

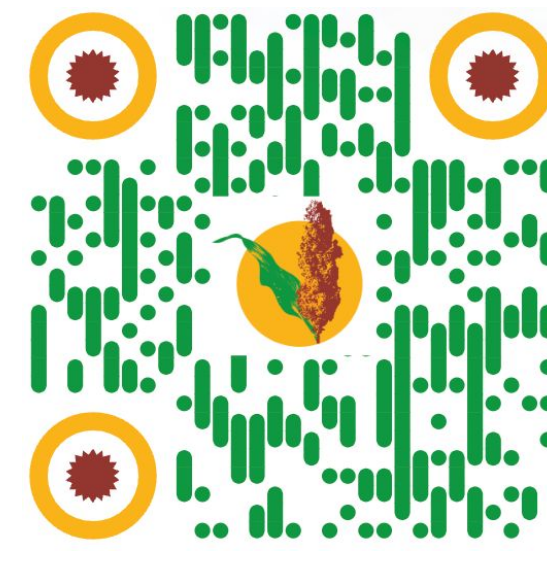
Kapeel Chougule<sup>1</sup>, Vivek Kumar<sup>1</sup>, Sunita Kumari<sup>1</sup>, Zhenyuan Lu<sup>1</sup>, Andrew Olson<sup>1</sup>, Audra Olson<sup>1</sup>, Marcela K. Tello-Ruiz<sup>1</sup>, Peter van Buren<sup>1</sup>, Sharon Wei<sup>1</sup>, Lifang Zhang<sup>1</sup>, Nicholas Gladman<sup>1,2</sup>, Doreen Ware<sup>1,2</sup>

<sup>1</sup> Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 11724, USA, <sup>2</sup> USDA ARS NEA, Plant Soil & Nutrition Laboratory Research Unit, Ithaca, NY, 14853, USA

## Project Targets

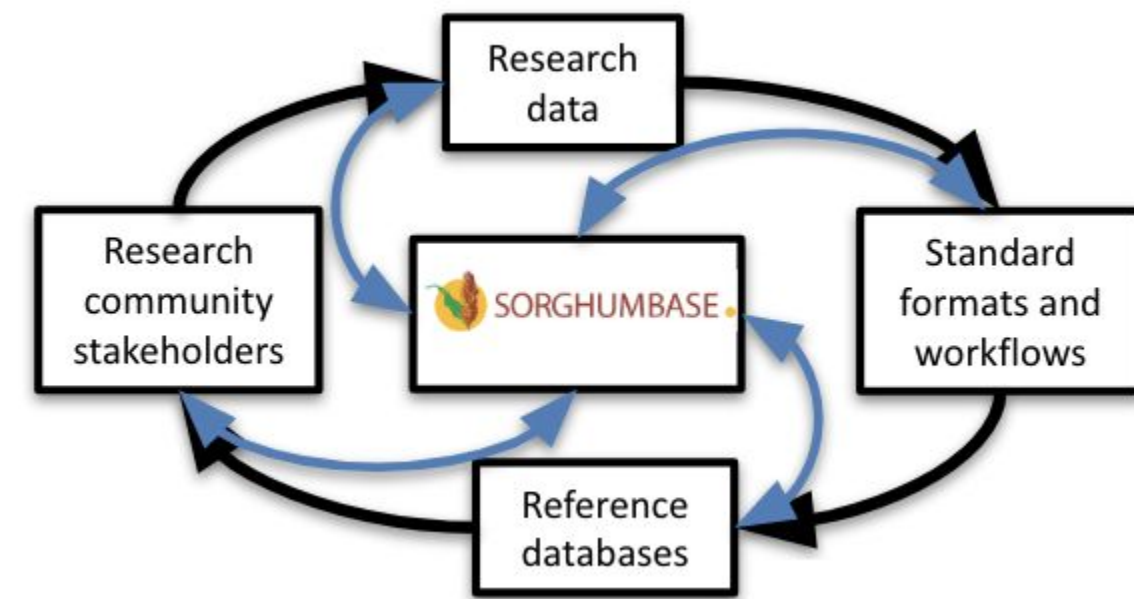
### SorghumBase comparative genomics resource

- Ingest public genomes and gene annotations
- Develop pan-gene/genome workflows
- Improve tools and visualizations
- API documentation and training



