Teaching Statement - C. Titus Brown

The future of biology and biomedicine lies at the intersection of data gathering, hypothesis generation, and hypothesis testing. It is increasingly easy to gather large amounts of relevant data, but **interpretation of that data and connection of that data to downstream hypothesis-driven experimentation is still difficult and rare**. Effectively integrating large scale data analysis with "wet" biology requires the development of new perspectives and new tools, and their integration with current research programs; it also requires biology-immersed scientists who are computationally expert.

The teaching, training, and education challenges associated with this shift in biology are immense and will be one of the major focuses in biology education moving forward. I have been and will continue to work to identify effective teaching and learning practice, effective technical approaches, and basic assessment materials to tackle these challenges.

Courses and workshops

During my 5 years as an assistant professor, I have developed and executed a number of new classes. In addition to a CS undergraduate class in Database Backed Web Programming, I ran an interdisciplinary graduate seminar course on Open Challenges in Bioinformatics (for two years); a two-week intensive summer course on Analyzing Next-Generation Sequencing Data (now in its 5th year, and NIH funded through 2016); and a graduate class in Computational Science for Evolutionary Biologists, arranged through our NSF Center for the Study of Evolution in Action and teleconferenced across 2 other institutions (now in its 4th year).

I have also run a number of workshops across the entire spectrum of interdisciplinary bioinformatics, from several "zero-entry" workshops for biologists through to advanced Software Carpentry workshops for bioinformatics researchers.

There are a number of cross-cutting themes and challenges in my efforts. First, interdisciplinary teaching is quite difficult, especially given the varied backgrounds of students: the differences in terminology, motivation, and expertise in undergraduates and beginning grad students can be substantial. Second, I tend to teach at the leading edge of technology, be it sequencing and genomics or Web development, which means that keeping course material up to date has been a constant effort. And, third, assessment and evaluation of teaching effectiveness is *particularly* difficult in these nascent disciplines, where little prior development of assessment material has occurred. This has led to a secondary interest in learning and education research, which I have as yet had little opportunity to indulge but that I regard as critical to my future teaching work.

All in all, there are no silver bullets that simply solve these problems, but there *are* teaching principles and technical solutions that can be brought to bear. I have brought a wide range of approaches into the classroom to motivate and drive the topics I teach, including hands-on algorithmic exercises in assembly, cloud computing as a platform to enable real-world software development and use, interaction with best-practice industry approaches for programming, and use of recent publications and news events. Since I often teach what I use for my own research, keeping the teaching materials up to date serves the secondary purpose of building a reusable research knowledge base for my graduate students, collaborators, and colleagues; increasingly, some of our materials serve as a standard repository for the field of bioinformatics. Assessment (which I discuss more below) has lagged behind, largely due to challenges in finding funding to develop discipline-specific assessment tools; however, I have recently been able to find seed research funding from the NSF for this purpose, and I anticipate substantial additional funding in the future, as

many grant agencies recognize the importance of technical training in my area of research.

As evidence of some success in my teaching endeavors, my students gave me the Computer Science Teaching Award for my first Web development course in 2008; and this year I received the MSU College of Natural Science Teacher-Scholar Award, and have been nominated for the MSU-wide Teacher-Scholar Award. More importantly I have repeatedly been told by my students that my enthusiasm for teaching, my interest in student development, and my engagement with the material at multiple levels of theory and practice, has been a strong positive force in their personal development.

Teaching philosophy: inductive learning

My overall teaching philosophy is to encourage the students to "learn by doing." While sometimes viewed as an overly practical approach, I have always found substantial opportunity for integrating theoretical considerations into courses. For example, my Web development course starts by having students implement their own Web server, together with automated tests and version control; at this point, a number of opportunities open up for discussing scalability, multiprocessing, and security, rooted in the implementation of an HTTP server but reaching out to all of the theoretical considerations inherent in widely distributed application services. Similarly, in sequence analysis, I discuss both the theoretically achievable results and then demonstrate what existing software can actually do; one is an upper bound, the other is practically achievable, and students need to be aware of both sets of limitations. Students seem to respond well to this approach, as I connect theory directly to relevant real-world examples.

Open materials and cloud computing

I believe strongly in making my educational materials openly and freely available, and have posted all of my materials and lectures online under a Creative Commons/0 requirements license. In practice this means that my materials serve as Google-searchable resources for my field; in particular, my NGS course materials (at http://ged.msu.edu/angus/) receive well over a hundred thousand Web hits a year.

My materials also rely largely on cloud computing, which has made them useful and reusable beyond our local compute/IT infrastructure. I am now working to produce a set of inverted lectures on YouTube that will enable more people to make use of my bioinformatics materials; I am also developing an approach to scale bioinformatics training to thousands of people, using cloud computing and a hybrid MOOC approach. The plan is to rely on local community building to identify technical resources that can help enable local learning based on globally available resources – an inverted MOOC, if you like.

