AG2PI SEED GRANT - PROJECT FINAL REPORT

PROJECT NAME	Developing a new machine learning tool for improved genomic selection

PROJECT PRINCIPAL INVESTIGATOR	TODAY'S DATE	PROJECT START DATE	DATE OF COMPLETION
James Polashock	6/29/23		5/30/23
TEAM MEMBERS (co-PI, co-I, personnel)		COLLABORATORS	
Joseph Kawash Iman Dehzangi			

NOTE: this report will only be shared with the AG2PI Executive Board and USDA.

ACCOMPLISHMENTS

Please provide a short summary of the conclusions (both successes and failures) made from your project. Include a description of how this project will provide benefits to the agricultural genome to phenome community and, possibly, to a broader audience. You should include both qualitative and quantitative details, as necessary, to support your conclusions. Include a short accomplishment statement in non-technical language and do not include names.

This is not a technical report. Please keep to no more than 6-8 sentences (e.g., 1-2 sentences per point, above).

This project was successful in developing a tool for genomic selection (GS) and fostering collaboration between USDA-ARS and Rutgers University research groups. Our collaboration raised interest and has expanded to include new members and projects to further explore other aspects surrounding GS including phenotyping and genetic evaluation.

-successfully built a tool that utilizes machine learning (ML) to generate markers for GS
--Used genotyping by sequencing (GBS) data from cranberry populations segregating from fruit rot and implemented it into a boosted regression tree structure to identify interacting markers to follow the fruit rot resistance trait. These markers were not able to be found previously using common QTL software.

-successfully validated genomic markers in a population
--The developed markers were then validated in a separate population segregating for fruit rot, that was not used during the training and development of the ML model. These markers successfully separated the population between rot susceptible and rot resistant individuals.

-engaged graduate, undergraduate students, and post docs in the research project
--During the project, graduate and undergraduate computer science majors worked with us to streamline the model and better parse through the imputed data.

-successful collaboration with between members that lead to further research
--This project generated a lot of interested between our USDA-ARS lab and Rutgers University lab groups. We have since expanded the project goals and aims to include improving the phenotyping component needed for GS and characterizing genome heterozygosity in cranberry. An additional 2 Rutgers Ph.D. students have joined the project, and it is expected that at least 1 more will be joining in the fall.

-engaged with other members of the USDA community on the project
--On the USDA-ARS side, another member is interested in using the tool for their own research and providing feedback for its development. We are currently adapting his projects and the tool for ML based GS for broader utility.

(HINT: You can expand sections as necessary by pulling down on the square in the lower right corner of each box)

Products

Please list any products from this project. This may include (but not limited to) publication, concept/white paper, workshop, conference presentation, website, publicly available data or pipelines, etc. Reminder: you are required to make your products available to the broader stakeholder community using standard USDA practices, open source, FAIR, or other models. Metrics may include number of participants or times accessed, etc. Include links to recordings, DOI, etc. when possible. For presentations and posters, provide authors, date, location and presentation title.

ACTIVITY / PRODUCT	DESCRIPTION (include URL, if applicable)	OUTCOME / METRICS
American Cranberry Association Meeting	Presentation: Identifying genetic markers for defense mechanisms in cranberry fruit Joseph K. Kawash, James J. Polashock, Jennifer Johnson-Cicalese, Nicholi Vorsa 1/20/22	Presented to USDA scientists, university researchers, and grower groups
AG2PI	Poster: Developing a new machine learning tool for improved genomic selection Joseph K. Kawash, Jeffery Neyhart, Lindsay Erndwein, James J. Polashock 6/15/23	Presented to AG2PI members
American Cranberry Association Meeting	Presentation: Updating Approaches to Fruit Rot Resistance Research Joseph K. Kawash, James J. Polashock 1/19/23	Presented to USDA scientists, university researchers, and grower groups
International Vaccinium Symposium	Presentation: Identifying epistatic loci for fruit rot resistance in cranberry using machine learning Joseph K. Kawash, James J. Polashock, Jennifer Johnson-Cicalese, Nicholi Vorsa 9/1/21	Presented to USDA scientists, university researchers, and scientists from other countries working on Vaccinium spp.

Audience

With whom has this work been targeted to and shared? Please describe how this project and its products have been disseminated to a community of interest. Include any outreach activity or information sharing as well as training or professional development opportunities provided in this project.

This project was targeted to researchers, both within the USDA-ARS and University colleagues, including graduate students and post docs. We have presented our work at international research conferences, national conferences, and growers' groups. We are in the process of publishing the results of this tool in cranberry.

CONTINUATION OF WORK

Next steps

How do you/your team plan to continue moving this project forward? Include how AG2PI can assist in your forward momentum.

Our team plans to move forward expanding the development of our GS tool as well as branching other projects. A large collaboration that has spun off from this is the development of AI/ML tools for crop phenotyping both in the field and lab based. Our team has onboarded 2 new graduate students who are very interested in the phenotyping aspects of our continued collaboration. Another graduate student will be joining in the fall and is interested in the genetics of GS. Specifically, they will be working with us to investigate the unbalanced effects of inbreeding that can be experienced during the process of generating populations for GS. We additionally gained another USDA member who is interested in using our tool for cranberry and blueberry GS projects and will be providing feedback on tool use.