



# Getting involved with Gramene

Marcela Karey Tello-Ruiz, PhD (CSHL)

July 16<sup>th</sup>, 2018



# Transnational collaboration



Doreen Ware

Irene Papatheodorou

Pankaj Jaiswal




Paul Kersey  
EMBL-EBI



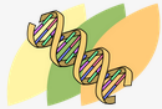
Crispin Taylor



# Outline

- Overview & e-learning
- Integrated Search
- Genomes
  - Variation
  - Expression
  - Pan-Genome Browsers
- Pathways: the Plant Reactome
- Gramene in  **Plantae**

## Gramene Portals



### Genome Browser

Browse genomes with annotations, variation and comparative tools



### Plant Reactome

Browse and analyze metabolic and regulatory pathways



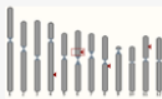
### Tools

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



## Welcome to the Gramene Outreach Site!

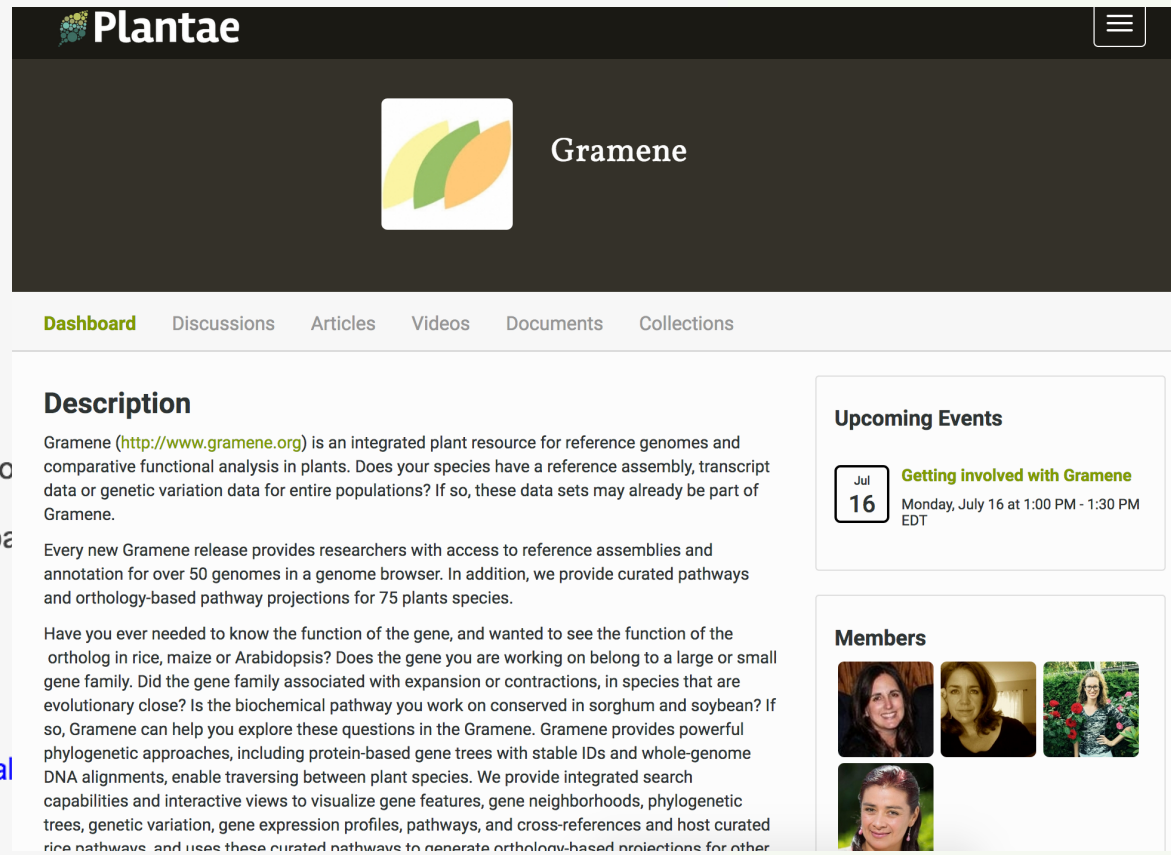
- [Gramene News](#)
- [Gramene Collaborators](#)
- [Contact Form](#)
- [Mailing List](#)
- [Facebook](#)
- [Twitter](#)
- [Gramene's YouTube channel](#) - Look here for past and upcoming events
- [Gramene's FTP](#) - Look here for past and upcoming events
- [Gramene's Outreach Wiki](#)

### Upcoming Conferences

- [Plant Biology 2018](#)
- [American Society for Horticultural Science](#)


### Past Conferences

- [The Biology of Genomes 2018](#)
- [Maize Genetics Conference 2018](#)



The screenshot shows the Gramene website interface. At the top, there is a navigation bar with the Gramene logo and a search bar. Below the navigation bar, there is a main content area with a dark header featuring the word "Plantae" and the Gramene logo. The main content area is divided into several sections: a navigation menu with links for Dashboard, Discussions, Articles, Videos, Documents, and Collections; a Description section; an Upcoming Events section; and a Members section. The Description section contains text about the Gramene resource and its capabilities. The Upcoming Events section features a calendar entry for July 16, 2018, titled "Getting involved with Gramene". The Members section displays three profile pictures of individuals.

## Plantae



## Gramene

[Dashboard](#) [Discussions](#) [Articles](#) [Videos](#) [Documents](#) [Collections](#)

### Description

Gramene (<http://www.gramene.org>) is an integrated plant resource for reference genomes and comparative functional analysis in plants. Does your species have a reference assembly, transcript data or genetic variation data for entire populations? If so, these data sets may already be part of Gramene.

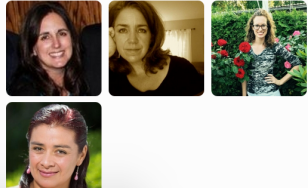
Every new Gramene release provides researchers with access to reference assemblies and annotation for over 50 genomes in a genome browser. In addition, we provide curated pathways and orthology-based pathway projections for 75 plants species.

Have you ever needed to know the function of the gene, and wanted to see the function of the ortholog in rice, maize or Arabidopsis? Does the gene you are working on belong to a large or small gene family. Did the gene family associated with expansion or contractions, in species that are evolutionary close? Is the biochemical pathway you work on conserved in sorghum and soybean? If so, Gramene can help you explore these questions in the Gramene. Gramene provides powerful phylogenetic approaches, including protein-based gene trees with stable IDs and whole-genome DNA alignments, enable traversing between plant species. We provide integrated search capabilities and interactive views to visualize gene features, gene neighborhoods, phylogenetic trees, genetic variation, gene expression profiles, pathways, and cross-references and host curated rice pathways. and uses these curated pathways to generate orthology-based projections for other

### Upcoming Events

Jul 16 **Getting involved with Gramene**  
Monday, July 16 at 1:00 PM - 1:30 PM EDT

### Members

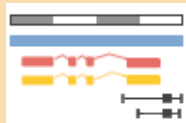


53 reference  
Genomes





2.1M gene  
annotations



67K gene  
trees



318 whole-  
genome  
alignments

ATCGAGCT  
ATCCAGCT  
ATCGAGAT

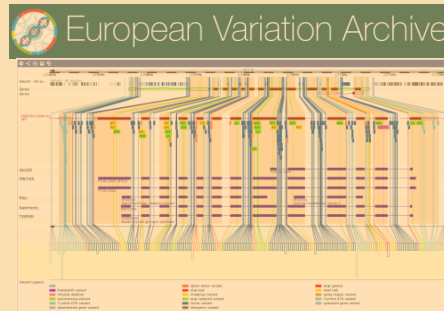
## Expression



734 experiments in  
18 species

**Integrated search  
and visualization**

## Variation



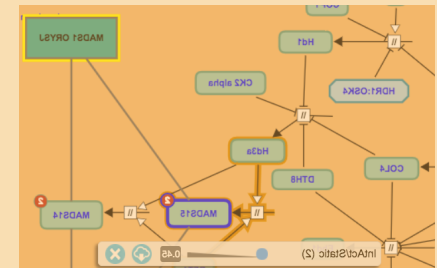
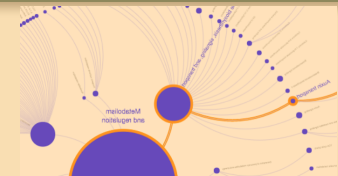
~231M SNPs in  
9 genomes

## Pathways

264 curated rice  
projected to  
74 species



*Plant* REACTOME



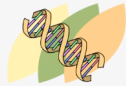
# Search



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

2075935 genes in 53 genomes

## Gramene Portals



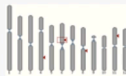
### Genome Browser

Browse genomes with annotations, variation and comparative tools



### Tools

Tools for processing both our data and yours



### BLAST

Query our genomes with a DNA or protein sequence



### Track Hub Registry

A global centralised collection of publicly accessible track hubs



### Bulk Downloads

FTP download of our data



### Plant Reactome

Browse and analyze metabolic and regulatory pathways



### Plant Expression ATLAS

Browse plant expression results at EBI ATLAS



### Gramene Mart

An advanced query interface powered by BioMart



### Outreach and Training

Educational resources and webinars



### Archive

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

## Latest News

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

The Gramene and MaizeCode Projects launched the first Maize Annotation Jamboree  
Tue, 12 Dec 2017

