



Rice Data & Resources

Joshua Stein / Doreen Ware
Cold Spring Harbor Laboratory
19 May 2015

Navigation

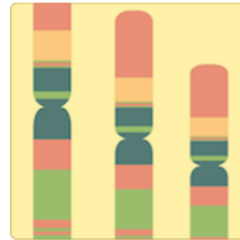
- Current Release (45)
- Search
- Genomes
- Pathways
- BLAST
- Gramene Mart
- News
- Archive (Build 39)
- Download
- Web Services
- Contact
- Mailing list
- Tools

Recent blog posts

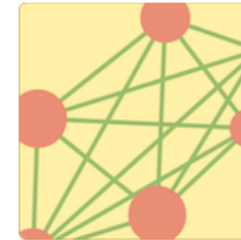
- Webinar: Rice Data & Resources in Gramene on Tues, May 19, 2015 @ 2:00 pm EDT
- NSF Funds Planteome Project
- Gramene database build 45 released
- Webinar: Maize Data & Resources in Gramene on Tues, April 7, 2015 @ 2:00 pm EDT
- Webinar: Gene Family Trees on Tues, Mar 17, 2015 @ 2:00 pm EST

[More](#)

Gramene: A comparative resource for plants



Genomes



Pathways

Gramene is a curated, open-source, integrated data resource for comparative functional genomics in crops and model plant species. Our goal is to facilitate the study of cross-species comparisons using information generated from projects supported by public funds. Gramene currently hosts annotated whole genomes in over two dozen plant species and partial assemblies for almost a dozen wild rice species in the Ensembl browser, genetic and physical maps with genes, ESTs and QTLs locations, genetic diversity data sets, structure-function analysis of proteins, plant pathways databases (BioCyc and Plant Reactome platforms), and descriptions of phenotypic traits and mutations.

Gramene Portals

- **Genome Browser:** Browse gene annotations & diversity data
- **BLAST:** Align DNA & protein sequences
- **Plant Reactome:** Browse metabolic & regulatory pathways
- **Pathways databases:** BioCyc based cellular metabolic networks for 10 plant species
- **Gramene Mart:** Customized data queries
- **Bulk downloads**
- ARCHIVE - Markers, Proteins and Ontology databases, QTLs, Comparative Maps

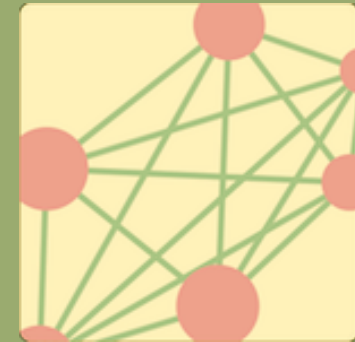
www.gramene.org

GENOMES



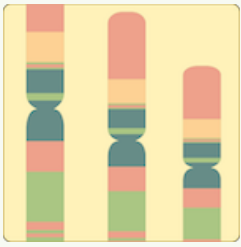
Genome Browser
BLAST
BioMart

PATHWAYS



BioCyc
Plant Reactome

Simple Search
Helpdesk
FTP



Genomes Platform



Accessioned
Genomes +
Community
Annotation +
Variation

Annotation Pipelines

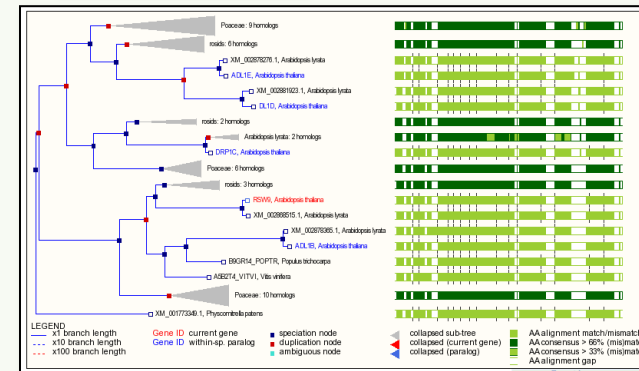
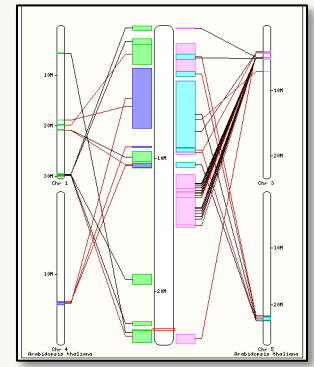
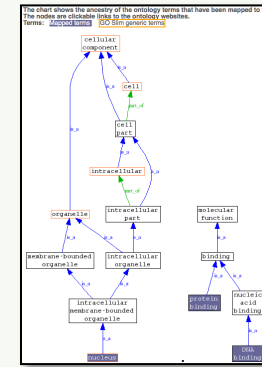
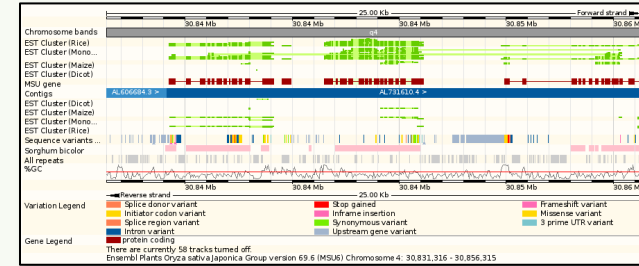
- Repeats/TE's
- Genes
- EST/cDNA
- InterPro domain
- Gene Ontology (GO)
- Variant Effect Prediction

