Gramene - Exploring Function through Comparative Genomics & Network Analysis

Transnational collaboration

Doreen Ware
USDA ARS, CSHL – USA

Bruno Contreras-Moreira
Ensembl Plants
EMBL-EBI – UK

Irene Papatheodorou
Expression Atlas
EMBL-EBI – UK

Pankaj Jaiswal
OSU – USA

Crispin Taylor
ASPB – USA

- Gramene is a Comparative Database for Plants since 2001
- 61 plant genomes with gene family trees, sequence alignments, synteny, homologs (EBI-Ensembl)
- Gene expression (EBI-Atlas)
- Pathways (OSU-Reactome)
- Automated ingestion of functional annotations from publications (ASPB)

Cold Spring Harbor Laboratory

EMBL-EBI

OICR Ontario Institute for Cancer Research

OSU Oregon State University

USDA

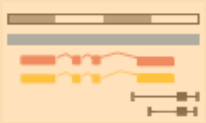
NSF

Gramene

61 reference Genomes

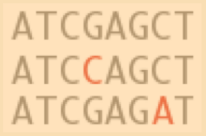
Ensembl Plants

2.3M gene annotations



92.7K gene trees

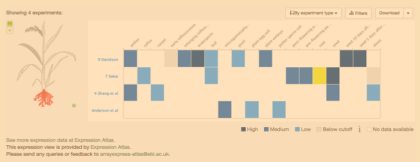
278 whole-genome alignments



65 syntenies

Expression

EMBL-EBI Expression Atlas

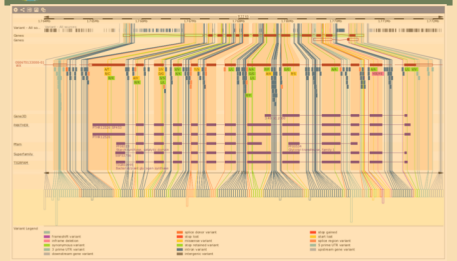


804 experiments in 24+ species

Integrated search and visualization

Variation

European Variation Archive

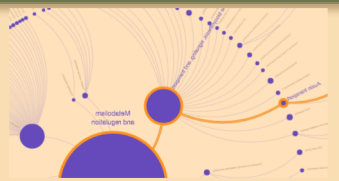


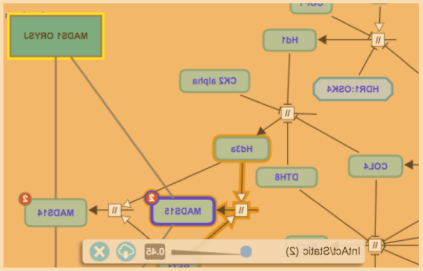
~223.7M SNPs in 12 genomes

Pathways

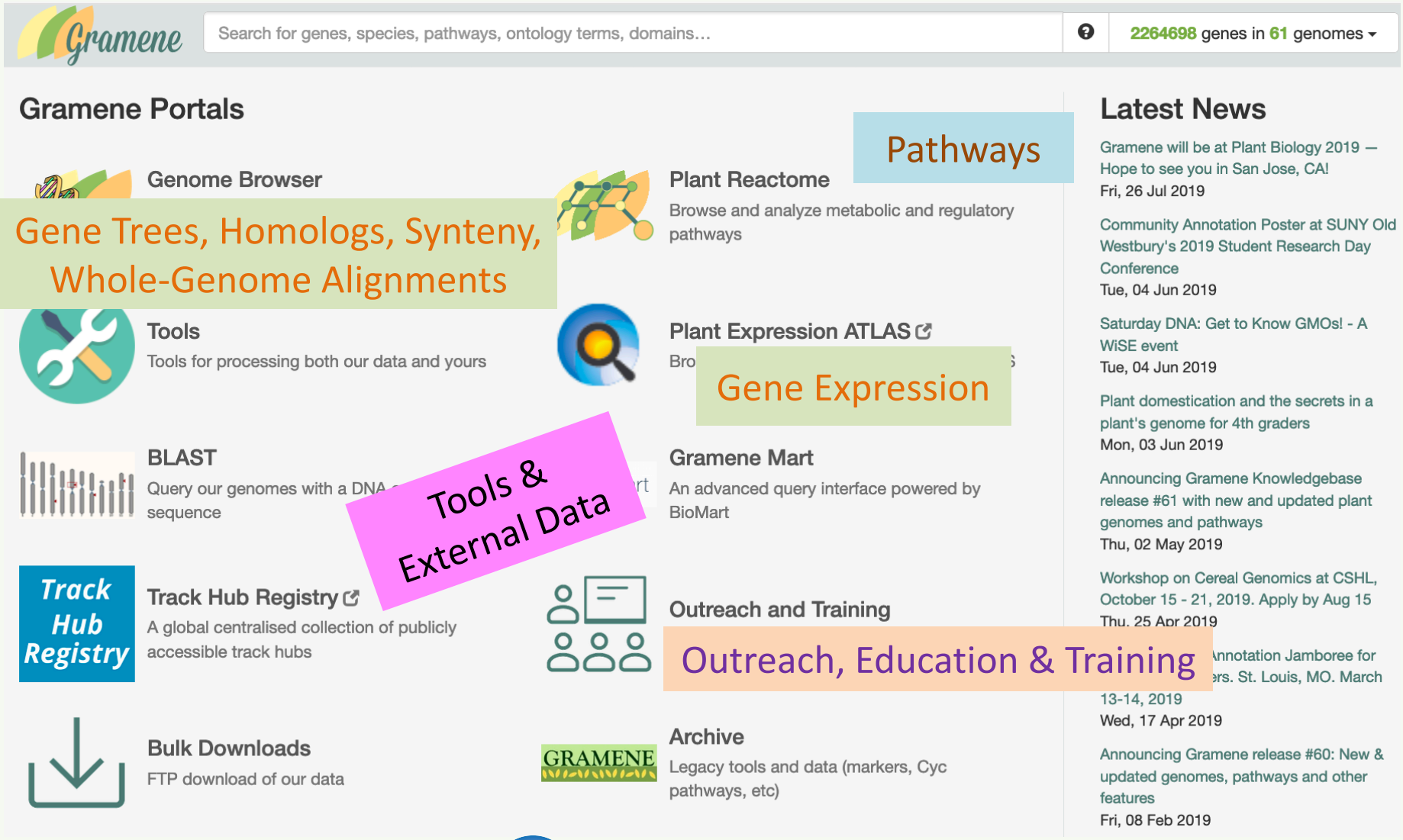
298 curated rice projected to 82 species

