Statement of Teaching

Nicholas Santantonio

Plant breeders have traditionally been generalists, combining genetics with a range of plant sciences to identify farmers' needs and produce products to meet those needs. However, the range of skills required in the field is rapidly increasing. Competitive plant breeders are now expected to be proficient in statistics, programming, bioinformatics and computational biology along with 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 post-genomics 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.

Due to the wide breadth of the plant breeding discipline, it may be pertinent to develop multiple paths of instruction. While I believe all future scientists will require some degree of graduate level statistics and computational coursework, a more structured series of quantitative and computational courses would benefit students seeking specialization. Students who chose this path would finish graduate school with the comprehension and ability to effectively use the latest computational techniques in plant breeding, including the use of genome-wide information for decision making.

Teaching Philosophy

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. In addition, regularly assigned homework and hands-on labs are crucial for estimation of the pace and overall understanding of the material presented, such that adjustments can be made where necessary. The 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, and often large, 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 consist of a hands-on computational component, where students either find a dataset or use their own 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. For courses without a significant quantitative aspect, term projects may be formulated as a research proposal.

Courses

I propose to teach the Advanced Plant Breeding course indicated in the position announcement every year to meet the teaching requirement of the position. I would also like to develop an advanced quantitative genetics course that would be offered every other year.

The Advanced Plant Breeding course would target 21st century plant breeding concepts, focusing on the use of genome-wide information to drive breeding decisions. Starting with basic probability theory and the single locus model, the course would advance through genome-wide association, genomic prediction and selection, as well as selection theory and breeding program optimization. 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 consist of groups of 2-4 students finding a genotype-phenotype dataset, and working to develop a genotype to phenotype map, assess genomic predictability and determine an optimal breeding scheme throughout the semester. By the end of the semester, students would be able to demonstrate critical thinking of how to apply quantitative genetics theory to plant breeding problems, with the ability to synthesize when given new plant systems or breeding goals.

The quantitative genetics course would be based on primary literature review, and offered as a 1 or 2 credit hour course that would meet once a week and cover 1-2 papers per week. Advanced concepts to be covered would including coalescent theory, hierarchical Bayesian models, spatial variation, G$\times$E, multivariate and selection indices, longitudinal models and optimal contribution. Specific computational examples would be developed for key topics and provided to students prior to meeting so that methods can be discussed in detail in class.

Students would be evaluated using a gain in understanding approach. This approach would require students to submit responses to questions developed to guide understanding of the reading materials prior to class, with the ability to resubmit revised answers after attending. Concepts and questions would be sufficiently complex that prior effort and in class learning can be assessed.

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 CSES, statistics, engineering and computer science 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, more and more the community is 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. Dealing with big data requires skills in programming, linear algebra, statistics and machine learning.

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 science, and Virginia Tech must be at the forefront of preparing the individuals who will usher in this new era.