BIO-ANALYTIC RESOURCE FOR PLANT BIOLOGY

Asher Pasha
Nicholas Provart

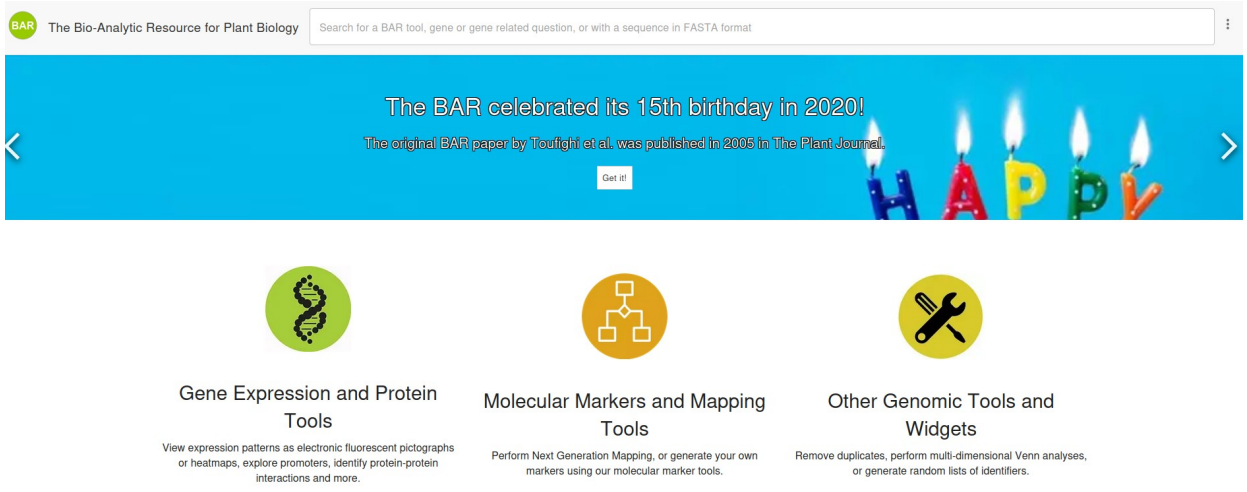
Bio-Analytic Resource, CAGEF
University of Toronto, Canada
Plant Biology 2023

August 6, 2023

@vegan_asher@genomic.social

@BAR_PlantBio@genomic.social

Bio-Analytic Resource



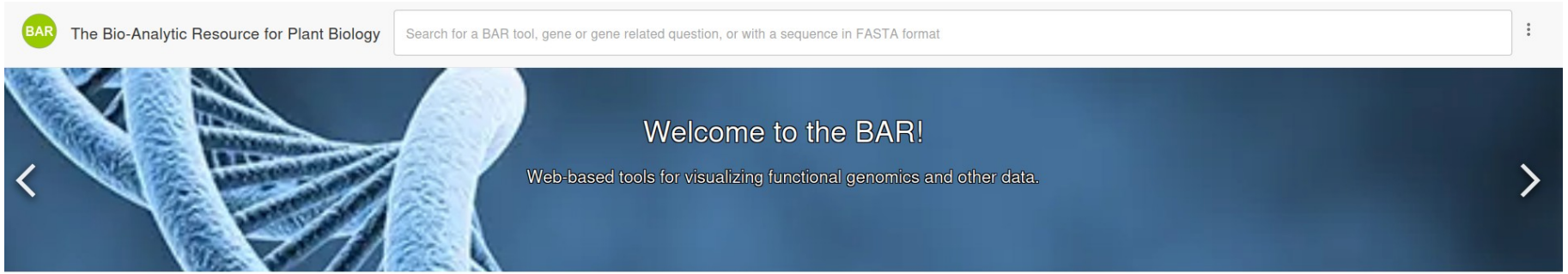
https://bar.utoronto.ca/

- 175M (100M *A. th.*)
- 11.7K (SUBA4)
- 80K predicted
  100+K experimental
  2.8M PDIs
- 29K Phyre2
  26k AlphaFold2
  885 from PDB
- 10+M nsSNPs (1135 ecotypes)

BAR and FAIR Principles

- Findable: BAR datasets can be easily searched on the BAR homepage and GAIA app.
- Accessible: BAR datasets do not require authentication or authorization and are publicly available.
- Interoperable: BAR and ThaleMine webservices provide data in JSON and XML formats that can be easily integrated into workflows.
- Reusable: BAR includes datasets that are publicly available and do not require a subscription. So the data can be reused in any external apps.

New BAR Homepage (with GAIA search input)



Gene Expression and Protein Tools

View expression patterns as electronic fluorescent pictographs or heatmaps, explore promoters, identify protein-protein interactions and more.

