



FIFA WORLD CUP
RUSSIA 2018

Thank You, Merci, Gracias!



Y el Campeón es...



FRANCIA

_ R _ _ C I A



CROACIA

The last 3 World Cups have all gone to extra-time...

Gramene - Exploring Function through Comparative Genomics and Network Analysis



Transnational collaboration



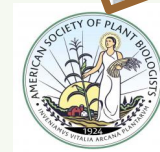
Doreen Ware, PI (USDA ARS, CSHL)
Yinping Jiao, Sunita Kumari, Demetri Muna, Andrew Olson,
Joshua Stein, Marcela K. Tello-Ruiz, Bo Wang, Sharon Wei

Pankaj Jaiswal, Co-PI (OSU)
Noor Al-Bader, Justin Elser, Matthew Geniza, Parul Gupta,
Sushma Naithani, Justin Preece

Paul Kersey / Irene Papatheodorou (EMBL-EBI)
Dan Bolser, Christopher Grabmuller, Chuang Kee Ong, Dan
Staines, Brandon Walts / Maria Keays, Alfonso Muñoz-
Pomer Fuentes, Laura Huerta Martínez

Lincoln Stein (OICR)
Peter D' Eustachio (NYU); Robin Haw, Joel Weiser, Sheldon
McKay; Antonio Fabregat (EBI)

Crispin Taylor (ASPB)
Jennifer Regala; Weijia Xu (TACC), Amit Gupta(TACC)



Where can I learn more about Gramene @ ASPB 2018?

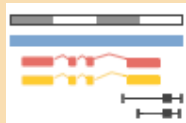
- [Monday July 16th, 1:00-1:30 PM @ *Plantae Pavillion* - Getting involved with Gramene](#)
- Poster #300-008
- AgBioData booth #1420

53 reference
Genomes





2.1M gene
annotations



67K gene
trees



318 whole-
genome
alignments

ATCGAGCT
ATCCAGCT
ATCGAGAT

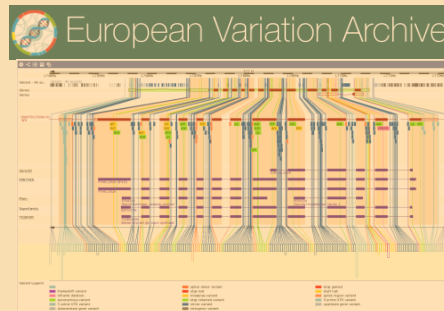
Expression



780 experiments in
22+ species

***Integrated search
and visualization***

Variation



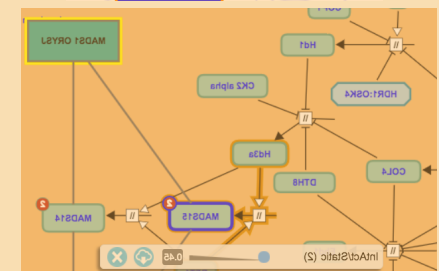
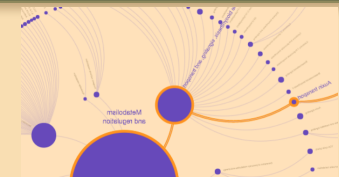
~231M SNPs in
9 genomes

Pathways

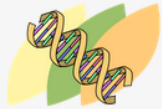
264 curated rice
projected to
74 species



Plant REACTOME



Gramene Portals



Genome Browser

Browse genomes with annotations, variation and comparative tools



Plant Reactome

Browse and analyze metabolic and regulatory pathways



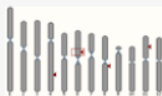
Tools

Tools for processing both our data and yours



Plant Expression ATLAS [↗](#)

Browse plant expression results at EBI ATLAS



BLAST

Query our genomes with a DNA or protein sequence



Gramene Mart

An advanced query interface powered by BioMart



Track Hub Registry [↗](#)

A global centralised collection of publicly accessible track hubs



Outreach and Training

Educational resources and webinars



Bulk Downloads

FTP download of our data



Archive

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

Latest News

Gramene will be at the Plant Biology Conference 2018 — Montreal, July 14-18
Fri, 22 Jun 2018

Strawberry DNA extraction with third graders
Thu, 14 Jun 2018

Mining Maize with Gramene - Free Webinar
May 22, 2018 @ 2 pm EDT
Wed, 16 May 2018

The Gramene Database build 57 is out with a new polyploid genomes view!
Thu, 03 May 2018

Gramene Workshop at the 2018 Maize Genetics Meeting in Saint-Malo, France
Tue, 06 Mar 2018

Gramene webinar Feb 27, 2018: Plant Reactome pathway updates and new features
Mon, 19 Feb 2018

