

Project Summary

There is a longstanding interest in understanding the genomic basis of reproductive isolation, as it can both inform how speciation occurs and the evolutionary forces occurring within lineages. Recently, an unusual type of hybrid incompatibility was characterized - a “hybrid decay” in which fitnesses of backcrosses progressively declined even when practically isogenic to one species.

In this work, I will characterize the genomic basis of this epistatic incompatibility, estimate its frequency in nature, and model its impacts on genome-wide patterns of introgression. I will also develop 2 teaching modules from this work: 1 for an undergraduate genetics course and 1 for a biostatistics course. My teaching will be guided by best pedagogical practices framed by my planned training as a data carpentry instructor.

Together, the training I receive in this fellowship will leave me well-prepared for a research career in plant and evolutionary genomics. This work will be done in collaboration with Dr. Yaniv Brandvain and Dr. Nathan Springer at the University of Minnesota.

Intellectual Merit

A unique case of hybrid incompatibility between maize and Mexican teosinte is characterized by no evidenced of reduced fitness in F1 hybrids but a sickly phenotype in hybrid descendants that are backcrossed to maize. Initial work implicates an uncoupling of transposable elements from their repressor systems, that halt their spread within the genome, in hybrids which triggers a cascade of causal downstream effects leading to phenotypic decay.

The role of repetitive DNA on driving speciation is well-characterized. However, the interaction between these elements and genetic machinery that stops their proliferation within the genome has largely been ignored. The proposed work tackles this problem directly by characterizing a demonstrated case of this interaction that triggers hybrid incompatibility. It then advances evolutionary theory by building upon prior knowledge to determine how these interactions emerge and persist in populations more broadly.

Broader Impacts

The application of computational biology, statistical genetics, and analytical modeling to large omic datasets hold significant potential for advancing research in plant biology. All pipelines and code generated will be shared publicly not only for reproducibility but also so that others may adapt these tools to their needs. In addition, I am committed to training the next generation of researchers in these techniques to expand our ability to explore a diverse set of questions.

First, I will advance my skills as a computational educator by enrolling in the Carpentries, a volunteer organization committed to training an inclusive community in the application of software and data in research. After this training, I will lead at least three software carpentry workshops throughout my postdoctoral training. The goal of these workshops will be to attract, train, and inspire a talented and diverse generation of trainees.

I will also develop two teaching modules for Dr. Brandvain and Dr. Springer's courses at the University of Minnesota. One will be an “active learning” exercise that will build upon the fundamentals of statistical methodology to highlight the real world applications of these methods for identifying genetic variants underlying pathogenic, agriculturally relevant, and evolutionarily conserved phenotypes. The other will connect population genetic theory with computer simulation to show how various evolutionary conditions drive the dynamics of genetic diversity.

This work will help biostatistic and genetic students scale up and see the connection between their classes in small data sets to modern genomic techniques and how they can be applied to their own research questions.