

## **A novel approach to E3 ubiquitin ligase-substrate interaction prediction in Arabidopsis**

**Jodi Callwood**<sup>1</sup> and Dior Kelley<sup>1</sup>

<sup>1</sup>Iowa State University

Ubiquitination, a post-translational modification, involves linking ubiquitin-protein to substrates via E3 ubiquitin ligases, altering protein function, location, and interactions. This process mainly impacts protein degradation, localization, activation, inactivation, and protein-protein interactions, influencing plant processes like development and stress responses. Despite >700 Arabidopsis E3 ligases with unknown substrates, predicting protein-protein interactions leading to ubiquitination lacks a method. Creating a prediction pipeline for ubiquitination could offer insights into protein homeostasis and conditions arising from these interactions. Arabidopsis thaliana possesses multiple E3s and ubiquitinated proteins as potential substrates, but few confirmed interactions. A Bayesian-based pipeline predicts E3-substrate interactions in Arabidopsis, extendable to crops like Zea mays and soybean. It employs confirmed E3-substrate interactions, annotated E3s, and candidate substrates for predictions. Leveraging protein characteristics from known E3-substrate pairs—protein family, domain(s), gene ontology terms, and co-expression data—a predictive model is developed. Positive and negative gold standard datasets curated from literature and Arabidopsis interactome data, respectively. Bayesian network analysis identifies potential E3-substrate pairs, yielding 46 interactions with a likelihood score threshold. Moreover, sequence- and structure-based tools, ESMFold and HDOCK, enhance consensus in interaction prediction signifying likely interactions. This pipeline can be applied to other species and be leveraged for better understanding of crop genomics.