

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

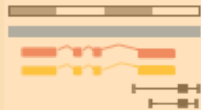
- Gramene is a Comparative Database for Plants since 2001
- 61 plant genomes with gene family trees, sequence alignments, synteny, homologs (EBI-Ensembl)
- Gene expression (EBI-Atlas)
- Pathways (OSU-Reactome)
- Automated ingestion of functional annotations from publications (ASPB)



61 reference
Genomes



2.3M gene
annotations



92.7K gene
trees



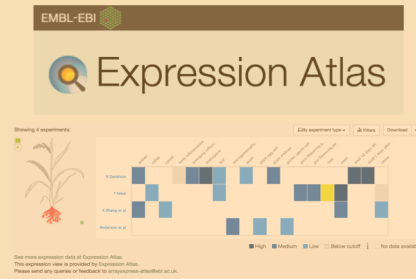
278 whole-
genome
alignments

ATCGAGCT
ATCCAGCT
ATCGAGAT

65
syntenies



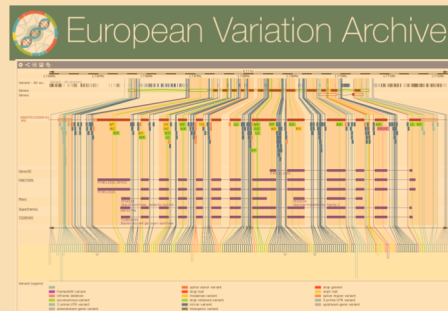
Expression



804 experiments in
24+ species

**Integrated search
and visualization**

Variation



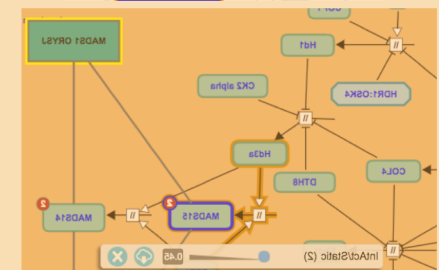
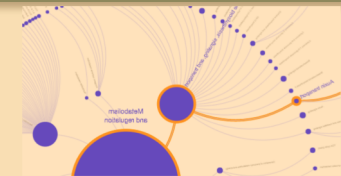
~223.7M SNPs
in 12 genomes

Pathways

298 curated rice
projected to
82 species



Plant REACTOME



Gramene.Org



Search for genes, species, pathways, ontology terms, domains...



2264698 genes in 61 genomes ▾

Gramene Portals



Genome Browser

Gene Trees, Homologs, Synteny,
Whole-Genome Alignments



Plant Reactome

Browse and analyze metabolic and regulatory pathways

Pathways



Tools

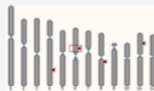
Tools for processing both our data and yours



Plant Expression ATLAS

Bro

Gene Expression



BLAST

Query our genomes with a DNA sequence

Tools &
External Data

Gramene Mart

An advanced query interface powered by BioMart



Track Hub Registry

A global centralised collection of publicly accessible track hubs



Outreach and Training

Outreach, Education & Training

Annotation Jamboree for
rs. St. Louis, MO. March



Bulk Downloads

FTP download of our data



Archive

Legacy tools and data (markers, Cyt pathways, etc)

Gramene will be at Plant Biology 2019 –
Hope to see you in San Jose, CA!
Fri, 26 Jul 2019

Community Annotation Poster at SUNY Old
Westbury's 2019 Student Research Day
Conference
Tue, 04 Jun 2019

Saturday DNA: Get to Know GMOs! - A
WiSE event
Tue, 04 Jun 2019

Plant domestication and the secrets in a
plant's genome for 4th graders
Mon, 03 Jun 2019

Announcing Gramene Knowledgebase
release #61 with new and updated plant
genomes and pathways
Thu, 02 May 2019

Workshop on Cereal Genomics at CSHL,
October 15 - 21, 2019. Apply by Aug 15
Thu, 25 Apr 2019

Annotation Jamboree for
rs. St. Louis, MO. March
13-14, 2019
Wed, 17 Apr 2019

Announcing Gramene release #60: New &
updated genomes, pathways and other
features
Fri, 08 Feb 2019



Gene & Protein
Expression



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



Search: *tassel seed1*

ms, domains...