Gramene in the Plant Genome Evolution 2017 Conference  
Mon, 02 Oct 2017

Early career Tansley Medal award for plant scientists!  
Wed, 27 Sep 2017

The Gramene Database build 55 is out with a brand new Jute genome!  
Wed, 20 Sep 2017

Gene annotation and Reactome pathway

# Genomes

## Gramene Portals



### Genome Browser

Browse genomes with annotations, variation and comparative tools



### Plant Reactome

Browse and analyze metabolic and regulatory pathways



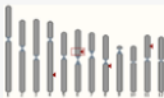
### Tools

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



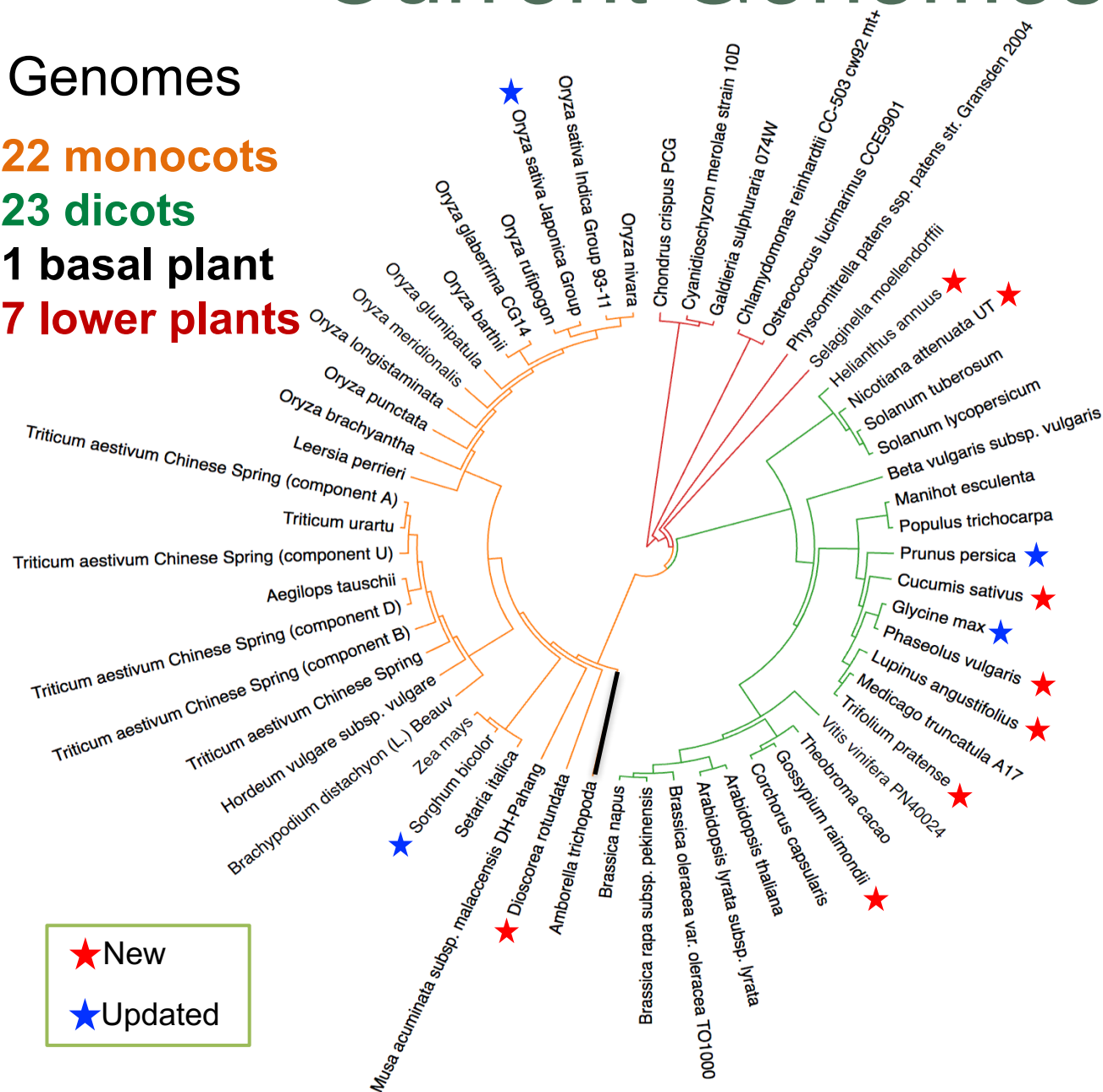
### Archive

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

# Current Genomes

## 53 Genomes

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



- NEW**
- Cucumber
  - Cotton
  - Sunflower
  - Blue lupin
  - Cassava
  - Coyote tobacco
  - Common bean
  - White Guinea yam
- UPDATED**
- Sorghum V3
  - Soybean V3
  - Peach V2
  - Japonica rice (models)



Search:  for

e.g. **Carboxy\*** or **chx28**

**Favourite genomes**



**Arabidopsis thaliana**  
TAIR10



**Oryza sativa Japonica**  
IRGSP-1.0



**Triticum aestivum**  
TGACv1



**Hordeum vulgare**  
Hv\_IBSC\_PGSB\_V2



**Zea mays**  
AGPv4



**Physcomitrella patens**  
ASM242v1

[Edit favourites](#)

**All genomes**

-- Select a species -- 



[View full list of all Gramene species](#)

**What's New in Release 56**

- New genomes
  - [Cucumis sativus](#) (cucumber, GCA\_000004075.2)
  - [Dioscorea rotundata](#) (white Guinea yam, GCA\_002240015.2)
  - [Gossypium raimondii](#) (cotton, GCA\_000327365.1)
  - [Helianthus annuus](#) (common sunflower, GCA\_002127321.1)
  - [Lupinus angustifolius](#) (blue lupin, GCA\_001865872.1)
  - [Manihot esculenta](#) (cassava, GCA\_001659215.1)
  - [Nicotiana attenuata](#) (coyote tobacco, GCA\_001879085.1)
  - [Phaseolus vulgaris](#) (common bean, GCA\_001938451.1)
- Updated genomes
  - [Sorghum bicolor](#) (sorghum V3)
  - [Glycine max](#) (soybean V3)
  - [Prunus persica](#) (peach V2)
  - Updated gene models for [Oryza sativa Japonica](#) (rice)

**Wheat polyploid views, rice gene synonyms, SIFT scores**

**Did you know...?**

You can search the [Track Hub Registry](#) to find more than 1,400 public RNA-Seq studies aligned to plants ([read more](#)) on the Ensembl Genomes site.



**Eight new genomes**

This release of [Gramene](#) brings eight new genomes, including [cassava](#), [cotton](#), [cucumber](#), [common bean](#), [sunflower](#) and [yam](#). In addition, we have updated to latest and greatly improved (v3.0) [sorghum](#) genome assembly from JGI as well as updated assemblies for both [soybean](#) and [peach](#). Users can now [view](#), [query](#), [download](#) and [analyse](#) our comparative analysis of nearly 2 million genes across 53 plant species. The current Gramene release includes 26 new synteny maps and whole-genome alignments (see a summary of all [genomic alignments and syntenies](#)). LastZ comparisons were ran between the new or updated genomes and each of *Oryza sativa Japonica*, *Arabidopsis thaliana*, and *Vitis vinifera*. A stats page for each pairwise LastZ alignment is available, see for example the one for [P. vulgaris vs V. vinifera](#) and the corresponding [synteny map for chromosome 9 of the common bean vs the common grape vine](#). Also available from Gramene's FTP are [Synteny tables](#).

**Updated genomes including Sorghum bicolor v3.1.1**

Three genome assemblies were updated, including *Sorghum bicolor* (sorghum v3.1.1), *Glycine max* (soybean V3), and *Prunus persica* (peach V2). For more details on the *Sorghum bicolor* v3.1.1 assembly, see the [Phytozome site](#). The release also includes updated gene models for *Oryza sativa Japonica*.

**Barley genome**

A ten-nation consortium reported the first high-quality reference genome sequence of *Hordeum vulgare* (barley). The barley genome was sequenced and assembled using an array of state-of-the-art methods, taking ten years. For the first time, scientists can now locate all genes precisely in the genome and analyze complex gene families that play a key role in malting and resilience.

**Arabidopsis variation data from the 1001 Genomes Project**

The representation of genetic variation of *Arabidopsis thaliana* incorporates the full data set from the [1001 Genomes Project](#), covering more than 10 million variant loci across 1,135 samples ([Cell 2016](#)).

**Bread wheat variation data**

A total of 80,829 variation markers from the iSelect 90k array and 13.8 million Inter-Homoeologous Variants (IHVs) have been added to the new genome assembly of *Triticum aestivum* cv. Chinese Spring produced by the [Earlham Institute](#).

This data is viewable alongside the existing Axiom 35k and 820k SNP marker sets provided by [CerealsDB](#) and located on the new assembly and the publicly available EMS mutant lines from tetraploid (cv. Kronos) and hexaploid (cv. CadENZA) TILLING populations ([read more](#)).