### Findable Accessible Interoperable Reusable community data (FAIR)

- Adopt/develop standards for data sharing
  - Germplasm
  - Genomes
  - Genetic variation
  - Genes
  - Gene function
  - Gene expression
  - Traits/phenotypes
- Integrate access via data federation
- Documentation and training



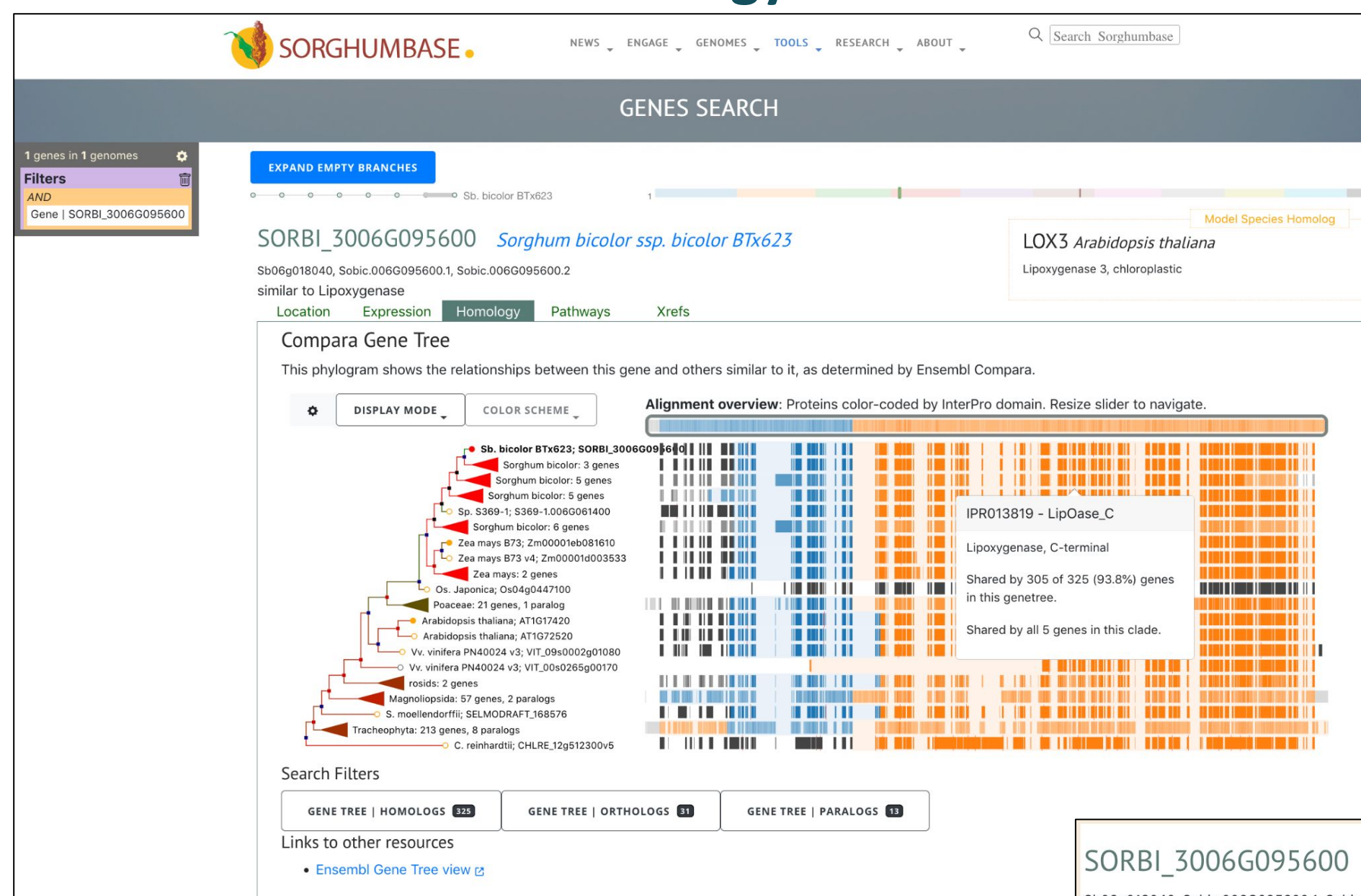
### Community information hub

- News
- events
- publications
- projects
- datasets
- people

Contact us to highlight your research, integrate your data & get personalized training and support:  
<https://www.sorghumbase.org/feedback>