Gramene release # 56b: updates to Plant Reactome
Thu, 01 Feb 2018

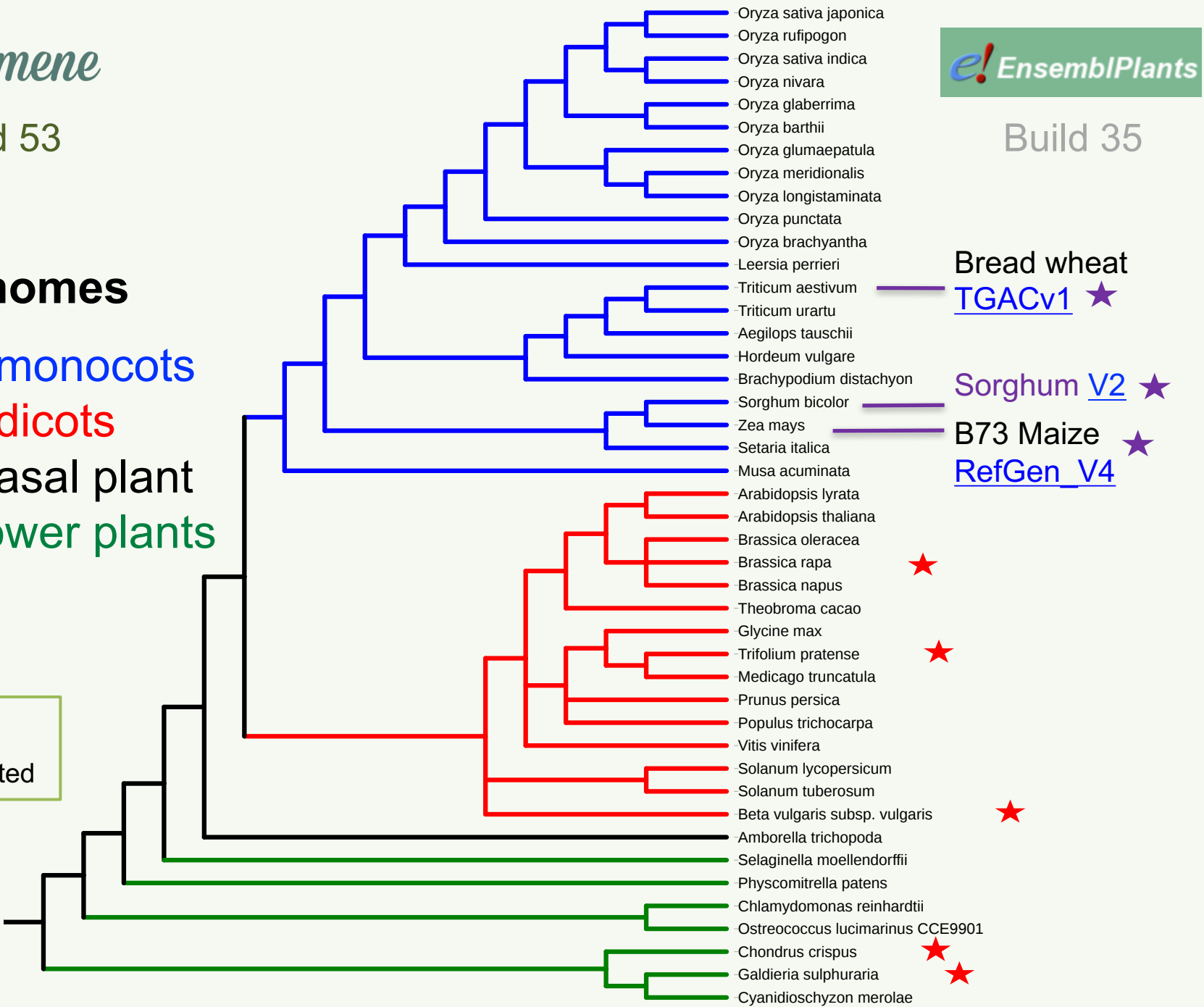
The Gramene Database build 56 is out with 8 new plant genomes!
Mon, 15 Jan 2018

Gramene will be at the PAG Conference 2018 — See you in San Diego!
Mon, 08 Jan 2018

53 Genomes

- 22 monocots
- 23 dicots
- 1 basal plant
- 7 lower plants

★ New
★ Updated



Bread wheat
[TGACv1](#) ★

Sorghum [V2](#) ★

B73 Maize
[RefGen_V4](#) ★

The Banana Genome Hub

Cacao Genome Database



www.panzea.org



Whole is greater than the sum of its parts



Brachypodium.org



Amborella



Gramene Adds Value to Plant Genomes

Annotation Pipelines

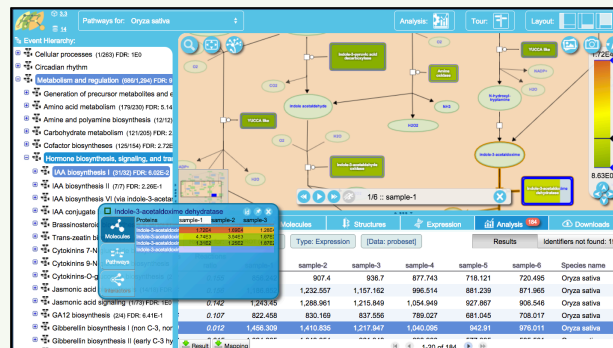
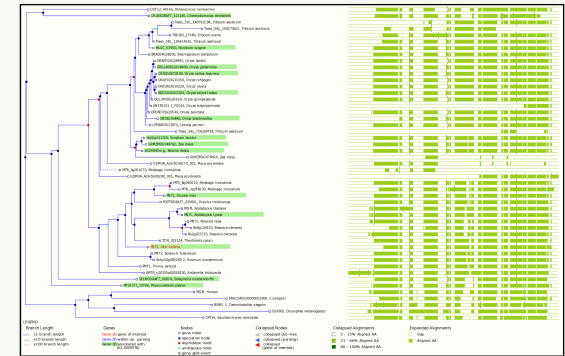
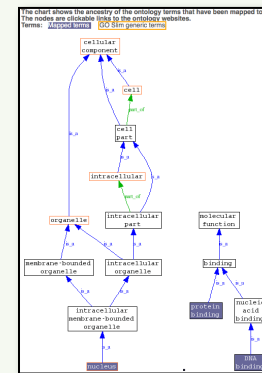
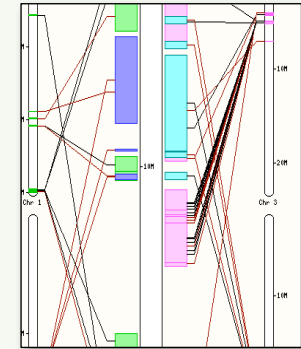
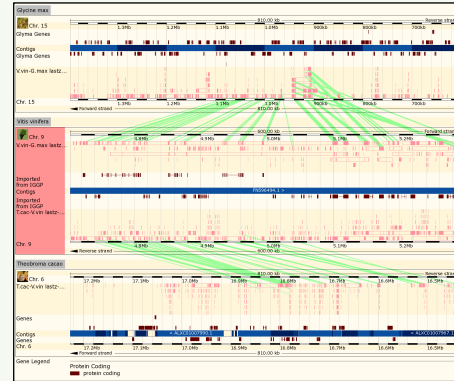
- Repeats/TEs
- Genes
- EST/cDNA
- InterPro domains
- Gene Ontology (GO)
- Variant Effect Prediction
- Expression (EBI Atlas)
- Pathway curation & projection

Comparative Analysis

- Whole Genome Alignment
- Phylogenetic Gene Trees
- Ortholog/Paralog calling
- Synteny mapping
- Pathway visualization

Programmatic Access:

- Gramene API
- Ensembl & Reactome – API & RESTFUL interface
- BioMart
- Public MySQL server

