



Gramene: It's Not Just For Grasses Anymore

Joshua C. Stein ¹, Ed Buckler ^{2,3}, Terry Casstevens ⁴, Charles Chen ⁴, Genevieve DeClerck ⁴, Palitha Dharmawardhana ⁵, Pankaj Jaiswal ⁵, AS Karthikeyan ⁴, Marcela Monaco ¹, Susan R McCouch ⁴, Will Spooner ¹, Jim Thomason ¹, Sharon Wei ¹, Ken Youens-Clark ¹, Shiran Pasternak ¹, Doreen Ware ^{1,3}

¹Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA; ²Institute for Genomic Diversity, Cornell University, Ithaca, NY, 14853, USA; ³USDA ARS NAA Robert W. Holley Center for Agriculture and Health, Cornell University, Ithaca, NY, 14853, USA; ⁴Department of Plant Breeding and Genetics, 240 Emerson Hall, Cornell University, Ithaca, NY, 14853, USA; ⁵Dept of Botany and Plant Pathology, 3082 Cordley Hall, Oregon State University, Corvallis, OR, 97331, USA

Integrating functional, genetic & comparative information for discovery and hypothesis based research



Home

GRAMENE Home

Release #33
April 2011

News

- Gramene at grape conference
- Gramene is sending a member to the International Grass Genome Program's (IGGP) 3rd Annual Grape Conference.
- Marcela Monaco, who works on Gramene's pathway analysis, will be presenting a poster on Gramene release 33 out.
- The Gramene team is happy to announce the 3rd annual release available at www.gramene.org. This build...
- Subscribe to RSS Feed
- Have Questions...?
- Quick Search Help
- Ask questions through Feedback or Email.
- See FAQ's

Outreach calendar

Presentation materials

Gramene is a curated, open-source, data resource for grasses and other plants. Our goal is to facilitate the study of grasses and other plants by providing access to information derived from public projects and databases, including protein structure and function analysis, gene expression, QTL analysis, protein-protein interaction, and phylogenetic analysis.

Arabidopsis Entry Page

GRAMENE Arabidopsis thaliana (TAIR10)

Search Arabidopsis thaliana

Description

About Arabidopsis thaliana

Arabidopsis thaliana is a small flowering plant that is widely used as a model organism in plant biology. Arabidopsis is a member of the mustard (Brassicaceae) family, which includes cultivated species such as cabbage and radish. Arabidopsis is not major crop, but it has been used extensively for its ease of growth and its small genome size. The genome of Arabidopsis thaliana has a genome size of approx. 135 Mbase, and a haploid chromosome number of 5. (From TAIR)

Genome Sequencing and Gene Prediction (Release TAIR10)

The complete genome sequence of Arabidopsis thaliana was first published by the Arabidopsis Genome Initiative in 2000 [1] and was determined by a BAC-by-BAC sequencing strategy anchored to chromosomes using a variety of genetic and physical maps. Gene annotations use cDNA and EST data as well as manual updates informed by cross-species alignments, peptides and community input regarding missing and incorrectly annotated genes. The genome contains approximately 25,000 genes and 100,000 transcripts. This browser is based on data from version 10 of The Arabidopsis Information Resource (TAIR) database, released in November 2010. [more]

Gramene/Ensembl Genomes Annotation

Additional annotations generated by the Gramene/Ensembl Genomes projects include:

- The standard set of Gramene/Ensembl details here.
- Phylogenetic tree and whole-genome alignments with several plant and metazoan species, using Ensembl's Compara pipeline.
- The Arabidopsis variation database contains data from the screening of 1179 strains using the Affymetrix 250K Arabidopsis SNP chip, and an update of the trait database through a collaboration with the Wellcome Trust Centre for Human Genetics in Oxford. Paula Kopriva at the University of Illinois, Urbana-Champaign, funded by the BBSRC, which involved screening of 1878 strains and 18 strains from the 1001 Genomes Project. Phenotype data has also been added from a GWAS study of 107 phenotypes in 95 inbred lines performed by the Atwell et al. study.
- Mappings for probes from the following arrays were also added to the functional genomics schema:
 - Arabidopsis ATH1 Genome Array
 - Agilent G2519F-021186
 - Agilent G2519F-015059
 - Agilent G2519F-015060
 - ATM

Links

- The Arabidopsis Information Resource

References

1. Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*.

