

Statement of Research

Nicholas Santantonio

Small grains comprise an important part of Virginia agriculture, providing flexibility and diversity to the agricultural system. To compete with private small grains seed companies, the public breeding program must implement the latest breeding technology to achieve population improvement alongside varietal development and maximize the genetic response to Virginia environments.

I intend to develop the foundational capacity to implement 21st century breeding technology in the Small Grains Breeding Program at Virginia Tech, providing ample opportunities for extramural funding. This will include the collection, storage and accessibility of genome-wide information, pedigrees, high-throughput phenotypes collected through proximal sensing, and traditional phenotypes from field trials. Ground-breaking breeding technologies will be implemented to demonstrate the potential of these new resources and pique public interest in crop improvement at Virginia Tech.

Short-term gains

The most easily exploited terms in the breeders equation are cycle time and selection intensity. The number of lines evaluated can be drastically increased by reducing or eliminating replication in early-stage trials, thus increasing selection intensity. Genotyping some lines within each family will increase reliability and allow for estimation of genetic correlations of locations and prediction of genetic merit within locations. All materials that are advanced to the second year yield trials will be genotyped. Aerial phenotypes obtained through proximal sensing will be used in multi-trait models to substantially increase trial sizes and reduce the number of plots that are harvested. High-throughput phenotypes will also be used to account for spatial variation within the field for better estimation of genetic values of unreplicated lines, as well as monitor diseases such as scab.

Initially, the breeding cycle will be shortened by recycling materials after the second year of yield trials, using genotypic information and mathematical optimization to drive crossing decisions. A Smith-Hazel economic selection index including grain yield, days to maturity, deoxynivalenol (DON) content, and test weight will be used as the genetic merit criterion to transition to an optimal contribution genomic selection program. Other properties, including milling and baking quality in wheat, forage and malting quality in barley, straw content and quality, as well as other disease resistances will all be monitored and included in the decision making process.

Long-term sustainability

The infrastructure to support regular genotyping and proximal sensing must be developed before the program is overwhelmed by incoming data. While I intend to host a local breeding program database, I would like to work with the T3 database to host proximal sensing images along with phenotypic trial data, pedigrees and genotypes. Quality control of genetic markers, pedigrees and phenotypes must be implemented and standard operating procedures must be developed. Genotyping will allow for construction of a multi-trait, multi-environment genomic prediction model that will be updated yearly. Field experimental design will then be optimized to leverage all phenotypic information available, using genotypic information to link otherwise unrelated trials and traits.

Once genotyping and rapid recycling has been hammered out and sufficient genotypic and phenotypic information is generated (≤ 5 years), a recurrent population will be constructed using current breeding materials. This population will be subjected to rapid cycling twice a year to increase the genetic merit of the population, generate unique meiotic recombinations and minimize inbreeding using optimal contribution. Materials will be pulled out of the recurrent population

on a yearly basis for two generations of high intensity genomic selection and two generations of inbreeding in the greenhouse. Family sizes will vary based on predicted family means and variances. Materials will also be pulled directly out of the recurrent population and phenotyped for generating information highly connected to the recurrent population to increase accuracy of selection.

Breeding for synergistic organism interactions

The wheat-soybean double cropping system has allowed farmers to profit from growing winter wheat by production of a late season soybean crop, but breeding efforts have largely focused on improving each crop independently. I envision an exciting collaboration between the small grains and soybean breeding programs to breed for synergistic combinations suited for the double cropping system.

While the number of potential wheat-soybean pairs is very large and generally intractable, clever testing of select pairs will allow for genomic prediction of unobserved combinations. This is similar to the $G \times E$ problem, where one crop (e.g. wheat) can be thought of as an environment to which the genotype (e.g. soybean) is subjected to. In this case, the genetic correlation of genotypes *and* environments would be estimated with genome-wide markers. All pairwise combinations would then be predicted using multivariate linear mixed models. Promising predicted pairs would then be evaluated the following season, and crosses made to maximize the beneficial interaction effect.

These synergistic genomic selection models could also be adapted to other goals, allowing for much wider collaborations between animal breeders, microbiologists and soil scientists. Examples include forage/animal and forage/rumen-microbiome using forage barley as a model, root/soil microbiome using wheat as a model, and even barley/yeast interactions using malting barley as a model. Products specifically improved for one another could be marketed as a package (e.g. maltsters could offer specific yeasts with certain barley malts), or a service (e.g. genotyping a farmers field to determine optimal varieties for their soil microbiome). Synergistic improvement of multiple organisms has only just recently become feasible due to cost-effective genotyping platforms, and has the potential to change the way we breed within our food systems.

Research Philosophy

In the era of big data, the sheer amount of testable hypotheses is seemingly limitless. A shift away from small designed experiments to large observational studies at the breeding program or whole organism scale is inevitable. A traditional breeding program generates a plethora of phenotypic data that is used to make yearly breeding decisions, and subsequently discarded. If genotyped, these materials become treasure troves of data for asking questions, as well as making breeding decisions. This does not mean that we should cease the design and execution of experiments to address specific hypotheses, but we cannot ignore the valuable resource of observational data being collected, typically at great expense. Genotyping at this scale is feasible given the drastic reduction in costs and availability of third party services, and can be offset by clever experimental design that trades replication at an individual level for replication at a genetic level.

I believe in the collaborative model, where breeding programs do not operate in isolation. They share germplasm, resources, expertise, and most importantly, ideas. Unlike germplasm, ideas also have the merit of being species flexible. I intend to build a collaborative effort at Virginia Tech to aid all breeding programs to build foundational capabilities to increase efficiency of varietal development. As climate change progresses, Virginia will likely experience rising temperatures and more severe weather events, making crop production more difficult. Heat, drought, intense storms and hard frosts will be the new norm, and we must work together to do our part in defending our food security through accelerated genetic improvement.