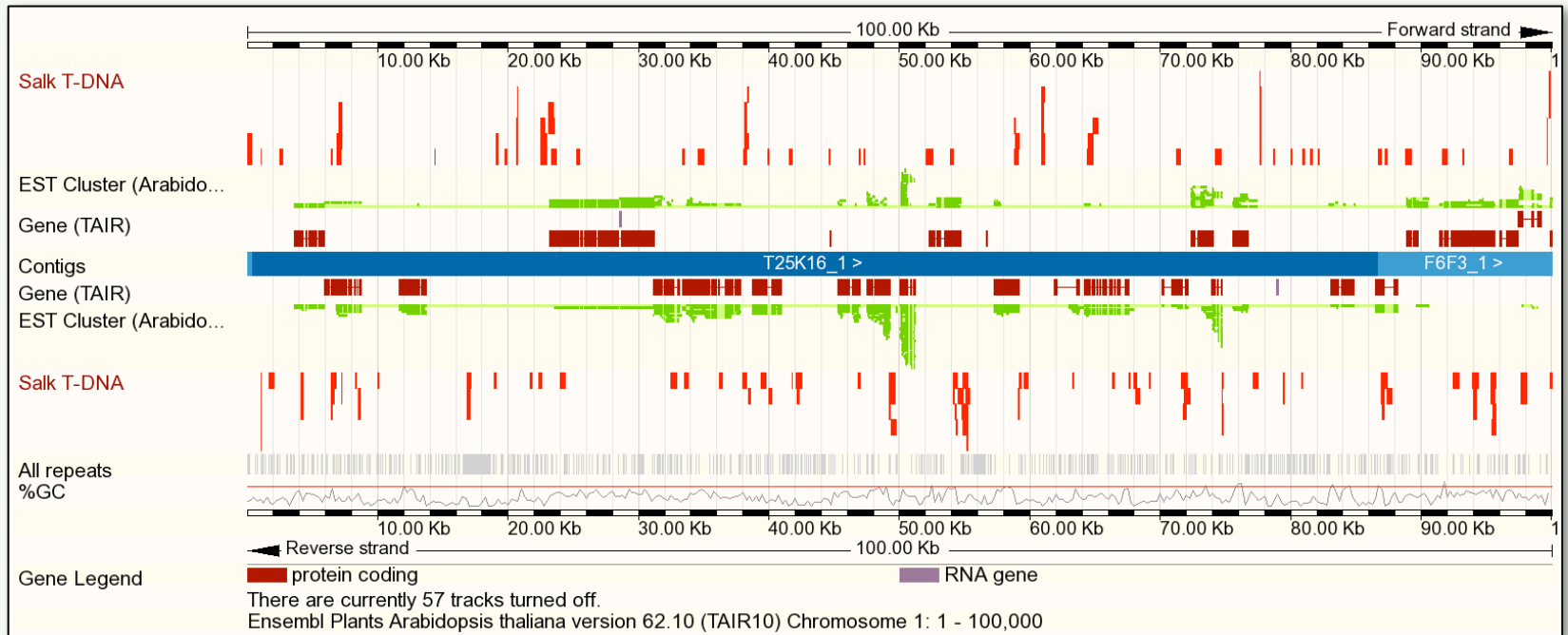


Tools:

- Search
- BLAST



Genome Browsers



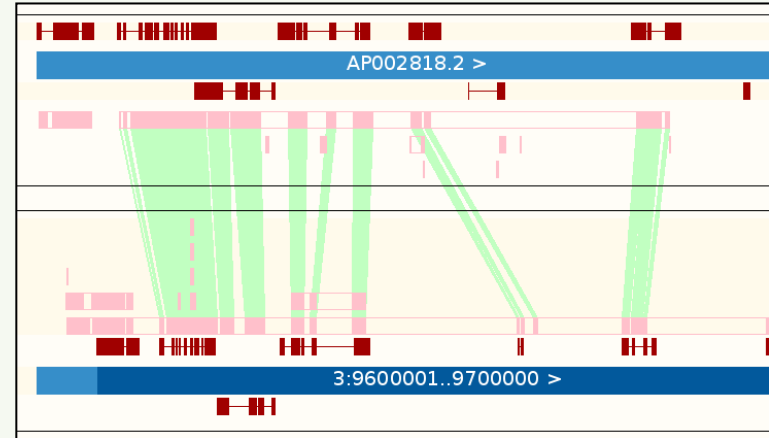
- Visualize genomic sequence and annotations
- Upload your own data (e.g., distribution of Salk T-DNA mutant insertion lines in GFF file format)
- Share data via Tracks Hubs

Ensembl Comparative Phylogenomics

Reconstructing evolutionary histories

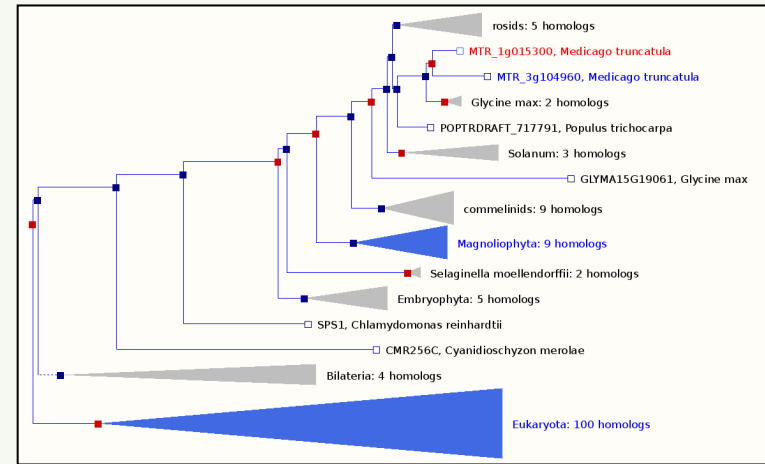
Whole Genome DNA Alignment

- Pairwise alignments
- B/LASTZ-CHAIN-NET method



Phylogenetic Gene Trees

- Gene families
- Infers orthologs and paralogs
- Taxonomic dating
- Synteny mapping