Molecular Markers and Mapping Tools

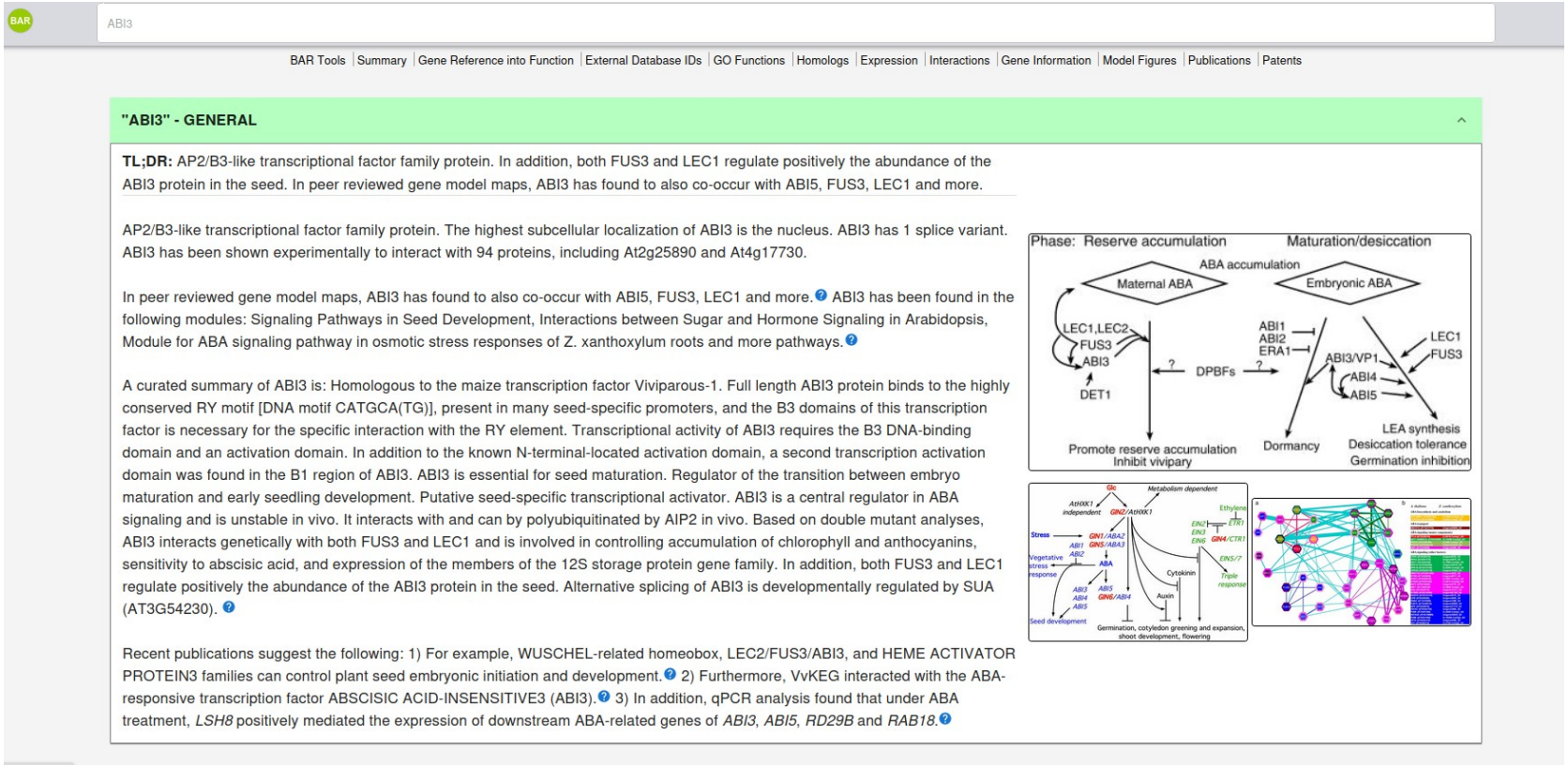
Perform Next Generation Mapping, or generate your own markers using our molecular marker tools.

Other Genomic Tools and Widgets

Remove duplicates, perform multi-dimensional Venn analyses, or generate random lists of identifiers.