3 genes in 2 genomes ▾

Gene | *tassel seed1* ✕

Show All Species

Pick Species from List

tassel seed1 Zm00001d003533 GRMZM2G104843 *Zea mays* ♂

Lipoxygenase

Location Expression **Homology** Pathways X-refs

Compara Gene Tree

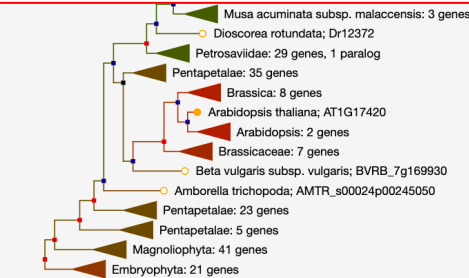
This phylogram shows the relationships between this gene and others similar to it, as determined by Ensembl Compara ⌵

Display mode ▾ ⚙

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

Zea mays: Zm00001d003533

Sorghum ortholog: **SORBI_3006G095600** [*msd2*]



Search Gramene

Show All Homologs 205

Show Orthologs 104

Show Paralogs 3

Links to other resources

Ensembl Gene Tree view ⌵

LOX3 *Arabidopsis thaliana*
Lipoxygenase 3, chloroplastic

Model Species Homolog

For specific tasks,
check out our YouTube
video tutorials

Os04g0447100 *Oryza sativa Japonica Group* ♂

LOX3 *Arabidopsis thaliana*

Model Species Homolog



Gene Trees: Focus on species of interest

=> *Prune unwanted species*

tassel seed1 Zm00001d003533 GRMZM2G104843Zea mays

Lipoxygenase

Location Expression Homology Pathways X-refs

12 genes in 8 genomes

Show All Species

Pick Species from List

LOX3 Arabidopsis thaliana
Lipoxygenase 3, chloroplastic

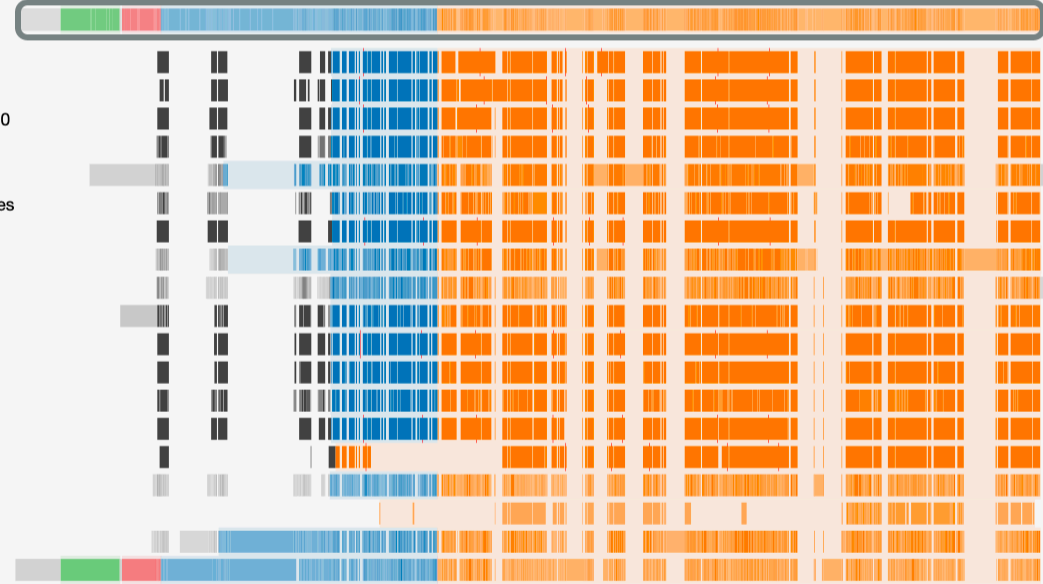
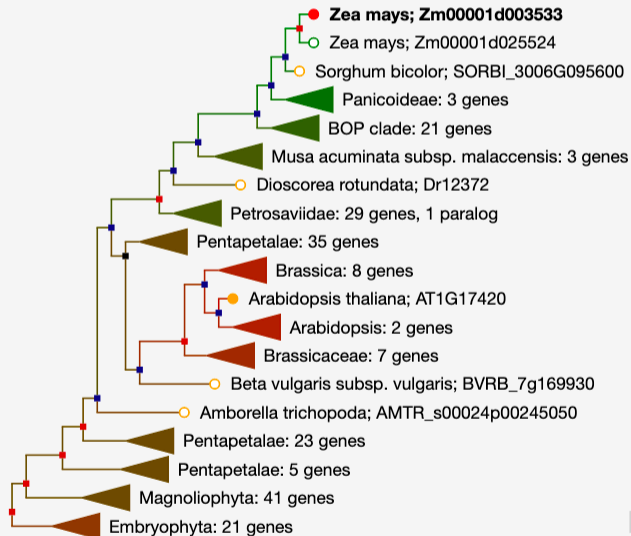
Compara Gene Tree