Vilella A.J., et al. (2008). *Genome Res.* doi:10.1101/gr.073585.107

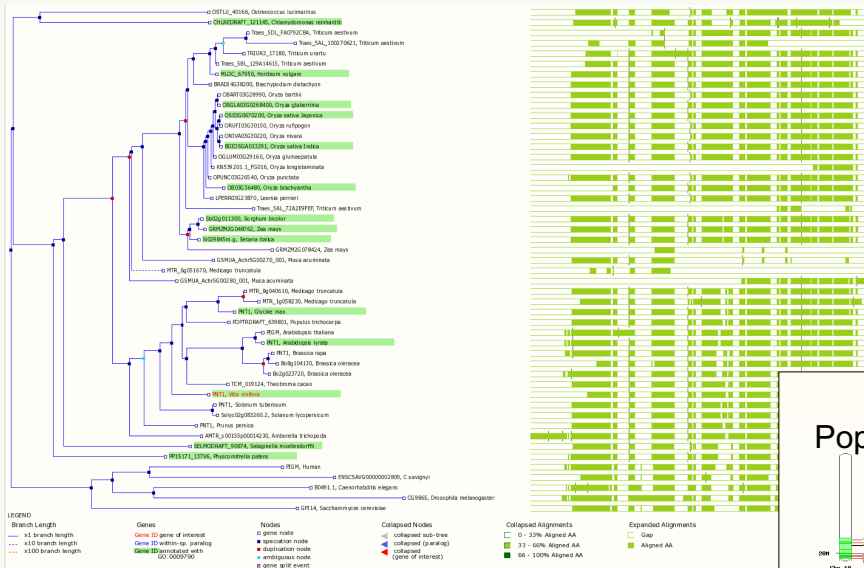
Schwartz S et al., *Genome Res.*;13(1):103-7, Kent WJ et al., *Proc Natl Acad Sci USA.*, 2003;100(20):11484-9

http://useast.ensembl.org/info/docs/compara/homology_method.html

Comparative Genomics Illustrations

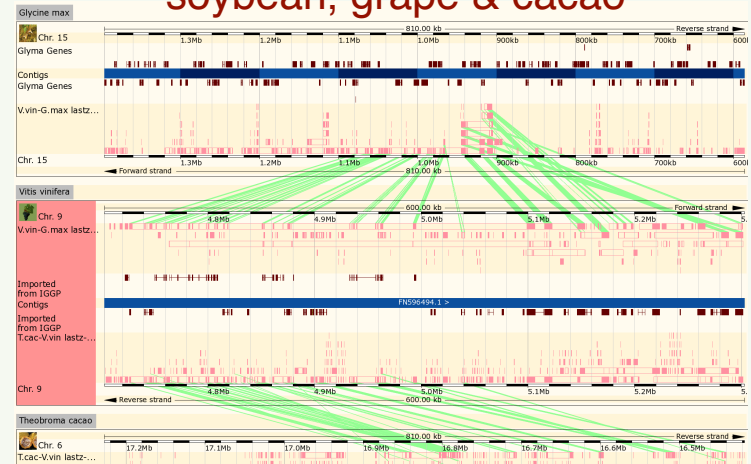
Phylogenetic tree

Glycosyltransferase genes associated with embryo development (GO:0009790) show conservation throughout the eukaryotic lineage



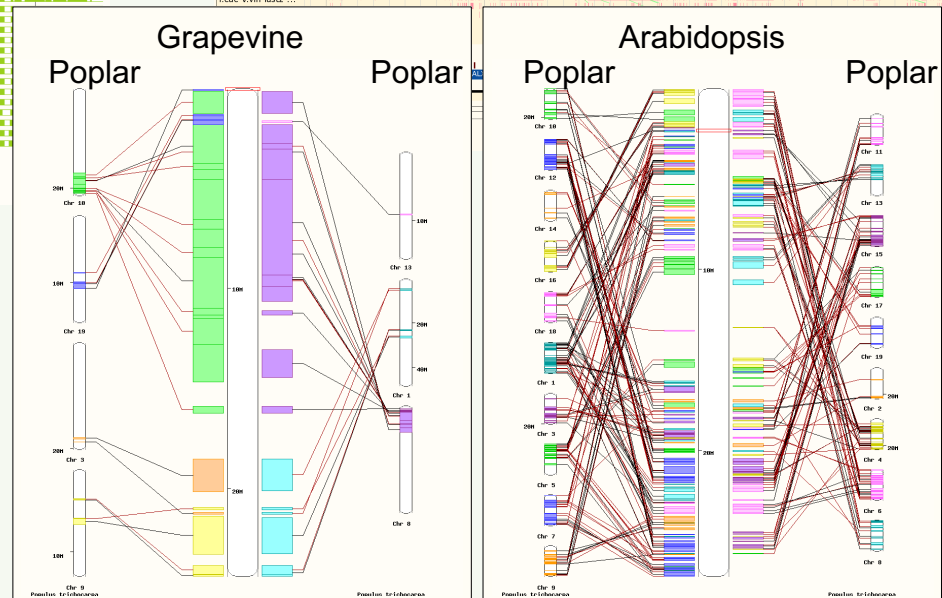
Multi-species DNA alignments

Conservation of micro-synteny between soybean, grape & cacao



Synteny

Whole-genome duplication in poplar & no duplication in grape make grape an excellent evolutionary Eudicot reference



231 Million Genetic Variants

SNPs & indels

EMS-induced

ENSVATH00125659 SNP

Most severe consequence | **Stop gained** | See all predicted consequences

Alleles | **A/T** | Highest population MAF: 0.22

Location | Chromosome 1:24628699 (forward strand) | [View in location tab](#)

HGVS names | This variant has 3 HGVS names - [Show](#)

Synonyms | **Perlegen PERL0223135**

Original source | Variation features from Affy_250k, Perlegen_1M, WTCHG and 1001 Genomes, with Ensembl identifiers

About this variant | This variant overlaps 2 transcripts, has 200 sample genotypes and is associated with 2 phenotypes.

Explore this variant

tmp_3_55893324_G_A SNP

Most severe consequence | **Stop gained** | See all predicted consequences

Alleles | **G/A** | Highest population MAF: < 0.01

Location | Chromosome 3:55905987 (forward strand) | [View in location tab](#)

HGVS names | This variant has 3 HGVS names - [Show](#)

Original source | [A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses](#) | Contact [Dr. Zhanguo Xin](#) for seeds

About this variant | This variant overlaps 1 transcript and has 251 sample genotypes.