**Ensembl Plants Archive Site**

Alongside release 32, Ensembl Plants launched a new [archive site](#), where selected previous releases of Ensembl Plants are being kept publicly available. The first release available on the archive site is release 31, and includes the previous assemblies for wheat and maize.



This work is a joint project between [EnsemblGenomes](#) at the [European Bioinformatics Institute](#) and the group of [Doreen Ware](#) at the [Cold Spring Harbor Laboratory](#), who have developed the [Gramene](#) database, a resource for plant comparative genomics based on Ensembl technology. A common set of databases are available through EnsemblGenomes and Gramene, and the two groups are collaborating on the integration of content, quality control and the development of new features.





53 Species





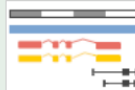
Show All entries		Show/hide columns				Filter				
Name	Classification	Taxon ID	Assembly	Accession	Variation database	Regulation database	Whole genome alignments	Other alignments	In peptide compara	In pan-taxonomic compara
<b>Aegilops tauschii</b>	Liliopsida	<a href="#">37682</a>	ASM34733v1	<a href="#">GCA_000347335.1</a>	-	-	✓	✓	✓	-
	Amborellales	<a href="#">13333</a>	AMTR1.0	<a href="#">GCA_000471905.1</a>	-	-	✓	✓	✓	✓
<b>Arabidopsis lyrata</b>	eudicotyledons	<a href="#">81972</a>	v.1.0	<a href="#">GCA_000004255.1</a>	-	-	✓	✓	✓	-
<b>Arabidopsis thaliana</b>	eudicotyledons	<a href="#">3702</a>	TAIR10	<a href="#">GCA_000001735.1</a>	✓	✓	✓	✓	✓	✓
<b>Beta vulgaris subsp. vulgaris</b>	eudicotyledons	<a href="#">3555</a>	RefBeet-1.2.2	<a href="#">GCA_000511025.2</a>	-	-	-	✓	✓	-
<b>Brachypodium distachyon</b>	Liliopsida	<a href="#">15368</a>	v1.0	<a href="#">GCA_000005505.1</a>	✓	-	✓	✓	✓	-
<b>Brassica napus</b>	eudicotyledons	<a href="#">3708</a>	AST_PRJEB5043_v1	<a href="#">GCA_000751015.1</a>	-	-	-	✓	✓	-
<b>Brassica oleracea</b>	eudicotyledons	<a href="#">109376</a>	v2.1	<a href="#">GCA_000695525.1</a>	-	-	✓	✓	✓	-
<b>Brassica rapa</b>	eudicotyledons	<a href="#">51351</a>	IVFCAASv1	<a href="#">GCA_000309985.1</a>	-	-	✓	✓	✓	-
<b>Setaria italica</b>	Liliopsida	<a href="#">4555</a>	JGlv2.0	<a href="#">GCA_000263155.1</a>	-	-	✓	✓	✓	-
<b>Solanum lycopersicum</b>	eudicotyledons	<a href="#">4081</a>	SL2.50	<a href="#">GCA_000188115.2</a>	✓	-	✓	✓	✓	✓
<b>Solanum tuberosum</b>	eudicotyledons	<a href="#">4113</a>	SoITub_3.0	<a href="#">GCA_000226075.1</a>	-	-	✓	✓	✓	-
<b>Sorghum bicolor</b>	Liliopsida	<a href="#">4558</a>	Sorghum_bicolor_NCBIv3	<a href="#">GCA_000003195.3</a>	✓	-	✓	✓	✓	-
<b>Theobroma cacao</b>	eudicotyledons	<a href="#">3641</a>	Theobroma_cacao_20110822	<a href="#">GCA_000403535.1</a>	-	-	✓	✓	✓	-
<b>Trifolium pratense</b>	eudicotyledons	<a href="#">57577</a>	Trpr	<a href="#">GCA_900079335.1</a>	-	-	-	✓	✓	-
<b>Triticum aestivum</b>	Liliopsida	<a href="#">4565</a>	TGACv1	<a href="#">GCA_900067645.1</a>	✓	-	✓	✓	✓	-
<b>Triticum urartu</b>	Liliopsida	<a href="#">4572</a>	ASM34745v1	<a href="#">GCA_000347455.1</a>	-	-	✓	✓	✓	-
<b>Vitis vinifera</b>	eudicotyledons	<a href="#">29760</a>	IGGP_12x	<a href="#">GCA_000003745.2</a>	✓	✓	✓	✓	✓	✓
<b>Zea mays</b>	Liliopsida	<a href="#">4577</a>	AGF							

<http://ensembl.gramene.org/species.html>




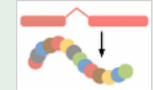


**Search**e.g. [Zm00001d048577](#) or [1:109000-145001](#) or [Carboxypeptidase](#)**About *Zea mays***




*Zea mays* (maize) has the highest world-wide production of all grain crops, yielding 875 million tonnes in 2012 (<http://faostat.fao.org/>). Although a food staple in many regions of the world, most is used for animal feed and ethanol fuel. Maize was domesticated from wild teosinte in Central America and its cultivation spread throughout the Americas by Pre-Columbian civilizations. In addition to its economic value, maize is an important model organism for studies in plant genetics, physiology, and development. It has a large genome of about 2.4 gigabases with a haploid chromosome number of 10 (Schnable *et al.*, 2009; Zhang *et al.*, 2009). Maize is distinguished from other grasses in that its genome arose from an ancient tetraploidy event unique to its lineage.

Taxonomy ID [4577](#)Data source [wareLab](#) [More information and statistics](#)**Genome assembly: [AGPv4](#)** [More information and statistics](#) [Download DNA sequence \(FASTA\)](#) [Convert your data to AGPv4 coordinates](#) [Display your data in Gramene](#)[View karyotype](#)[Example region](#)**Gene annotation**


**What can I find?** Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs.

 [More about this genebuild](#) [Download genes, cDNAs, ncRNA, proteins - FASTA - GFF3](#) [Update your old Ensembl IDs](#) [Gene function summary](#) [Transposon annotation download](#)[Example gene](#)[Example transcript](#)**Comparative genomics**

**What can I find?** Homologues, gene trees, and whole genome alignments across multiple species.

 [More about comparative analyses](#) [Phylogenetic overview of gene families](#) [Download alignments \(EMF\)](#) [Genomic alignments \[7\] \[\\[Show\\]\]\(#\)](#)[Example gene tree](#)**Variation**


**What can I find?** Short sequence variants.

 [More about variation in Zea mays](#) [More about variation in Gramene](#) [Download all variants - VGF - VCF - VEP](#)

Variant Effect Predictor

[Example variant](#)**Regulation**

**What can I find?** Microarray annotations.

 [More about regulation in Zea mays](#) [More about the Gramene microarray annotation strategy](#)**Links**

- [Maize B73 RefGen v4 related data at NCBI \(raw data included\)](#)
- [MaizeSequence.org FTP site](#)
- [Legacy Gramene Maize RefGen V4 genome browser](#) (includes RefGen\_V3)
- [MaizeGDB - Look here for RefGen V3 on the generic genome browser](#)
- Description of the B73 *Zea mays* RefGen\_V4 assembly in [MaizeGDB](#)

# Navigating Browsers

**Zea mays (AGPv4)** **Zea mays (AGPv4)**

Location: 6:85,061,295-85,064,797 Gene: yellow endosperm1 Trans: Zm0001d036345\_T001 Variant: PZE0680878460 Jobs ▾

Location: 6:85,061,295-85,064,797

**Variant displays**

- Explore this variant
- Genomic context
  - Genes and regulation
  - Flanking sequence**
  - Genotype frequency
  - Phenotype data
  - Sample genotypes
  - Linkage disequilibrium
  - Phylogenetic context
  - Citations

**PZE0680878460 SNP**

**Most severe consequence** | [See all predicted consequences](#)  
**missense variant**

**G/C** | Highest population MAF: 0.09

**Location** [Chromosome 6:85061564](#) (forward strand) | VCF: 6 85061564 PZE0680878460 G C

**HGVS names** This variant has 5 HGVS names - [Hide](#)

- 6:g.85061564G>C
- Zm0001d036345\_T001:c.155G>C
- Zm0001d036345\_P001:p.Trp52Ser
- Zm0001d036345\_T002:c.155G>C
- Zm0001d036345\_P002:p.Trp52Ser

**Original source** This variant overlaps [2 transcripts](#).

**About this variant**

**Flanking sequence**

[Download sequence](#) [BLAST this sequence](#)

**Flanking sequence**

The sequence below is from the **reference genome** flanking the variant location. The variant is shown in **red** text. Neighbouring variants are shown with highlighted letters and ambiguity codes.

To change the display of the flanking sequence (e.g. hide the other variants, change the length of the flanking sequence), use the **"Configure this page"** link on the left.