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*



tassel seed1 Zm00001d003533 GRMZM2G104843Zea mays

Lipoxygenase

Model Species Homolog

LOX3 Arabidopsis thaliana

Lipoxygenase 3, chloroplastic

Location Expression **Homology** Pathways X-refs

Compara Gene Tree

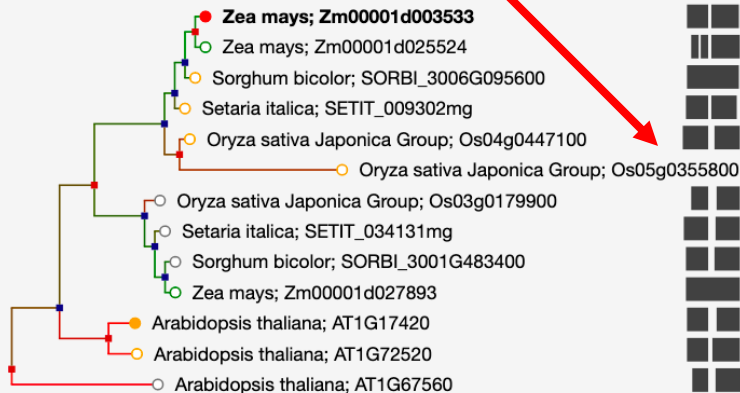
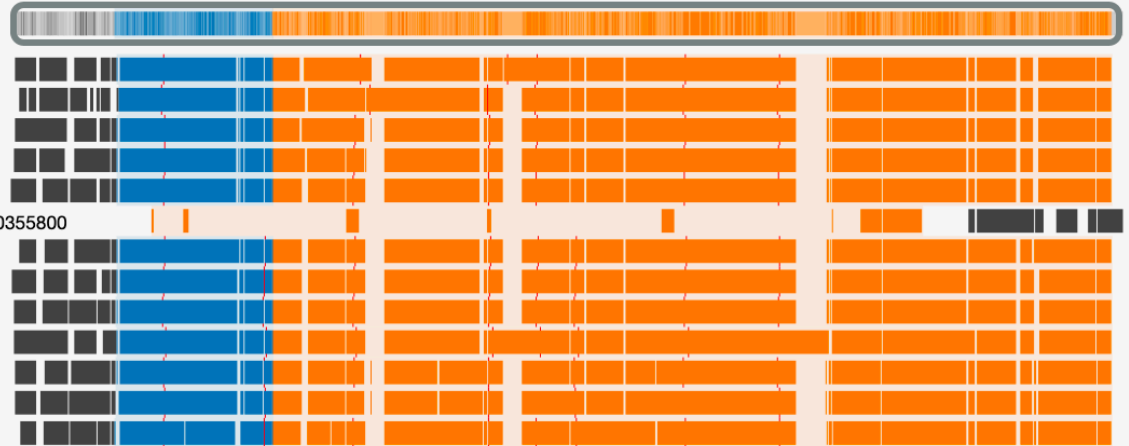
Artifact or Biology?

This phylogram shows the relationship between genes identified by Ensembl Compara