Explore this variant

Structural variants (CNP, RFLP, SSR)

QTLs

Structural variant: ENSVATH00000001

Variation class | deletion (SO:0000159) [g?](#)

Source | Ensembl - Variation features from Affy_250k, Perlegen_1M, WTCHG and 1001 Genomes, with Ensembl identifiers

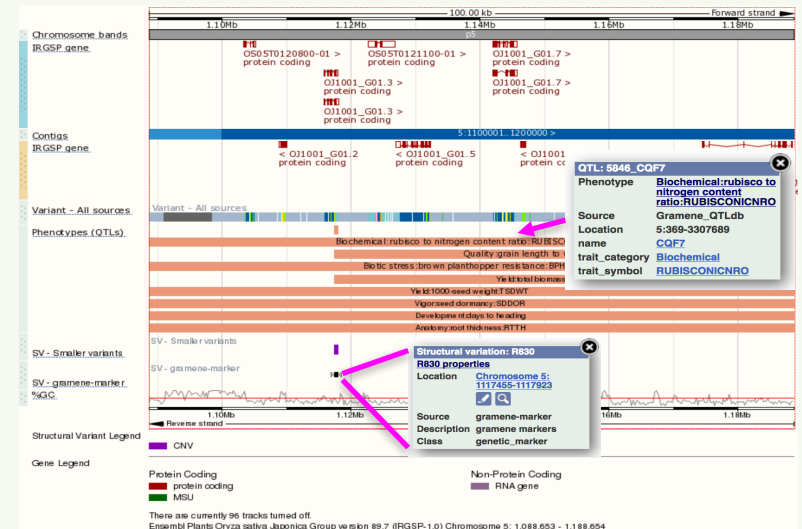
Study | [Structural variation](#) - unknown

Location | Chromosome 1:10019676-10020065 (forward strand) | [View in location tab](#)

Genomic size | 390 bp

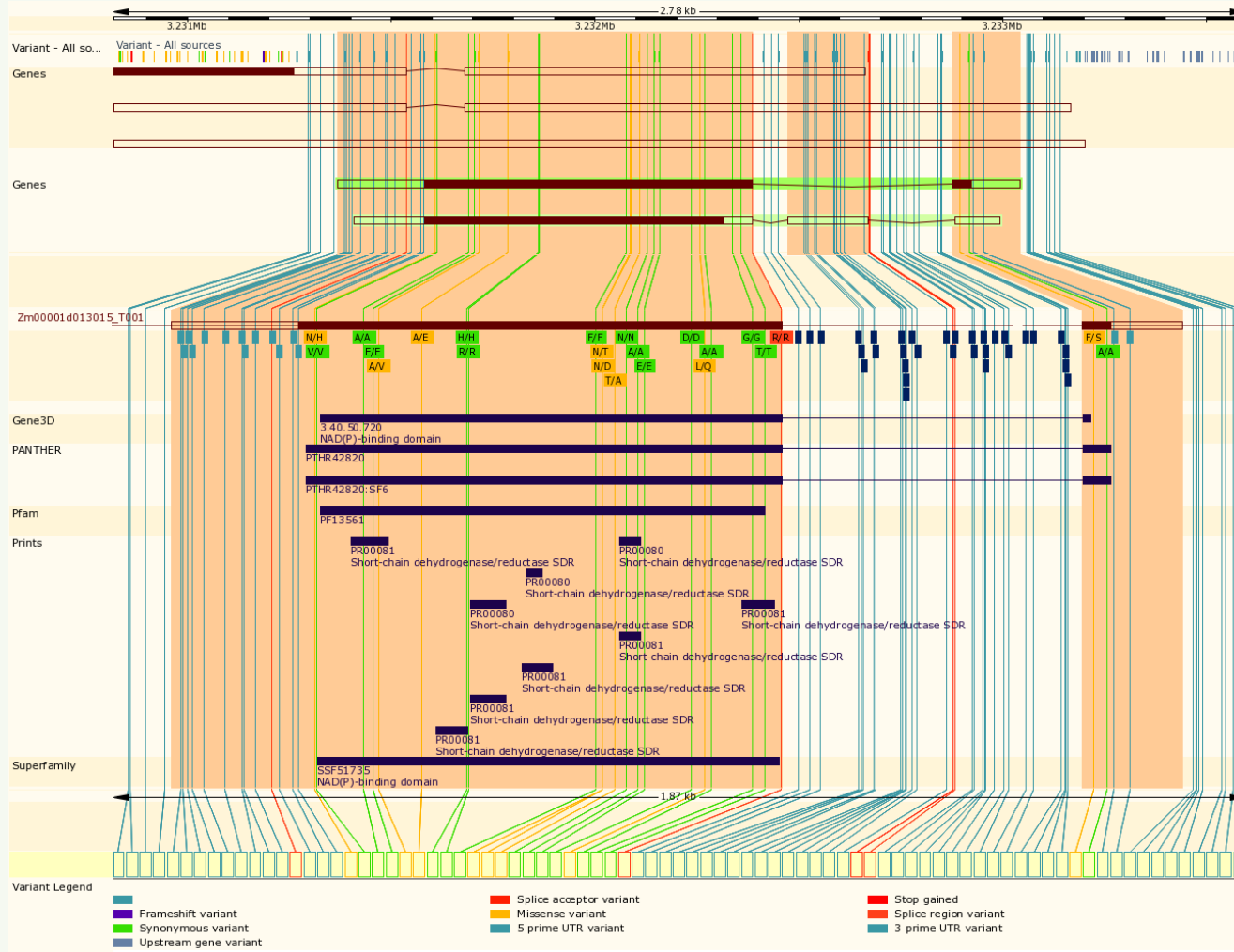
About this structural variant | This structural variant overlaps 1 transcript.