Complete Genomes: *Arabidopsis thaliana*, *Arabidopsis lyrata*, grapevine, poplar, rice (3), Brachypodium, sorghum, & maize

EnsemblPlants

Custom queries and downloads



Export to Excel or text

Orthologs in lyra, grape, poplar, rice, Brachypodium, sorghum maize, & moss

View 20 → rows as: HTML | XML | Plain text | CSV | TSV | PDF | Unique results only

Ensembl Gene ID: (ID-list)
AT1G08680

Ensembl Gene ID: (ensembl gene ID)
AT1G08680

Ensembl Gene ID: (ensembl gene ID)
AT1G08680

All transmembrane-targeted genes, showing InterPro domains, GO terms, and AFFY id's

AT1G08680 Trimm Sign IP500118 Carbonyl_anhydride_cat, cat

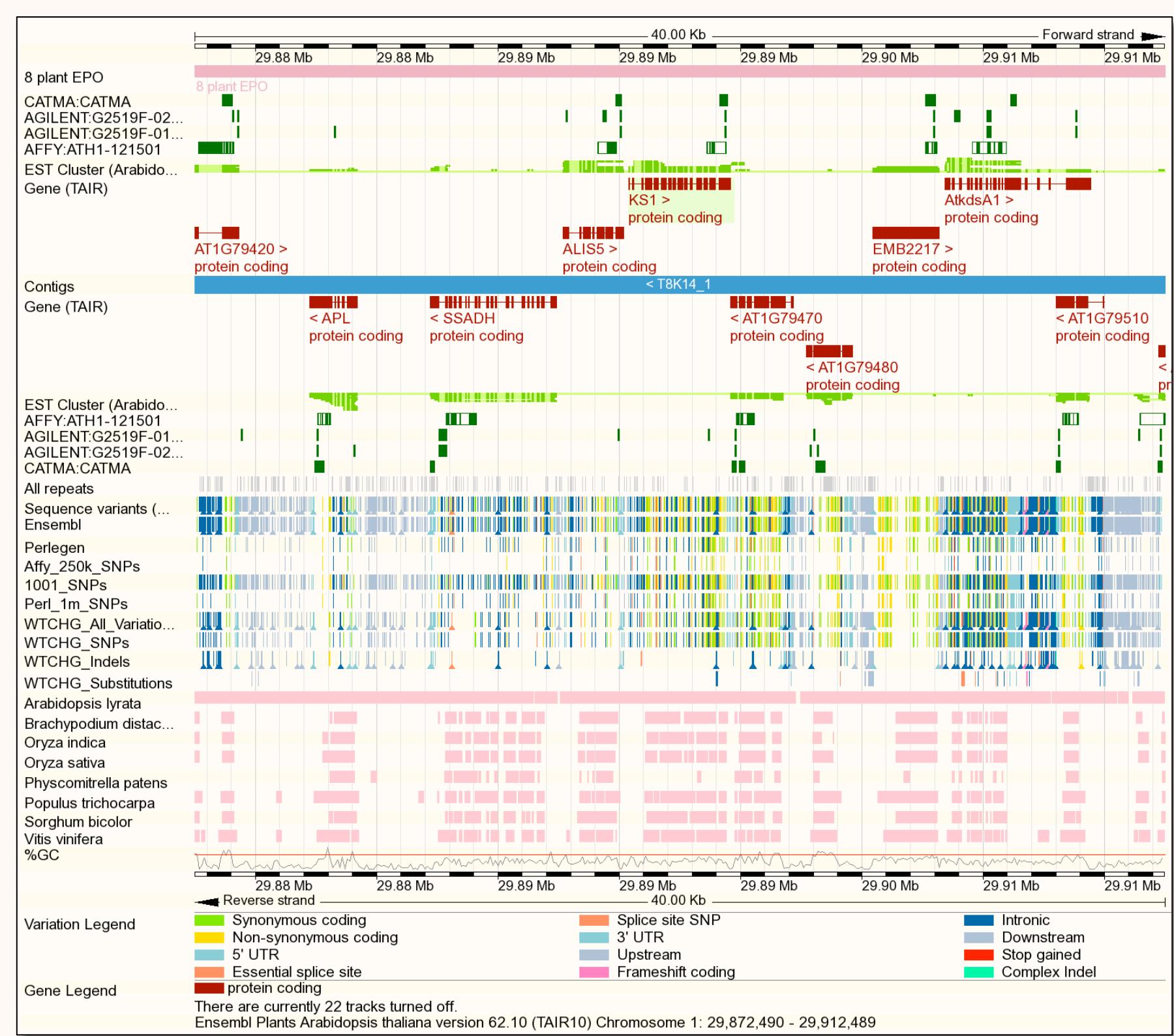
All Myb-domain genes with "STOP_GAINED" SNP allele in germplasm