Comparative Analysis

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

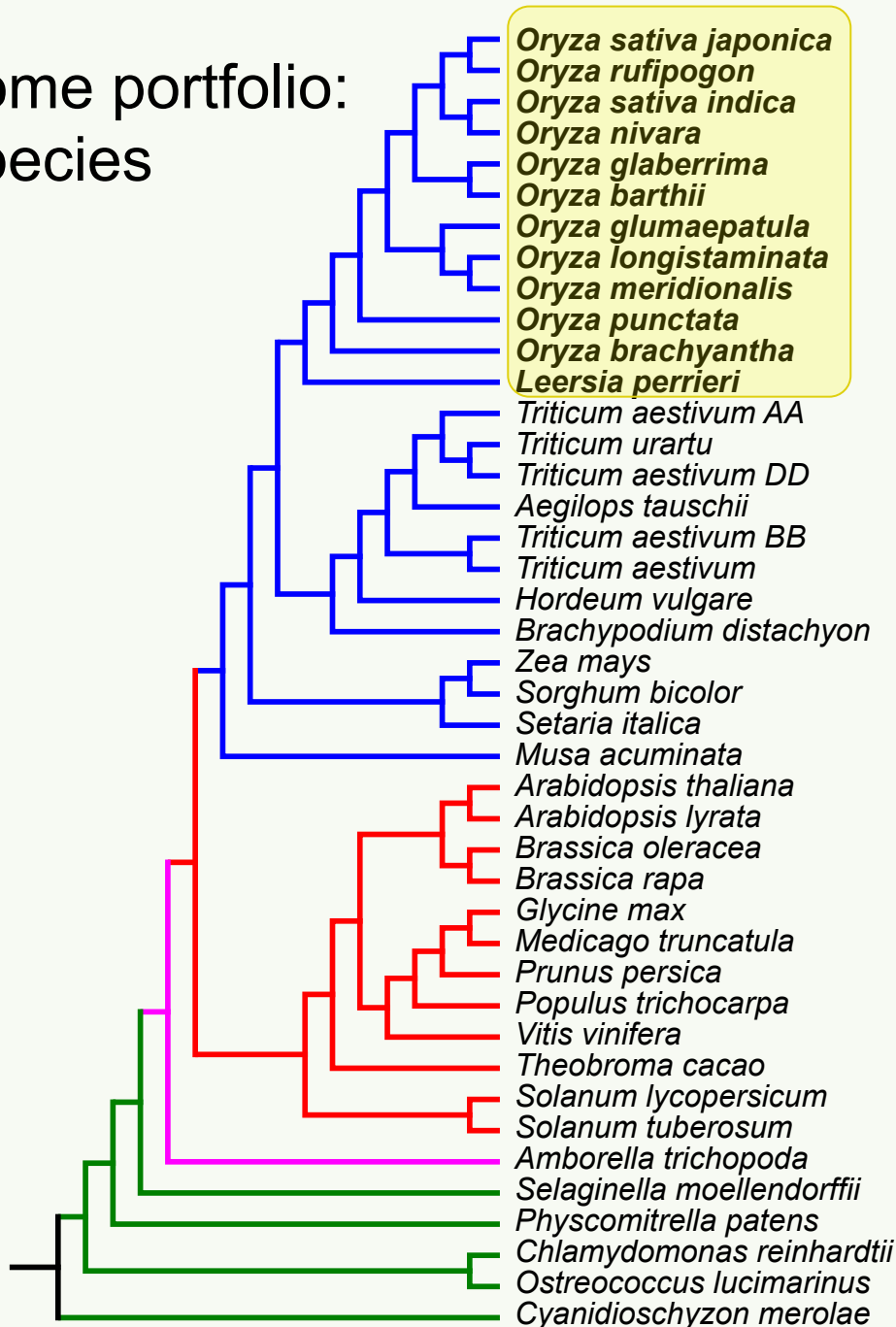
BLAST

Programmatic Access:
Ensembl & RESTFUL API



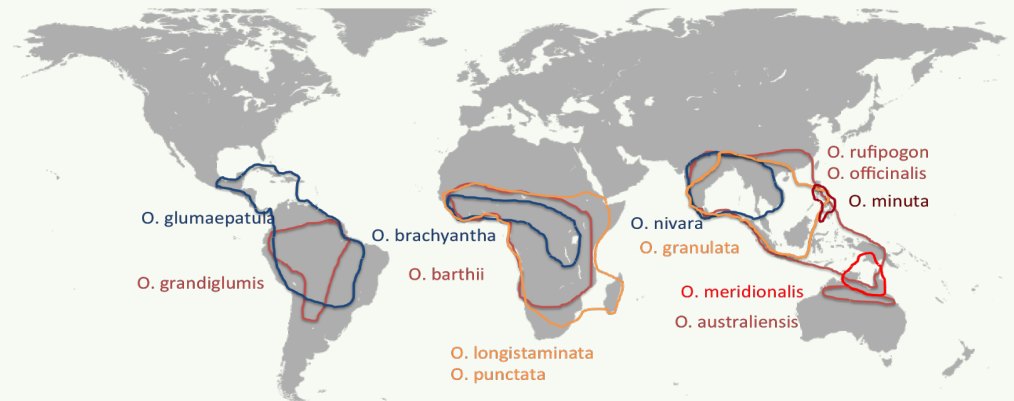
Genome portfolio:
39 species

The Oryzeae



Oryza Genus

- 23 Species
- 2 cultivated species
- 10 Genome types
- Polyploidy
- World-wide distribution
- Untapped diversity



Adapted from: Rice Knowledge Bank (www.knowledgebank.irri.org)

Adapted from: Rice Knowledge Bank (www.knowledgebank.irri.org)

Oryza Genome Evolution Project and I-OMAP

Rod Wing –PI  Arizona Genomics Institute



OGE Co-PI:

Manyuan Long (U. Chicago)

Scott Jackson (U. Georgia)

Carlos Machado (U. Maryland)

Mike Sanderson (U. Arizona)

I-OMAP Consortium → Genome Sequencing

Bin Han (CAS, China)

Yue-ie Hsing (Academia Sinica, Taiwan)

Nori Kurata (NIG, Japan)

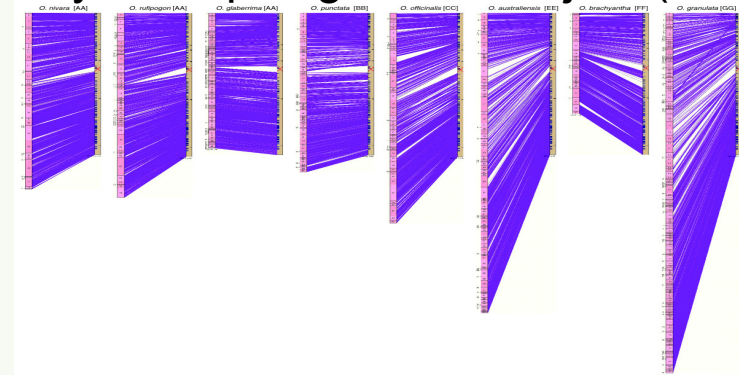
Antonio Costa de Oliveira (UFPEL, Brazil)

Mingsheng Chen (CAS, BGI, China)

Olivier Panaud (CNRS-IRD, France)

Robert Henry (UQ, Australia)

Oryza Map Alignment Project (OMAP)



Genome portfolio: 12 *Oryzaeae*

Shared under
Fort Lauderdale
Agreement

	Sequencing Group	Published?
<i>O. sativa ssp. japonica [AA]</i>	IRGSP	1
<i>O. rufipogon [AA]</i>	B. Han	in prep
<i>O. sativa ssp. indica [AA]</i>	BGI	2
<i>O. nivara [AA]</i>	Y. Hsing /AGI	in prep
<i>O. glaberrima [AA]</i>	AGI	3
<i>O. barthii [AA]</i>	AGI	in prep
<i>O. glumaepatula [AA]</i>	A. Oliveria/AGI	in prep
<i>O. longistaminata</i>	CASK/BGI China	in prep
<i>O. meridionalis [AA]</i>	O. Panaud R. Henry/AGI	in prep
<i>O. punctata [BB]</i>	AGI	in prep
<i>O. brachyantha [FF]</i>	M. Chen	4
<i>Leersia perrieri</i>	AGI	in prep

1. Kawahara et al. (2013). Rice 6:4. 2. Yu et al. (2002). Science 296:79. 3. Wang et al. (2014). Nature Genetics 46:982. 4. Chen et al. (2013). Nature Communications 4:1595

Browser View

EST alignment

RNA-seq

Panicle

Leaf

Roots

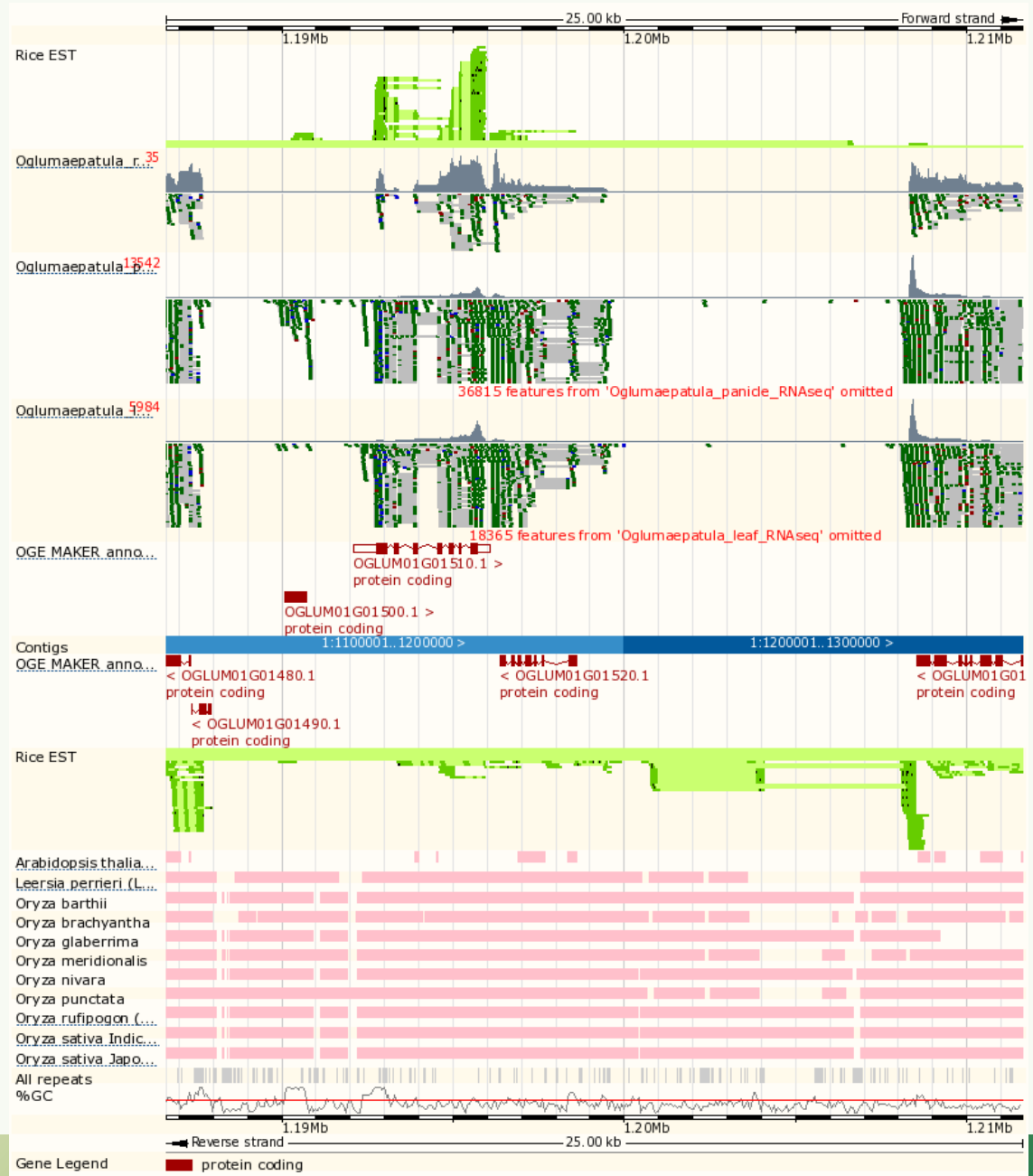


OGE Annotation
(MAKER-P)