Variants [5 prime UTR](#) [Focus variant](#) [Intronic](#) [Missense](#) [Synonymous](#) [Upstream](#)

Markup loaded

```
TCCTGTCCAAGCCCGCCCGCTTATCTACCGTGCCGGGCTCGGACCGGGCCCAAAAAGCG
GGCTTCGTGCCGGCTCACGGGCTCGTGCTTTTGGCCATCTAAGCCGACACTTAG
CATACATACGCAAGAAGAGGAGAGGCCGGAGGTGCGCGTGCTCCTTGCTTCTGCTGAC
TGGTCTCAGCATCTCATCCACCACCACCACCACCAACCATCTTAGGATAAGATAGC
AAATATATGGCCATCATACTCGTAGAGCAGCGTSGCYGGGGCTCTCCGCCCGGACAGC
ATCAGCCACCAGGGGACTCTCCAGTGCTCCACCCTGCTCAAGACAAAGAGCCGGCGGC
CGCCGGTGGATGCCGTGCTCGCTCTTTGGCTCCACCCTGSGGAGGCTGGCCGTCCCTCC
CCGGCGTACTCCAGCCTCGCGTCAACCCGGCGGAGAGGCGCTCGTCTCGTCCGAG
CAGAAGGTCTACGACGTCTGCTCAAGCAGGCCGATGCTCAAACGCCAGTGCACAGC
CCGTCCTCGACCCAGGCCAGGACATGGACATGCCACGCAACGGGCTCAAGGAAGCC
TACGACCGCTGCGGCGAGATCTGTGAGGAGTATGCCAAGACGTTTTACCTCGGTACGTAC
KRATATATATGGGATCCATCTTCTCTCCAATCCACAATCTCATCGTTTCAGTCTGCT
TCGATATATATGGGATCCATCTTCTCTCCAATCCACAATCTCATCGTTTCAGTCTGCT
```

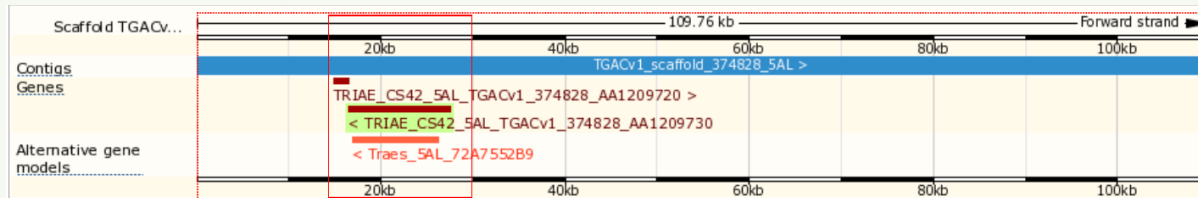


# Configuring Browser Views

The image displays the Gramene browser interface with several key elements highlighted for configuration:

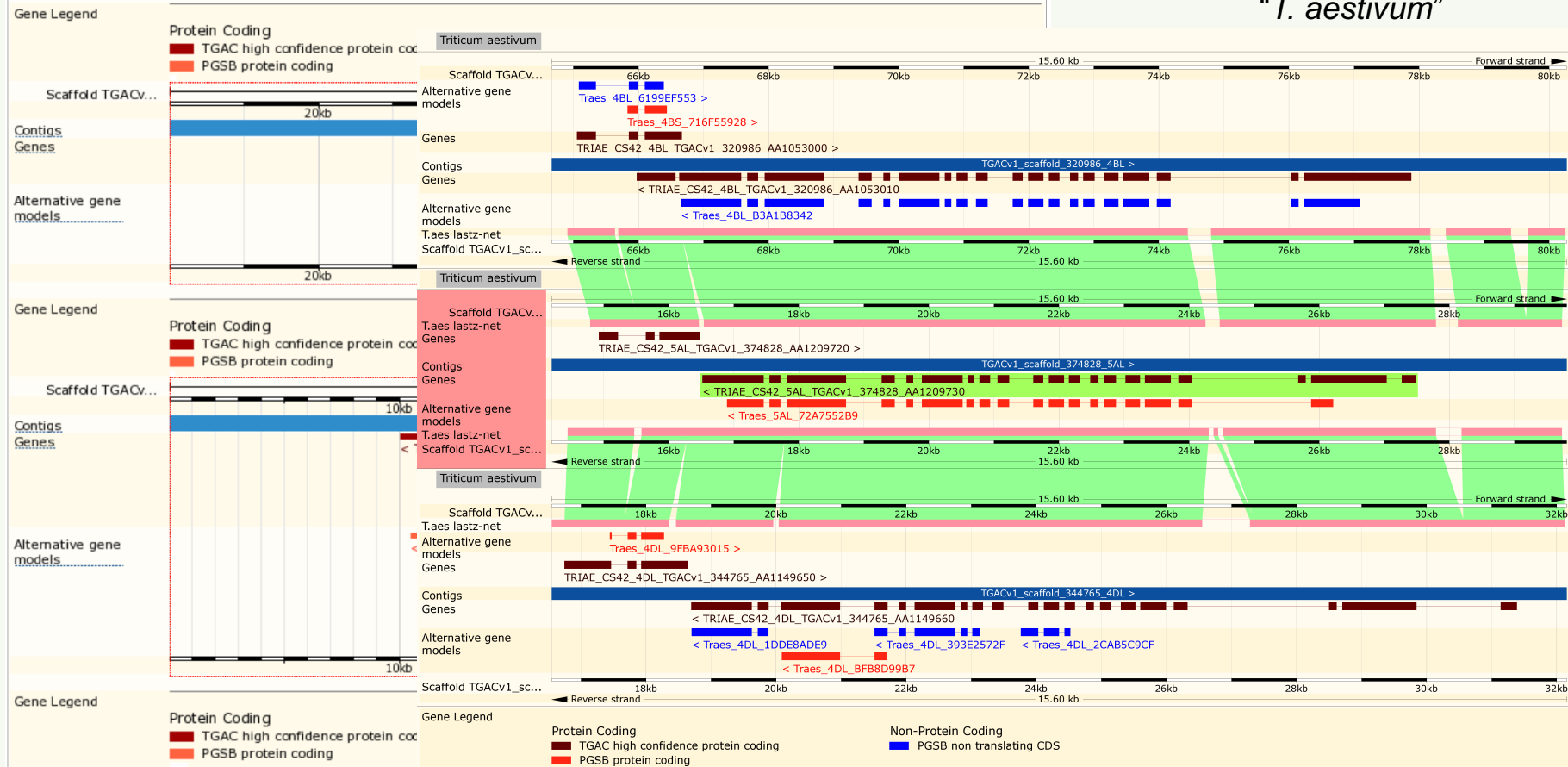
- Left Sidebar:** A list of browser categories including "Resequencing", "Linkage Data", "Markers", "Other genome browsers", and "Phytozome". The "Configure this page" option is circled in red.
- Top Panel:** Contains navigation controls such as "Scroll:", "Track height:", and "Drag/Select:". A red box highlights a track name, "ctg78", which is the subject of a configuration callout.
- Gene Legend:** A legend at the bottom of the top panel shows a red square representing "protein coding".
- Location and Gene Fields:** A search bar with "Location: 2:45809642-45885359" and a "Gene:" field.
- Track Configuration:** A callout box titled "Mouse over track name & configure symbol" points to a track named "B73\_CpG\_coverage". A secondary callout box titled "Click on 'Configure this image' symbol" points to a gear icon in the track's header.
- Configuration Menu:** A pop-up menu titled "Change track style:" offers options: "Off", "Wiggle plot", "Gradient", and "P-value".
- Main View:** A detailed genomic track showing various data layers like "B73\_CpG\_ratio", "B73\_CpG\_coverage", "cDNA", "EST.Cluster (Maize)", "MAKER-P.genes", "Gramene.gene", "Contigs", "mRNA alignments", "HapMap2", "Panzea\_2.7GBS", and "%GC".

# Wheat Polyploid Views



Wheat homologs of Arabidopsis  
**DCL1** (Dicer-like 1)

Hint: Search for “dicer” in  
“*T. aestivum*”



# Variation

## SNPs

**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, WTCBG 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**

- Genomic context
- Genes and regulation
- Flanking sequence
- Genotype frequency
- Phenotype data
- Sample genotypes
- Linkage disequilibrium
- Phylogenetic context
- Citations

## Structural variants (Arabidopsis & sorghum)

**Structural variant: ENSSVATH00000001**

**Variation class** deletion ([SO:0000159](#))

**Source** Ensembl - Variation features from Affy\_250k, Perlegen\_1M, WTCBG 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**

- Genomic context
- Genes and regulation
- Supporting evidence
- Phenotype data

## EMS-induced (sorghum & wheat)

**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**

- Genomic context
- Genes and regulation
- Flanking sequence
- Genotype frequency
- Phenotype data
- Sample genotypes
- Linkage disequilibrium
- Phylogenetic context
- Citations