Plant Reactome





Gramene.Org



e!Ensembl

Gene & Protein Expression

reactome

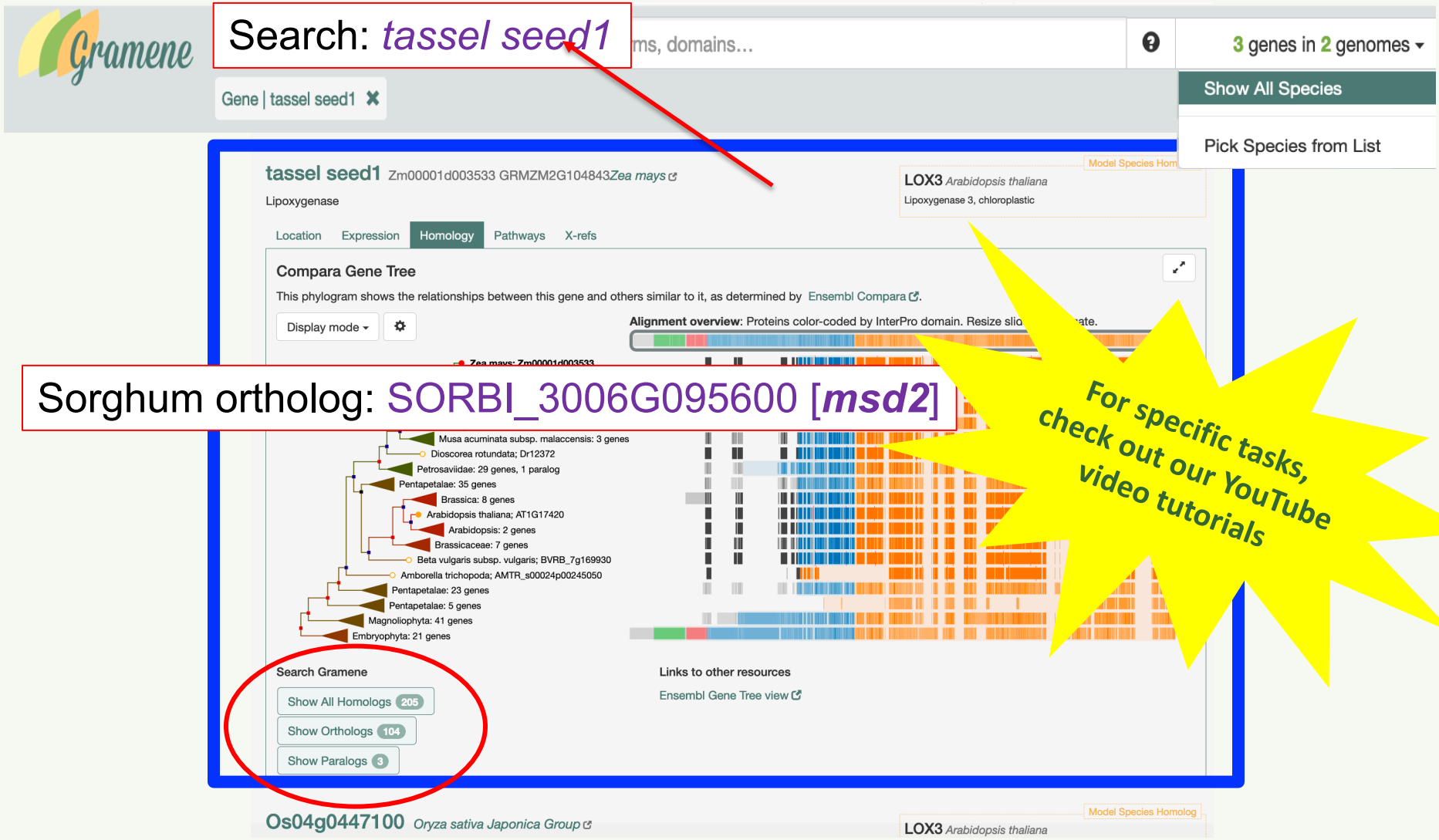
What can I do with Gramene?

- Is my gene conserved in other species? Can I identify function of a novel gene through homology?
- Where and when is my gene expressed?
- Is my gene in a metabolic pathway? Is the pathway conserved in other crops?
- Have genetic variants been identified in my gene? Can I access germplasm lines with naturally occurring or introduced genetic variants in my gene?

Work Example

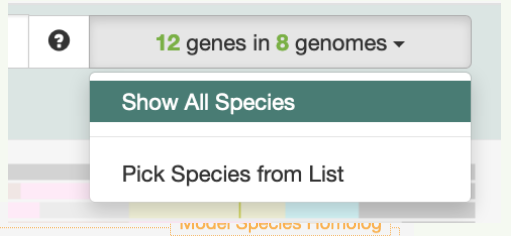
1. Identify the sorghum ortholog of the maize *tassel seed1* (*ts1*) gene
2. Reduce the complexity of the gene's family tree, focusing on species of interest
3. Visualize synteny between species
4. Explore gene structure and identify functional annotations associated with this gene family
5. Explore patterns of gene expression
6. Identify pathways associated with the gene
7. Explore genetic variation in *ts1* and find induced EMS-mutants in sorghum

