Statement of Research

Nicholas Santantonio

Winter wheat comprises a substantial portion of Colorado's agriculture output, which relies on the public wheat breeding program at Colorado State University to provide improved varieties. I intend to develop the foundational capacity to implement 21$^{\text{st}}$ century breeding technology in this program, 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. Integration of these technologies with advanced statistical and machine learning methodology should maximize product development for Colorado growers and provide long-term sustainability of the breeding program. Demonstration of these new techniques and resources should help pique public interest in crop improvement at CSU.

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 will be used in multi-trait models to substantially increase trial sizes and reduce the number of plots that need to be harvested.

Initially, the breeding cycle will be shortened by recycling materials after the second year of yield trials, using genotypic information to drive crossing decisions. A Smith-Hazel economic selection index including yield and test weight will be constructed through consultation with local grain elevator operators. This index will be used as the genetic merit criterion to transition to a genomic selection program. Optimal contributions will then be used to maximize genetic gain while minimizing the reduction in genetic variance through time.

I also aim to target specialty flour markets, providing farmers with unique products that demand a premium price from millers. Selection indices can be modified to include milling and baking quality traits in the decision making process. These indices would be built in coordination with millers, to optimize selection for the specific flour qualities necessary for specialty products.

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. High-throughput phenotypes may also be used to account for spatial variation within the field, observe growth through time and monitor response to biotic and abiotic stress. These phenotypes may also be used in longitudinal random regression models to build genotype specific growth curves. Plasticity in growth curves may shed light on how response to stress allows for stability or sensitivity to varying environments.

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

A legume rotational crop with wheat could drastically improve soil health, nitrogen and long-term sustainability. The wheat-soybean double cropping system has allowed many farmers in the Midwest 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 intend to investigate the implementation of a legume double cropping system for Colorado using field peas. By breeding both organisms for synergistic combinations well suited for the tight double cropping system, I hope to produce variety pairs that mature early, yield well and sustain adequate soil moisture for the more valuable wheat crop.

While the number of potential wheat-pea genotype 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×E problem, where one crop (e.g. pea) can be thought of as an environment to which the genotype (e.g. wheat) 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 further beneficial G×G interactions.

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 root/soil microbiome using wheat as a model, forage/animal and forage/rumen-microbiome using alfalfa 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).

Research Philosophy

In the era of big data, a shift away from small designed experiments to large observational studies at the breeding program scale is inevitable. When genotyped and made publicly available with FAIR data principals, the vast amount of data generated in a breeding program becomes a treasure trove for asking questions and informing breeding decisions. Clever experimental design will offset genotyping costs by trading replication at an individual level for replication at the genetic level.

I believe in a collaborative model, where breeding programs do not operate in isolation. They share germplasm, resources, expertise, and most importantly, ideas. I intend to contribute to the collaborative effort at CSU and across the globe to build the foundational capabilities needed to deploy the latest technology for variety development. As climate change progresses, heat, drought, intense storms and hard frosts will be the new norm, and we must work together to help defend our food security through accelerated genetic improvement.