Explore this SV

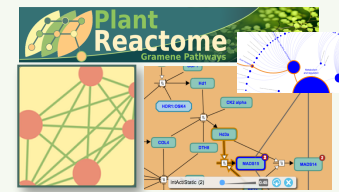


SNP Variant Effect Prediction in the context of InterPro domains

Short chain alcohol dehydrogenase 1 maize SNPs with predicted functional effects & projected to protein domains



<http://plantreactome.gramene.org>



Pathways

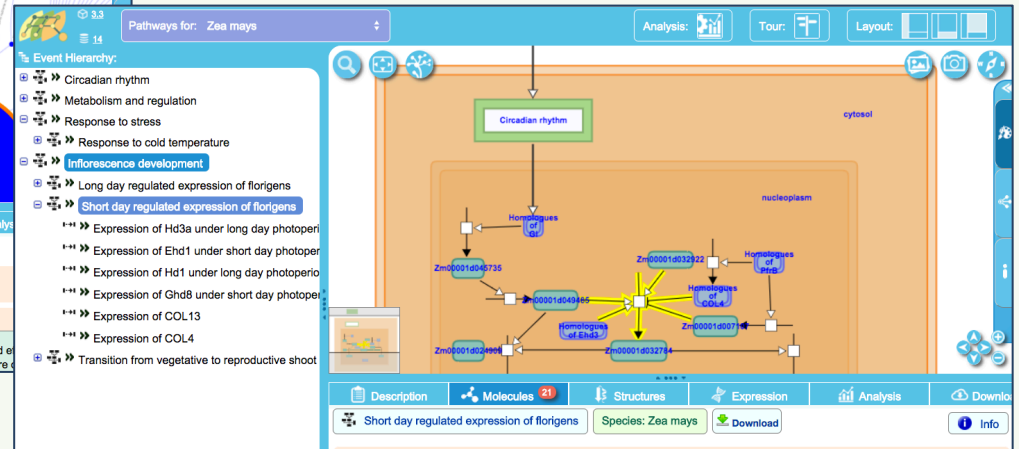
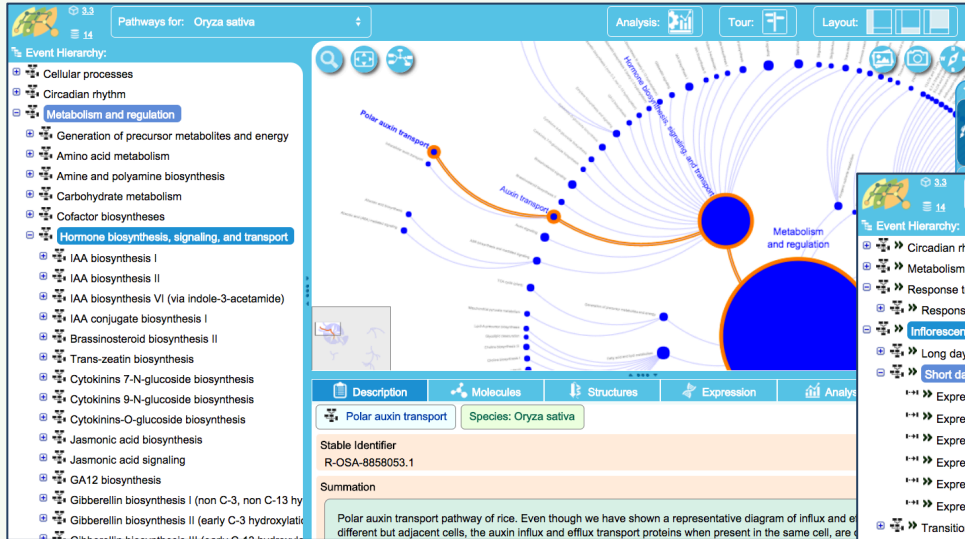
- 74 plant species
- Curated reference:
Oryza sativa japonica
- Super-cluster orthology projections:

Compara & Inparanoid

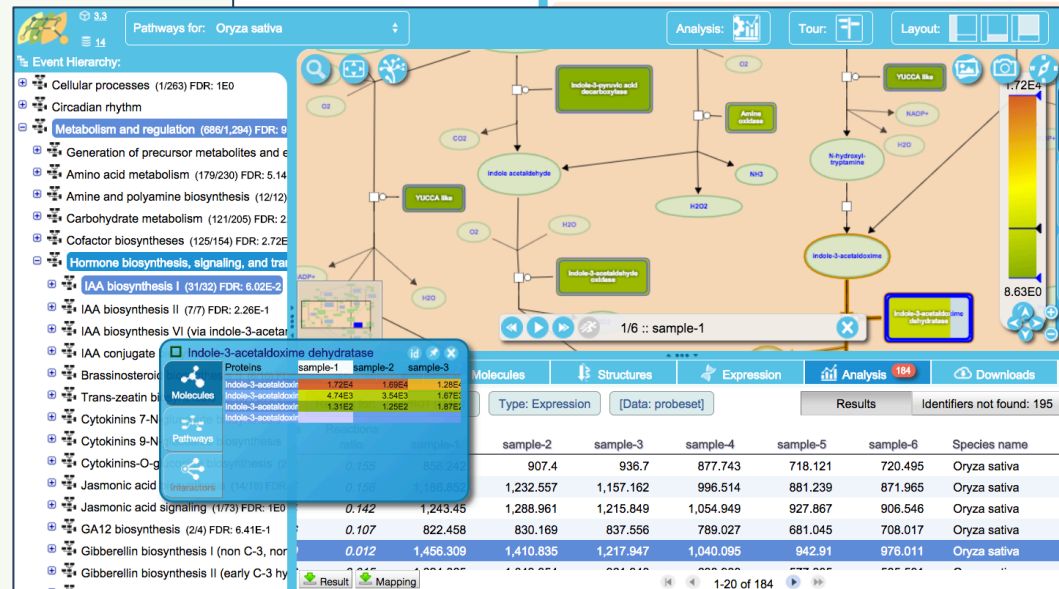
“Fireworks” pathway hierarchy: explore pathways by functional category



Updated browser interface



Visualization overlays
for data analysis
Users can upload
their own -omics data



data from 780 curated experiments. Baseline expression in Gramene

search.gramene.org

ensembl.gramene.org

Oryza sativa Japonica (IRGSP-1.0)
Location: 5:29,200,200-29,205,715 | Gene: GH3.5 | Transcript: GH3.5

Gene-based displays

- Summary
- Splice variants
- Transcript comparison
- Gene alleles
- Sequence
- Secondary Structure
- Gene families
- Literature
- Plant Compara
- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues
- Pan-taxonomic Compara
- Gene Tree
- Orthologues
- Ontologies
- Phenotypes
- Genetic Variation
- Variant table
- Variant image
- Structural variants
- Gene expression**
- Regulation
- External references
- Supporting evidence
- ID History
- Gene history