Display mode

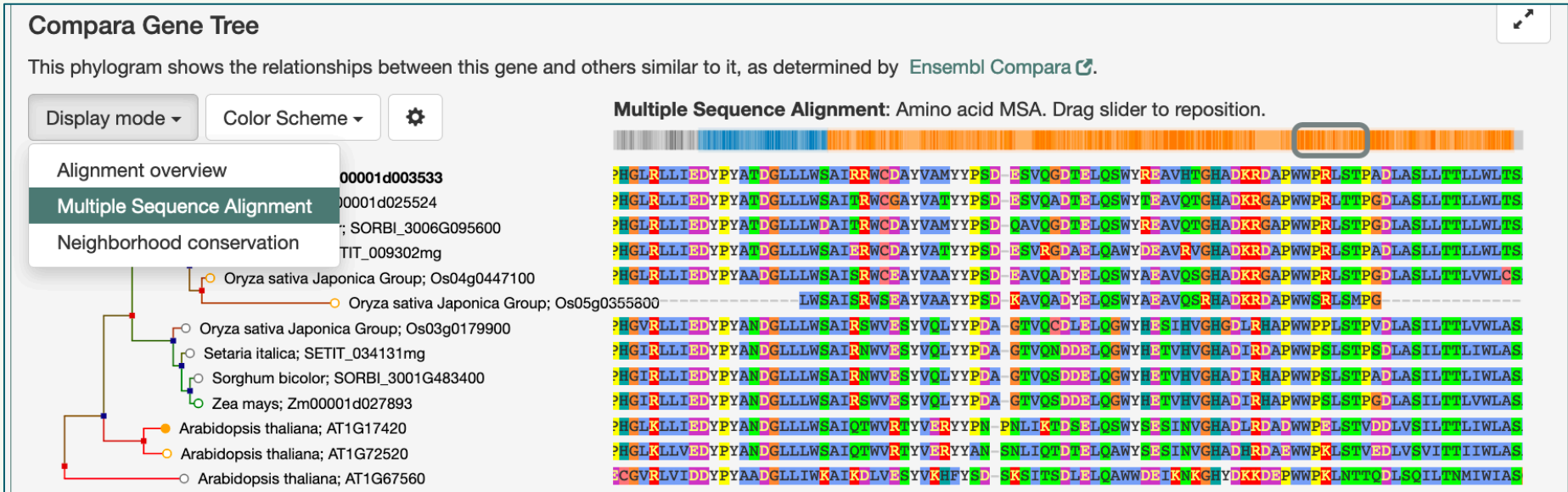


Alignment overview: Proteins color-coded by InterPro domain. Resize slider to navigate.

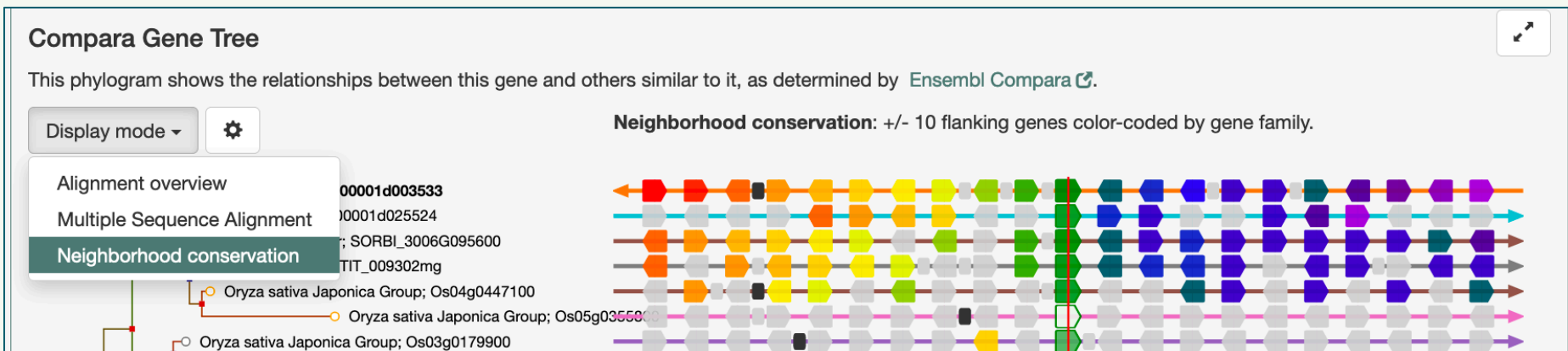


Explore a gene family tree

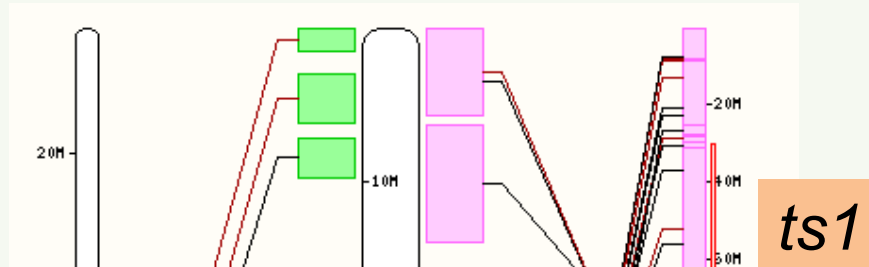
=> *Zoom In to view aligned amino acids*



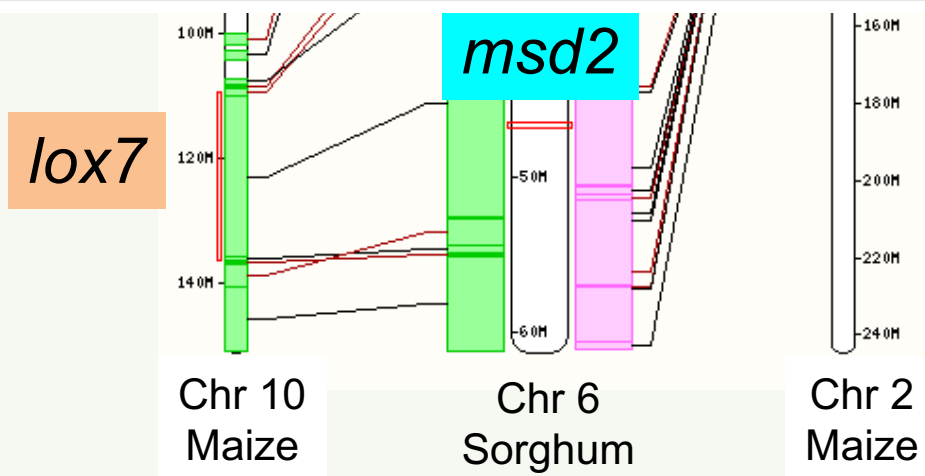
=> *Zoom Out to view conserved gene neighborhoods*