AT1G08680 Variation ID: STOP_GAINED

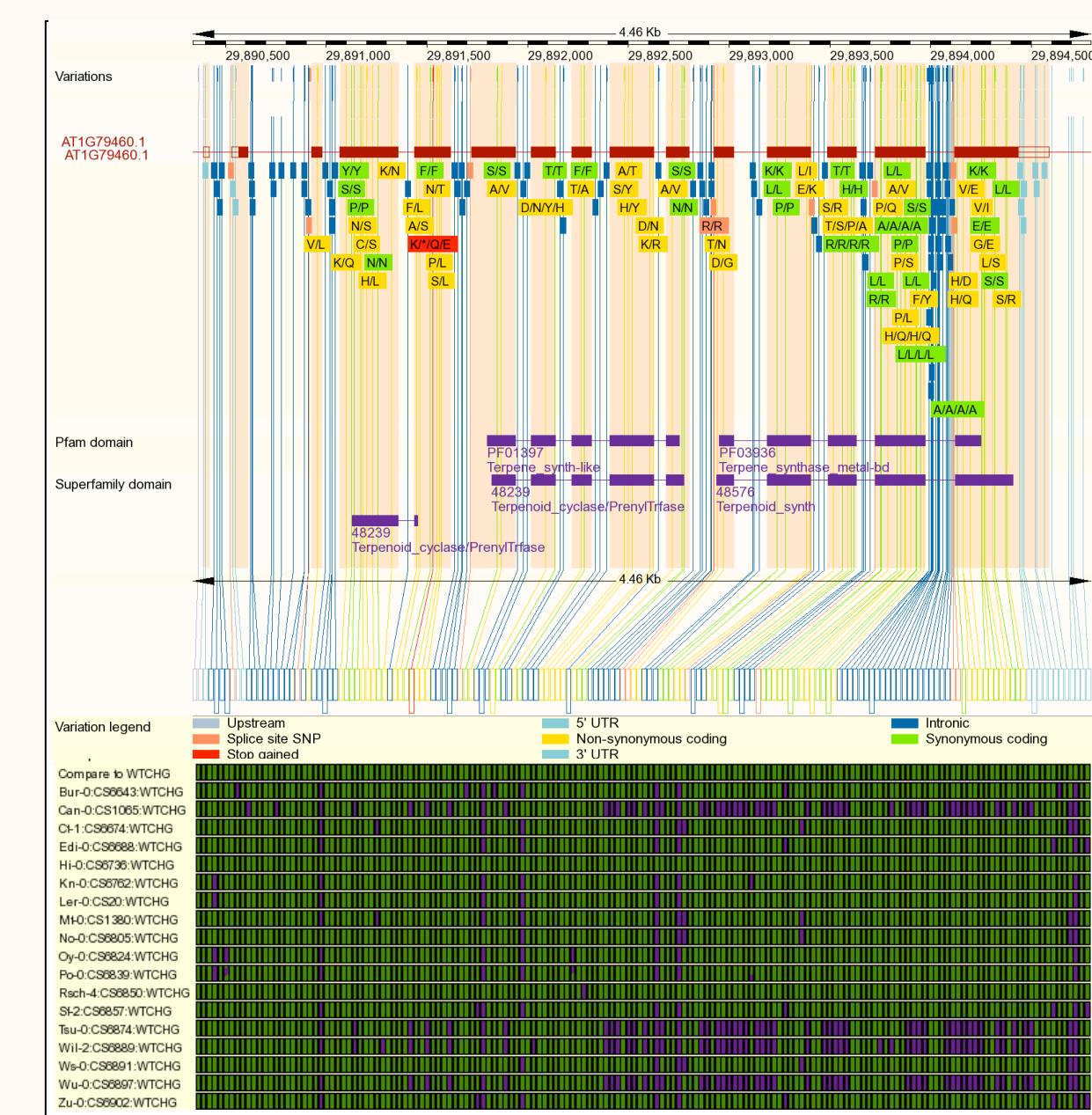
Taxon of origin for paralog pairs of cyclin-domain genes that have an ortholog in Physcomitrella