Gene: GH3.5 OS05G0586200

Description
Jasmonic acid-amido syn

Location
Chromosome 5: 29,200,200-29,205,715

About this gene
This gene has 1 transcript

Transcripts
Show transcript table

Gene expression
Showing 4 experim

9 Davidson
7 Sakai
4 Zhang et al
Anderson et al

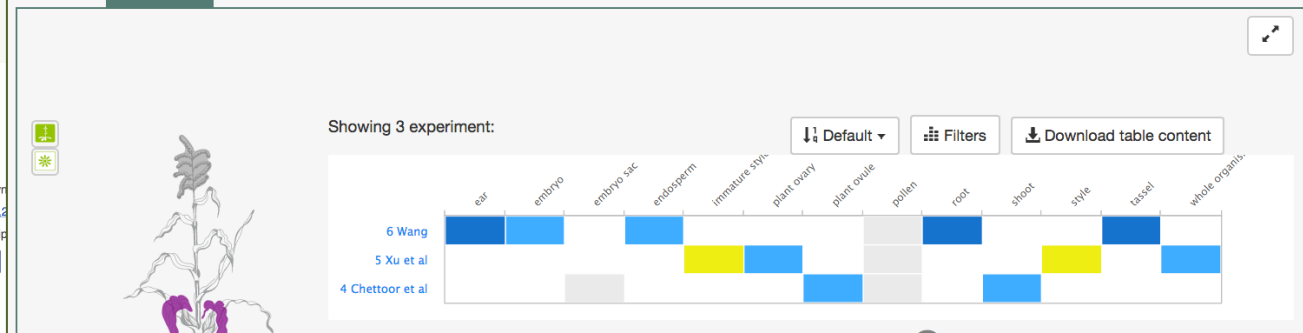
See more expression data at Expression Atlas.
This expression view is provided by Expression Atlas.
Please direct any queries or feedback to arrayexpress-atlas@ebi.ac.uk

Gramene release 53 - May 2017 © EMBL-EBI

Zm00001d033339 GRMZM2G081915 *Zea mays*
Zm00001d033339

Model Species Homolog: **SRK2E** *Arabidopsis thaliana*
Serine/threonine-protein kinase SRK2E

Location **Expression** Homology Pathways X-refs



plantreactome.gramene.org

Pathways for: *Oryza sativa*

Event Hierarchy:

- Cellular processes
- Circadian rhythm
- Metabolism and regulation**
 - Generation of precursor metabolites
 - Amino acid metabolism
 - Amine and polyamine biosynthesis
 - Carbohydrate metabolism
 - Cofactor biosyntheses
- Hormone biosynthesis, signaling, and response
 - IAA biosynthesis I
 - IAA biosynthesis II
 - IAA biosynthesis VI (via indole-3-acetamide)
 - IAA conjugate biosynthesis I
 - Brassinosteroid biosynthesis II
 - Trans-zeatin biosynthesis
 - Cytokinins 7-N-glucoside biosynthesis
 - Cytokinins 9-N-glucoside biosynthesis
 - Cytokinins-O-glucoside biosynthesis
 - Jasmonic acid biosynthesis
 - Jasmonic acid signaling
 - GA12 biosynthesis
 - Gibberellin biosynthesis I (non C-3)
 - Gibberellin biosynthesis II (early C-3)
 - Gibberellin biosynthesis III (early C-3)
 - ABA biosynthesis and mediated signaling
 - Ethylene biosynthesis and signaling
 - Strigolactone biosynthesis
 - Strigolactone signaling

Expression

Showing 1 gene:

Gene	callus	leaf	post-flowering ...	pre-flowering p...	root	seed	shoot
GH3.5	Low	Low	Low	Low	Low	High	Low
EPIOSAG0000004322	Low	Low	Low	Low	Low	High	Low
OS01G0187500	Low	Low	Low	Low	Low	High	Low
OS02G0798600	Low	Low	Low	Low	Low	High	Low
OS01G0767000	Low	Low	Low	Low	Low	High	Low

Experiment: OS01G0767000
Organism part: seed
Expression level: 58 FPKM

Users can also project differential expression from Atlas onto the Gramene browser

Integrated Search



PAD4 3 genes in 1 genomes

TOP

Taxagenomic distribution

GENE

PAD4 1 | PAD4 4 | PAD4 1-1 | All genes that contain the word "PAD4"

All genes that contain a word that starts with "PAD4"

PLANT REACTOME REACTION

Formation of EDS1/PAD4 complex

Zm00001d0047
Lipase-like PAD4

Location Expression

- Metabolism and reg
- Hormone biosynth
- Salicylic acid sig

Transcription c

Formation of E

Location Expression Homology X-refs

Showing 1 experiment:

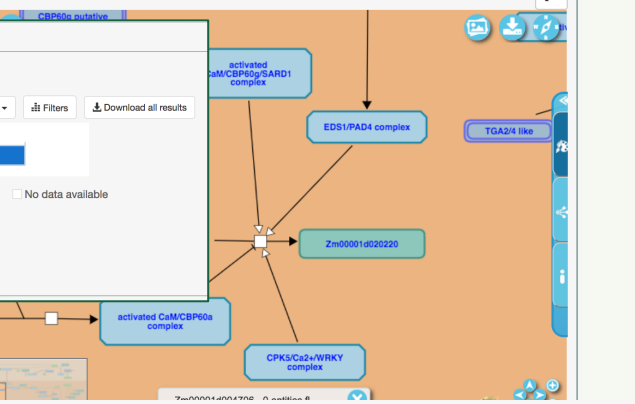
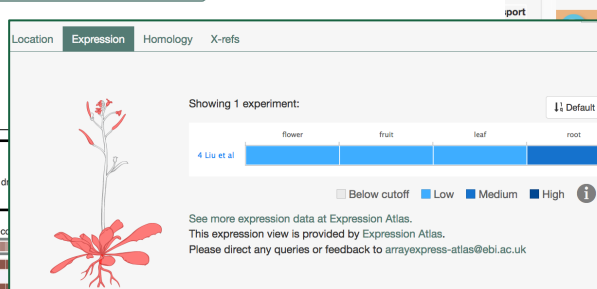
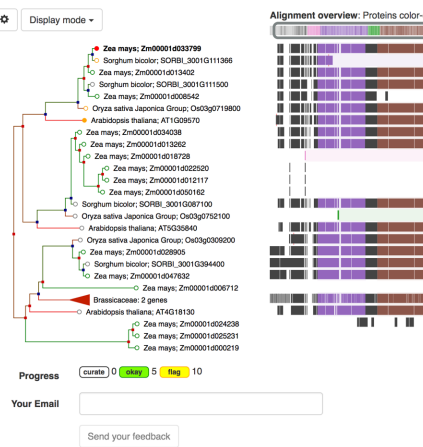
4 Liu et al

flower fruit leaf root

Below cutoff Low Medium High No data available

See more expression data at Expression Atlas.
This expression view is provided by Expression Atlas.
Please direct any queries or feedback to arrayexpress-atlas@ebi.ac.uk