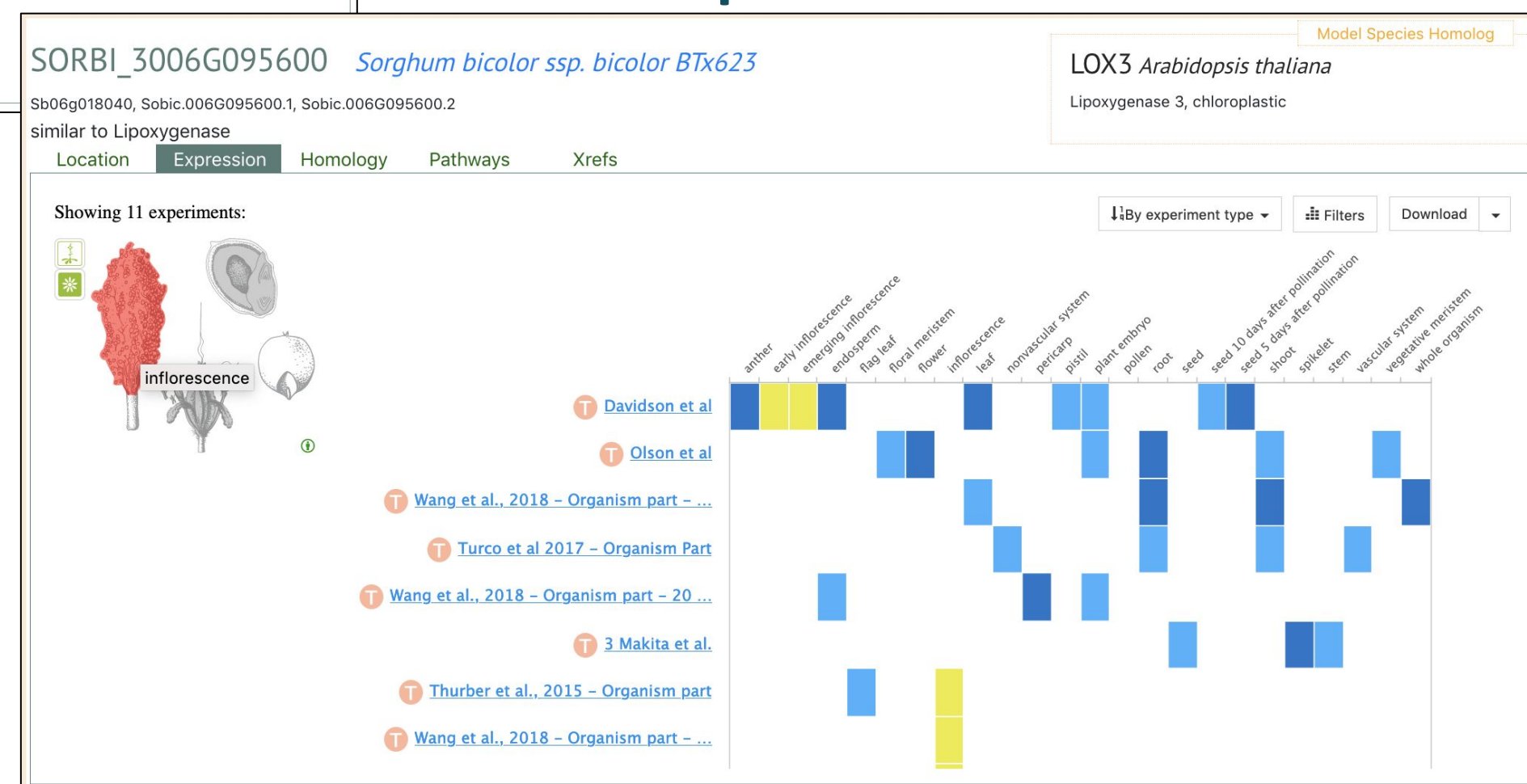
## Search Interface

### Homology View



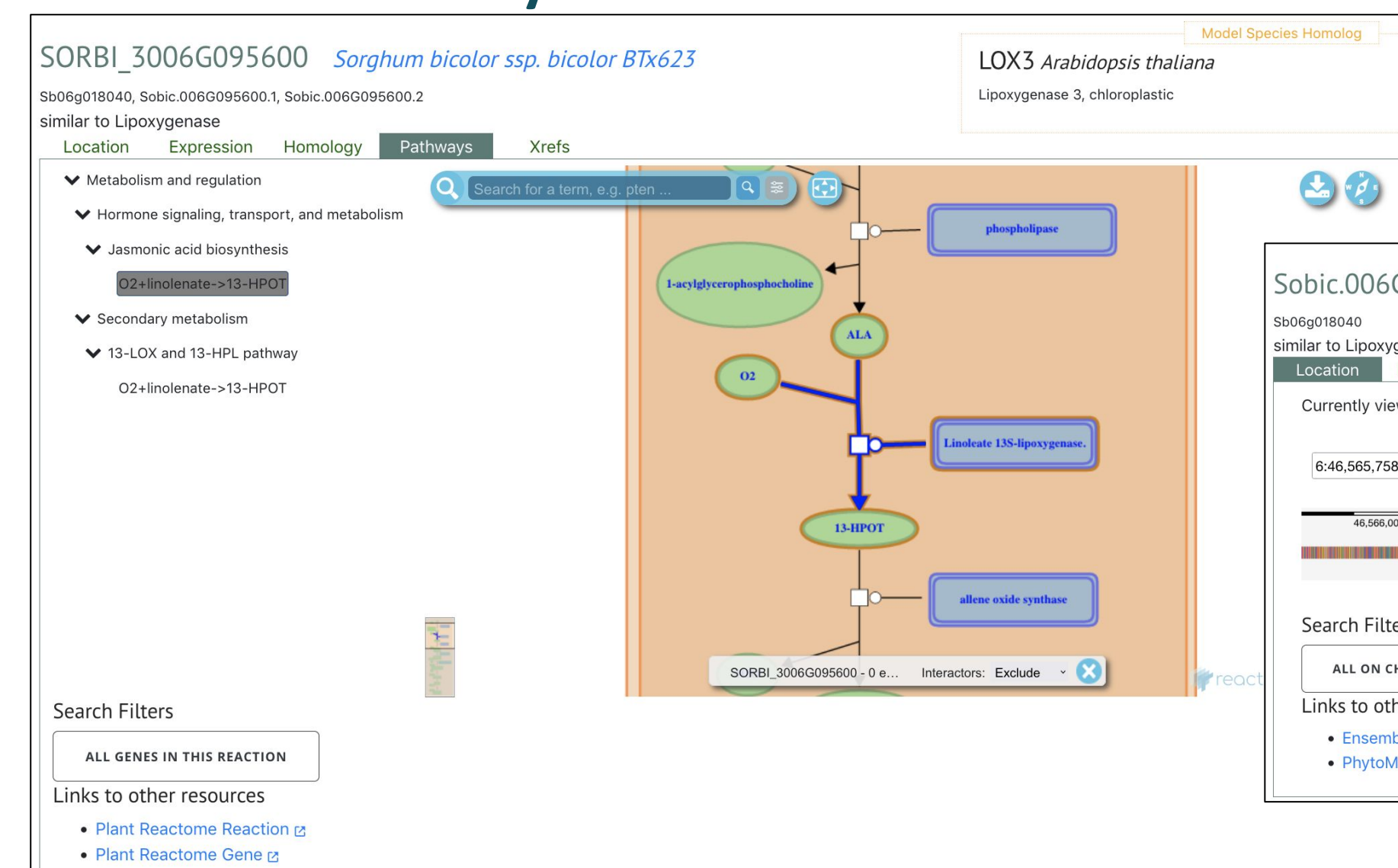
SorghumBase search capabilities include searching for common gene names, gene IDs, functional characterization (i.e., gene description, protein domain, ontology term, pathway), publication author name, and key words.

### Expression View



Available views of the results include expression & pathways panels, pruned gene trees to show selected species, highlighted protein domain structure, aligned amino acid sequences and neighborhood conservation.