Explore synteny



Show/hide columns		Filter		
<i>Sorghum bicolor</i> genes	Location	<i>Zea mays</i> homologues	Location	
SORBI_3006G095600 (SORBI_3006G095600)	6:46566240-46571064	→ tassel seed1 (Zm00001d003533)	2:47105187-47109372	Region Comparison
		→ lipoxygenase7 (Zm00001d025524)	10:121268606-121272616	Region Comparison



- Comparative Region View
- Download Gene Lists

Explore gene structure

=> Identify protein domains



tassel seed1 Zm00001d003533 GRMZM2G104843Zea mays [↗](#)

Lipoxygenase

Model Species Homolog

LOX3 *Arabidopsis thaliana*

Lipoxygenase 3, chloroplastic

Location Expression **Homology** Pathways X-refs

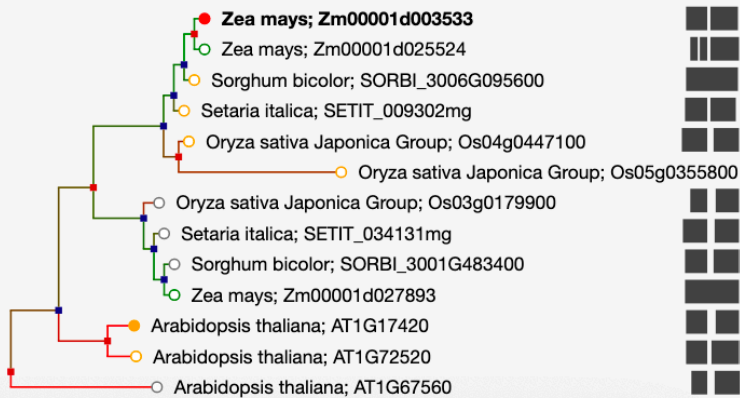
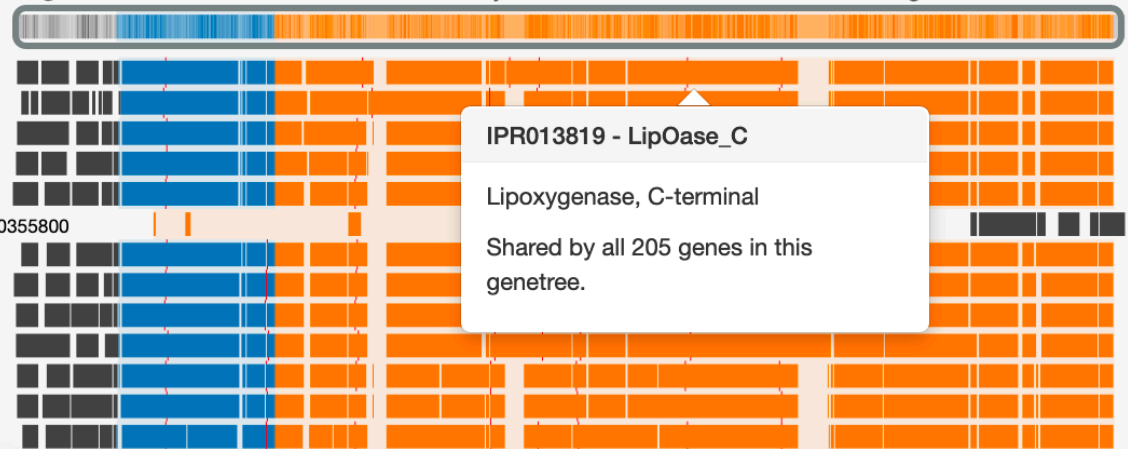
Compara Gene Tree

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.