This is the gene tree containing Zm00001d033799. Gene name: phytochromeA1.
Mark genes as okay or flag genes that you think might have issues and choose a reason from the d



Display mode

Alignment overview

Multiple Sequence Alignment

Neighborhood conservation

Zm00001d044330

Zm00001d011252

RBL_3003G202600

403m.g

Triticeae: 3 genes

TRIAE_CS42_3AL_TGACv1_196916_AA0003740

TRILUR3_08672

HORVU3Hr1G051190

BRAD2G42240

Oryzinae: 12 genes

Poaceae: 33 genes, 2 paralogs

Musa acuminata subsp. malaccensis: 4 genes

Dioscorea rotundata: 2 genes

AMTR_s00007p00193000

fabids: 12 genes

Malvaceae: 2 genes

AT1G49950

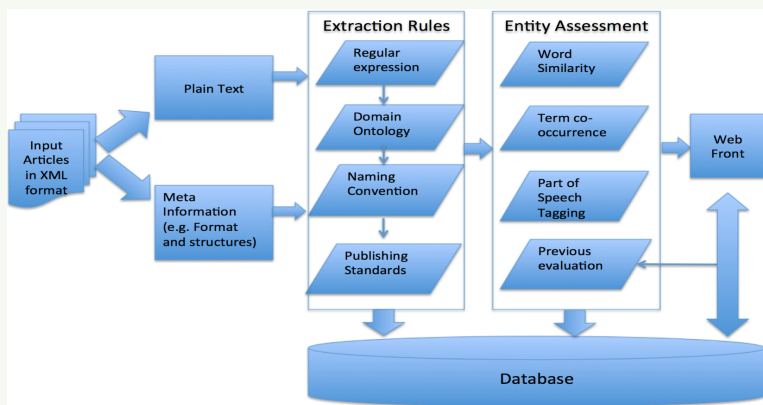


How do we move from manual curation to automated ingestion of published functional annotations



- Challenge: Automate and Scale Manuscript Curation
- Approach: Capture information at the time of submission
- Objective: Prototype bi-directional links between journal articles & databases during production
- Components:

DIVE algorithm (automatic)



ArticleExpress (author curation)

Nonstructural [C1] carbohydrates (NSCs) play a crucial role in xylem formation and represent, with water, the main constraint to plant growth. We assessed the relationships between xylogenesis and NSCs in order to (1) verify the variance explained by NSCs and (2) determine the influence of intrinsic (tissue supplying carbon) and extrinsic (water availability and temperature) factors. During 2 years, wood formation was monitored in saplings of black spruce [C2] (*Picea mariana*) subjected to a dry period of about 1 month in June and exposed to different temperature treatments in a greenhouse. In parallel, NSC concentrations were determined by extracting the sugar compounds from two tissues (cambium and inner xylem), both potentially supplying carbon for wood formation. A mixed-effect model was used to assess and quantify the potential relationships. Total xylem cells, illustrating meristematic activity, were modeled as a function of water, sucrose, and pinitol (conditional r^2 of 0.79). Water availability was ranked as the most important factor explaining total xylem cell production, while the contribution of carbon was lower. Cambium stopped dividing under water deficit, probably to limit the number of cells remaining in differentiation without an adequate amount of water. By contrast, carbon factors were ranked as most important in explaining the variation in living cells (conditional r^2 of 0.49), highlighting the functional needs during xylem development, followed by the tissue supplying the NSCs (cambium) and water availability. This study precisely demonstrates the role of carbon and water in structural growth expressed as meristematic activity and tissue formation.

Genotypic, phenotypic, and functional terms have been associated with your paper by the Gramene database. Please follow this link to confirm their accuracy or to edit: <http://archive.gramene.org/>...

Once you have reviewed the terms applied, please click on the Gramene button below. A paragraph listing these terms with links will be added to the end of your article.

FIGURES
TABLES
COMMENTS
FILE ATTACHMENT
TUTORIALS

Gramene

Accept Edits Save Submit Why can't I submit?

Take Home Messages

- Gramene is an international collaboration with quarterly coordinated regular release cycles
- Built on best-of-class software platforms integrated via a powerful search engine thus enabling easy access to analysis tools
- Resource for hosting and browsing Phylogenetic Trees, Genomes, Expression & Pathways data
- Contributes value-added annotations from Comparative Analysis between & within species
- Interoperability & FAIR principles:
 - Fits larger bioinformatics ecosystem by relying on standard inter-compatible formats & federation among data resources
 - Enables traversing to other curated, clade-specific & species-specific resources



Scholarships for PUI Faculty to Develop Student Projects in Genome Annotation



Apply to join us in November 2018 for a
Maize Genome “Annotation Jamboree”
and the CSHL Biological Data Science Meeting

We invite PUI faculty to apply to a Maize genome annotation jamboree on November 10-11, 2018 at Cold Spring Harbor Laboratory (CSHL). Faculty will learn how they can involve students in authentic research by making meaningful contributions to genome annotation while improving reference Maize genomes at PanMaize. We will introduce tools developed by MaizeCODE and Gramene and provide help and support for faculty bringing these tools into the classroom. Resources are available to support travel, room and board, and registration for interested faculty to the adjacent CSHL Biological Data Science meeting (November 7-10). See our website for more about our current outreach project in annotation and to learn more about MaizeCODE and Gramene.

Visit www.maizecode.org/outreach to register your interest
or send questions: feedback@gramene.org

www.maizecode.org

www.gramene.org

How to reach us

- Gramene Network in



- Social media & news blog



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

[@GrameneDatabase](#)

- feedback@gramene.org & mailing list
- <http://www.gramene.org/contact>