Ensembl Gene ID	Paralogous Gene ID	% Identity	Orthologous Thaliana Gene ID	Annotation
AT1G08680	ENSGA000000000000	60	Arabidopsis	
AT1G08680	ENSGA000000000001	58	Arabidopsis	
AT1G08680	ENSGA000000000002	55	Arabidopsis	
AT1G08680	ENSGA000000000003	53	Arabidopsis	
AT1G08680	ENSGA000000000004	41	Magnaporthe	
AT1G08680	ENSGA000000000005	22	Emoryella	
AT1G08680	ENSGA000000000006	22	Emoryella	
AT1G08680	ENSGA000000000007	48	rice	
AT1G08680	ENSGA000000000008	28	rice	
AT1G08680	ENSGA000000000009	41	Erythrina	

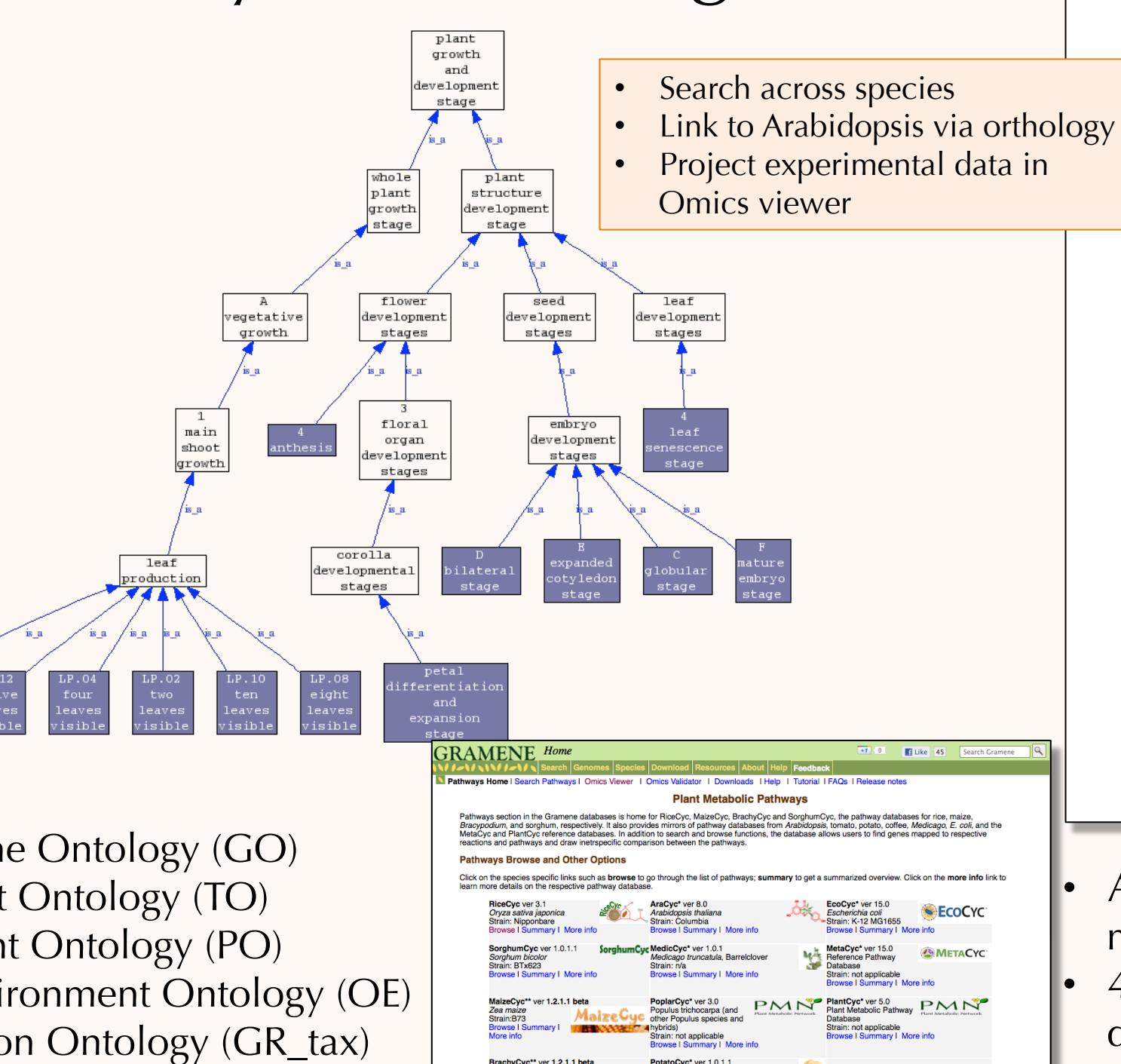
The Browser: Focus on Diversity



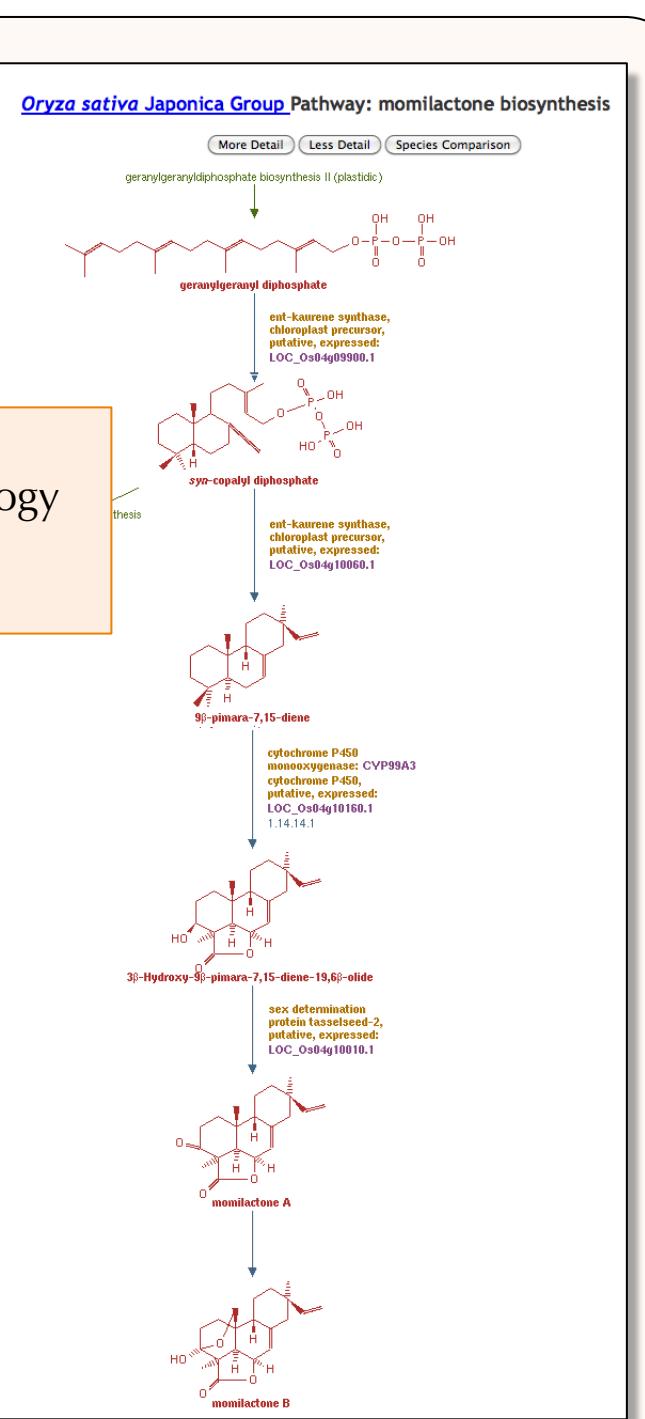
- Millions of SNPs, thousands of ecotypes
- View functional consequences within gene and protein domains
- Genotypes displayed in graphical & tabular form
- Order seed from ABRC via direct links to TAIR



