bioinformatics **9**, 3, 232–42 (2008).

5. D. Counsell, Briefings in Bioinformatics **4**, 1, 7–21 (2003).

6. M.S. Shamsir and Z.A.M. Hussein, AJTLHE: ASEAN Journal of Teaching and Learning in Higher Education **2**, 1, 30–40 (2010).

7. M. Gerstein, D. Greenbaum, K. Cheung, and P.L. Miller, Journal of Biomedical Informatics **40**, 1, 73–9 (2007).

8. R.B. Altman and T.E. Klein, Journal of Biomedical Informatics **40**, 1, 55–8 (2007).

9. M.P. Cummings and G.G. Temple, Briefings in Bioinformatics **11**, 6, 537–43 (2010).

10. N. Wingreen and D. Botstein, Nature Reviews. Molecular Cell Biology **7**, 11, 829–32 (2006).

11. S.R. Gallagher, W. Coon, K. Donley, A. Scott, and D.S. Goldberg, PLoS Computational Biology **7**, 10, e1002244 (2011).

12. S.H. Wefer and K. Sheppard, CBE Life Sciences Education **7**, 1, 155–62 (2008).

13. J. McQueen, J.J. Wright, and J.A. Fox, PLoS Computational Biology **8**, 8, e1002636 (2012).

14. A.E. Bednarski, S.C.R. Elgin, and H.B. Pakrasi, Cell Biology Education **4**, 3, 207–20 (2005).

15. N.B. Centeno, J. Villà-Freixa, and B. Oliva, Biochemistry and Molecular Biology Education **31**, 6, 386–391 (2003).

16. S. Cooper, Biochemistry and Molecular Biology Education **29**, 4, 167–168 (2001).

17. A.L. Feig and E. Jabri, Biochemistry and Molecular Biology Education **30**, 4, 224–231 (2002).

18. J.E. Honts, Cell Biology Education **2**, 4, 233–47 (2003).

19. B.S. Chapman, J.L. Christmann, and E.F. Thatcher, Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology **34**, 3, 180–6 (2006).

20. J.L. Ditty, C.A. Kvaal, B. Goodner, S.K. Freyermuth, C. Bailey, R.A. Britton, S.G. Gordon, S. Heinhorst, K. Reed, Z. Xu, E.R. Sanders-Lorenz, S. Axen, E. Kim, M. Johns, K. Scott, and C.A. Kerfeld, PLoS Biology **8**, 8, e1000448 (2010).

21. J.A. Boyle, Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology **32**, 4, 236–8 (2004).

22. D. Weisman, Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology **38**, 1, 4–9 (2010).

23. L.L. Furge, R. Stevens-Truss, D.B. Moore, and J.A. Langeland, Biochemistry and molecular biology education : a bimonthly publication of the International Union of Biochemistry and Molecular Biology **37**, 1, 26–36 (2009).

24. B. Krilowicz, W. Johnston, S.B. Sharp, N. Warter-Perez, and J. Momand, CBE Life Sciences Education **6**, 1, 74–83 (2007).

25. P.A. Pevner, Bioinformatics **20**, 14, 2159–2161 (2004).

26. T.W. Tan, S.J. Lim, A.M. Khan, and S. Ranganathan, BMC Genomics **10**, Suppl 3, S36 (2009).