# Variation Data Summary

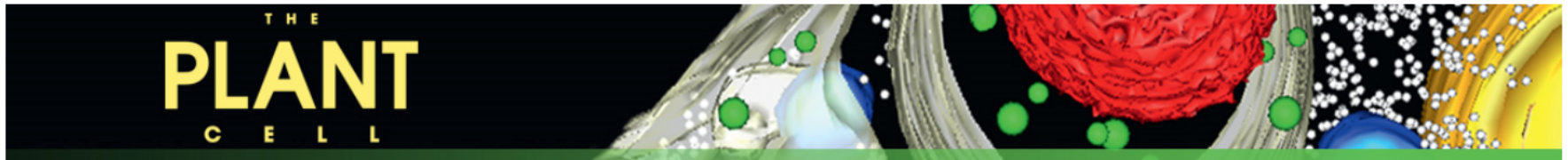


~231 M plant SNPs in 9 species

- **Arabidopsis** –1001 Genomes (3 studies; includes **structural**), Horton *et al* , BBSRC-funded collaboration, Atwell *et al* [**phenotypes**]
- **Barley** – Barley Consortium, Mascher *et al*, Comadran *et al*
- **Brachypodium** – *B. sylvaticum* (Fox *et al*)
- **Grapevine** – USDA germplasm collection (Myles *et al*)
- **Maize** – HapMap 1 & 2 (Gore *et al*, Chia *et al*), Panzea 2.7 GBS, HapMap 3 (coming soon)
- **Rice** [*Oryza sativa* Japonica & Indica, *O. glumaepatula*, *O. glaberrima*, *O. barthii*] – dbSNP (Yu *et al*), McNally *et al*, Zhao *et al*, Duitama *et al*, OGE< 3000 Genomes IRRI (coming soon)
- **Sorghum** – Morris *et al*, Mace *et al*, Xin *et al* [**EMS**]; DGVa (**structural**)
- **Tomato** – 100 Tomato Genome Sequencing Consortium
- **Wheat** – HapMap (Jordan *et al*), CerealsDB, Krasileva *et al* [**EMS**]

# The Power of Comparative Genomics

## Sorghum as a functional genomics platform for maize



Other | Large-Scale Biology Article

256 mutant lines sequenced  
~1.5 M point mutations

### A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses

Yinping Jiao, John J. Burke, Ratan Chopra, Gloria Burow, Junping Chen, Bo Wang, Chad Hayes, Yves Emendack, Doreen Ware, Zhangguo Xin

Published June 2016. DOI: <https://doi.org/10.1105/tpc.16.00373>



# Variation Use Case

Find “stop\_gained” alleles in transcription factors

**Step 1: Use Gramene BioMart**  
Search genes with GO:0003700 (transcription factor activity) with predicted stop-gain variants

**Step 2: Use Genome Browser to explore structure, expression, conservation, and other features in gene of interest**

bio::mart Login/Register

Search all species...

Dataset 77 / 34567 Genes  
Sorghum bicolor genes (Sorb1 (2007-12-JGI))

Filters  
GO term accession [e.g. GO:0050789] : [ID-list specified]  
Variation type : stop\_gained  
Germline variation source : Sorghum EMS\_mutants

Attributes  
Gene stable ID  
InterPro ID  
InterPro short description

Dataset 7447 / 8189230 SNPs  
Sorghum bicolor variations (Sorb1 (2007-12-JGI))

Filters  
Consequence type : stop\_gained  
Variation source : Mace\_2013, Morris 2013, Sorghum EMS\_mutants

Attributes

Gene stable ID	InterPro ID	InterPro short description	Variation ID	Consequence to transcript	Variation source	Variant allele
Sb04g000300	IPR004827	TF_bZIP	tmp_4_112236_G_A	stop_gained	Mace_2013	G/A
Sb08g004360	IPR000679	Znf_GATA	tmp_8_5166294_C_G	stop_gained	Mace_2013	C/G
Sb08g004380	IPR013088	Znf_NHR/GATA	tmp_8_5166294_C_G	stop_gained	Mace_2013	C/G
Sb08g003700	IPR007126	Histone_core_D	tmp_8_4256587_G_A	stop_gained	Mace_2013	G/A
Sb08g003700	IPR007126	Histone_core_D	tmp_8_42565818_G_A	stop_gained	Mace_2013	G/A
Sb08g003700	IPR009072	Histone-fold	tmp_8_4256587_G_A	stop_gained	Mace_2013	G/A
Sb08g003700	IPR009072	Histone-fold	tmp_8_42565818_G_A	stop_gained	Mace_2013	G/A
Sb07g004580	IPR001471	TF_ERF_DNA-bd	tmp_7_5865063_G_A	stop_gained	Mace_2013	G/A
Sb07g004580	IPR016177	DNA-bd_integrase-tyr	tmp_7_5865063_G_A	stop_gained	Mace_2013	G/A
Sb07g004580	IPR031112	AP2-like_transcript_factor	tmp_7_5865063_G_A	stop_gained	Mace_2013	G/A
Sb08g005635	IPR004827	TF_bZIP	tmp_8_7821495_C_T	stop_gained	Mace_2013	C/T
Sb01g038370	IPR002589	A1pp	tmp_1_61895020_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g038370	IPR011146	His_triad-like_motif	tmp_1_61895020_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g038370	IPR018808	Histidine_triad_CS	tmp_1_61895020_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g038370	IPR026963	Apratxin	tmp_1_61895020_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g038370	IPR032566	zf-C2HE	tmp_1_61895020_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g030770	IPR004827	TF_bZIP	tmp_1_53392629_C_T	stop_gained	Sorghum EMS_mutants	C/T
Sb01g039810	IPR000943	RNA_pol_sigma70	tmp_1_63256928_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g039810	IPR000943	RNA_pol_sigma70	tmp_1_63256775_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g039810	IPR007624	RNA_pol_sigma70_r3	tmp_1_63256928_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g039810	IPR007624	RNA_pol_sigma70_r3	tmp_1_63256775_G_A	stop_gained	Sorghum EMS_mutants	G/A
Sb01g039810	IPR007627	RNA_pol_sigma70_r2	tmp_1_63256928_G_A	stop_gained	Sorghum EMS_mutants	G/A



**Step 3: Identify germplasm containing genotype of interest:**

tmp\_1\_53392629\_C\_T SNP

Original source: A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses. Contact Dr. Zhanguo Xin for seeds

Alleles: C/T | Ambiguity code: Y

Location: Chromosome 1:53392629 (forward strand) | View in location tab

Most severe consequence: Stop gained | See all predicted consequences (Genes and regulation)

HGVS names: This variant has 3 HGVS names - Hide

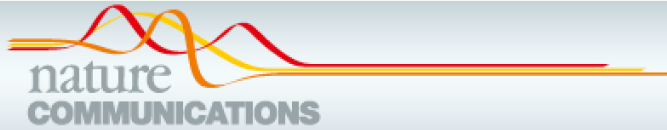
- 1:g.53392629C>T
- Sb01g030770.1:c.943C>T
- Sb01g030770.1:p.Gln315Ter

Genotypes for Sorghum EMS\_mutant\_population

Sample	Genotype (forward strand)
ARS230	CIT
ARS1	CIC
ARS10	CIC
ARS100	CIC
ARS101	CIC

Order seeds from USDA or contact Dr. Xin

# Sorghum *msd1*



ARTICLE

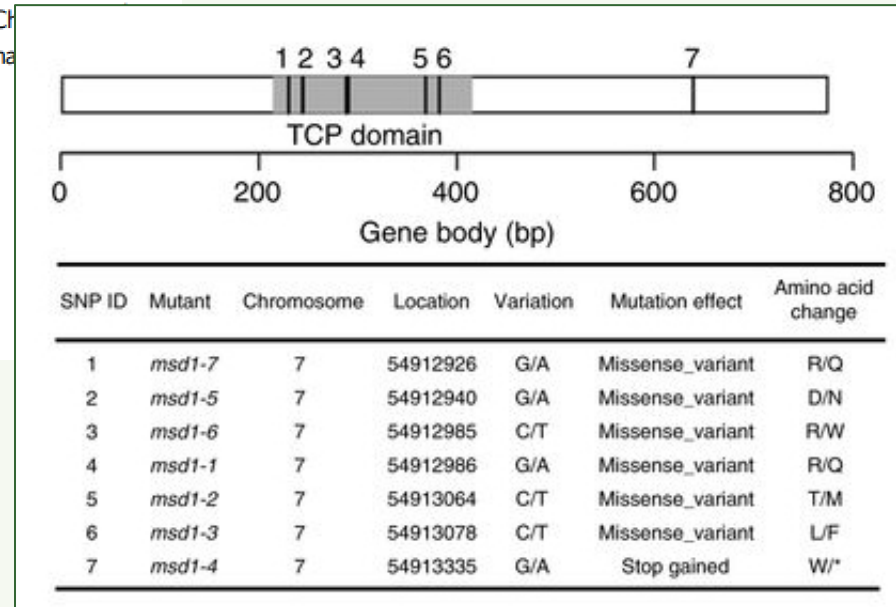
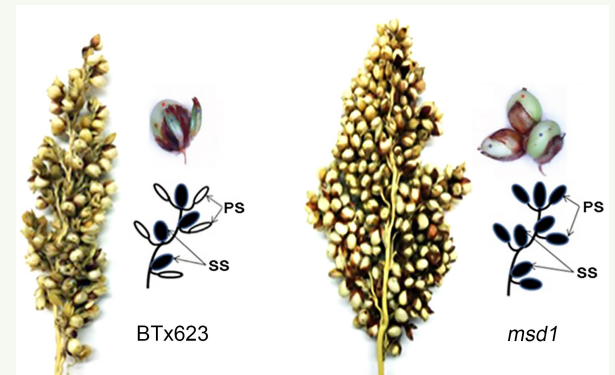
DOI: 10.1038/s41467-018-03238-4