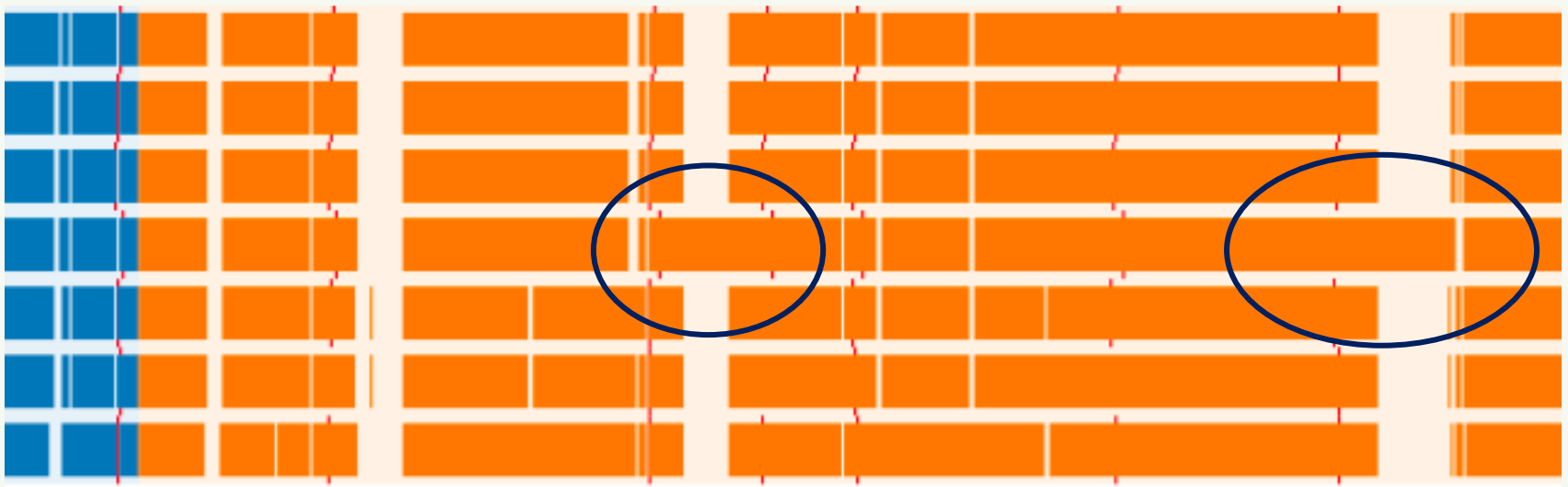


Explore gene structure

=> *Identify splice junctions*

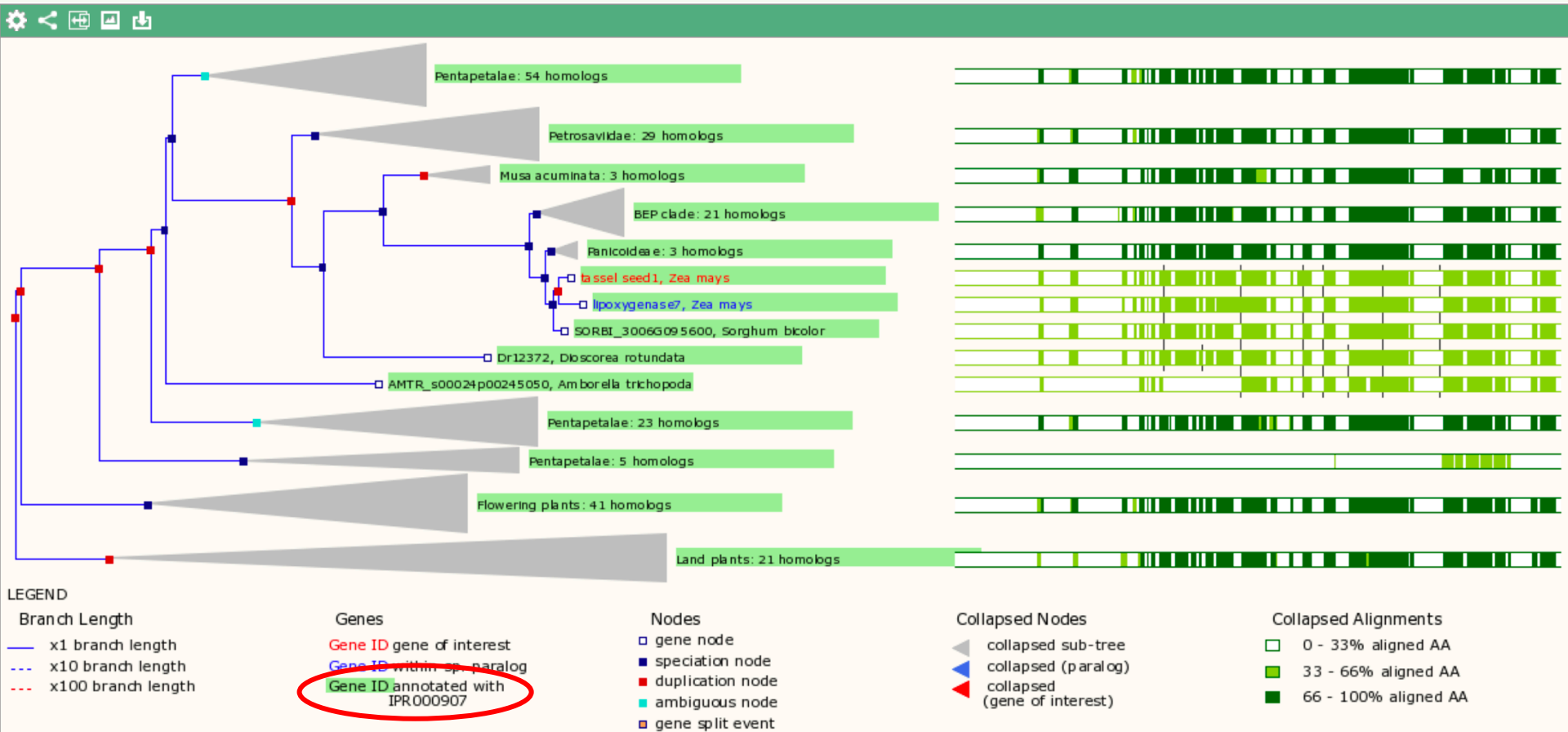


Artifact or Biology?



Explore functional gene structure

=> Highlight InterPro & GO annotations



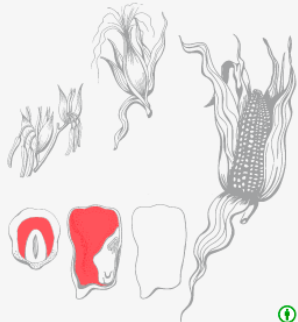
View patterns of gene expression

