Statement of Teaching

Nicholas Santantonio

Plant breeders have traditionally been generalists, combining genetics with a range of plant sciences to identify farmers' needs and turn out products to meet those needs. However, the range of skills required in the field is rapidly increasing. Proficiency in statistics, programming, bioinformatics and computational biology are now expected in addition to the traditional skills in physiology, pathology, agronomy and genetics. As educators, we must decide how to update our curriculum to best prepare students for careers in the new era of digital agriculture.

Both generalists and specialists will be needed for integrating genomics and digital agriculture into modern breeding programs. Generalists may require instruction in team management, whereas specialists may need more rigorous courses in areas previously considered outside plant breeding, such as engineering and computer science. Specialists will be vital for collecting and storing large amounts of data that will be used to model and predict crop growth and development. Generalists will coordinate those activities to lead a team that produces new varieties to meet farmers' needs.

The widening breadth of the plant breeding discipline may demand multiple paths of instruction. While I believe all future scientists will require some graduate level statistics and computational coursework, a more structured series of quantitative courses would benefit students seeking specialization. Students who chose this path would finish graduate school with the comprehension and ability to effectively adapt the latest computational techniques to meet future breeding goals.

Teaching philosophy

Inclusive teaching tactics are key to ensuring equal access to knowledge in the classroom. Expectations must be made clear, and reinforced throughout the semester so that students do not get behind and fail to meet milestones. It is important for students to be exposed to plant breeding ideas from multiple perspectives, and they must be given the opportunity to demonstrate critical thinking on different levels. While some students may show analytical thinking and synthesis on exams, others may shine in more hands-on projects. Most mathematical and computational learning occurs through doing, not through watching. The lecture is important to present material in a concise, structured manner, but concepts are cemented when the student can reconstruct the ideas on their own time.

Longer term projects provide students the opportunity to practice and apply concepts in a more autonomous environment, and are invaluable for assessing comprehension and critical thinking. Regularly assigned homework and hands-on labs are crucial for evaluation of the pace and overall understanding of the material presented. Greenhouse and field facilities on campus provide opportunities to get students out of the classroom to see plant breeding in practice. Public databases provide resources where students can get experience working with real datasets. Computer simulations are useful tools to evaluate comprehension, where in order to simulate a system correctly, the student must understand that system well.

All courses I instruct would contain a term project of relevant complexity to augment exams, in-class labs and homework assignments. For quantitatively oriented courses, these projects would be computational in nature, where students would use real or simulated data to explore the ideas covered in the course. They would then be asked to present their results in written and oral formats that mirror typical scientific communication. Projects may be team oriented to promote collaborative skills and project management.

I also intend to incorporate perspectives on how domestication, selection and seed systems work in different cultures through guest lectures and case studies. Review of how seed systems function

in other cultures will cultivate discussion on the benefits and drawbacks of these systems relative to the US and western European systems.

Courses

I intend to target 21$^{\text{st}}$ century plant breeding concepts for the Genetic Improvement of Crop Plants course (PLBRG4030), with a focus on the use of genome-wide information to drive breeding decisions. The course would include a hands-on computational component to reflect the skills currently desired in the field. In addition, I intend to develop a graduate course to augment PLBRG7170, that would be held in the alternate year, as well as continue to co-instruct PLBRG7420.

For PLBRG4030, computational labs would be used to augment student understanding of course material. Students would use available computational tools to analyze small example datasets with a focus on interpretation of results. For the term-project, students will be split into groups, and given a breeding scenario and a dataset. They will determine the genetic architecture of their trait and propose a breeding strategy based on their scenario and what they can learn from the data. By the end of the semester, students will be able to demonstrate critical thinking of plant breeding methods and ideas, with the ability to synthesize when given new plant systems or breeding goals.

The graduate course topic is flexible to the needs of the section, however, it is my observation that an introductory quantitative genetics course would be useful to help prepare students for further quantitative study. This course might also provide some necessary skills to those who choose not to follow a quantitative path, without the rigorous mathematical detail of the PLBRG7170 quantitative genetics course. I would work with other quantitative faculty to build a course that augments other courses current offered, with little overlap in content to avoid redundancy.

Concepts introduced in class would be reinforced through weekly computational labs and assignments that require students to write their own software to solve computational plant breeding problems. Assignments would be required to be submitted as typed documents in Markdown, LaTeX or similar format, to expose students to more effective modes of mathematical communication outside of Microsoft Office. The term project would then consist of groups of two to four students finding a genotype-phenotype dataset, and working to develop a genotype to phenotype map, assess genomic predictability and construct an optimized breeding scheme throughout the semester.

Curriculum

Currently, most life science students do not acquire in-depth statistics and programming skills until graduate school, impeding their progress while they learn to grapple with these new languages. Moving forward, I would like to work with faculty in SIPS, biometry, and computational biology to build a quantitative/computational genetics undergraduate curriculum. Genetics is a vast field; it is imperative that students get exposure to and training in the rapidly changing environment in which they will soon be seeking jobs. While single-gene genetics is still an important field, the community is ever more focused on the complex "omics" network that is the foundation of complex organisms. Students in essentially all sub-fields of genetics will need to be able to deal with large datasets, using tailored algorithms to make inferences, predict the unobserved, and guide decision making. Big data management requires skills in programming, linear algebra, statistics and machine learning that must be incorporated into the curriculum at earlier stages.

Genomics and digital agriculture are only just starting to change the landscape of food production. Quantitative skills are one of the specializations imperative in plant breeding, and Cornell must be at the forefront of best preparing the individuals who will usher in this new era.