

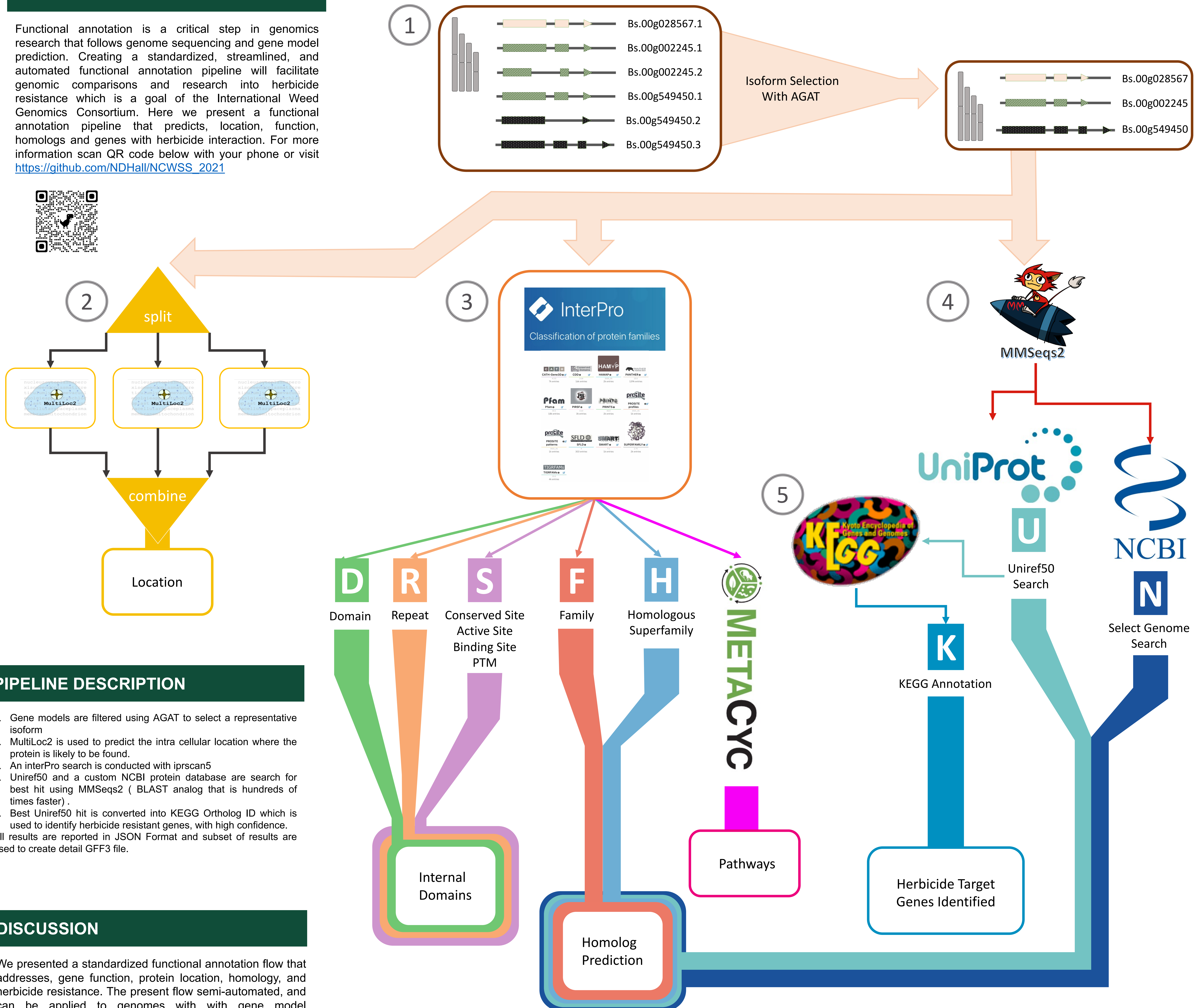
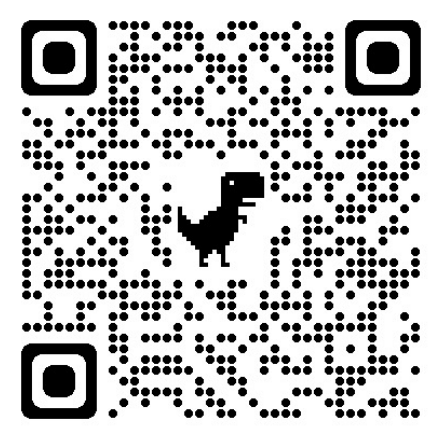
The Functional Annotation of Weedy Genomes for the International Weed Genomics Consortium

Nathan D. Hall and Eric Lloyd Patterson

**Department of Plant, Soil and Microbial Sciences, College of Agriculture & Natural Resources
Michigan State University**

INTRODUCTION

Functional annotation is a critical step in genomics research that follows genome sequencing and gene model prediction. Creating a standardized, streamlined, and automated functional annotation pipeline will facilitate genomic comparisons and research into herbicide resistance which is a goal of the International Weed Genomics Consortium. Here we present a functional annotation pipeline that predicts, location, function, homologs and genes with herbicide interaction. For more information scan QR code below with your phone or visit https://github.com/NDHall/NCWSS_2021



PIPELINE DESCRIPTION

1. Gene models are filtered using AGAT to select a representative isoform
 2. MultiLoc2 is used to predict the intra cellular location where the protein is likely to be found.
 3. An InterPro search is conducted with iprscan5
 4. Uniref50 and a custom NCBI protein database are search for best hit using MMSeqs2 (BLAST analog that is hundreds of times faster) .
 5. Best Uniref50 hit is converted into KEGG Ortholog ID which is used to identify herbicide resistant genes, with high confidence.
- All results are reported in JSON Format and subset of results are used to create detail GFF3 file.

DISCUSSION

We presented a standardized functional annotation flow that addresses, gene function, protein location, homology, and herbicide resistance. The present flow semi-automated, and can be applied to genomes with with gene model predictions.