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

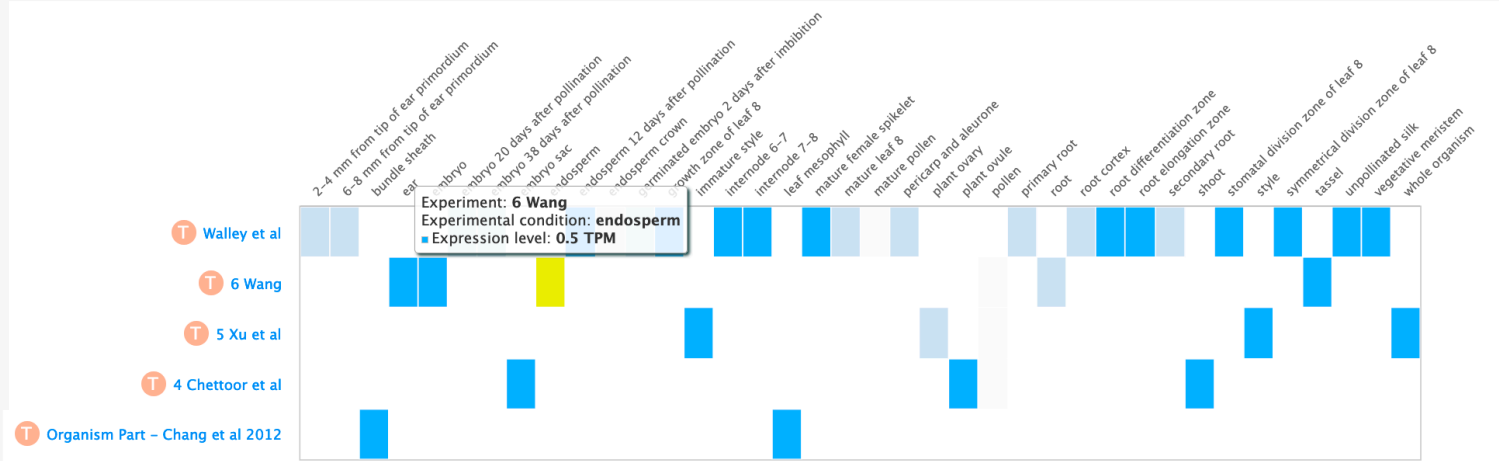
Location **Expression** Homology Pathways X-refs

Showing 5 experiments:

↓ By experiment type Filters Download



endosperm



- Walley et al
- 6 Wang
- 5 Xu et al
- 4 Chettoor et al
- Organism Part - Chang et al 2012