### Pathways View



### Location View



## Integrated Comparative Genomics Resources

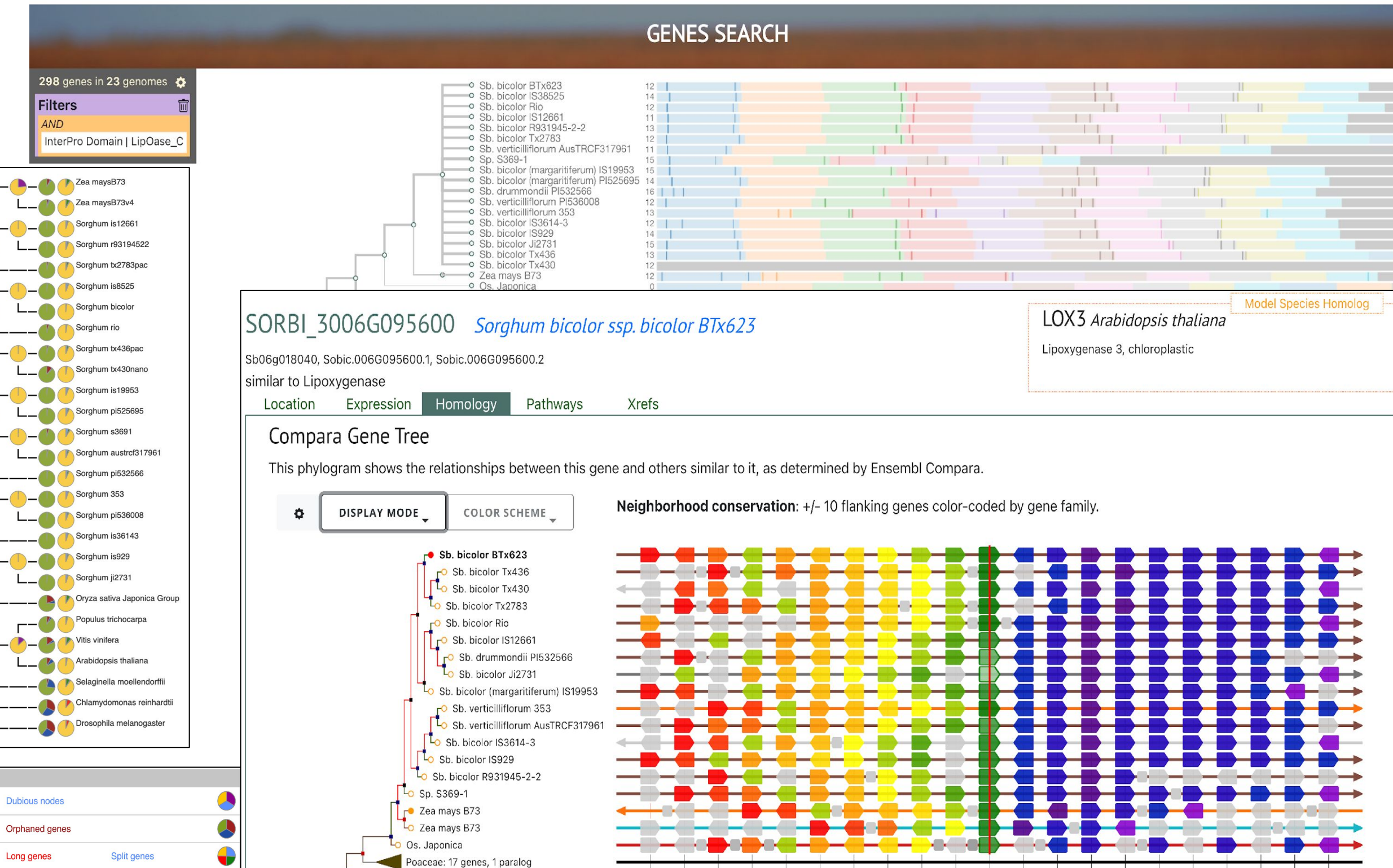
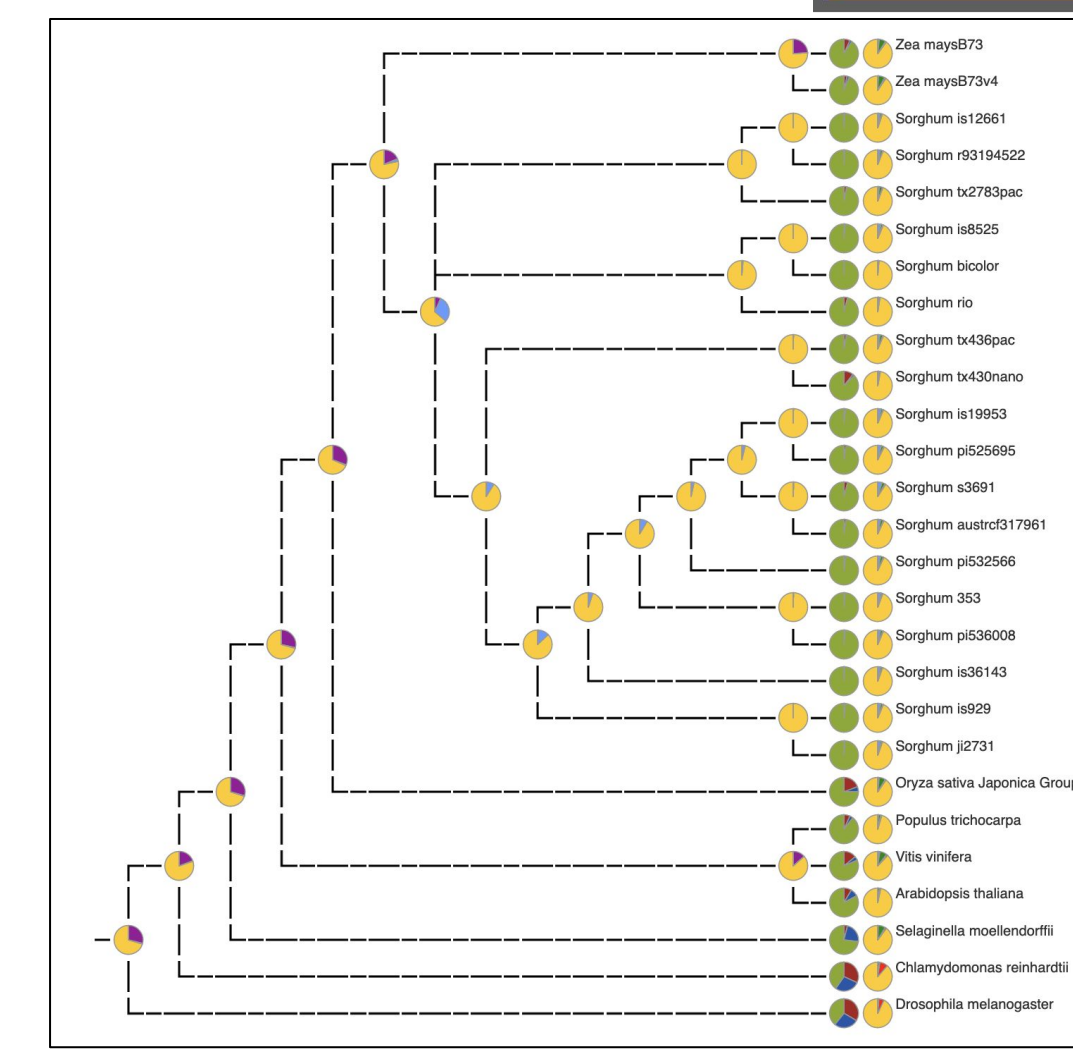
### Release # 5 (December 2022)