ncRNA

Whole genome alignments

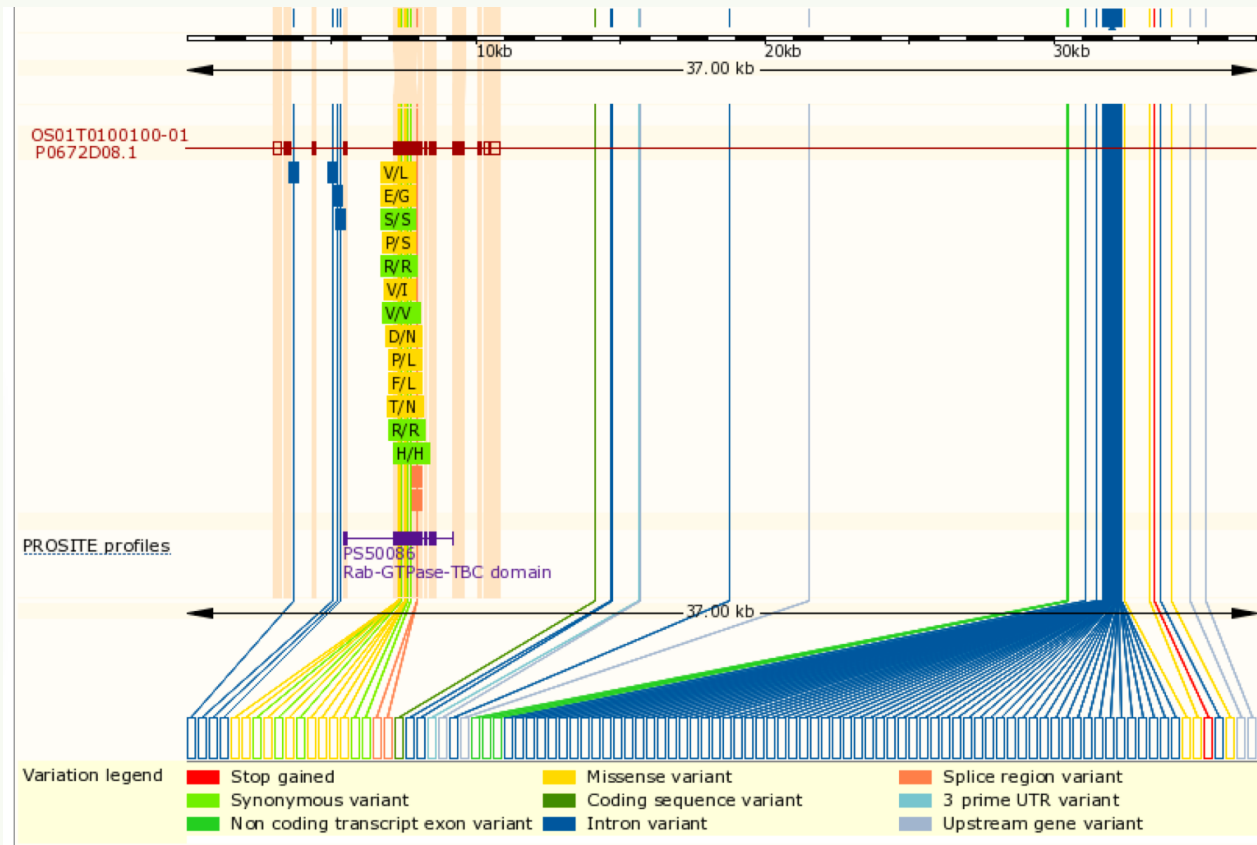
Repeats/TE



Variation Data



Species	Variants	Source	Studies
<i>Oryza sativa ssp japonica</i>	3,332,525	160K SNPs x 20 accessions 1311 SNPs x 395 accessions NCBI dbSNP	McNally et al. (2009). PNAS 106:12273-12278 Zhao et al. 2010. PLoS ONE. 5:e10780
<i>Oryza sativa ssp indica</i>	4,747,883	NCBI dbSNP	
<i>Oryza glaberrima</i>	7,704,409	Resequenced 20 accessions African rice & wild progenitor	Oryza Genome Evolution project

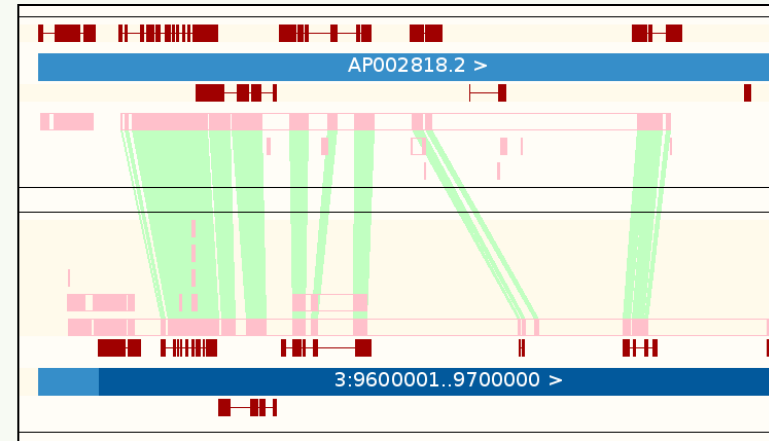


Comparative Phylogenomics

Reconstructing evolutionary histories

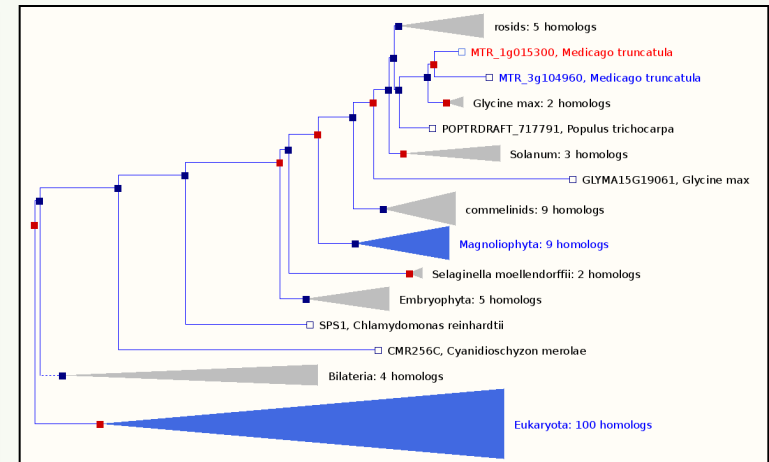
Whole Genome DNA Alignment

- Pairwise alignments
- LASTZ-CHAIN-NET method
- Cross-species browsing



Phylogenetic Gene Trees

- Infers orthologs and paralogs
- Taxonomic dating
- dN, dS, dN/dS
- Cross-species browsing



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

Region Comparison View

Whole genome alignments Ortholog links

- Insertions/deletions
- Inversions
- Duplications
- Assembly artifacts
- Annotation artifacts