Find orthologs & paralogs



Gene Trees: Focus on species of interest

=> Prune unwanted species



tassel seed1 Zm00001d003533 GRMZM2G104843 *Zea mays*

Lipoxygenase

LOX3 *Arabidopsis thaliana*

Lipoxygenase 3, chloroplastic

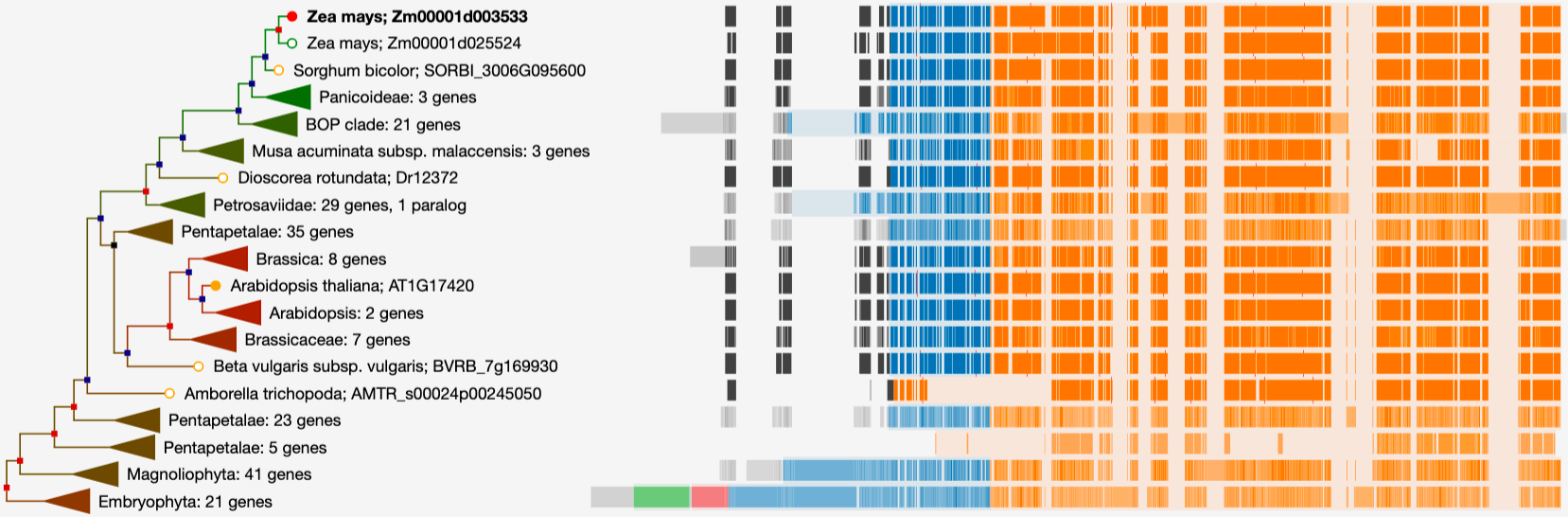
Location | Expression | Homology | Pathways | X-refs

Compara Gene Tree

This phylogram shows the relationships between this gene and others similar to it, as determined by Ensembl Compara.

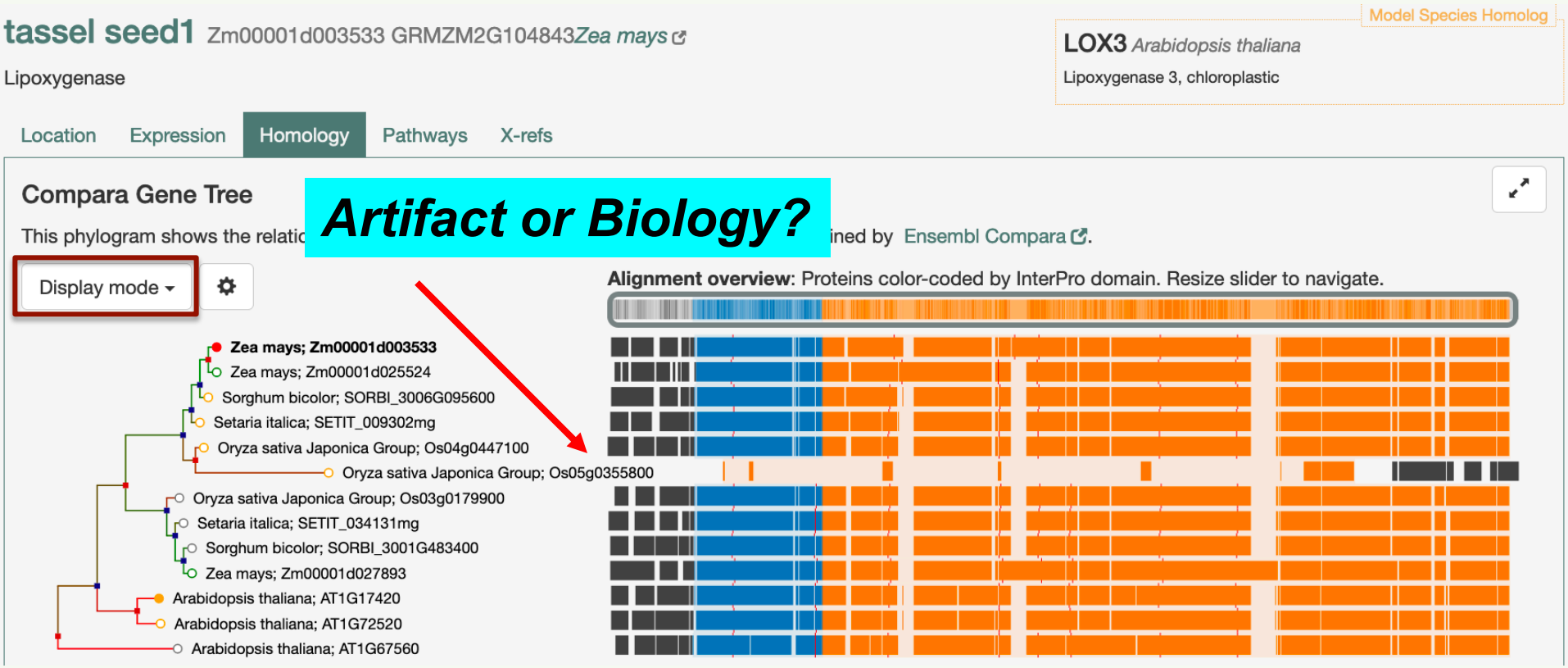
Display mode

Alignment overview: Proteins color-coded by InterPro domain. Resize slider to navigate.