Assessment

The most difficult aspect of my teaching has been to determine whether my students receive any lasting benefit from our courses or workshops. We have so far relied on multiple choice questionnaires and open-ended problem-based essays, which have shown that both concept familiarity and self-confidence levels have increased as a result of attending our workshops.

The most in-depth assessment of my teaching has been done as part of the two-week summer workshop, where it was funded by the NIH. We engaged StemEd LLC to develop assessment materials, apply them independently to the students, and provide evaluation reports each year. Three primary modes of assessment were applied in 2012. First, a pre- and post- written questionnaire

was given to the students to evaluate their baseline knowledge and gains in knowledge during the course. Second, several lectures were observed by personnel from StemEd LLC. And third, the post-mortem session was attended by evaluation personnel. This culminated in an assessment report that concluded that some relatively minor refocusing was needed in the background of an overwhelmingly successful course. The 2013 assessment report has not yet been completed.

We are also planning to conduct post-course interviews of admitted and rejected students. The primary goals of this component of the assessment will be to determine how accepted students are making use of their knowledge, and whether and how rejected students are learning material on their own. We also hope to develop a better understanding of how advanced professionals learn computational practice.

I have also begun to collaborate more closely with our local Center for Engineering Education Research on developing more in-depth qualitative assessment materials for my graduate course. In collaboration with the Software Carpentry initiative (discussed below), I am also working on a "Driver's License" for data-oriented computing that will let us quickly assess and evaluate applicant levels so as to ascertain the appropriate teaching level for workshops.

Collaborative teaching

Much of my teaching has been driven by personal incentives, but I collaborate with a wide range of people in my actual teaching. In addition to co-teaching my summer workshops with Dr. Ian Dworkin (a statistician at MSU) and Dr. Istvan Albert (a bioinformatics director at Penn State) I routinely work with other instructors to deliver workshops and lectures. I am particularly interested in developing new tandem-taught biology courses at the introductory level, e.g. exploring microbiology, developmental biology, and/or evolution by bringing in discussions of genomics, bioinformatics, and computational inference.

Enhanced learning in the REAL Classroom

In spring 2013, I taught my upper-division undergraduate class in Web dev in a REAL classroom at MSU; these are classrooms that enable and encourage active learning approaches. I am again teaching in a REAL classroom in spring 2014. My preliminary experience is that these classrooms requires significant adjustment by the teacher and the students; one entertaining observation made by my TA was that the seniors in the class preferred lecture, as they had become used to being able to tune out during lectures! I am redesigning my course for spring 2014 with more online collaborative activities as well as more inverted lectures.

Software Carpentry

I have been a Software Carpentry instructor for over 7 years. Software Carpentry (http://software-carpentry.org) is an initiative and organization that develops material to teach scientists to do better and more efficient computing; it has run almost 100 workshops in the past year. I am a long-time supporter, sit on the advisory board, and help seek out funding and arrange workshops. I also teach several workshops a year at MSU and elsewhere. Some of my current funding has the goal of connecting the more advanced Software Carpentry material with the more introductory needs of molecular biologists.

Future plans

I will continue working to improve the approaches I and others have developed to teach at the intersection of biology and computing. My goals are to build reusable materials with demonstrated

effectiveness and partnered with standard assessment instruments; the challenge is to do this material development in a sustainable way and scale delivery to more trainees. I now have funding to do this from the NSF, and am anticipating applying for an NSF RCN in this direction in 2014. A number of funding agencies, including the NSF, the NIH, the USDA, the Moore Foundation, and the Sloan Foundation, have all become increasingly interested in data science training, and I am in contact with all of them; the goal of the RCN will be to coordinate between different communities and user bases and develop focused initiatives.

I am particularly interested bringing underrepresented minorities into the field of computational biology; I have been a part of our NSF Center's minority requirement efforts for several years, and helped teach an introductory workshop for undergraduates from HBCUs. Several of my graduate students have been involved in similar efforts that involve traveling to other Center partners, with my support and encouragement. However, there has been relatively little scope for engagement with underrepresented minorities at MSU during the school year. My last NSF CAREER proposal would have enabled me to fund undergraduate training during the school year for this purpose, but it was declined; I plan to resubmit next year.

My biggest personal challenge is to become more reflective in my teaching. I spend a significant amount of time actually teaching, but not nearly enough time exploring alternative teaching models and techniques. To address this, I plan to ask a teaching coach to observe some of my future workshops and classes, to give me feedback on my performance and approach.

Appendix – Material reference

The majority of our teaching materials are linked to from the ANGUS site, at http://ged.msu.edu/angus/.

- The full 2013 NGS course schedule and linked materials: http://ged.msu.edu/angus/tutorials-2013/
- The "zero-entry" U. Washington workshop for biologists: http://2013-uw-zero-entry.readthedocs.org
- Software Carpentry materials are all available at http://software-carpentry.org/
- An example "advanced bioinformatics/software development" workshop is online at http://2013-norwich-bioinfo.readthedocs.org/.
- My lab Web site at http://ged.msu.edu/interests.html contains links to several funded grant proposals for courses and material development.