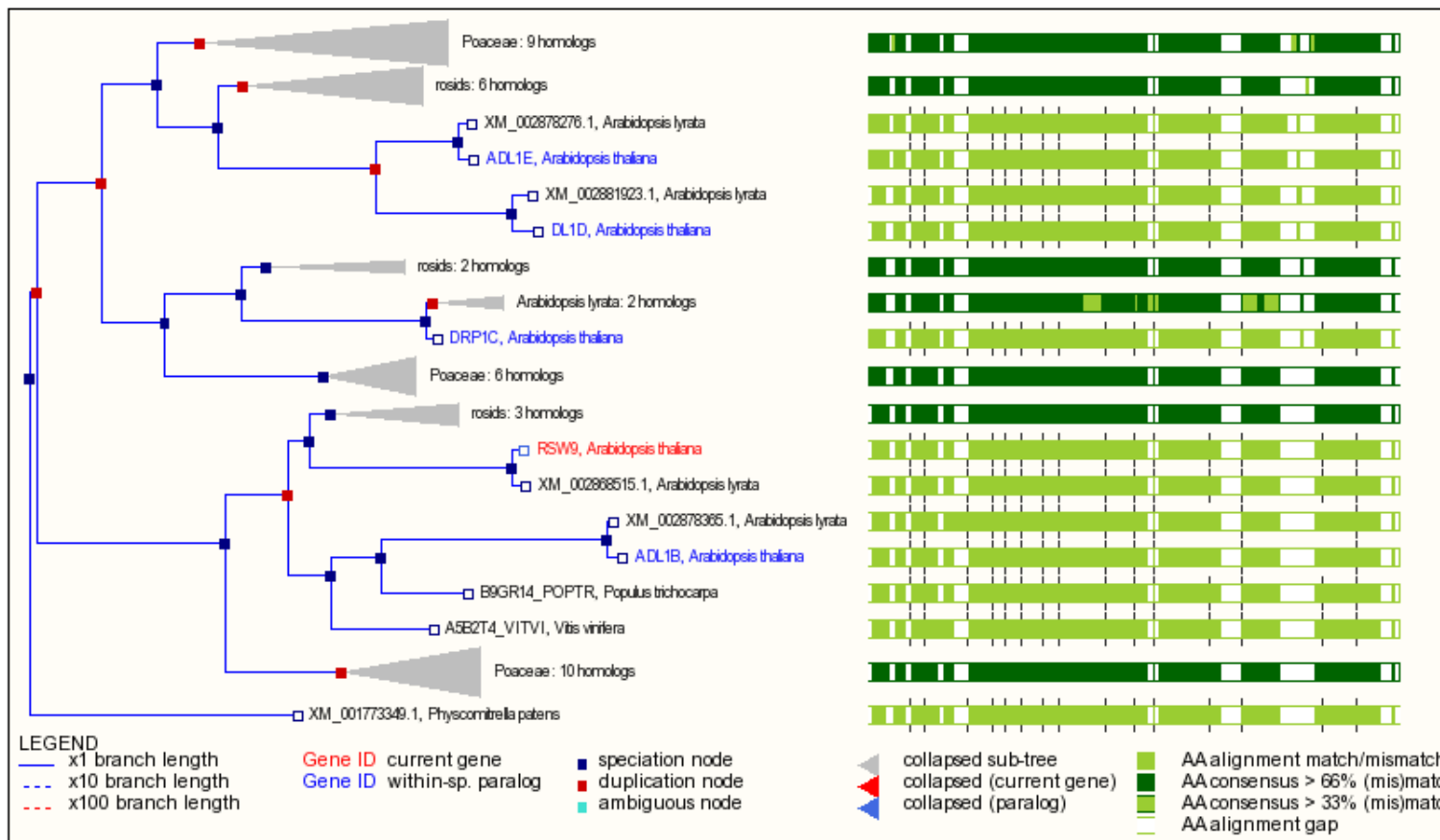
Tree Viewer

Taxon: Poaceae X

Gene_Count 10
 Branch_Length 0.120504
 Bootstrap 100
 Type Duplication (confidence 0.667)
 Image expand this sub-tree
 Image expand all sub-trees
 Image collapse other nodes
 Comparison Jump to Multi-species view
 View Sub-tree Alignment: FASTA
 View Sub-tree Tree: New Hampshire
 View Sub-tree Expand for Jalview

Gene-based displays

- Gene summary
- Splice variants (3)
- Supporting evidence
- Sequence
- External references
- Regulation
- Expression
- Plant Compara
 - Genomic alignments
 - Gene Tree (image)**
 - Gene Tree (text)
 - Gene Tree (alignment)
 - Orthologues (15)
 - Paralogues (4)
 - Protein families
- Pan-taxonomic Compara
 - Gene Tree (image)
 - Gene Tree (text)
 - Gene Tree (alignment)
 - Orthologues (94)
 - Paralogues (10)
 - Protein families (1)
- Genetic Variation
 - Variation Table
 - Variation Image
- External Data
- ID History
 - Gene history



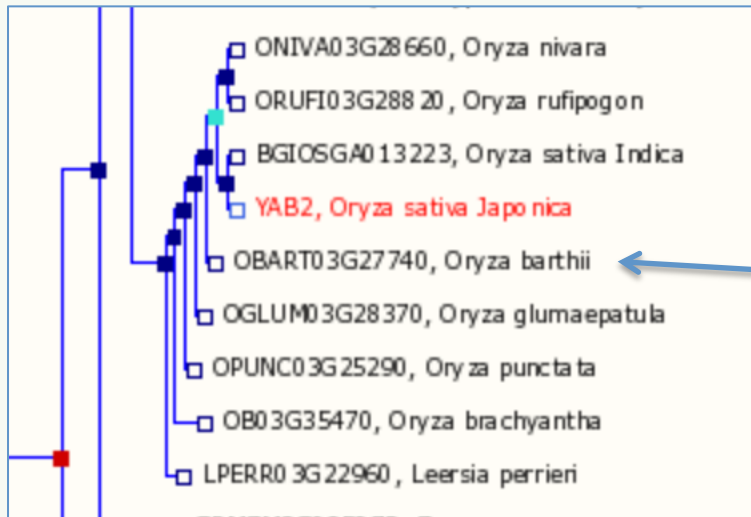
■ Speciation node = ortholog
 ■ Duplication node = paralog


Use Case: Domestication Genes

Sh1 locus (YABBY TF): controls seed dispersal



Asian rice has ~4 Kb insertion limiting expression compared to progenitor
Gene also disrupted in cultivated Sorghum



 African rice (*O. glaberrima*)
Progenitor (*O. barthii*)

Lin, Z. *et al.* Parallel domestication of the *Shattering1* genes in cereals. *Nat. Genet.* **44**, 720–724 (2012).

