# Impact Statement

## Inari Project

Victor Avram, Sergio Miguel Moya Jimenez, M. Eagon Meng, Wenhan Zhang

Our project aids Inari in building genetic knowledge through computational methods. Overall, Inari brings CRISPR gene editing technology to engineering plant seeds, so as to increase crop diversity and productivity, and to make more efficient use of water, land and fertilizer. Studies predict that food production has to increase as much as two times by 2050 to meet the food demands of the world’s growing population. Increasing crop yields can thus increase food security and reduce the burden on the environment and natural resources.

Our aim was to develop a model that allows us to predict the impact of perturbation of genes in the other genes. Even though this could bring benefits in the gene editing work of Inari, there is always the uncertainty of the effect of a bad prediction and its consequences on the consumers of maize. For example, given specific gene perturbations, if our computational methods fail to predict severe implications in other parts of the maize genome and lead Inari researchers to think that the perturbation is safe, it would lead to inadvertent complications in crop healthy, crop diversity, or worse, human health. Possible consequences of genetically modified crops for humans could be cancer and allergic reactions. As a result of the butterfly effect, many more future scenarios could unfold from genetic modifications.

As we are studying the complex graph of interactions between genes for potential modification, the interpretation and derivation of insights gleaned from data-determined connections must be valued as much as or more so than traditional metrics. A central theme behind our research process has been a focus on correctness from first principles to the best of our ability, which is reflected in our diverse approach in progressing along both traditional machine learning techniques, and novel approaches in non-Euclidean geometric techniques such as graph neural networks. Much of the value of our work is not in presenting a closed-box solution, but rather in the body of research we have cultivated about dead-end approaches, or identifying problematic framings of the problem that conflate causation with correlation.

We encourage the user of our computational methods to carefully examine the use case and use them as an initial first screen, rather than the sole source of truth. When in doubt that certain genetic perturbations might have more profound implications, experimentation and real-life observation of the crops are indispensable.