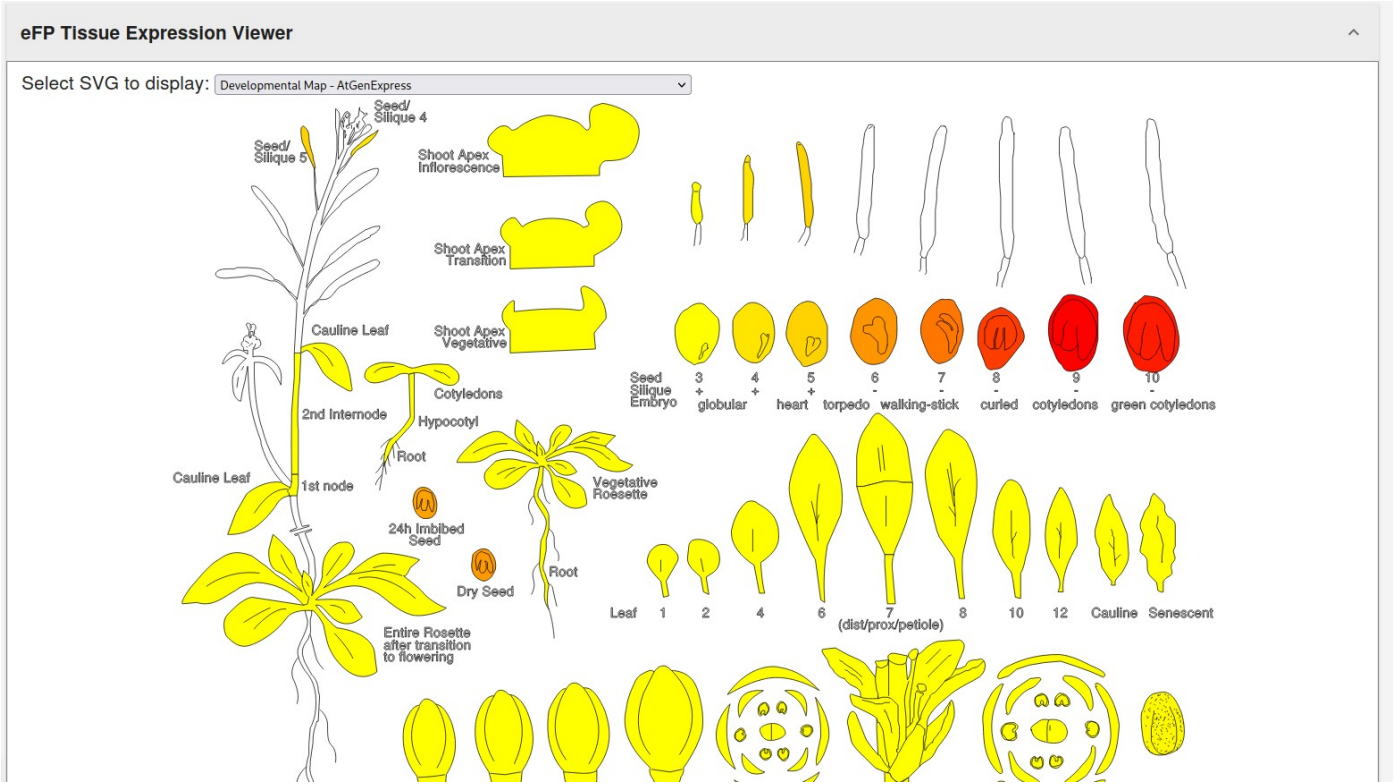
https://bar.utoronto.ca/

GAIA: Gene Search



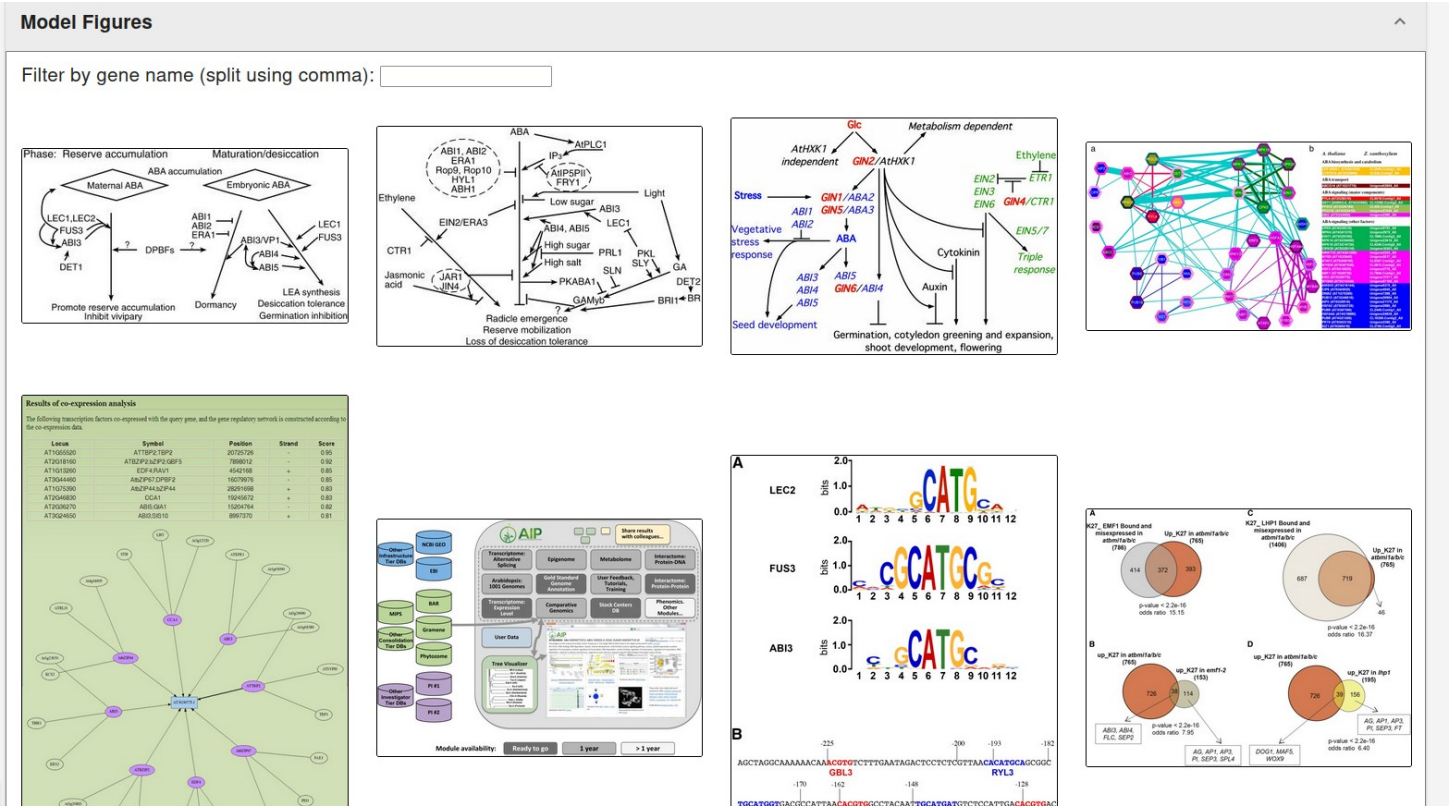
https://bar.utoronto.ca/gaia/ (Sullivan, *et al.,* in prep)

GAIA: Gene Search



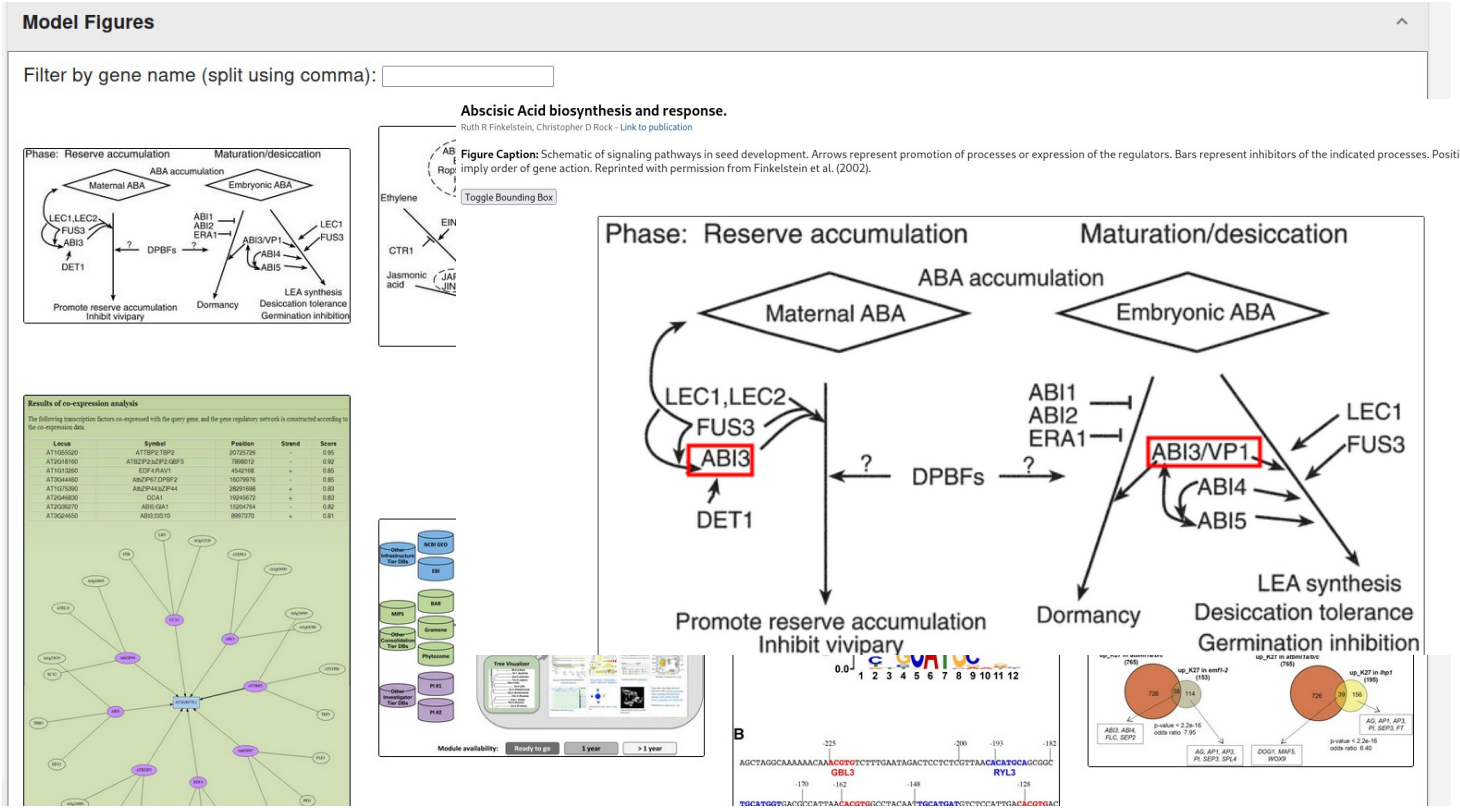
https://bar.utoronto.ca/gaia/ (Sullivan, *et al.,* in prep)