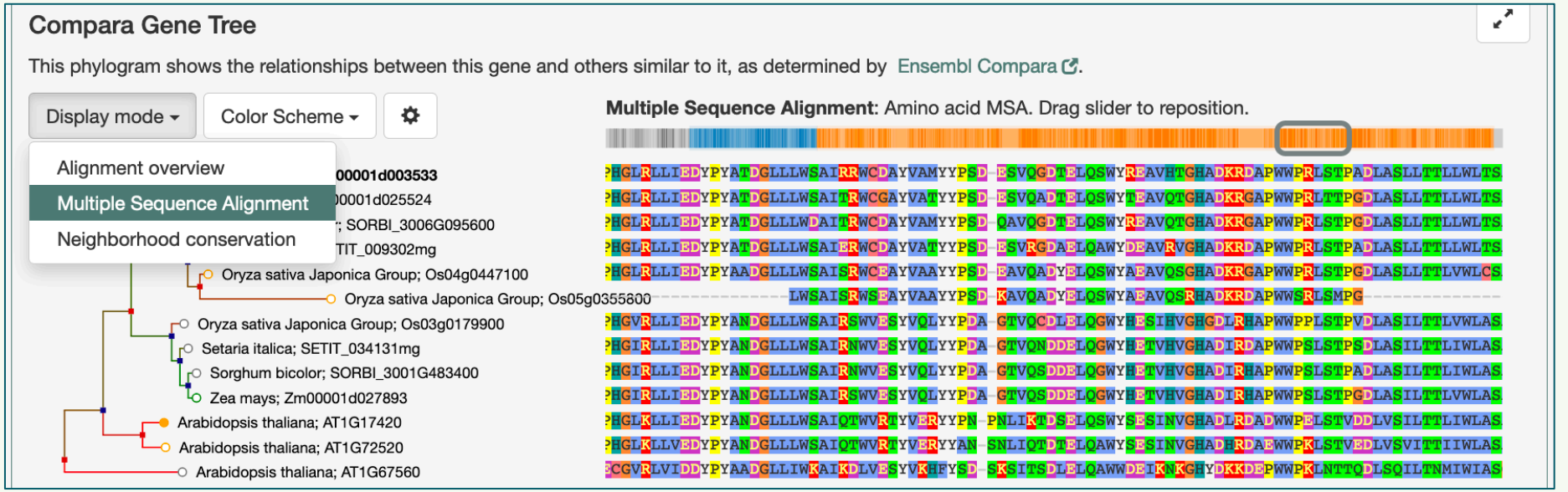
Gene Trees: Focus on species of interest

=> Prune unwanted species

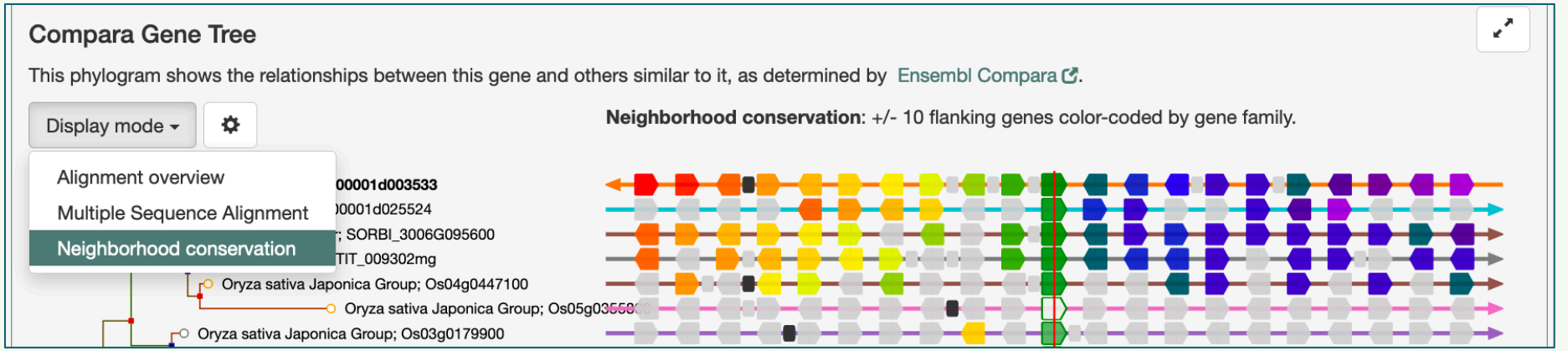


Explore a gene family tree

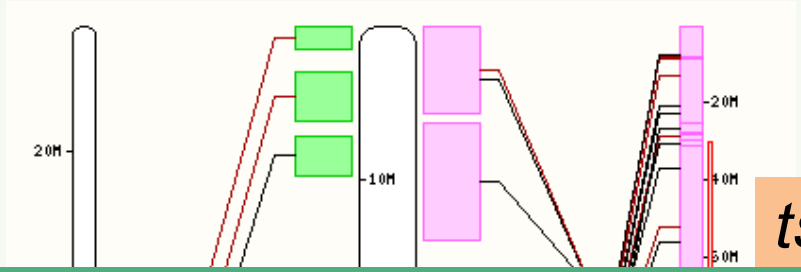
=> Zoom In to view aligned amino acids



=> Zoom Out to view conserved gene neighborhoods



Explore synteny



ts1

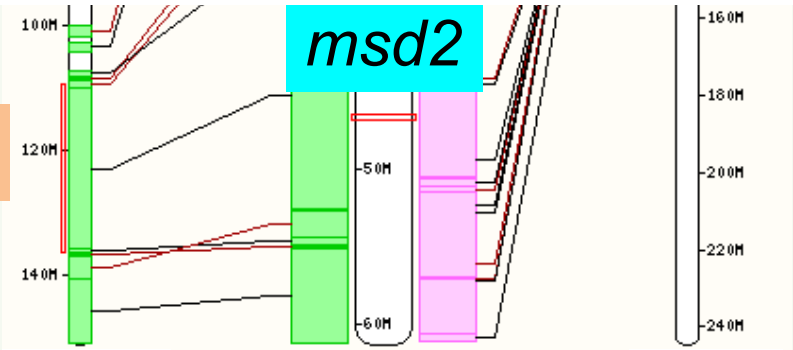
Show/hide columns

Filter

Sorghum bicolor genes	Location		*Zea mays* homologues	Location	
SORBI_3006G095600 (SORBI_3006G095600)	6:46566240-46571064	→	tassel seed1 (Zm00001d003533)	2:47105187-47109372	Region Comparison
		→	lipoxygenase7 (Zm00001d025524)	10:121268606-121272616	Region Comparison