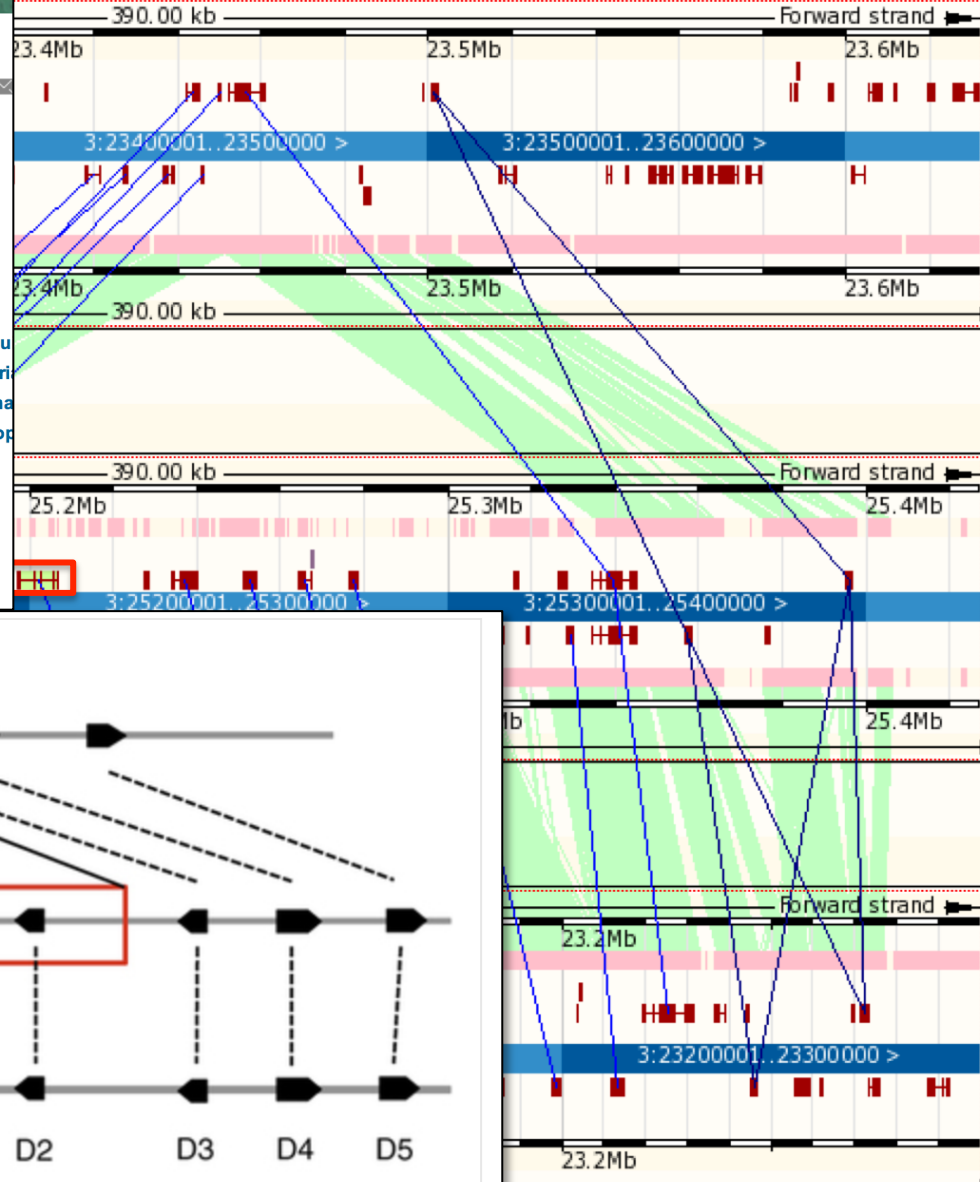
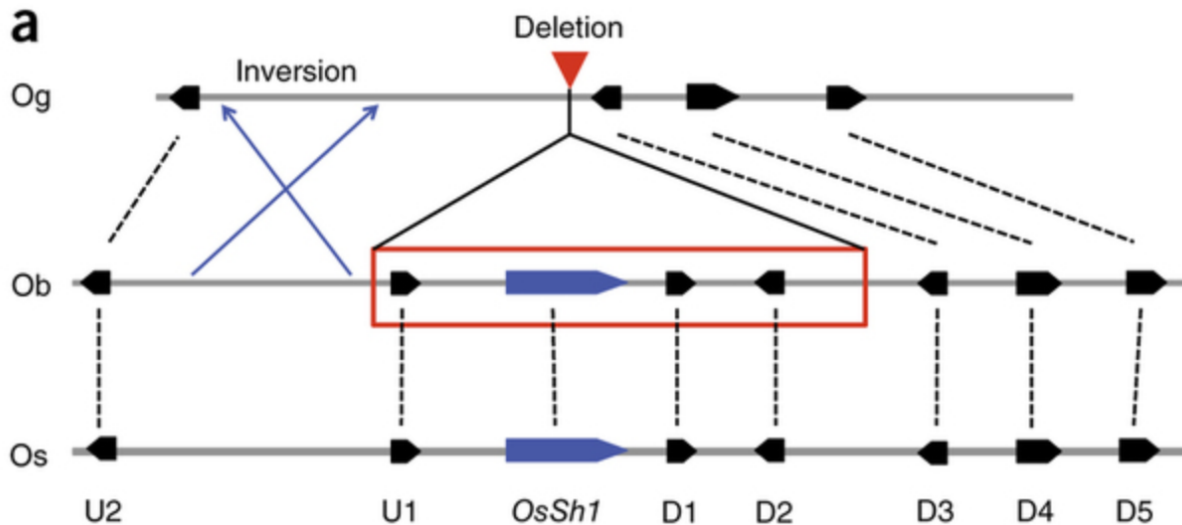
The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication

Muhua Wang, Yeisoo Yu, Georg Haberer, Pradeep Reddy Marri, Chuanzhu Fan, Jose Luis Goicoechea, Andrea Zuccolo, Xiang Song, Dave Kudrna, Jetty S S Ammiraju, Rosa Maria Cossu, Carlos Maldonado, Jinfeng Chen, Seunghee Lee, Nick Sisneros, Kristi de Bayona, Wolfgang Golser, Marina Wissotski, Woojin Kim, Paul Sanchez, Marie-Noelle Ndjiondjop, Kayode Sanni, Manyuan Long, Judith Carney, Olivier Panaud *et al.*

Affiliations | Contributions | Corresponding authors

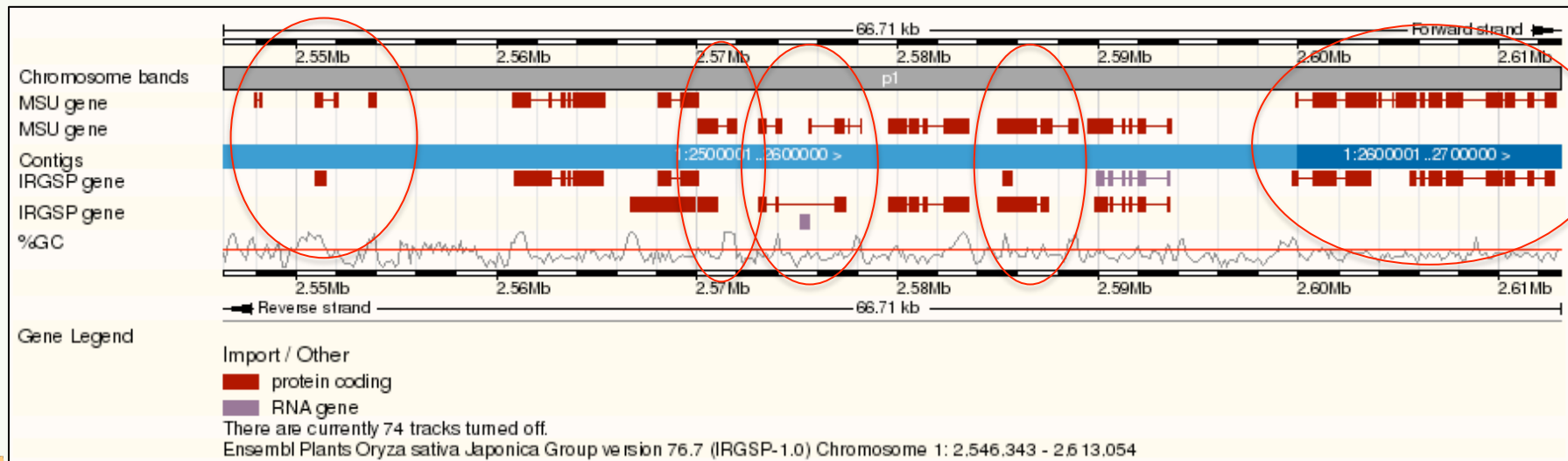
Nature Genetics 46, 982–988 (2014) | doi:10.1038/ng.3044

Received 19 January 2014 | Accepted 30 June 2014 | Published online 27 July 2014



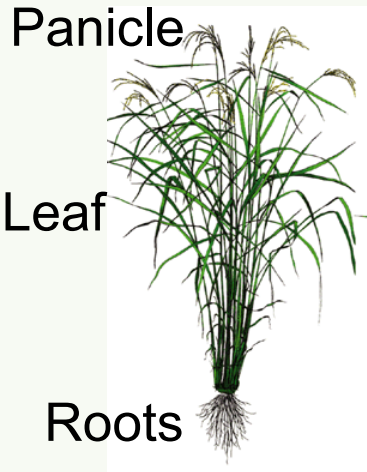
Need Consistent Annotation

- Existing community annotations applied different methods & evidence
- *O. sativa ssp. japonica* (IRGSP1.0) annotated twice:
 - RAP-DB vs. MSU inconsistency
- *O. sativa ssp. indica* annotated by BGI
- *O. glaberrima* annotated by MIPS
- *O. brachyantha* annotated by CAS





MAKER-P Annotation



- RNA-Seq: 3 tissues
 - Trinity assembly
 - Cufflinks assembly

- NCBI collections
 - FL-cDNA (61 K)
 - EST (1.3 M)

- Prior annotations
 - CDS, Protein
 - *O. sativa*, *O. glaberrima*, *Brachypodium distachyon*

- Repeat libraries
 - Customized per species
Dario Copetti (AGI)

- Ab initio* prediction
 - SNAP
 - FGENESH



MAKER-P engine



Protein-coding gene annotations



InterPro/GO

Screen out lingering TE

34-263 M reads/sample
Yeisoo Yu (AGI)



Kapeel Chougule (AGI/CSHL)