GAIA: Gene Search



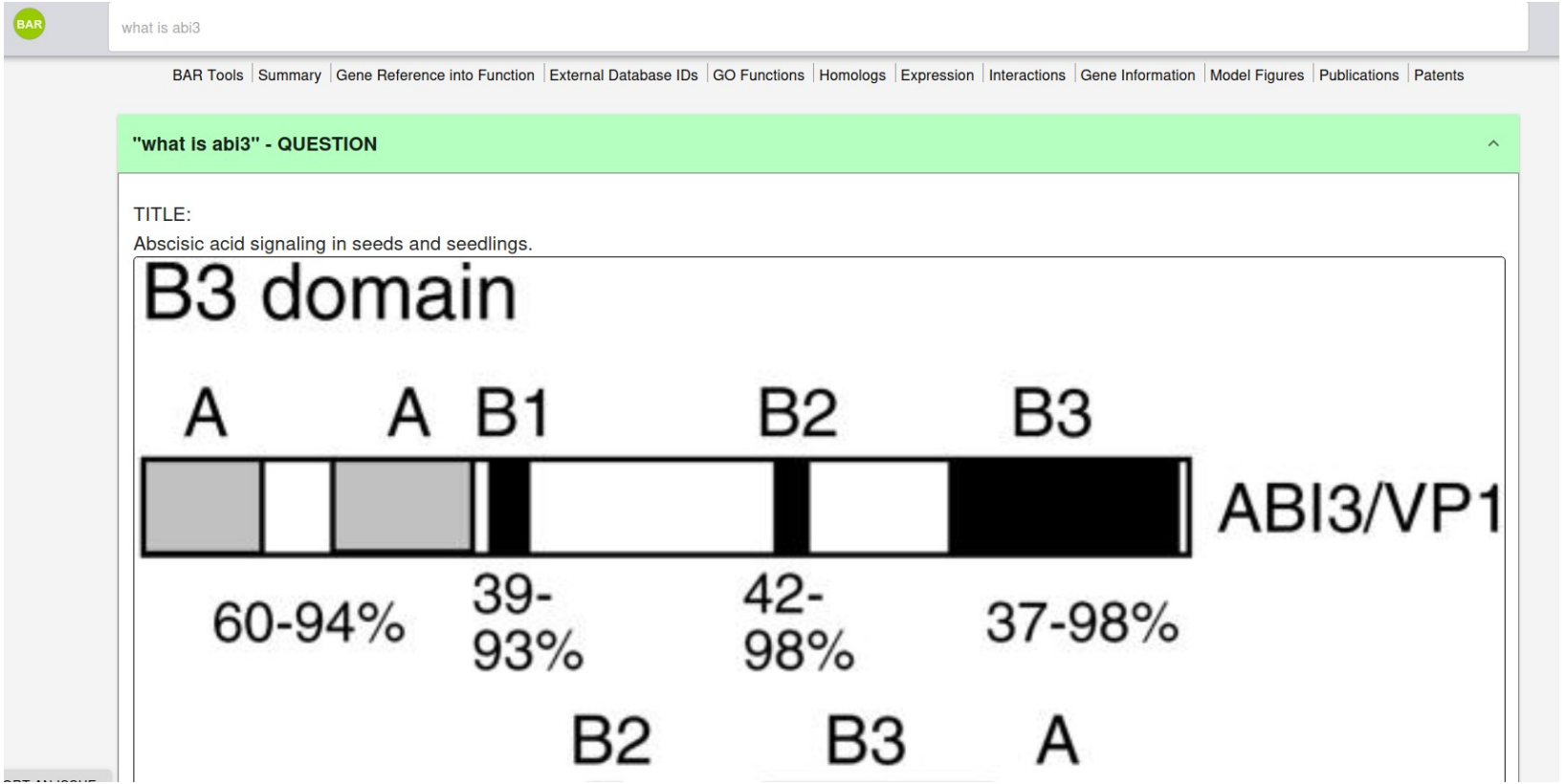
https://bar.utoronto.ca/gaia/ (Sullivan, *et al.,* in prep)