lox7

msd2



Chr 10 Maize

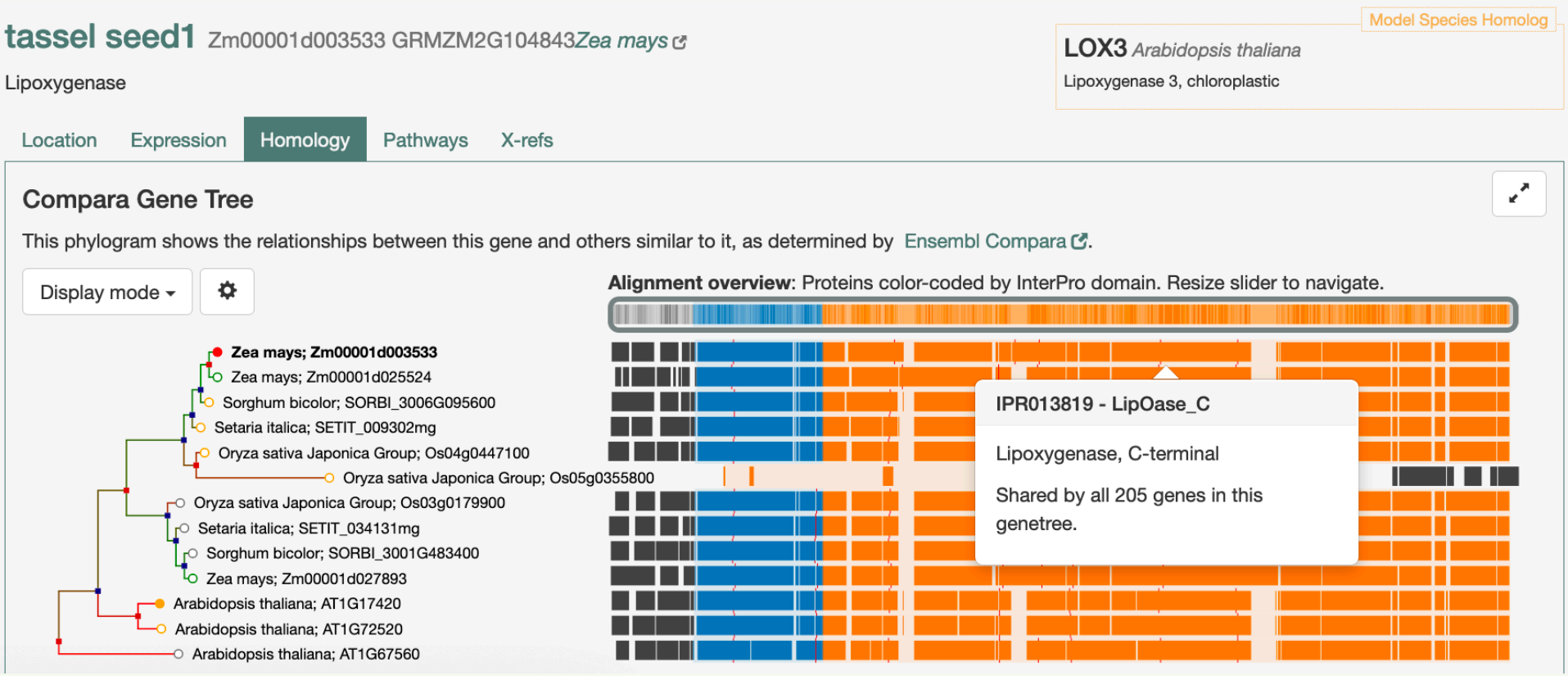
Chr 6 Sorghum

Chr 2 Maize

- Comparative Region View
- Download Gene Lists

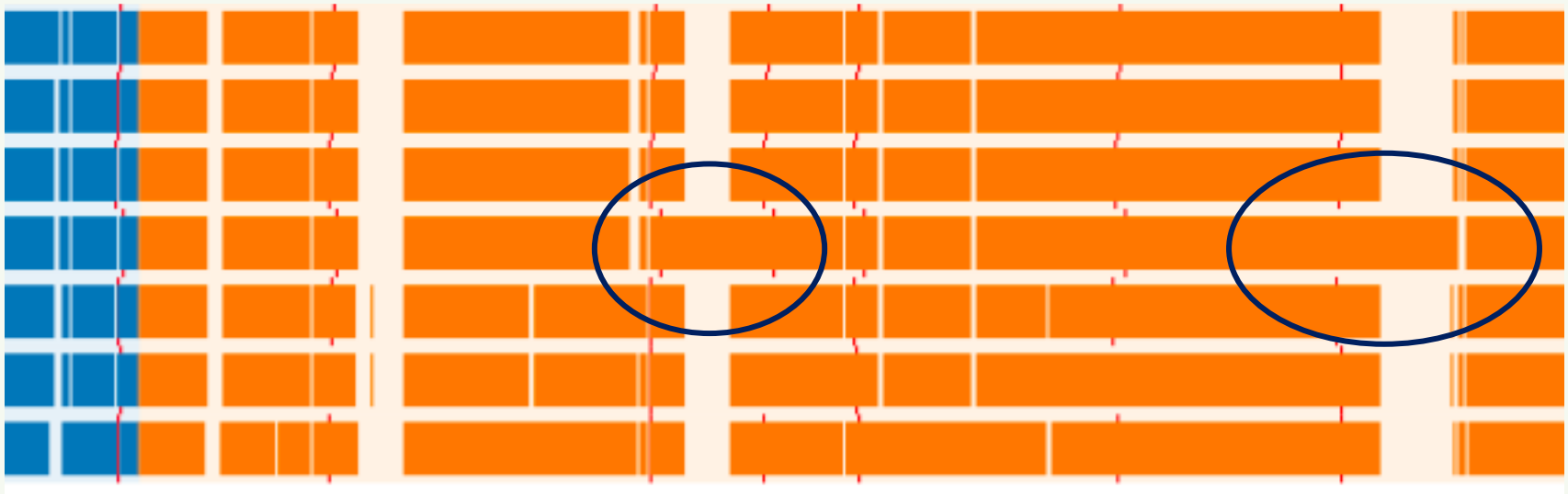
Explore gene structure

=> Identify protein domains



