



GRAMENE DIVERSITY MODULE

Terry Casstevens³, Peter Bradbury¹, Payan Canaran⁴, Dallas Kroon³, David Matthews^{1,2}, Susan McCouch^{2,3}, Junjian Ni², Doreen Ware^{1,4}, Immanuel Yap², Wei Zhao⁴, Ken Youens-Clark⁴, Isaak Teclé², Shuly Avraham⁴, Greg Wilson², Frank You⁵, Edward Buckler^{1,2,3}

¹ USDA-ARS ² Dept. of Plant Breeding and Genetics, Cornell University ³ Institute for Genomic Diversity, Cornell University, Ithaca, NY ⁴ Cold Spring Harbor Laboratory, NY, USA ⁵ Dept. of Plant Sciences, University of California, Davis

<http://www.gramene.org>

Introduction

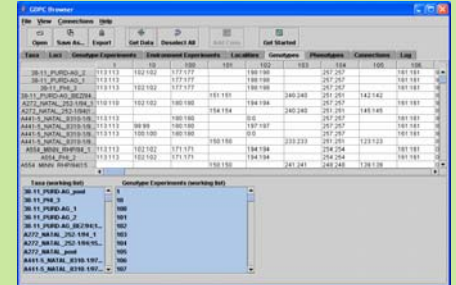
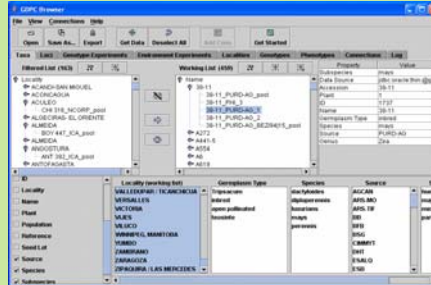
Many studies have produced large quantities of genetic and phenotypic data (i.e. QTL, germplasm, molecular). The Gramene Diversity project is building an infrastructure to share this basic molecular and phenotypic diversity data. This infrastructure includes a database schema (GDPDM; www.maizegenetics.net/gdpdm), a Java XML-SOAP middleware (GDPC; www.maizegenetics.net/gdpc), a sequence alignment-SNP viewer (www.panzea.org/db/snp_alignment/snp_cgi.dev.pl), a pedigree viewer, and an association diversity analysis tool (TASSEL;

www.maizegenetics.net/bioinformatics/tasselindex.htm). All these components are free open source tools. Species specific tools to diversity data sets are currently available for maize (www.panzea.org) and rice (rice-evolution.plbr.cornell.edu). We are currently integrating diversity data from Maize, Wheat, and Rice diversity projects into one database. Diversity data in that database and the USDA Germplasm Resources Information Network (GRIN; www.ars-grin.gov) will soon be publicly shared via the GDPC middleware. Much of the documentation and source code is now available in sourceforge (see projects gdpdm, gdpc, and tassel). We encourage community input and collaboration on this effort, so that the largest possible community can access and productively use diversity data.

This research is supported by the NSF Gramene Grant, the USDA-ARS, and several NSF project grants on maize, wheat, and rice.

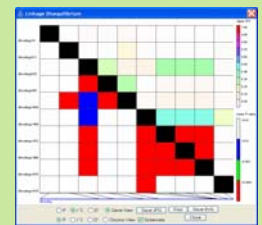
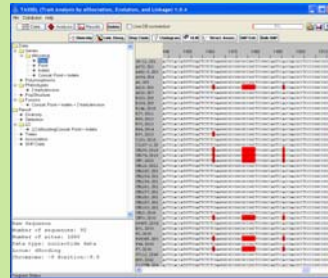
GDPC Browser

The GDPC Browser is a front-end software tool that uses the GDPC API. It retrieves genomic and phenotypic data based on selected criteria. Data can be saved, exported, grouped, etc. <https://sourceforge.net/projects/gdpc>



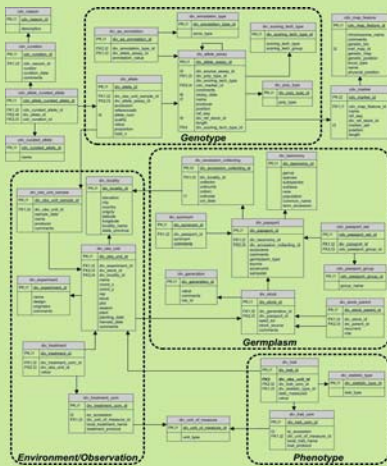
TASSEL

Quantitative trait analysis, Diversity and linkage disequilibrium analyses, Association analyses, SNP Assay Extractor, Wide range of visualizations. <https://sourceforge.net/projects/tassel>



GDPDM Schema

<https://sourceforge.net/projects/gdpdm>



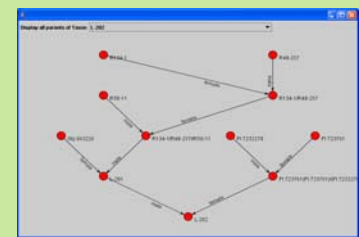
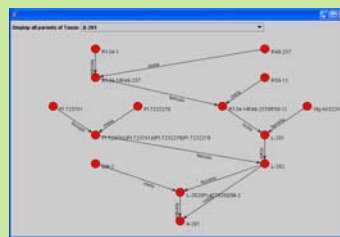
Germplasm – Basic Passport. Details link out to germplasm banks. Can curate groups of germplasm. Supports complex genetic pedigrees. Records seed lots and plants.

Genotype – Captures wide range of approaches (SNP, SSR, sequence alignments, isozyme, etc.). Supports full range of ploidy. Can capture background on methodology. Connects to genomic map data.

Phenotype – Quantitative or qualitative traits. Supports ontology integration. Range of statistics can be recorded (mean, measure, median, etc.).

Environment – Full field design can be recorded to the plant if desired. Basic planting and treatment information. Integrates with locality information. Observations can be grouped into experiments.

Pedigree Viewer



Web based portals & Sequence Alignment-SNP Viewer

www.panzea.org, future access on www.gramene.org