Functional Annotation: Pathways & Ontologies

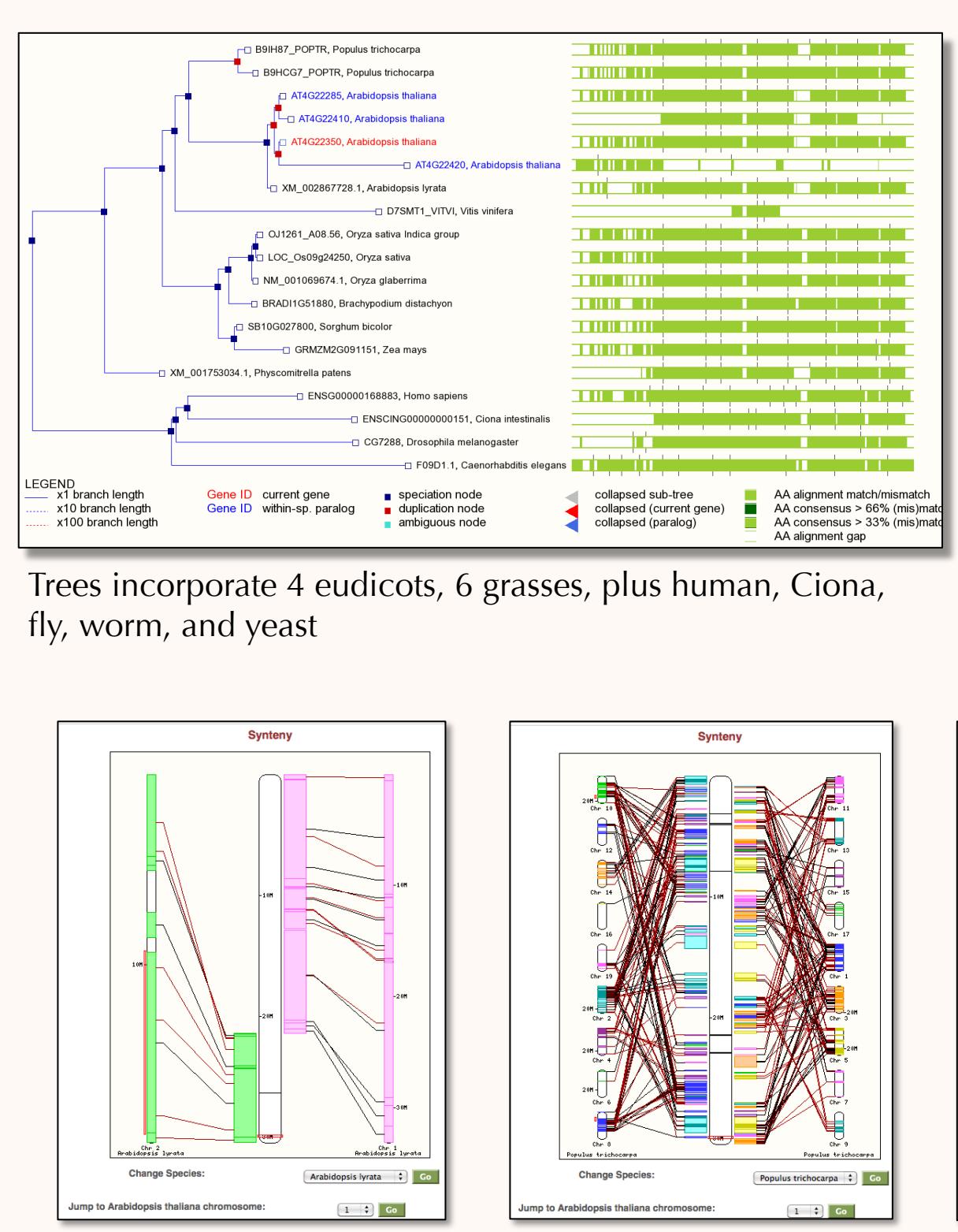


- Gene Ontology (GO)
- Trait Ontology (TO)
- Plant Ontology (PO)
- Environment Ontology (OE)
- Taxon Ontology (GR_tax)