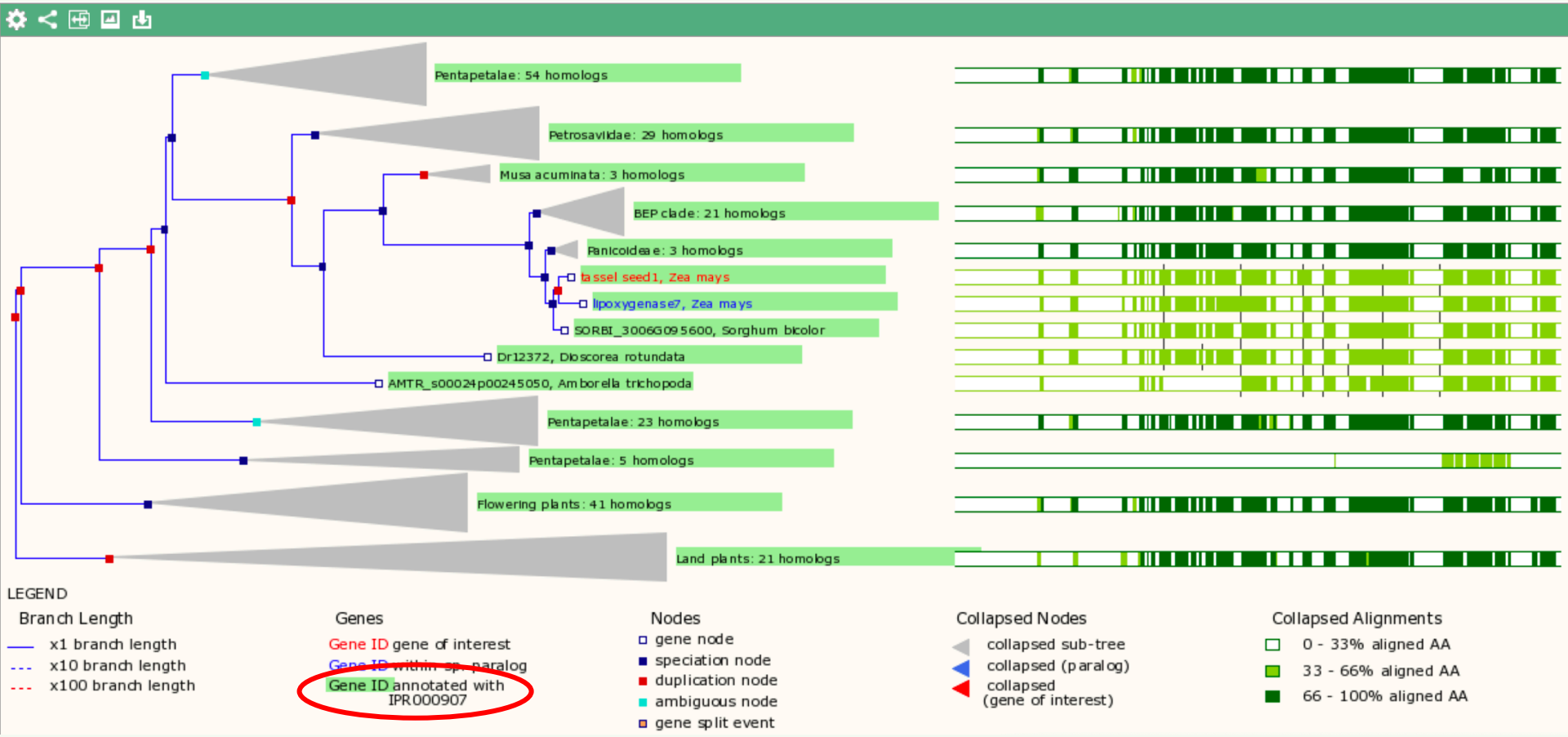
Explore gene structure

=> Identify splice junctions



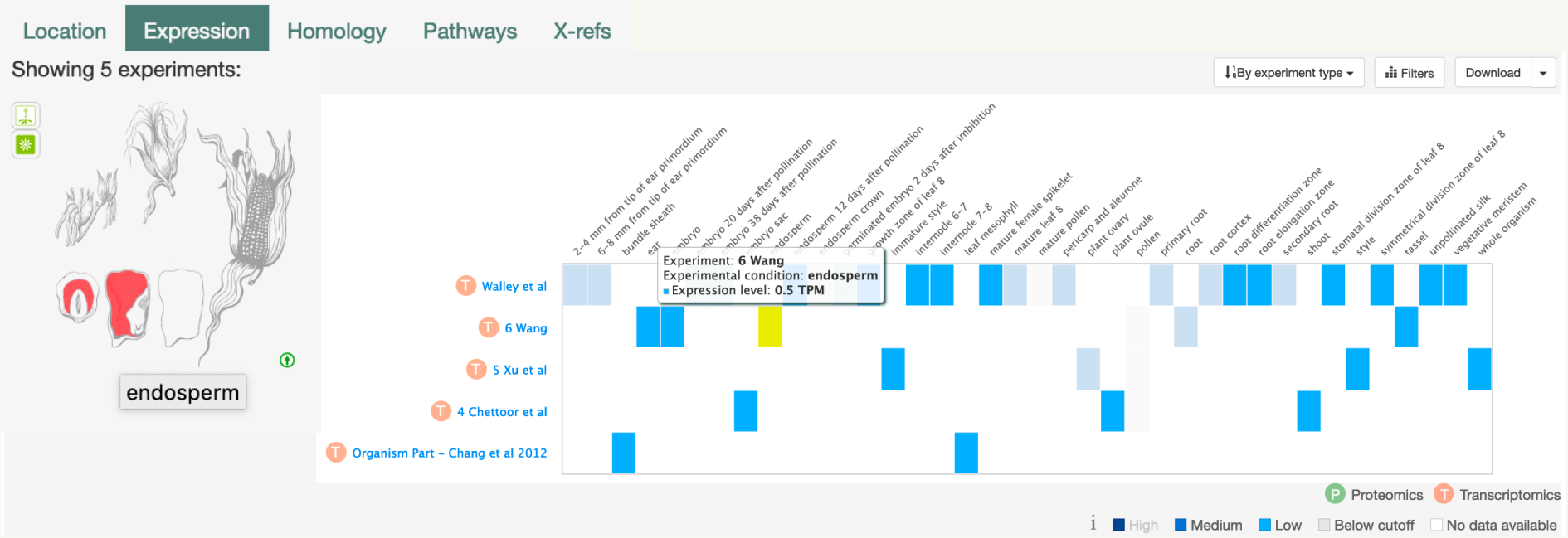
Explore functional gene structure

=> Highlight InterPro & GO annotations



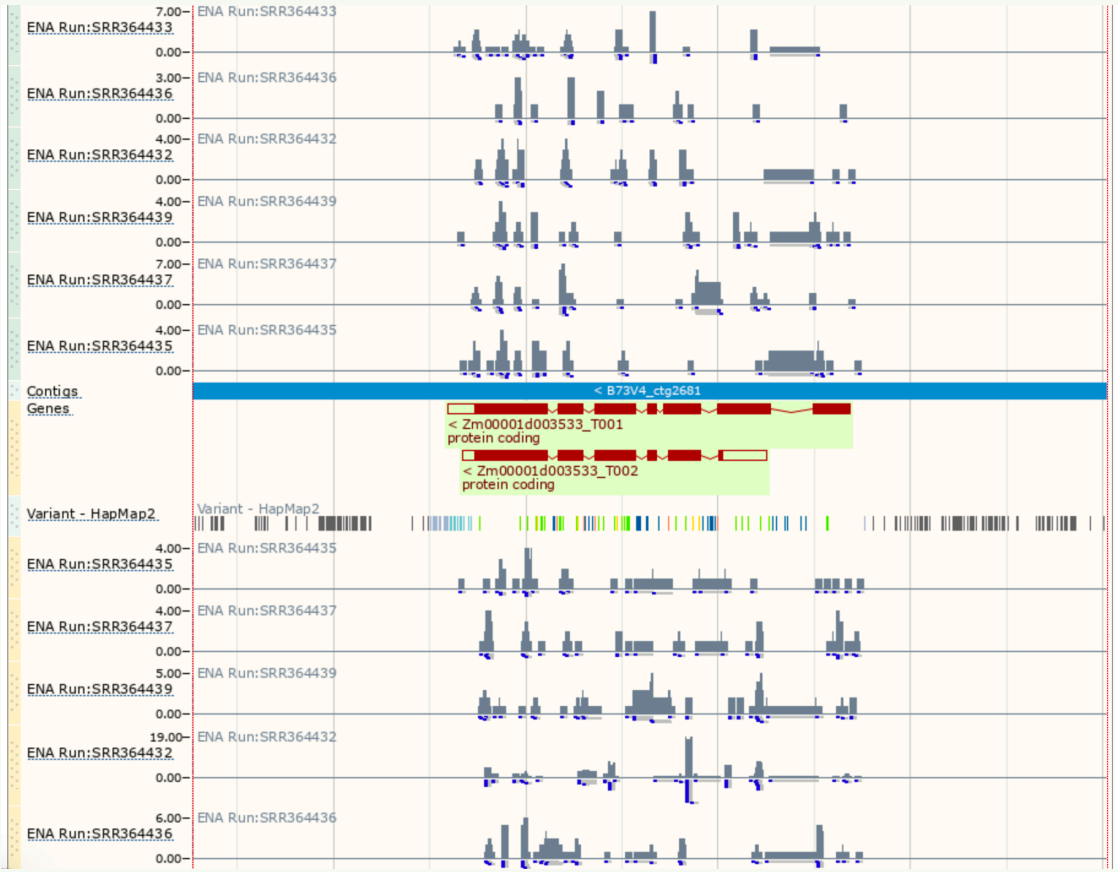