- 28 sorghum genomes (10 CP-NAM), 8 outgroups
- 44K gene family trees
- 61M genetic variants (SNPs & EMS mutations) affecting 23.7K LOF genes:
  - 44M SNPs, 400 SAP (Boatwright, 2022)
  - 13M SNPs, 499 samples (Lozano, 2022)
  - >2.5M EMS, 486 samples (Addo-Quaye, 2018)
  - >1.7M EMS, 252 samples (Jiao, 2018)
- Variant effect prediction
- Atlas QTLs: 6K from (223 traits, >150 studies)
- Functional domain annotation
- Gene expression
- Pathways



### Views & Tools

- Integrated search interface
- Gene family tree viewer
- Ensembl Genome browser
- BLAST (sequence search)
- EBI Expression Atlas widget
- Plant Reactome Pathway diagram
- Genome pages cross-link to GRIN-Global, ICRISAT, FAO, GeneSys germplasm



### Phylogenetic overview

View the evolutionary history of protein-coding gene families.

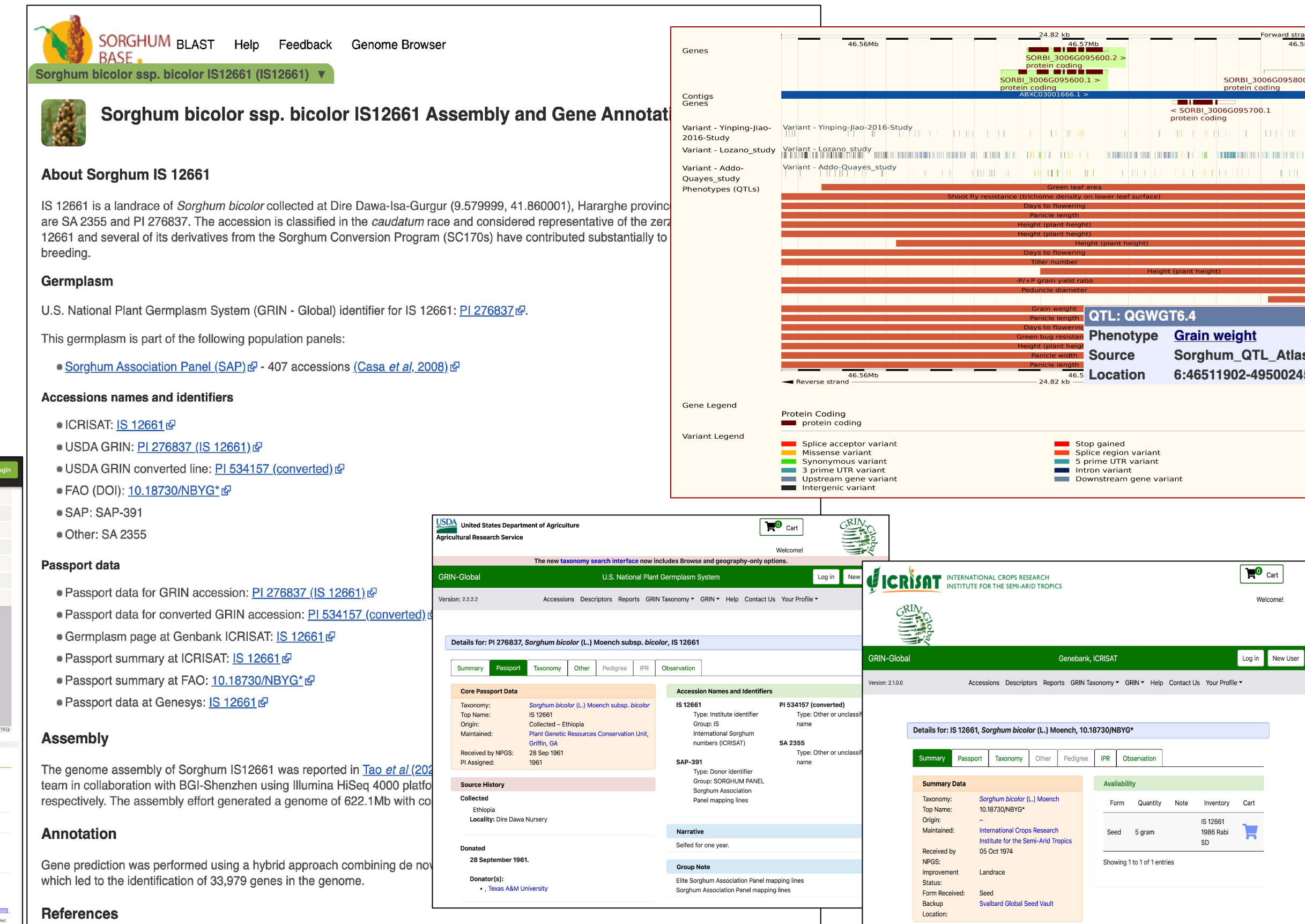
### Neighborhood conservation

Visualize gene copy variation among sorghum genomes and outgroup species in the context of a set of homologous genes.

Disease, Trait and Phenotype QTL annotations

### Genome pages

Access information on the germplasm, genome assembly, and gene annotation for each genome in related sites including GRIN, ICRISAT, FAO & Genesys.