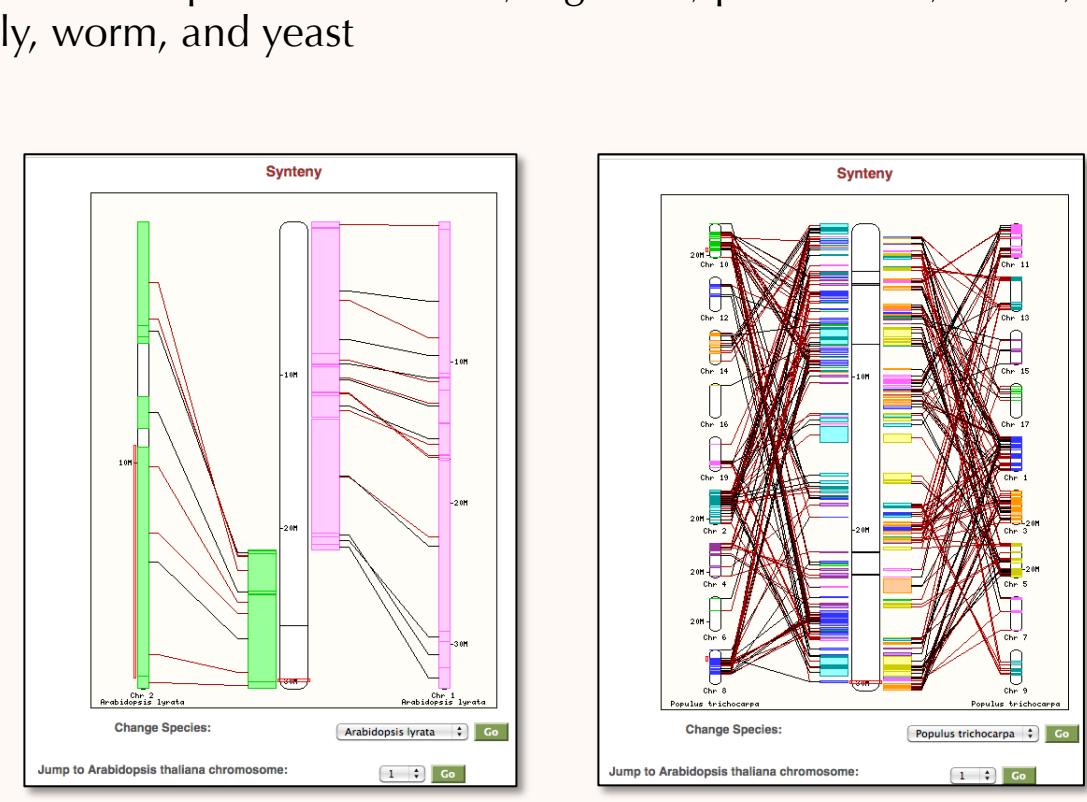


- AraCyc pathway database mirrored at Gramene
- 4 pathway databases developed and manually curated by Gramene
- 8 Additional databases mirrored

Comparative & Phylogenomics



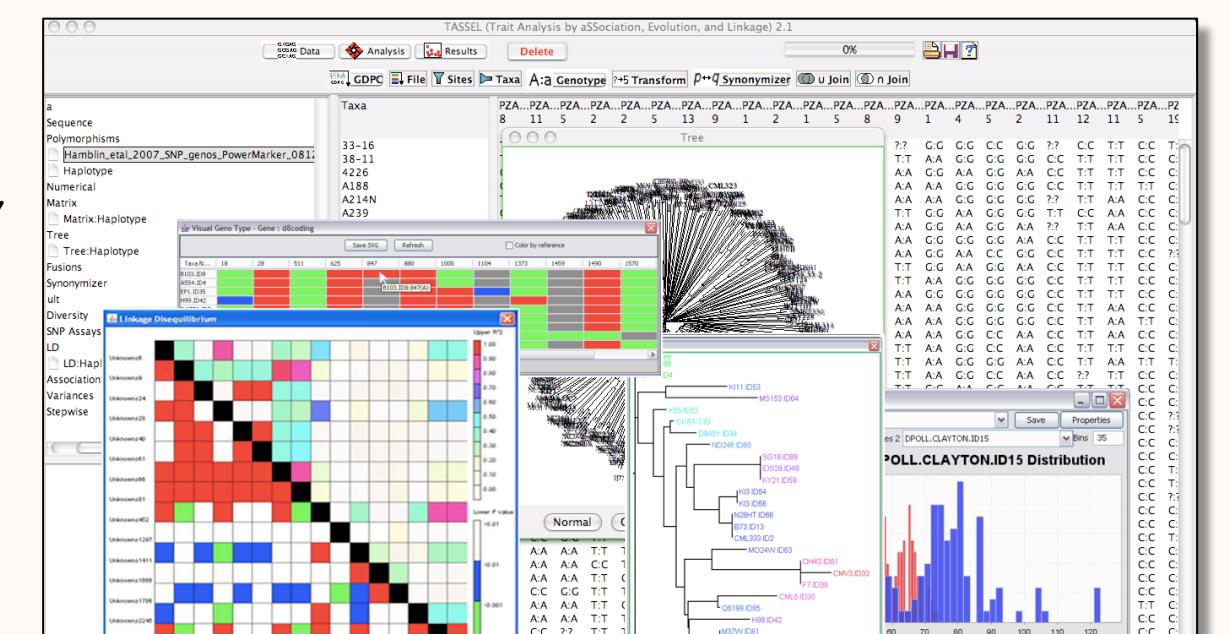
- Whole genome alignments vs 8 reference genomes (BLASTZ-CHAIN-NET)
- Compara gene trees trace evolutionary histories in over 35,000 families
- Ortholog & paralog classification
- Taxon dating of duplication nodes
- Interspecies browsing in Multi-species view and Synteny-view