- Long term: quarterly releases
- Community annotation
 - IRGSP, BGI, CAS, MIPS
- 39 species → trees
- Plant Reactome
- RiceCyc pathways
- BLAST
- BioMart



- Short term: This experiment
- MAKER-P annotation
- 14 species → 21,450 gene trees
 - 11 Oryzaeae + 3 outgroups
- 5 additional Oryza genomes (partial or scaffold level)
- RNA-seq from OGE project

- In this section
- Species List
 - Scientific Publications
 - Outreach
 - Legal Notices
 - Software Licence
 - Browser Genome Release Agree
 - Privacy Statement
 - About the Ensembl Genomes pr
 - Acknowledgements
 - Job Vacancies
 - Release Cycle
 - Scientific Advisory Board

Documentation | About the Ensembl Genomes project | Species List

Find a Species

Gramene Species

Liliopsida



Brachypodium distachyon
[Brachypodium.org](#) | *Brachypodium distachyon* (L.) Beauv | [15368](#)



Oryza granulata3s
OGE | *Oryza granulata3s* | [110450](#)



Oryza nivara
OGE | *Oryza nivara* | [4536](#)



Leersia perrieri
OGE | *Leersia perrieri* | [77586](#)



Oryza indica
OGE | *Oryza indica* | [39946](#)



Oryza officinalis3s
OGE | *Oryza officinalis3s* | [4535](#)



Oryza barthii
OGE | *Oryza barthii* | [65489](#)



Oryza longistaminata
OGE | *Oryza longistaminata* | [4528](#)



Oryza punctata
OGE | *Oryza punctata* | [4537](#)



Oryza brachyantha
OGE | *Oryza brachyantha* | [4533](#)



Oryza meridionalis
OGE | *Oryza meridionalis* | [40149](#)



Oryza rufipogon
OGE | *Oryza rufipogon* | [4529](#)



Oryza glaberrima
OGE | *Oryza glaberrima* | [4538](#)



Oryza minutabb3s
OGE | *Oryza minutabb3s* | [63629](#)



Oryza sativa Japonica
OGE | *Oryza sativa* | [39947](#)



Oryza glumaepatula
OGE | *Oryza glumaepatula* | [40148](#)



Oryza minutacc3s
OGE | *Oryza minutacc3s* | [63629](#)



Sorghum bicolor
JGI | *Sorghum bicolor* BTX623 | [4558](#)

eudicotyledons

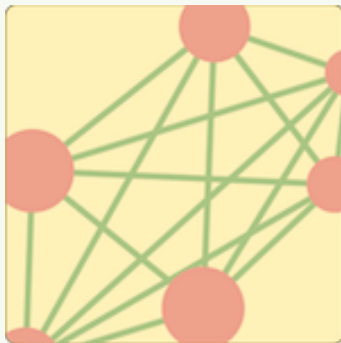


Arabidopsis thaliana
TAIR | *Arabidopsis thaliana* | [3702](#)

Sharon Wei



Pathways Platform



Home About Documentation Tools Contact Us



▼ Search examples...

- LOC_Os01g45760.1
- YUC4
- cytokinins
- glucose

Advanced search

Browse Pathways

Retrieve SBML

Explore...

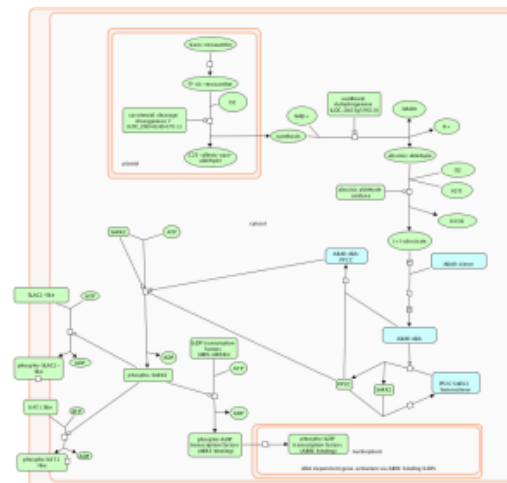
a sampling of metabolic and regulatory pathways in *Oryza sativa*:

- Proline biosynthesis I
- Lysine biosynthesis I
- Jasmonic acid biosynthesis
- Cytosolic glycolysis
- COPII Mediated Vesicle Transport
- Gibberellin biosynthesis II

About Plant Reactome

Plant Reactome (now available in *beta* version) is a freely accessible plant pathway database which hosts plant metabolic and regulatory pathways. Plant Reactome pathways are constructed by manual curation of pathways and reactions reported in the published literature or derived by orthology-based computational projections from curated pathways in the MetaCyc, Plant Metabolic Network, and Human Reactome databases. Pathways, reactions and gene entries in Plant Reactome are cross-referenced to many bioinformatics databases... [More]

Featured pathway: ABA biosynthesis and mediated signaling



Click image to see pathway

News and Notes

• New rice pathways

Plant Reactome has released a new batch of rice pathways associated with cofactor biosynthesis, hormone biosynthesis, and fatty acids & lipids biosynthesis. Examples include: **pyridoxamine anabolism**, **ABA biosynthesis and mediated signaling** and **glycolipid desaturation**.

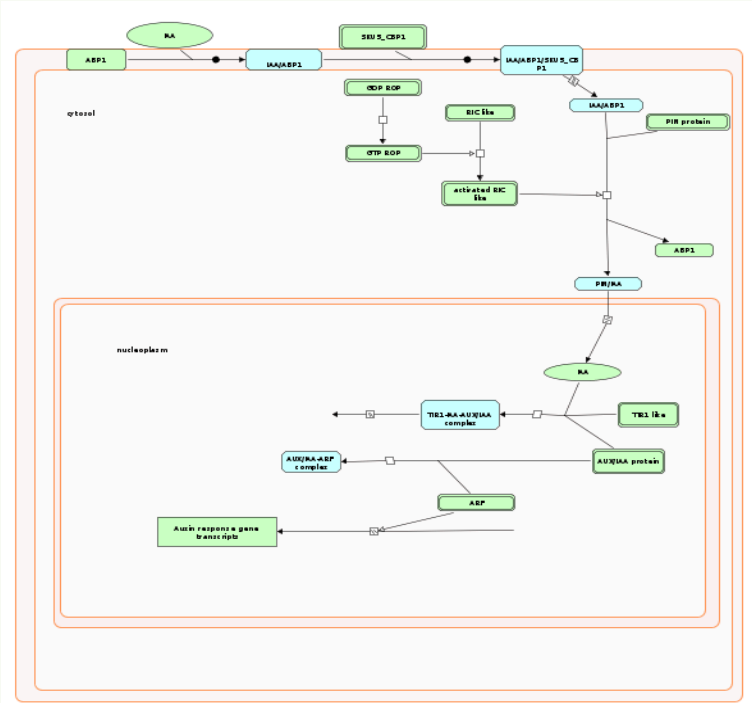
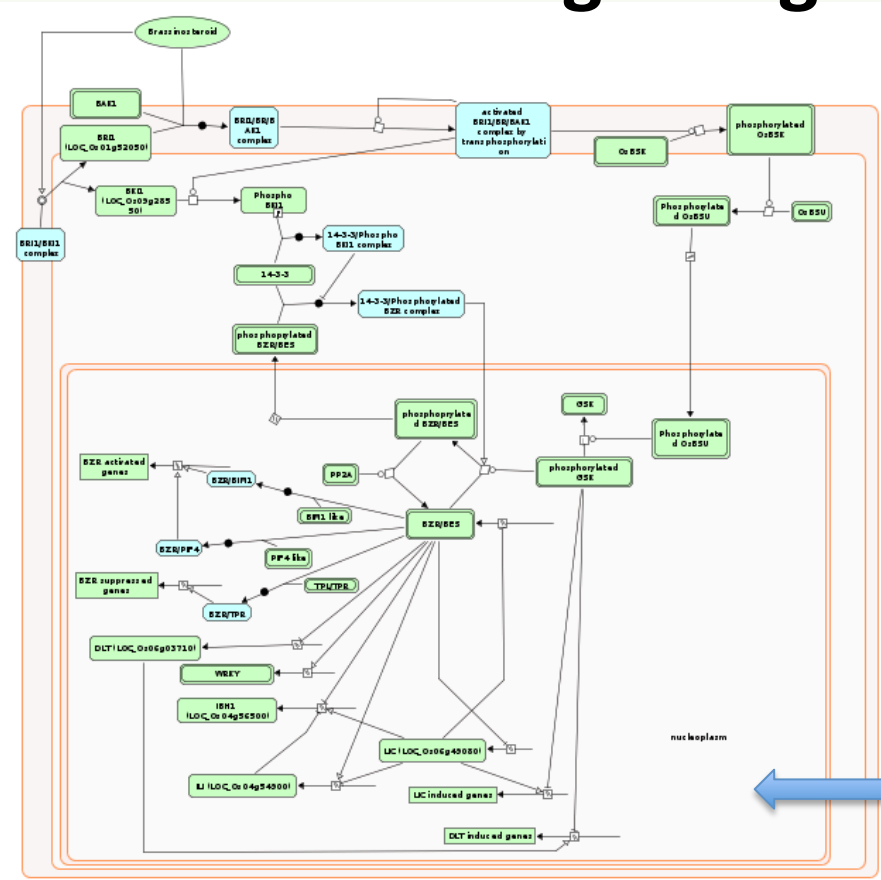
Pankaj Jaiswal,
Co-PI (OSU)
Vindhya
Amarasinghe,
Justin Elser,
Sushma Naithani,
Justin Preece