## Community Engagement

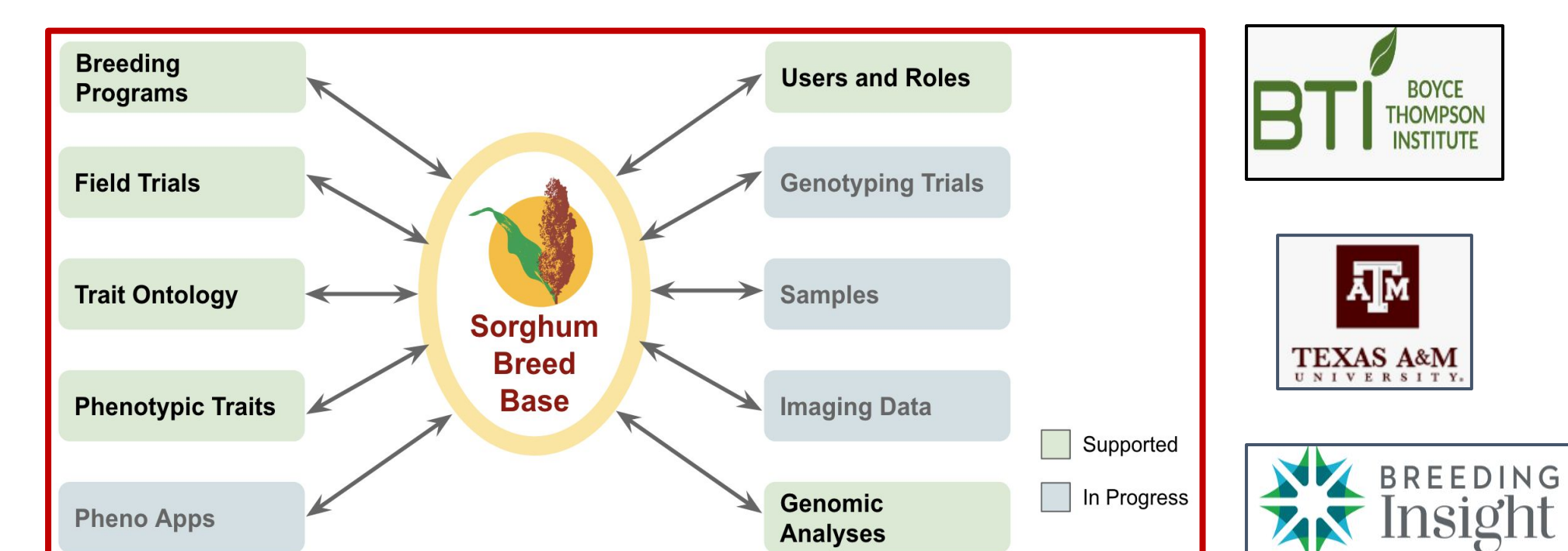
### Resources

- News, Meetings & Events
- Publications
- Research highlights
- User guides
- Webinars

### Communication

- Breeders
- Stakeholder meetings
- Slack channels
- Mailing list
- Contact form

### Collaboration with breeders, researchers and stakeholders



Planta (2022) 255:35  
<https://doi.org/10.1007/s00425-022-03821-6>

**SHORT COMMUNICATION**

**SorghumBase: a web-based portal for sorghum genetic information and community advancement**

Nicholas Gladman<sup>1</sup>, Andrew Olson<sup>1</sup>, Sharon Wei<sup>1</sup>, Kapeel Chougule<sup>1</sup>, Zhenyuan Lu<sup>1</sup>, Marcela Tello-Ruiz<sup>1</sup>, Ivar Meijs<sup>1</sup>, Peter Van Buren<sup>1</sup>, Yinying Jiao<sup>1</sup>, Bo Wang<sup>1</sup>, Vivek Kumar<sup>1</sup>, Sunita Kumari<sup>1</sup>, Lifang Zhang<sup>1</sup>, John Burke<sup>2</sup>, Junping Chen<sup>2</sup>, Gloria Burov<sup>2</sup>, Chad Hayes<sup>2</sup>, Yves Emendack<sup>2</sup>, Zhanguo Xin<sup>2</sup>, Doreen Ware<sup>1,4</sup>