GAIA: Gene Search



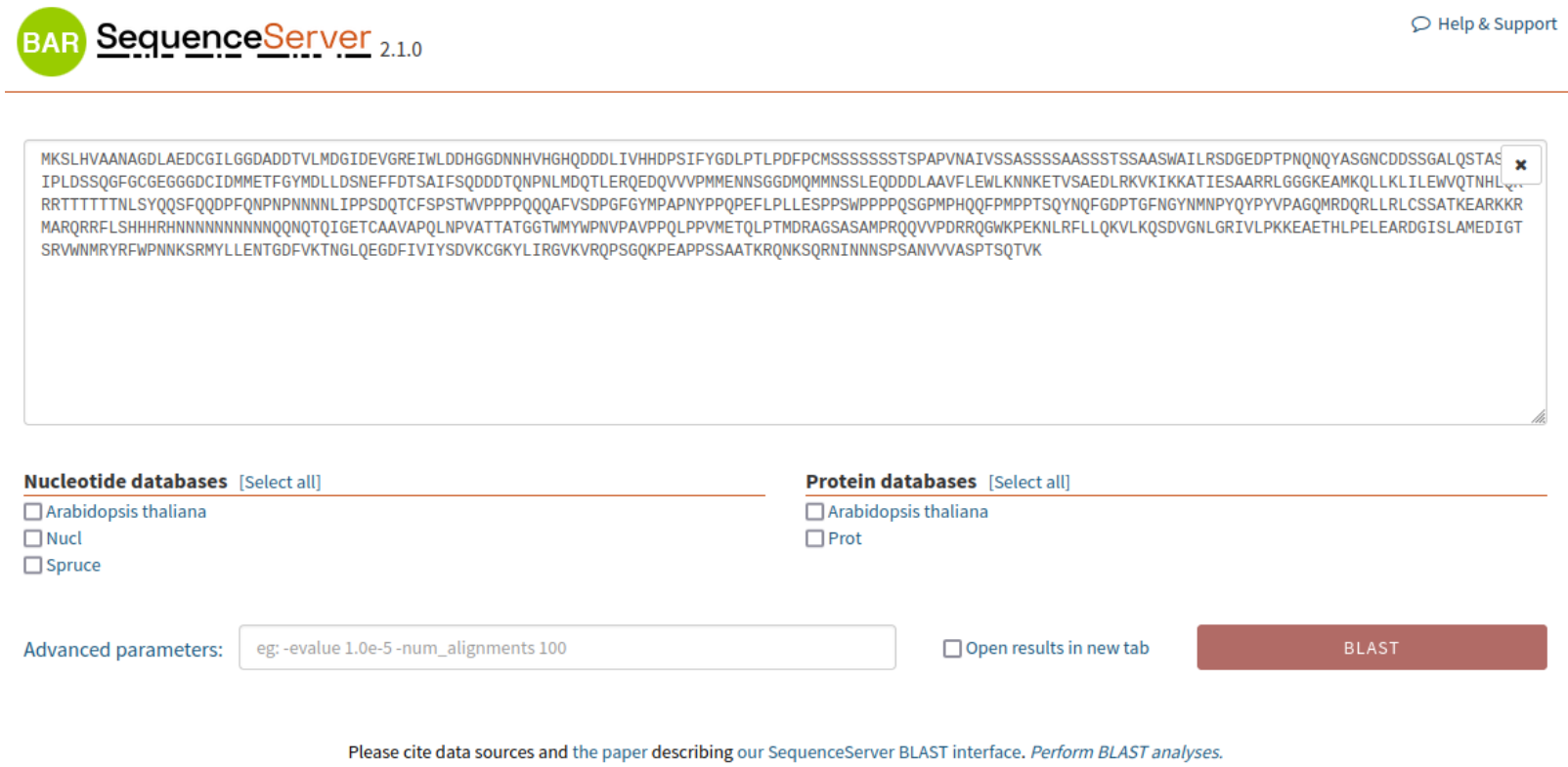
https://bar.utoronto.ca/gaia/ (Sullivan, *et al.,* in prep)

GAIA: Question Search



https://bar.utoronto.ca/gaia/ (Sullivan, *et al.,* in prep)

Homepage: BLAST Search



https://bar.utoronto.ca/blast/ (Sullivan, *et al.,* in prep)

Homepage: BLAST Search

BAR SequenceServer 2.1.0

Help & Support

BLASTP: 1 query, 2 databases
Edit search | New search

Download FASTA, XML, TSV
FASTA of all hits
FASTA of selected hit(s)
Alignment of all hits
Alignment of selected hit(s)
Standard tabular report
Full tabular report
Full XML report

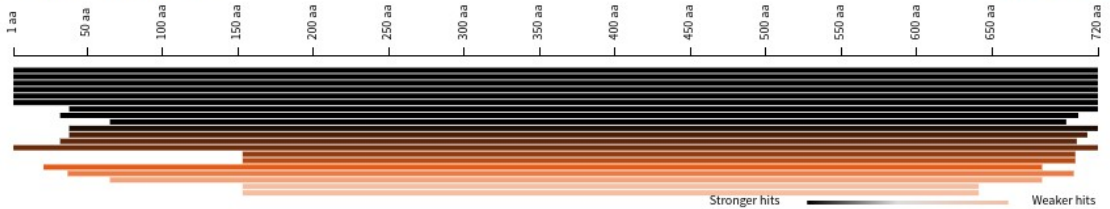
Share results
Share to cloud

SequenceServer 2.1.0 using **BLASTP 2.14.0+**, query submitted on 2023-08-03 21:39:45 UTC
Databases: Arabidopsis thaliana, Prot (985386 sequences, 364523406 characters)
Parameters: evalue 1e-05, matrix BLOSUM62, gap-open 11, gap-extend 1, filter F
Please cite: https://doi.org/10.1093/molbev/msz185

Queries and their top hits: chord diagram

Query= Query_1 length: 720

Graphical overview of hits SVG | PNG



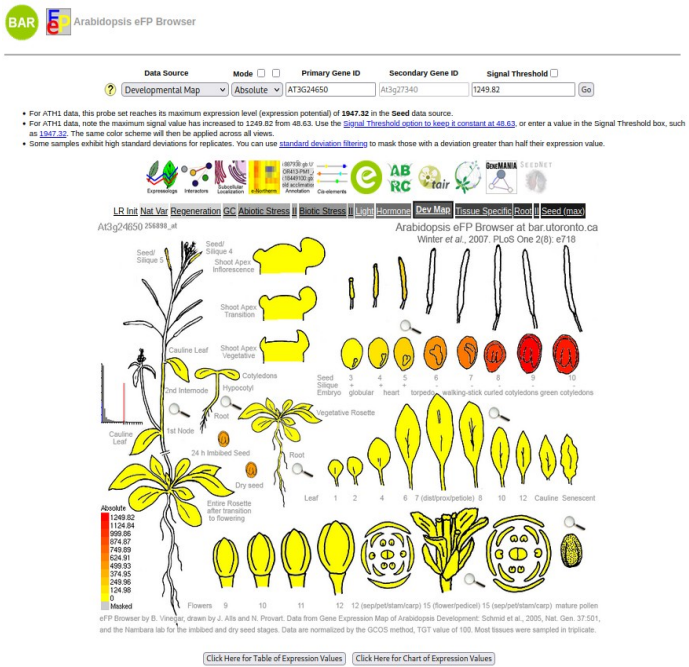
View More

Length distribution of matching sequences
Sequences producing significant alignments