Ongoing rice curation

Currently >200 curated rice pathways

Notable signaling pathways in recent releases:

Auxin signaling →



← Brassinosteroid signaling

Pathway Example: Cytosolic glycolysis

Pathway details on bottom panel:

- Reference info
- Compartment
- Pathway download
- Participating molecules

Home & Search
Switch Species: *Oryza sativa* ▾
Pathways Help

Rice pathways
Rice cellular processes
Rice metabolism
Generation of precursor metabolites ar
Amino acid biosyntheses
Carbohydrate metabolism
CMP-KDO biosynthesis II (from D-a
Cytosolic glycolysis
Plastid glycolysis
dolichyl-diphosphooligosaccharide
stachyose biosynthesis
cellulose biosynthesis

Download pathway in one of the formats [SBML][SBGN][BioPAX2][BioPAX3][PDF][Word][Protege]

List Download options

Molecule types (select as required):
Proteins (34) Chemical compounds (19)

Fields (select as required):
Type Name Uniprot ID Gene Name ChEBI ID

Formats (select one):
CSV TSV XML Excel

Type	Name	Uniprot ID
Proteins	6-phosphofructokinase (LOC_OS01G09570.1), ATP-dependent [cytosol]	Q5SNH5
Proteins	6-phosphofructokinase (LOC_OS01G53680.1), ATP-dependent [cytosol]	Q5JNC6
Proteins	6-phosphofructokinase (LOC_OS09G30240.1), ATP-dependent [cytosol]	Q69QS5
Proteins	HXX1 [cytosol]	Q0D6S1
Proteins	HXX10 [cytosol]	Q0DIN3
Proteins	HXX2 [cytosol]	B9FH43
Proteins	HXX3 [cytosol]	Q2KNB4
Proteins	HXX5 [cytosol]	Q5W676
Proteins	HXX6 [cytosol]	Q8LQ68

View Download Clear Select all

Reaction-level Data Example

Home & Search

Switch Species: *Oryza sativa*

Pathways

- Fructose 1,6-bisphosphate <=>
- dihydroxyacetone phosphate <=>
- D-glyceraldehyde 3-phosphate <
- D-glyceraldehyde 3-phosphate +
- 1,3-bisphospho-D-glycerate + ADP
- 3-phospho-D-glycerate <=> 2-ph
- 2-phospho-D-glycerate <=> phos
- phosphoenolpyruvate + ADP =>
- Plastid glycolysis
- dolichyl-diphosphooligosaccharide bi
- stachyose biosynthesis
- cellulose biosynthesis
- starch biosynthesis

fructose-bisphosphate aldolase

Name	fructose-bisphosphate aldolase
Cellular compartment	cytosol GO
Organism	<i>Oryza sativa</i>
Cataly	fructose-bisphosphate aldolase (LOC_OS01G67860.1) [cytosol] RUUUU
GO m	fructose-bisphosphate aldolase (LOC_OS05G33380.1) [cytosol] RUUUUUU
May represent entities	fructose-bisphosphate aldolase (LOC_OS08G02700.1) [cytosol] RUU
	fructose-bisphosphate aldolase (LOC_OS10G08022.1) [cytosol] RUUU
Preced	open all close all show/hide hierarchy types
Hierarchical view of the components	<input checked="" type="checkbox"/> fructose-bisphosphate aldolase [cytosol]
Follow	fructose-bisphosphate aldolase activity GO
Biochemical activities	Fructose 1,6-bisphosphate <=> dihydroxyacetone phosphate + glyceraldehyde 3-phosphate [cytosol] [<i>Oryza sativa</i>]
Catalyses events	
Compare human pathway to	

Protein (green box) Small Molecule (green circle)

phosphoglycerate mutase (2,3-bisphosphoglycerate-independent)

ATP

phosphoglycerate kinase

ADP

phospho-D-glycerate

three rice proteins annotated with

Plant Reactome Data

Pathway
projection by
orthology
with Rice

Species	Pathways	Reactions	Gene Products
Oryza sativa japonica (curated reference)	194	583	867
Oryza australiensis	163	423	732
Oryza barthii	166	466	691
Oryza brachyantha	166	458	669
Oryza glaberrima	168	461	691
Oryza glumaepatula	167	468	697
Oryza granulata	164	428	789
Oryza kasalath	115	198	165
Oryza longistaminata	168	452	619
Oryza meridionalis	160	426	610
Oryza minuta	163	430	553
Oryza nivara	165	467	694
Oryza officinalis	168	416	473
Oryza punctata	164	459	687
Oryza rufipogon	166	462	703
Oryza sativa Indica	169	482	757
Leersia perrieri	166	458	658

Brachypodium distachyon	161	456	671
Hordeum vulgare	157	403	570
Aegilops tauschii (Wheat DDD)	157	413	665
Triticum aestivum (Wheat AABBDD)	165	464	1690
Triticum urartu (Wheat AA)	160	425	617
Zea mays	164	454	836
Sorghum bicolor	166	463	717
Setaria italica	165	459	740
Musa acuminata	159	432	894

Arabidopsis thaliana	164	446	731
Brassica rapa	163	443	1122
Glycine max	161	434	1304
Medicago truncatula	152	375	560
Populus trichocarpa	162	436	924
Prunus persica	164	447	631
Solanum lycopersicum	163	443	757
Solanum tuberosum	160	425	739