View patterns of gene expression

=> Visualize & download expression data from 800+ experiments in >24 plant species (EBI-Atlas)



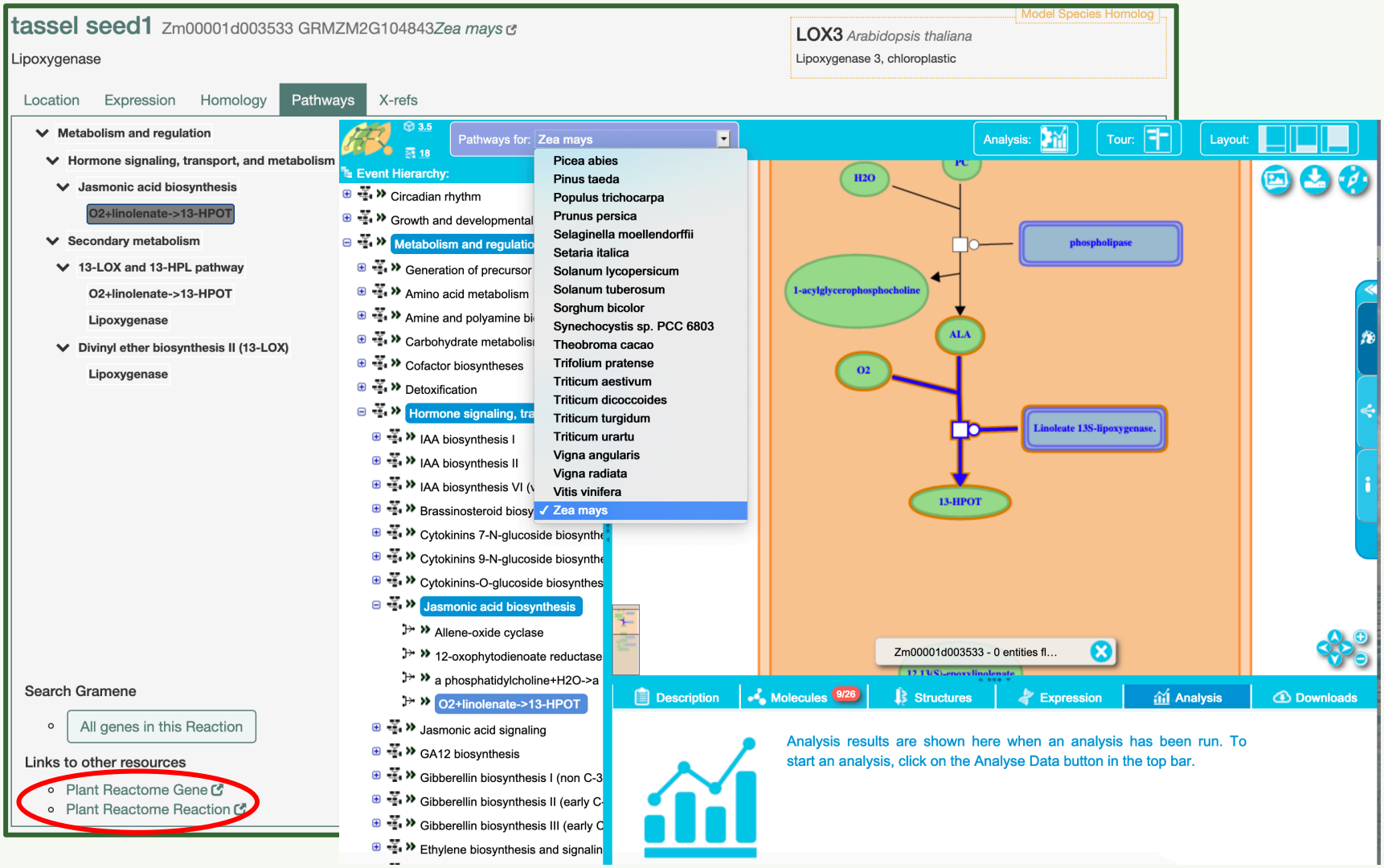
Project publicly available transcriptomics data