OPEN

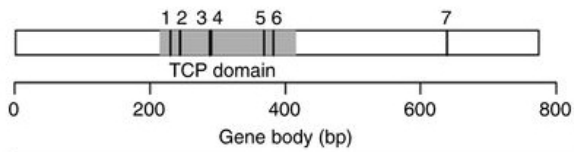
## MSD1 regulates pedicellate spikelet fertility in sorghum through the jasmonic acid pathway

Yinping Jiao<sup>1,2</sup>, Young Koung Lee<sup>2,3</sup>, Nicholas Gladman<sup>1,2</sup>, Ratan Chopra<sup>1</sup>, Shawn A. Chaffin<sup>1</sup>, Michael Regulski<sup>2</sup>, Gloria Burow<sup>1</sup>, Chad Hayes<sup>1</sup>, John Burke<sup>1</sup>, Doreen Ware<sup>1,2,5</sup> & Zhaohua

Grain number per panicle (GNP) is a major determinant of grain yield in cereals. However, the mechanisms that regulate GNP remain unclear. To address this issue, we isolate a series of sorghum [*Sorghum bicolor* (L.) Moench] *multiseeded* (*msd*) mutants that can double GNP by increasing panicle size and altering floral development so that all spikelets are fertile and set grain. Through bulk segregant analysis by next-generation sequencing, we identify *MSD1* as a TCP (Teosinte branched/Cycloidea/PCF) transcription factor. Whole-genome expression



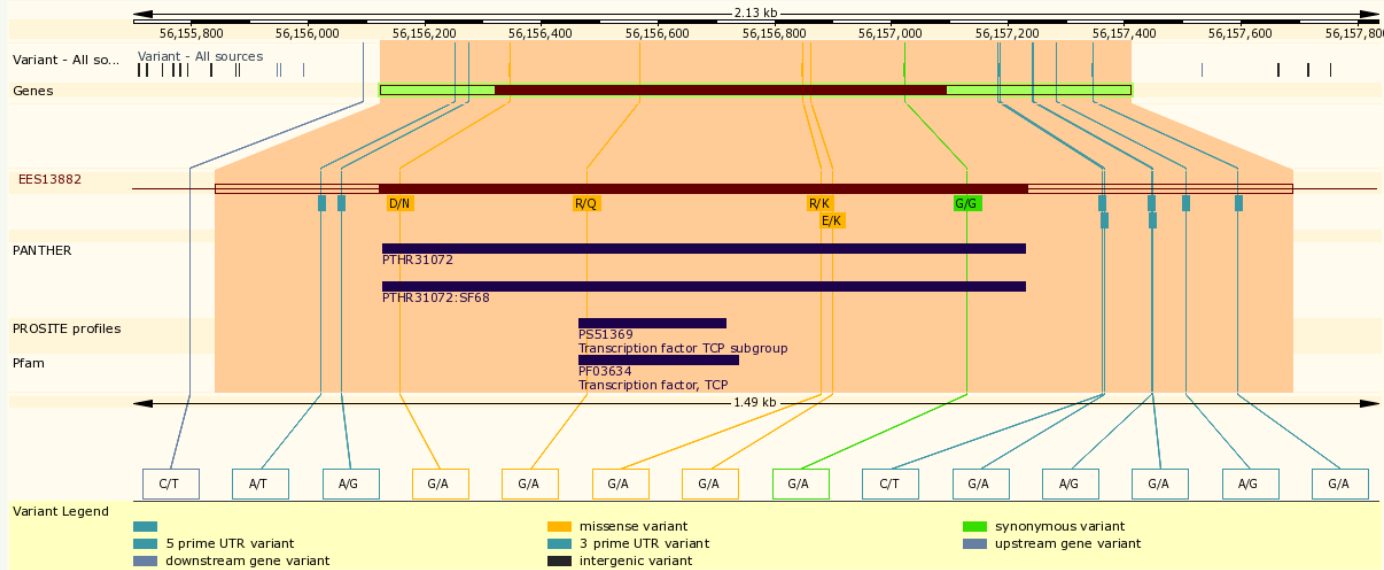
The secret to tripling the # of grains in sorghum & perhaps other staple crops



# *msd1* [SORBI\_3007G135700]

The *msd1* mutants are not in the sequenced population. But... Is the amino acid sequence conserved in other species like maize?

SNP ID	Mutant	Chromosome	Location	Variation	Mutation effect	Amino acid change
1	<i>msd1-7</i>	7	54912926	G/A	Missense_variant	R/Q
2	<i>msd1-5</i>	7	54912940	G/A	Missense_variant	D/N
3	<i>msd1-6</i>	7	54912985	C/T	Missense_variant	R/W
4	<i>msd1-1</i>	7	54912986	G/A	Missense_variant	R/Q
5	<i>msd1-2</i>	7	54913064	C/T	Missense_variant	T/M
6	<i>msd1-3</i>	7	54913078	C/T	Missense_variant	L/F
7	<i>msd1-4</i>	7	54913335	G/A	Stop gained	W/*



Variant ID	Chr: bp	Alleles	Class	Source	Evidence	Clin. Sig.	Conseq. Type	AA	AA coord	Transcript
<a href="#">tmp_7_54912651_A_G</a>	7:56156275	A/G	SNP	Mace_2013	-	-	5 prime UTR variant	-	-	EES13882
<a href="#">tmp_7_54912721_G_A</a>	7:56156345	G/A	SNP	Sorghum_EMS_mutants	-	-	missense variant	D/N	9	EES13882
<a href="#">tmp_7_54912944_G_A</a>	7:56156568	G/A	SNP	Sorghum_EMS_mutants	-	-	missense variant	R/Q	83	EES13882
<a href="#">tmp_7_54913223_G_A</a>	7:56156847	G/A	SNP	Sorghum_EMS_mutants	-	-	missense variant	R/K	176	EES13882
<a href="#">tmp_7_54913237_G_A</a>	7:56156861	G/A	SNP	Sorghum_EMS_mutants	-	-	missense variant	E/K	181	EES13882
<a href="#">tmp_7_54913398_G_A</a>	7:56157022	G/A	SNP	Sorghum_EMS_mutants	-	-	synonymous variant	G	234	EES13882
<a href="#">tmp_7_54913559_C_T</a>	7:56157183	C/T	SNP	Mace_2013	-	-	3 prime UTR variant	-	-	EES13882

# Tools

<http://ensembl.gramene.org/tools.html>



[BLAST](#)

[BioMart](#)

[Tools](#)

[Downloads](#)

[Help](#)







[Feedback](#)

[UploadData](#)

## Tools





We provide a number of ready-made tools for processing both our data and yours. We routinely delete results from our servers after 10 days, but if you have an [gramene genome account](#) you will be able to save the results for about 2-3 months until next release.

### Processing your data

Name	Description	Online tool	Upload limit	Download script	Documentation
<a href="#">Variant Effect Predictor</a> 	Analyse your own variants and predict the functional consequences of known and unknown variants via our Variant Effect Predictor (VEP) tool.		50MB*		
<a href="#">HMMER</a>	Quickly search our genomes for your protein sequence.				
<a href="#">BLAST/BLAT</a>	Search our genomes for your DNA or protein sequence.		50MB		
<a href="#">Assembly Converter</a>	Map (liftover) your data's coordinates to the current assembly.		50MB		

\* For larger datasets we provide an API script that can be downloaded (you will also need to install our Perl API, below, to run the script).

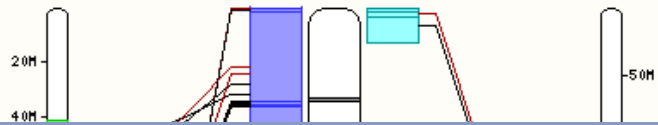
### Accessing Gramene data

Name	Description	Get it from:	Documentation
<a href="#">BioMart</a>	Use this data-mining tool to export custom datasets from Gramene.	<a href="#">Gramene BioMart</a>	
<a href="#">Ensembl Perl API</a>	Programmatic access to all Ensembl data using simple Perl scripts	<a href="#">GitHub</a> or <a href="#">FTP download</a> (current release only)	
<a href="#">Ensembl Genomes Virtual Machine</a>	Pre-configured VirtualBox virtual machine (VM) running the latest Ensembl Genomes browser.		
<a href="#">Ensembl Genomes REST server</a>	Access Ensembl data using your favourite programming language		