#	Similar sequences	Query coverage (%)	Total score	E value	Identity (%)	BAR Tools
1.	AT3G24650.1 | Symbols: ABI3, AtABI3, SIS10 | SUGAR INSENSITIVE ...	100	3881	0	100%	ePlant eFP
2.	Csa15g050420.1	100	2897	0	90.2%	ePlant
3.	Csa01g030760.1	100	2832	0	89.4%	ePlant
4.	Csa01g030760.2	100	2832	0	89.4%	ePlant
5.	Csa19g036630.1	100	2806	0	89.2%	ePlant

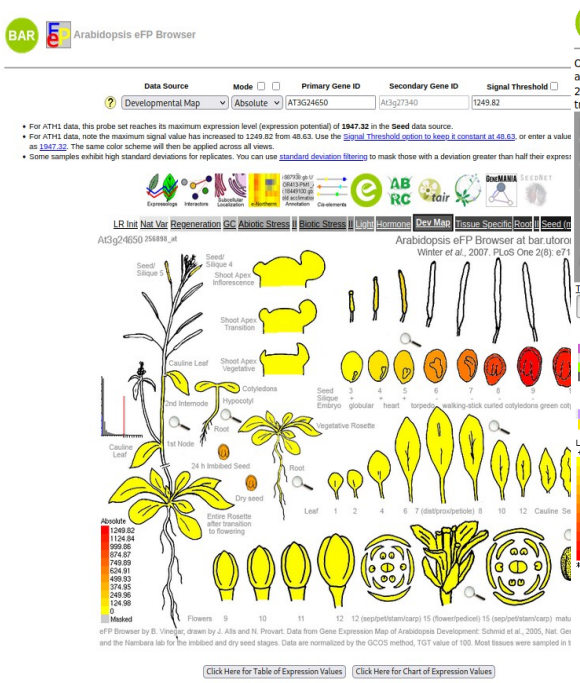
https://bar.utoronto.ca/blast/ (Sullivan, *et al.,* in prep)