=> Find patterns of gene expression from 4,100 public RNA-Seq studies in Track Hub



RNA-Seq Experiment: SRP009313 – Parent-of-origin effects on gene expression & DNA methylation in the maize endosperm.

Identify pathways associated with my gene



=> Explore Pathway & Overlay Omics Data]

Plant Reactome
Gramene Pathways

Explore genetic variation

=> Identify germplasm with genetic variants

Show/hide columns | Search...

Variant ID	Chr: bp	Alleles	Source	Conseq. Type	AA	AA co-ord	SIFT	Transcript
tmp_6_47486610_G_A	6:46568309	G/A	Sorghum_EMS_mutants	missens variant			0.02	EES10882
tmp_6_47487006_C_T	6:46568705	C/T	Sorghum_EMS_mutants	missens variant			0.03	EES10882
tmp_6_47487513_G_A	6:46569212	G/A	Sorghum_EMS_mutants	splice a variant				EES10882
tmp_6_47487541_C_T	6:46569240	C/T	Sorghum_EMS_mutants	missens variant			0.02	EES10882
tmp_6_47487823_G_A	6:46569522	G/A	Sorghum_EMS_mutants	missens variant			0	EES10882
tmp_6_47488026_C_T	6:46569725	C/T	Sorghum_EMS_mutants	missens variant			0	EES10882
tmp_6_47488163_G_A	6:46569862	G/A	Sorghum_EMS_mutants	missens variant			0.03	EES10882
tmp_6_47488256_G_A	6:46569955	G/A	Sorghum_EMS_mutants	missens variant			0.01	EES10882
tmp_6_47488613_C_T	6:46570312	C/T	Sorghum_EMS_mutants	missens			0.02	EES10882
tmp_6_474								
tmp_6_474		G/A	_mutants	missense variant		279	0.01	
tmp_6_47487006_C_T	6:46568705	C/T	Sorghum_EMS_mutants	missense variant	A/V	333	0.34	OQU81659

MSD2 phenotype – Jiao *et al* (2016)
"A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses."

EMS-mutant collections for sorghum & wheat

Thank you & Contact Us!

Posters # **300-011**
300-012

- Outreach: http://gramene.org/outreach
- Slides, exercises, brochures, posters: ftp://ftp.gramene.org/pub/gramene/outreach
- E-mail: feedback@gramene.org
- Collections in Plantae and YouTube
- Mailing list & social networks

AgBioData
Booth # **602**

facebook
https://www.facebook.com/Gramene

@GrameneDatabase

Workshop on Cereal Genomics

October 15 - 21, 2019 Applications Due: August 15

Cold Spring Harbor Laboratory
MEETINGS & COURSES PROGRAM

Cold Spring Harbor, New York
web: **meetings.cshl.edu/cereal19**

This one-week workshop in Cereal Genomics will enable participants to take advantage of emerging genetic tools and the completed cereal genome sequences of most of the major cereal crops.

The workshop will feature morning and evening lectures with afternoon lab exercises, including hands-on lab work and computer sessions in comparative anatomy, GWAS, Next Gen expression analysis, whole genome sequencing assembly, emerging model systems, genome editing and phenomics.

The faculty (instructors and invited lecturers) are active researchers in cereal genetics and genomics who have made significant contributions to the field, ensuring that the latest techniques and ideas will be presented. The course will be structured to provide time for informal discussions and exchange with instructors. Students from academic institutions are eligible to apply for partial scholarships.

INSTRUCTORS

Sarah Hake, USDA/UC Berkeley Plant Gene Expression Center

David Jackson, Cold Spring Harbor Laboratory

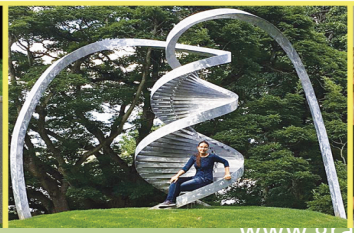
Doreen Ware, USDA/Cold Spring Harbor Laboratory

TOPICS

- **Comparative anatomy and phylogeny**
- **Cereal genomes, assembly, annotation and synteny**
- **Genetics and databases**
- **Quantitative trait locus mapping and genome wide association studies**
- **Genome wide expression analyses**
- **Reverse genetics and genome editing**
- **Phenomics**

2019 SPEAKERS

Edward Buckler, USDA-ARS
Katrien Devos, University of Georgia
Jorge Dubcovsky, University of California Davis
Andrea Eveland, Donald Danforth Plant Science Center
Lisa Harper, University of California Berkeley
Elizabeth Kellogg, Donald Danforth Plant Science Center
Devin O'Connor, Pairwise Plants
Uta Paszkowski, University of Cambridge, UK
Christopher Topp, Donald Danforth Plant Science Center
Bing Yang, Donald Danforth Plant Science Center



Post-doctoral positions available in the Ware Lab