GA Biosynthesis: Comparative view

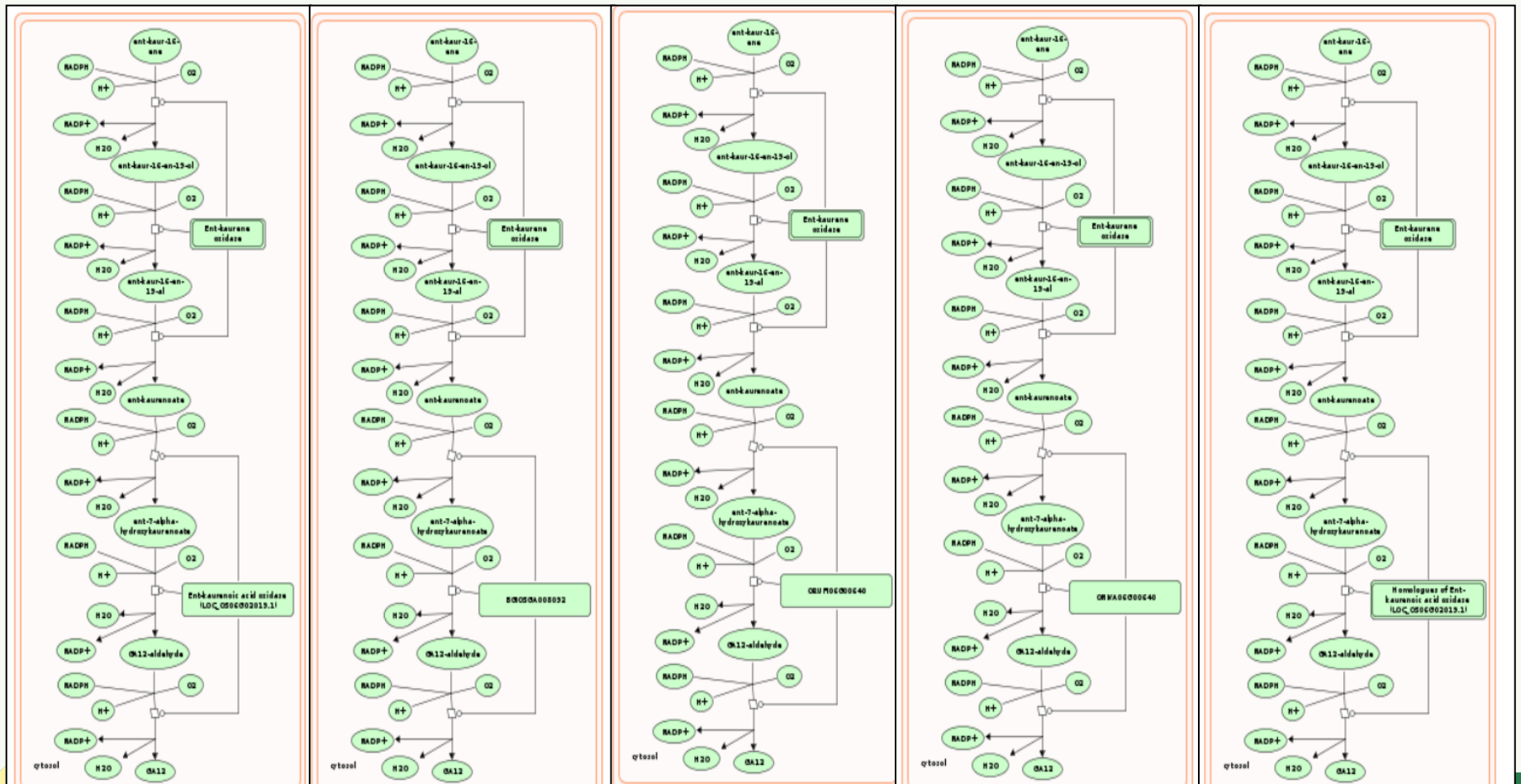
Os japonica

Os indica

O. rufipogon

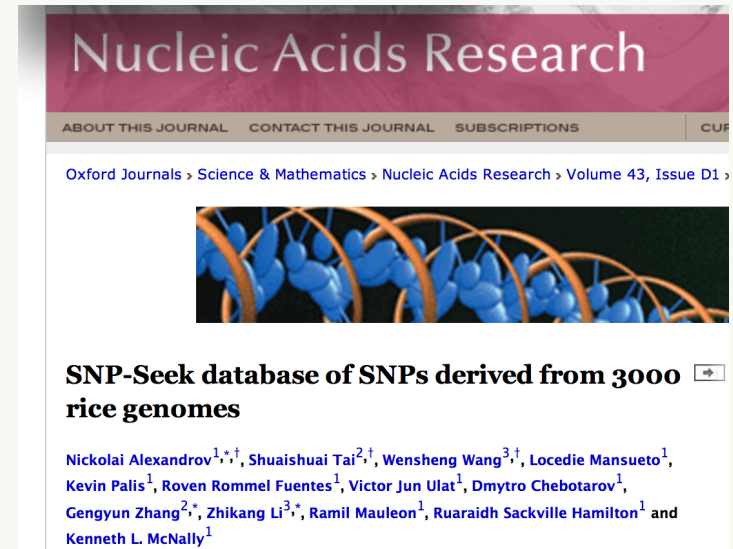
O. nivara

A. thaliana



Future Releases

- Variation Data
 - IRRI 3000 genomes ()
 - 104 rice genomes, Jorge Duitama (CIAT, LSU, USDA and NCGR)
- New Pathway curation



Gramene - Exploring Function through Comparative Genomics and Network Analysis

NSF IOS 1127112

Doreen Ware, PI (USDA ARS, CSHL)

Kapeel Cougule, Yinping Jiao, Sunita Kumari, Vivek Kumari, Marcela K. Monaco, Andrew Olson, Joshua Stein, Jim Thomason, Peter van Buren, Sharon Wei, Bo Wang, Zhenyuan Lu

Pankaj Jaiswal, Co-PI (OSU)

Vindhya Amarasinghe, Justin Elser, Sushma Naithani, Justin Preece

Paul Kersey / Helen Parkinson (EMBL-EBI)

Dan Bolser, Arnaud Kerhornou, Eugene Kulesha, Dan Staines, Brandon Walts / Nuno Fonseca, Maria Keays, Robert Petrysyk, Eleanor Williams

Lincoln Stein (OICR)

Robin Haw; Peter D' Eustachio (NYU); Guanming Wu; David Croft (EBI)

Crispin Taylor (ASPB)

Patty Lockhart



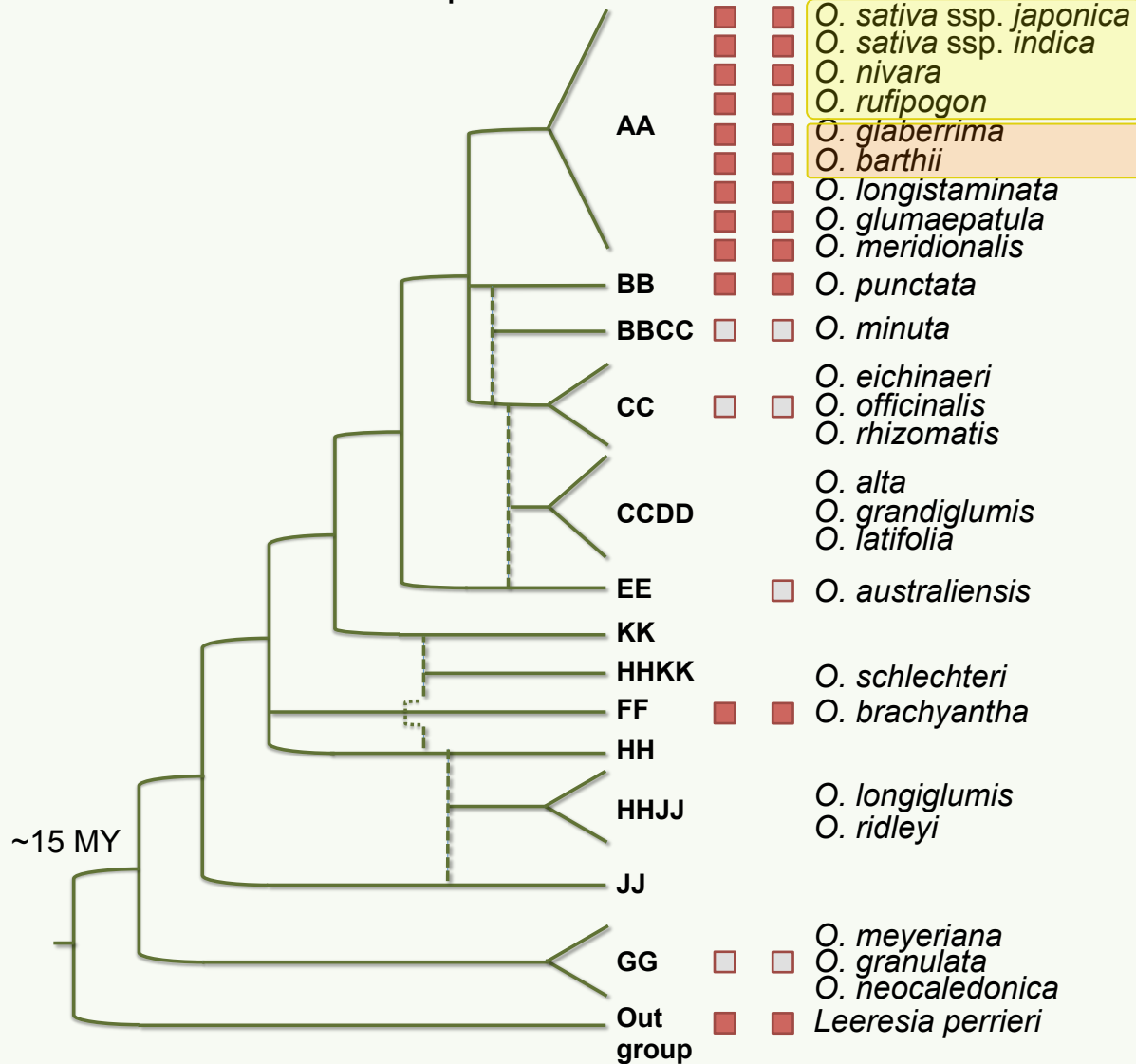
Cold
Spring
Harbor
Laboratory



END

- = complete reference assembly
- = chr-3short/transcriptome

Genomes
Pathways



Asian rice & progenitors

African rice & progenitor