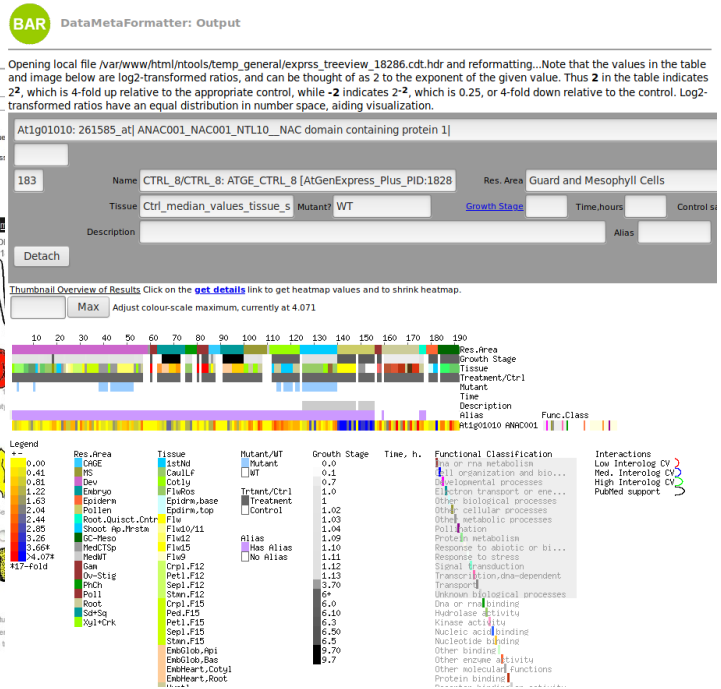
Bio-Analytic Resource



https://bar.utoronto.ca/

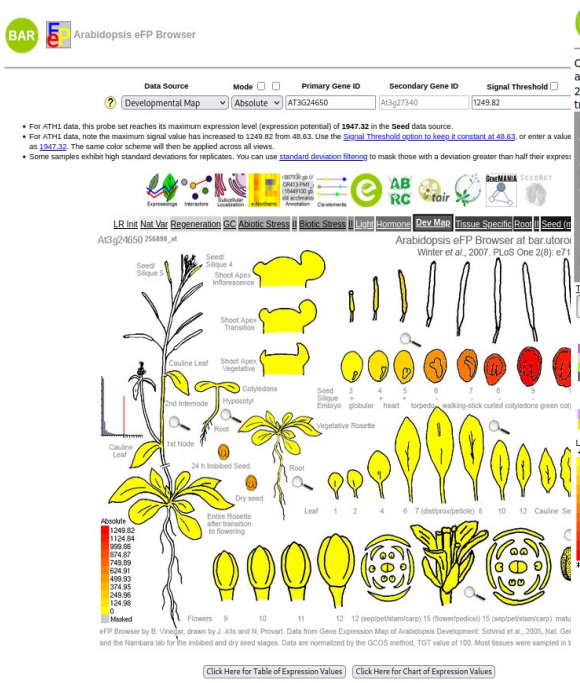
Bio-Analytic Resource

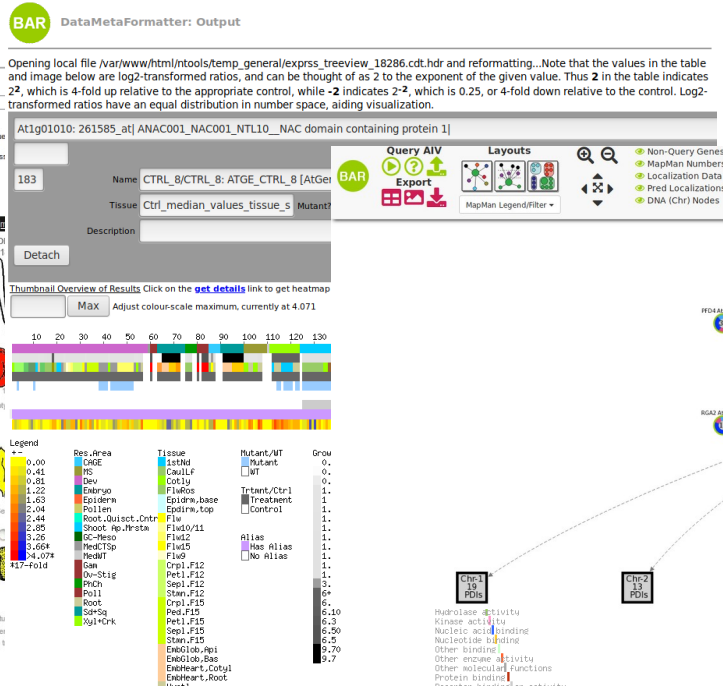


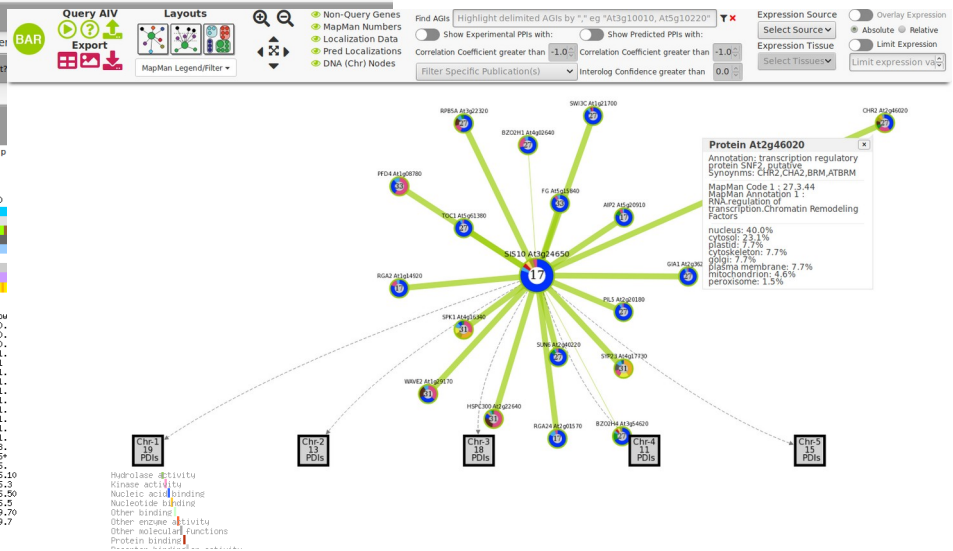


https://bar.utoronto.ca/

Bio-Analytic Resource

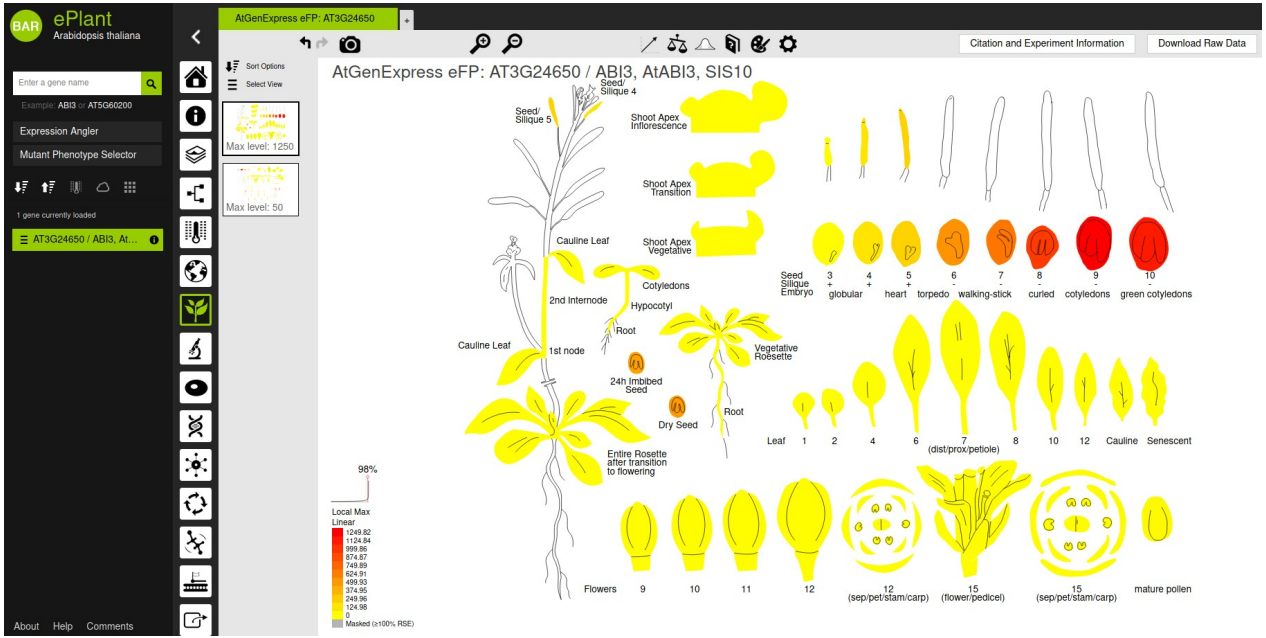