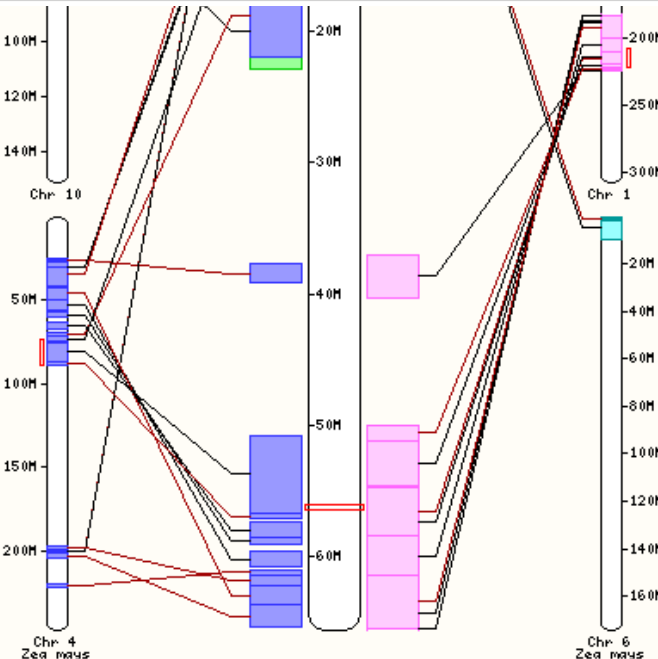


# Are the sorghum *msd1* variant sites conserved in maize?



**Sorghum bicolor genes**      **Location**      **Zea mays homologues**      **Location**

**SORBI\_3007G135700**      **7:56156125-56157410**      →      **TCP-transcription factor 11**      **1:217455034-217455756**  
 (SORBI\_3007G135700)                                                                  (Zm00001d032217)



Sorghum chr7

- Alignments (text)
  - Region Comparison
  - Genetic Variation
  - Resequencing
  - Linkage Data
  - Markers
  - Other genome browsers
  - Phytozome
- Configure this page
- Custom tracks
- Export data
- Share this page
- Bookmark this page
- Gramene is produced in collaboration with Ensembl Plants

**Alignments (text)**

Alignment: Zea mays - lastz

Location: 7:56156218-56157410      Gene:      Go

Download alignment

**Species Tree**

No tree is drawn when there is more than one block displayed because each block is represented by a separate tree

**Sorghum bicolor** > chromosome:Sorghum\_bicolor\_NCBIV3.7:56156218:56157410.1  
**Zea mays** > chromosome:AGPV4.4:85370657:85371903.1

**Alignment**

Currently showing full alignment. Please click the button below to show the alignment for first 120 columns.

Hide full alignment

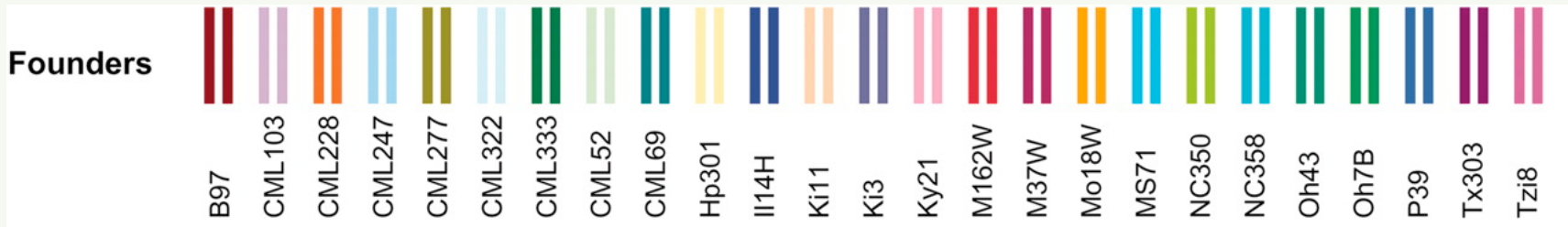
Sorghum bicolor	7:56156218	TCAATACCCCTAACAGACGCCGCTCTCATCACCTTCATGATCATTCGGCTTACCAGCAGAGACGATCA	---	TTTACCGAACCCCTACTGCAGCTGTTT	GTTCCTATGCGCTCCAGC	7:56156332
Zea mays	4:85370657	TCAATACCCCTAACAGACGCCGCTCTCATCACCTTCATGATCATTCGGCTTACCAGCAGAGACGATCA	---	TTTACCGAACCCCTACTGCAGCTGTTT	GTTCCTATGCGCTCCAGC	4:85370768
Sorghum bicolor	7:56156333	GC-----GATGCTTGGACGGGTACGGCGGCGAGATCTCCGGCCGACATCGTCTGCCACC	---	ACCAGGACACCTGGAGCGGTGTCGG	---	7:56156434
Zea mays	4:85370769	GC-----GATGCTTGGACGGGTACGGCGGCGAGATCTCCGGCCGACATCGTCTGCCACC	---	ACCAGGACACCTGGAGCGGTGTCGG	---	4:85370888
Sorghum bicolor	7:56156435	ACGGCCCTCTCCAGCCCGCCG	---	CTCTGAGGACGGGTCTCCGCTACCCGGTGGATCCGGGGTTCATGCCCGCCGCGAGCGGGCCG	---	7:56156554
Zea mays	4:85370889	ACGGCCCTCTCCAGCCCGCCG	---	CTCTGAGGACGGGTCTCCGCTACCCGGTGGATCCGGGGTTCATGCCCGCCGCGAGCGGGCCG	---	4:85370975
Sorghum bicolor	7:56156555	TTTCAGGACGGATTCACACAGCA	---	AGCCCGCCGATCGCGCTCGCTGGTGGGGTCCGGCCGAGGTTCTTCGGCTCAGGACCGCTGGGTCGAC	---	7:56156674
Zea mays	4:85370976	TTTCAGGACGGATTCACACAGCA	---	AGCCCGCCGATCGCGCTCGCTGGTGGGGTCCGGCCGAGGTTCTTCGGCTCAGGACCGCTGGGTCGAC	---	4:85371095
Sorghum bicolor	7:56156675	AAGCCAGCAAGAGC	---	AAGCCCGCCATCGACCGCTCGTGCAGCGCCGCG	---	7:56156779
Zea mays	4:85371096	AAGCCAGCAAGAGC	---	AAGCCCGCCATCGACCGCTCGTGCAGCGCCGCG	---	4:85371215
Sorghum bicolor	7:56156780	CCGACGGTGGTGAAG	---	AGCACTTGCTGTTGACGTGACTCGAG	---	7:56156890
Zea mays	4:85371216	CCGACGGTGGTGAAG	---	AGCACTTGCTGTTGACGTGACTCGAG	---	4:85371335
Sorghum bicolor	7:56156891	CGCGTACCGGTGTCCTGATGGCCACCG	---	GGGTCATGGGAAGCA	---	7:56156998
Zea mays	4:85371336	CGCGTACCGGTGTCCTGATGGCCACCG	---	GGGTCATGGGAAGCA	---	4:85371437
Sorghum bicolor	7:56156999	GCGCAGCCGACGAGTGGACGG	---	CTGGAGTACTACTACCGATTTGCCCTGCAGCTCGAGGAGATGA	---	7:56157106
Zea mays	4:85371438	ACGACCGCCGACGAGTGGACGG	---	CTGGAGTACTACTACCGATTTGCCCTGCAGCTCGAGGAGATGA	---	4:85371556
Sorghum bicolor	7:56157107	TATGCAAGATTATATATAGTCTATAAAGCTCC	---	AACCGCGCTGAGCA	---	7:56157199
Zea mays	4:85371557	TATGCAAGATTATATATAGTCTATAAAGCTCC	---	AACCGCGCTGAGCA	---	4:85371673
Sorghum bicolor	7:56157200	TAGTACATATAAATCCCAACTGGGCTTCTCTATCAAT	---	ATACCAATGGTGCFAATAGATATCGTATCTTTAT	---	7:56157304
Zea mays	4:85371674	TAGTACATATAAATCCCAACTGGGCTTCTCTATCAAT	---	ATACCAATGGTGCFAATAGATATCGTATCTTTAT	---	4:85371793
Sorghum bicolor	7:56157305	CGATATGCGTGTGTCATTTATTC	---	ATACCAATGGTGCFAATAGATATCGTATCTTTAT	---	7:56157410
Zea mays	4:85371794	CGATATGCGTGTGTCATTTATTC	---	ATACCAATGGTGCFAATAGATATCGTATCTTTAT	---	4:85371903

# Track Hubs in Gramene

EBI Atlas team: RNAseq-er API aligns all ENA short reads for genomes in Ensembl updated daily



# 24 Maize NAM Founder lines



Yu et al. (2008) Genetics 178:539-551; <https://doi.org/10.1534/genetics.107.074245>

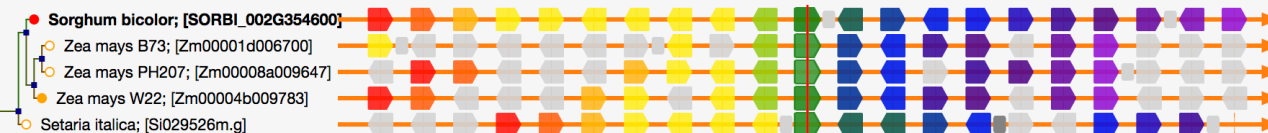
- Chromosome-level assembly
- Uniform annotation protocol
- Compara gene trees—ortholog identification
- All pair-wise synteny maps
- Pan-genome analysis