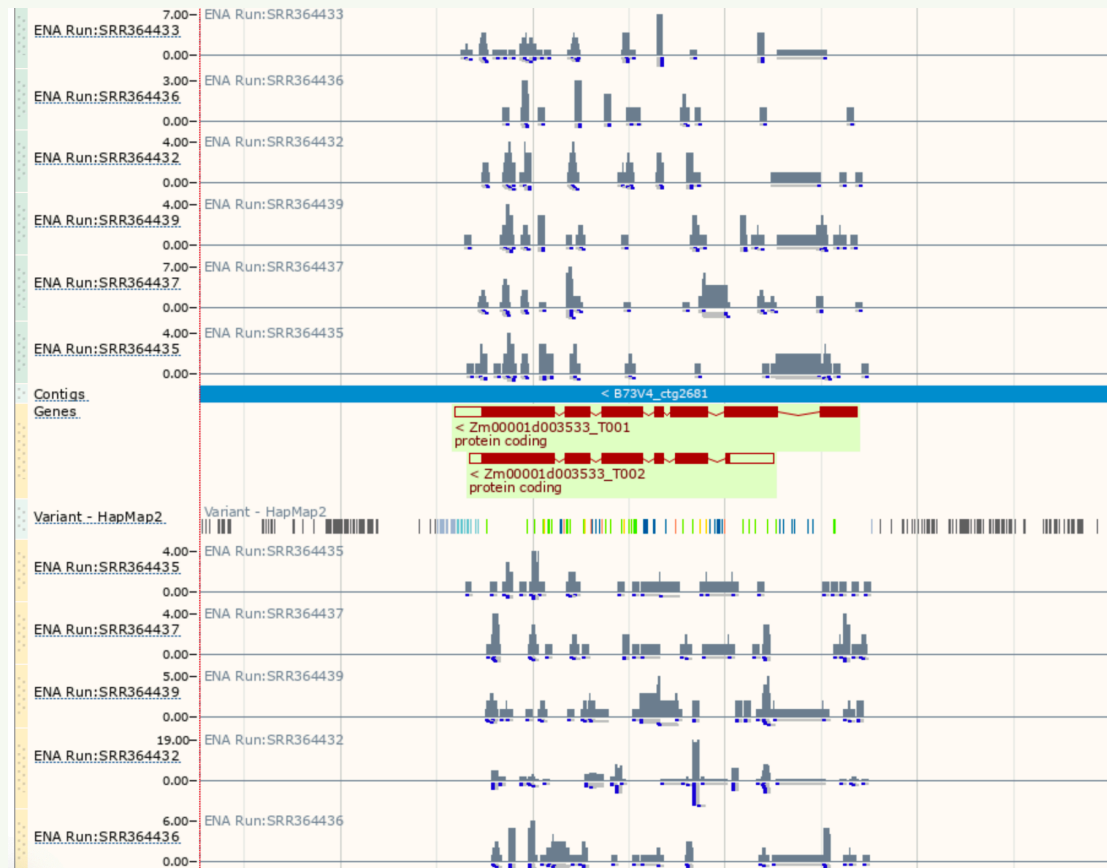
Experiment: 6 Wang
 Experimental condition: endosperm
 Expression level: 0.5 TPM

Legend: High (dark blue), Medium (medium blue), Low (light blue), Below cutoff (grey), No data available (white). Legend also includes Proteomics (P) and Transcriptomics (T).



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

tassel seed1 Zm00001d003533 GRMZM2G104843Zea mays [↗](#)

Lipoxygenase

Location Expression Homology **Pathways** X-refs

Model Species Homolog
LOX3 *Arabidopsis thaliana*
Lipoxygenase 3, chloroplastic

Metabolism and regulation

Hormone signaling, transport, and metabolism

Jasmonic acid biosynthesis

Q2+linolenate->13-HPOT

Secondary metabolism

13-LOX and 13-HPL pathway

O2+linolenate->13-HPOT

Lipoxygenase

Divinyl ether biosynthesis II (13-LOX)

Lipoxygenase

Search Gramene

All genes in this Reaction

Links to other resources

Plant Reactome Gene [↗](#)

Plant Reactome Reaction [↗](#)

Event Hierarchy:

3.5 18 Pathways for: Zea mays

Picea abies
Pinus taeda
Populus trichocarpa
Prunus persica
Selaginella moellendorffii
Setaria italica
Solanum lycopersicum
Solanum tuberosum
Sorghum bicolor
Synechocystis sp. PCC 6803
Theobroma cacao
Trifolium pratense
Triticum aestivum
Triticum dicoccoides
Triticum turgidum
Triticum urartu
Vigna angularis
Vigna radiata
Vitis vinifera
Zea mays

Analysis: Tour: Layout:

Zm00001d003533 - 0 entities fl...

Description Molecules 9/20 Structures Expression Analysis Downloads

Analysis results are shown here when an analysis has been run. To start an analysis, click on the Analyse Data button in the top bar.

=> Explore Pathway & Overlay Omics Data]

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**



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



[@GrameneDatabase](https://twitter.com/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

