

Public Abstract

First Name:Jiaojiao

Middle Name:

Last Name:Wang

Adviser's First Name:Trupti

Adviser's Last Name:Joshi

Co-Adviser's First Name:Dong

Co-Adviser's Last Name:Xu

Graduation Term:FS 2015

Department:Computer Science

Degree:MS

Title:Evolution of Soybean Knowledge Base

Soybean Knowledge Base (SoyKB), is a comprehensive web resource for knowledge about soybean genomics and multi-omics data. It is designed to give researchers easier access and better understanding of soybean traits and molecular breeding. In this thesis we have further expanded the analytics capabilities of SoyKB by developing new informatics tools including eFP Browser, SNPviz 2.0, WGCNA analysis and POP Select. The tools highlighted here provide users information ranging from genomics data to GWAS and its application in molecular breeding.

1) The eFP Browser was originally developed by the University of Toronto to visualize data intuitively. We have done a local standalone implementation in SoyKB with 16 transcriptomics expression datasets. Each dataset is represented by an image that will be recolored based on tissues' expression level. 2) SNPviz is a tool to analyze whole genome sequence SNP datasets for haplotypes of user-defined gene regions. SNPviz 2.0, developed in Javascript, is targeted to resolve the Java related security issues in SNPviz 1.0. It also includes several new features such as gene version control, neighbor joining cluster method and RGBY color scheme. At the same time, the cluster tree constructed in SNPviz 2.0 is dynamic which users can click a node to collapse or expand the sub-tree instead of just a static image. 3) WGCNA, is an open source R package for weighted gene co-expression network analysis and gene module detection, which we have incorporated in SoyKB as a new analysis feature in our Differential Expression Browser suite of tools. 4) Pop Select is a tool to help breeders analyze SNP population datasets and identify top scoring offspring with desired genomic information. It scores all offspring based on user specified region and parent type and then output top offspring information in charts and tables. These newly incorporated tools enriched SoyKB data visualization and analysis functionalities tremendously. In the future we will maintain these tools to make them more robust while exploring new application areas and developing new tools for the soybean research community.