**NSF Award # 1744001**

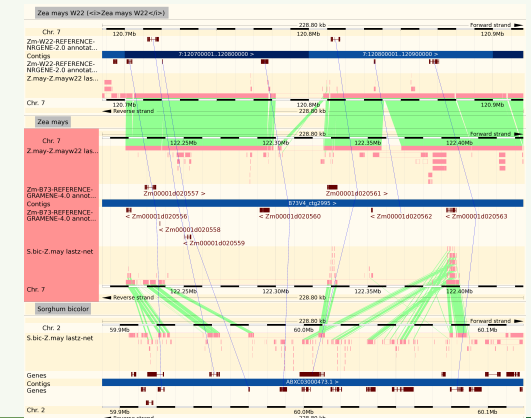
Preview at

[maize-pangenome.gramene.org](http://maize-pangenome.gramene.org)

Pangenome site: **B73, W22, PH207** maize + outgroups



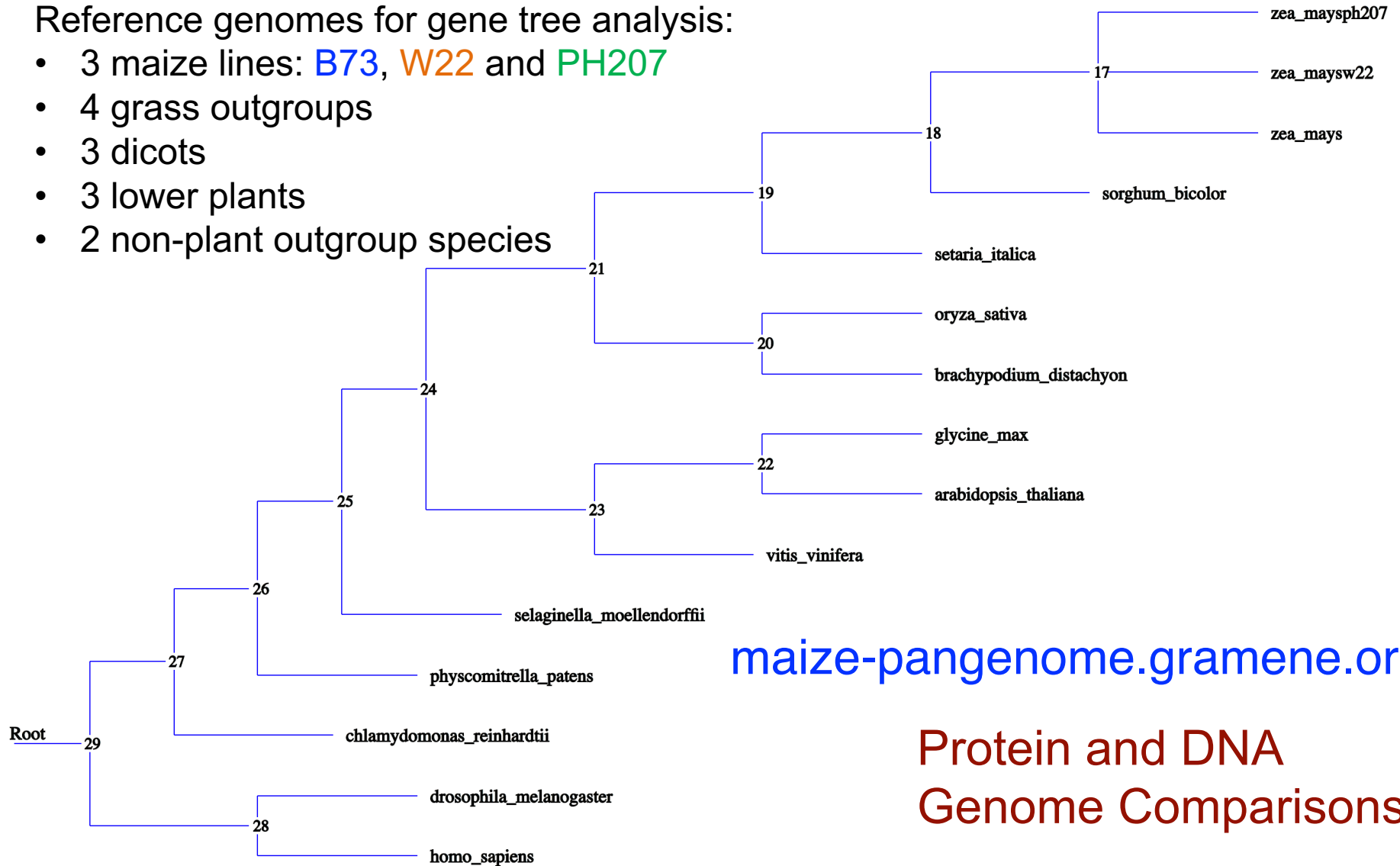
Variation: copy-number presence/absence



# Maize Genomes at PanMaize

Reference genomes for gene tree analysis:

- 3 maize lines: B73, W22 and PH207
- 4 grass outgroups
- 3 dicots
- 3 lower plants
- 2 non-plant outgroup species



# Pathways



## Plant Reactome: pathways, species, and analysis features

<http://plantreactome.gramene.org>

**Sushma Naithani**

Department of Botany and Plant Pathology

Oregon State University

naithans@science.oregonstate.edu





2018 Maize Genetics Conference 22-25 March


<https://www.youtube.com/watch?v=8BkwBTMY-3w&t=6s>



ASPB 2018

About | Contact | Join Home **Community** Jobs Log In

 **Gramene**

**Dashboard** Discussions Articles Videos Documents Collections

---

### Description

Gramene (<http://www.gramene.org>) is an integrated plant resource for reference genomes and comparative functional analysis in plants. Does your species have a reference assembly, transcript data or genetic variation data for entire populations? If so, these data sets may already be part of Gramene.

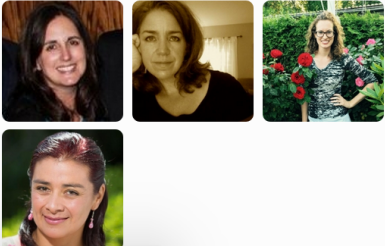
Every new Gramene release provides researchers with access to reference assemblies and annotation for over 50 genomes in a genome browser. In addition, we provide curated pathways and orthology-based pathway projections for 75 plants species.

Have you ever needed to know the function of the gene, and wanted to see the function of the ortholog in rice, maize or Arabidopsis? Does the gene you are working on belong to a large or small gene family. Did the gene family associated with expansion or contractions, in species that are evolutionary close? Is the biochemical pathway you work on conserved in sorghum and soybean? If so, Gramene can help you explore these questions in the Gramene. Gramene provides powerful phylogenetic approaches, including protein-based gene trees with stable IDs and whole-genome DNA alignments, enable traversing between plant species. We provide integrated search capabilities and interactive views to visualize gene features, gene neighborhoods, phylogenetic trees, genetic variation, gene expression profiles, pathways, and cross-references and host curated rice pathways, and uses these curated pathways to generate orthology-based projections for other

#### Upcoming Events

Jul <b>16</b>	<b>Getting involved with Gramene</b> Monday, July 16 at 1:00 PM - 1:30 PM EDT
------------------	--

#### Members





<https://community.plantae.org>

About | Contact | Join Home **Community** Jobs Log In

# Plantae

## Networks

View groups and information relating to topics of interest in the Plantae community.

Most Recent Name

Search...

★ Featured Plantae Bookshelf [Go to page](#)

★ Featured The Taproot: A Plant Podcast [Go to page](#)

★ Featured Big Data and Cyberinfrastructure [Go to page](#)

★ Featured Plant Phenomics [Go to page](#)

★ Featured Teaching Tools in Plant Biology [Go to page](#)

★ Featured Plant Science Research Network [Go to page](#)

★ Featured WAVETT: WordPress, Audacity, and Vid... [Go to page](#)

★ Featured Methods and Protocols [Go to page](#)

★ Featured Plant Science Graduate Programs [Go to page](#)

★ Featured Abiotic Responses [Go to page](#)

★ Featured Reproduction in Plant Science [Go to page](#)

★ Featured Plant Biology Block [Go to page](#)

# Gramene

[Go to page](#)

# Get in touch!

- Slides, exercises, brochures, posters:  
[ftp://ftp.gramene.org/pub/gramene/outreach/aspb\\_2018](ftp://ftp.gramene.org/pub/gramene/outreach/aspb_2018)
- Outreach: <http://gramene.org/outreach>
- E-mail: [feedback@gramene.org](mailto:feedback@gramene.org)
- Gramene collections in  **Plantae**
- Announcements mailing list



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



@GrameneDatabase

# Gramene - Exploring Function through Comparative Genomics and Network Analysis

## NSF IOS 1127112 (2012- 2017)

**Doreen Ware**, PI (CSHL, USDA ARS)

Kapeel Chougule, Yinping Jiao, Sunita Kumari, Demitri Muna, Andrew Olson, Joshua Stein, Marcela K. Tello-Ruiz, Peter van Buren, 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, Matthieu Muffato, Guy Naamati, Electra Tapanari, Dan Staines / Nuno Fonseca, Alfonso Munoz-Pomer Fuentes, Laura Huerta Martinez, Haider Iqbal, Maria Keays, Amy Tang

**Lincoln Stein** (OICR)

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

**Crispin Taylor** (ASPB)

Melanie Binder, Susan Cato, Jennifer Regala; Weijia Xu (TACC), Amit Gupta(TACC)