Navigate across duplications by centering on grape

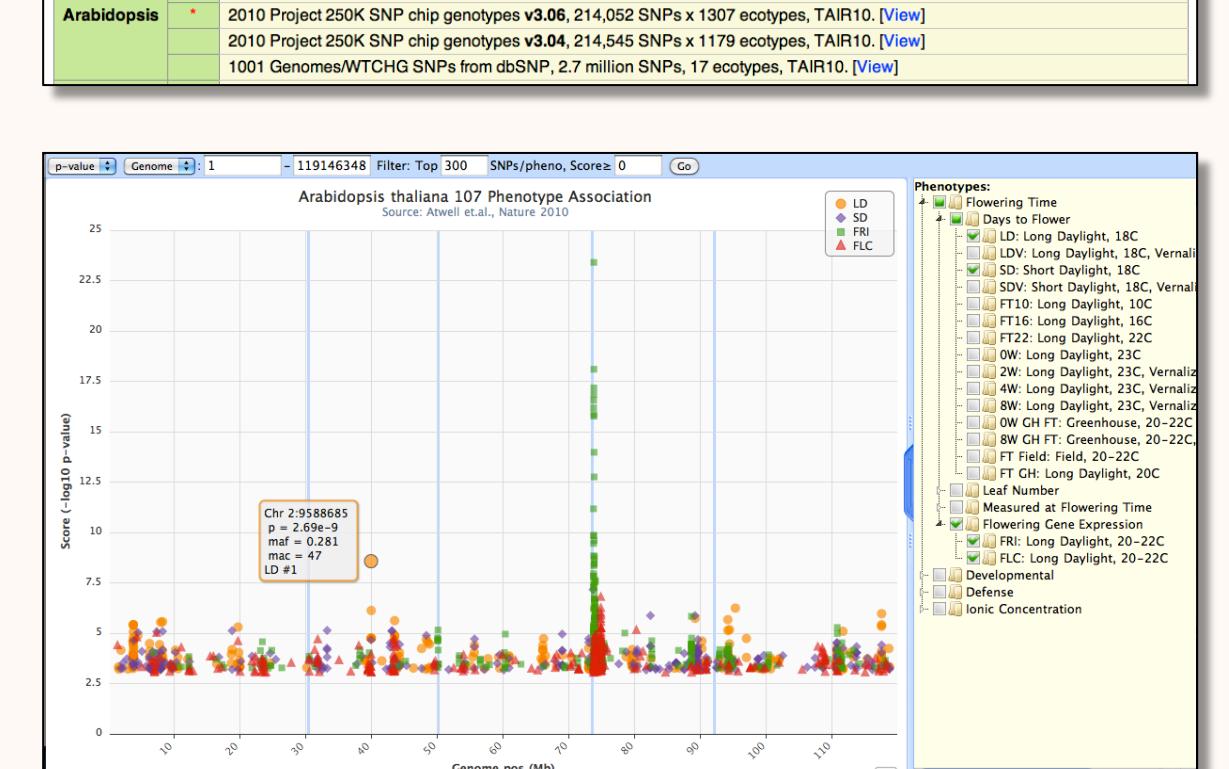
TASSEL

- Trait Analysis by aSSociation, Evolution, and Linkage
- Calculate and view linkage disequilibrium and evolutionary patterns
- Association analysis by generalized linear model and mixed linear model



In Development

- GWAS viewer to display Manhattan plots
- Data standardization to capture fine-grained details of phenotype and experiment in a logical and consistent manner
- Currently prototyping on the Atwell et al. 2010 study



Recent Publications:

- Jaiswal P. Gramene database: a hub for comparative plant genomics. *Methods Mol Biol*, 678:247-75.
- Youens-Clark K, et al. (2010) Gramene database in 2010: updates and extensions. *Nucleic Acids Research*, 39 (suppl 1): D1085-D1094.
- Ni J, et al. (2009) Gramene QTL database: development, content, and applications. *Database (Oxford)*, 2009:bap005. Epub 2009 May 8.