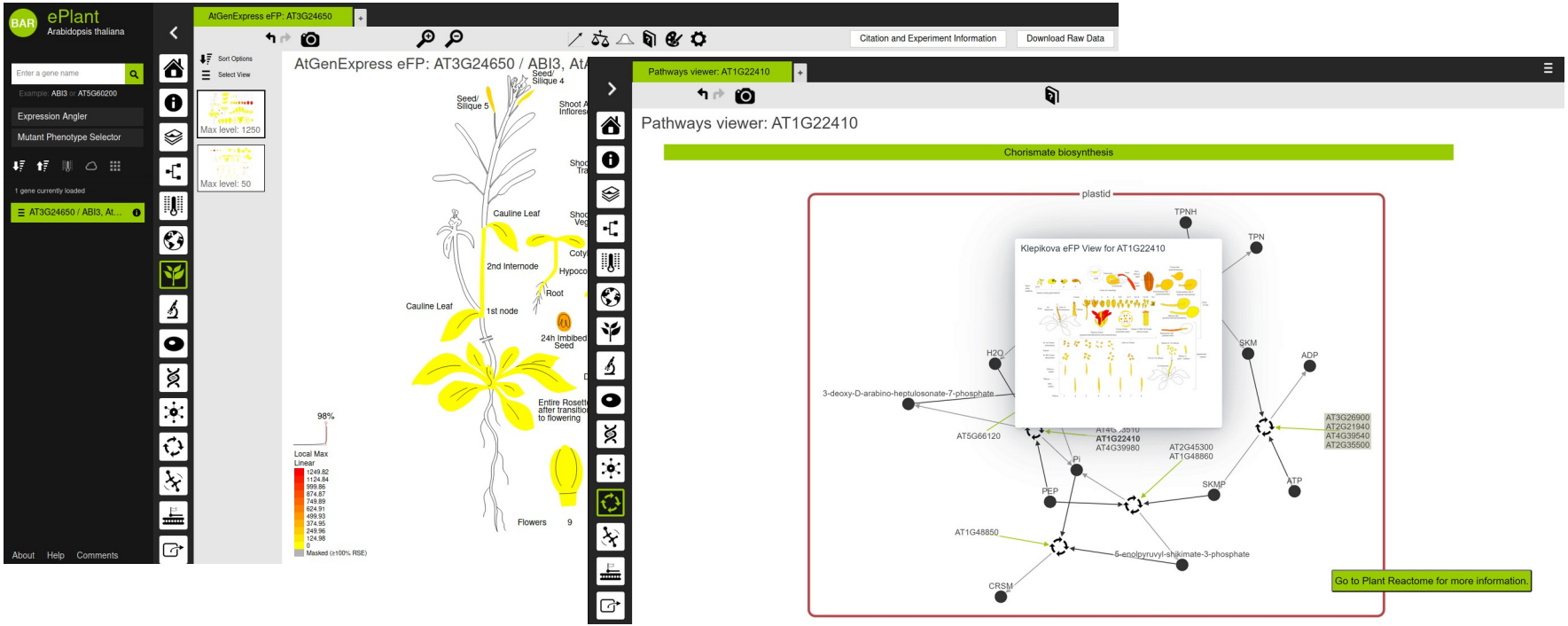
https://bar.utoronto.ca/

ePlant: Improved User Interface



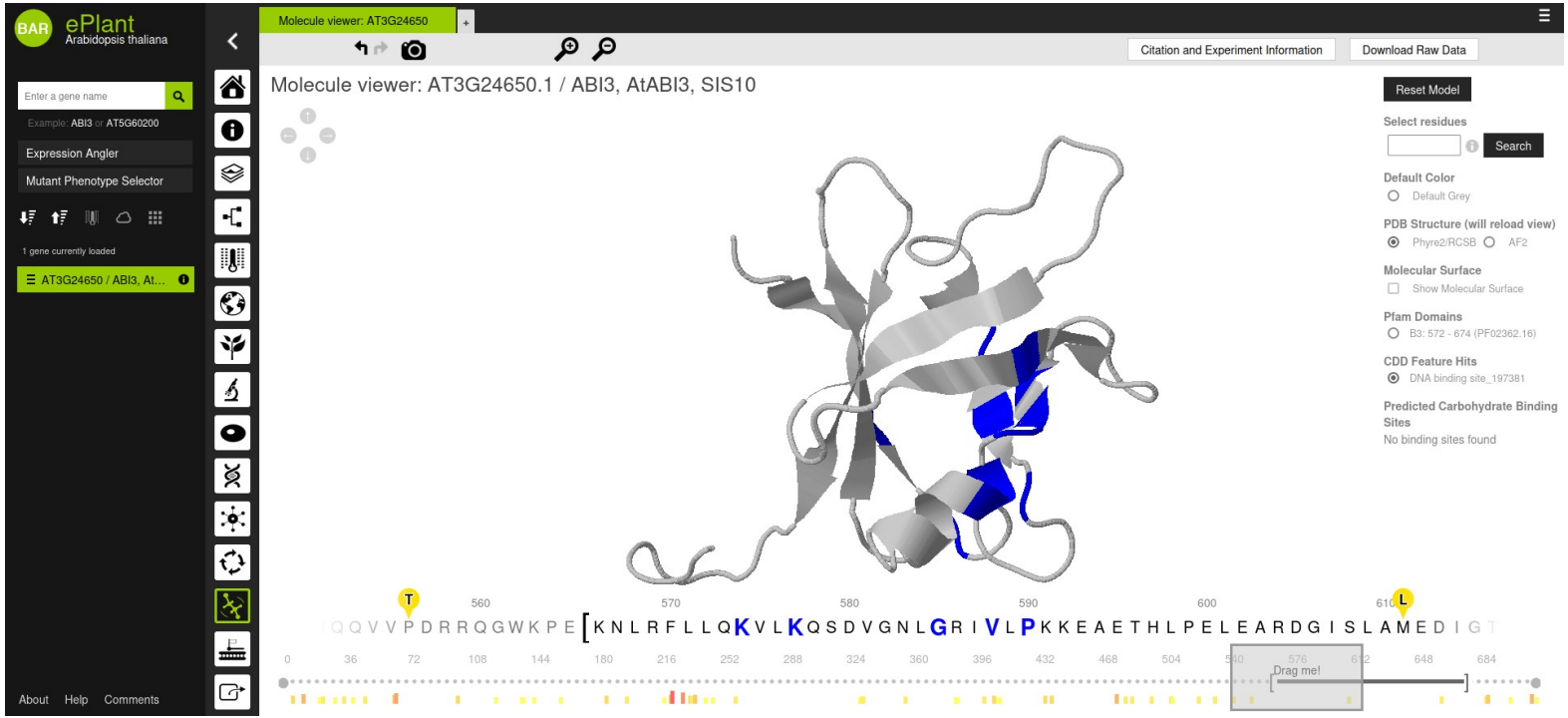
https://bar.utoronto.ca/eplant/ (Waese *et al., 2017, Waese-Perlman et al.* In prep.*)*
https://doi.org/10.1101/2021.04.28.441805

ePlant: Improved User Interface



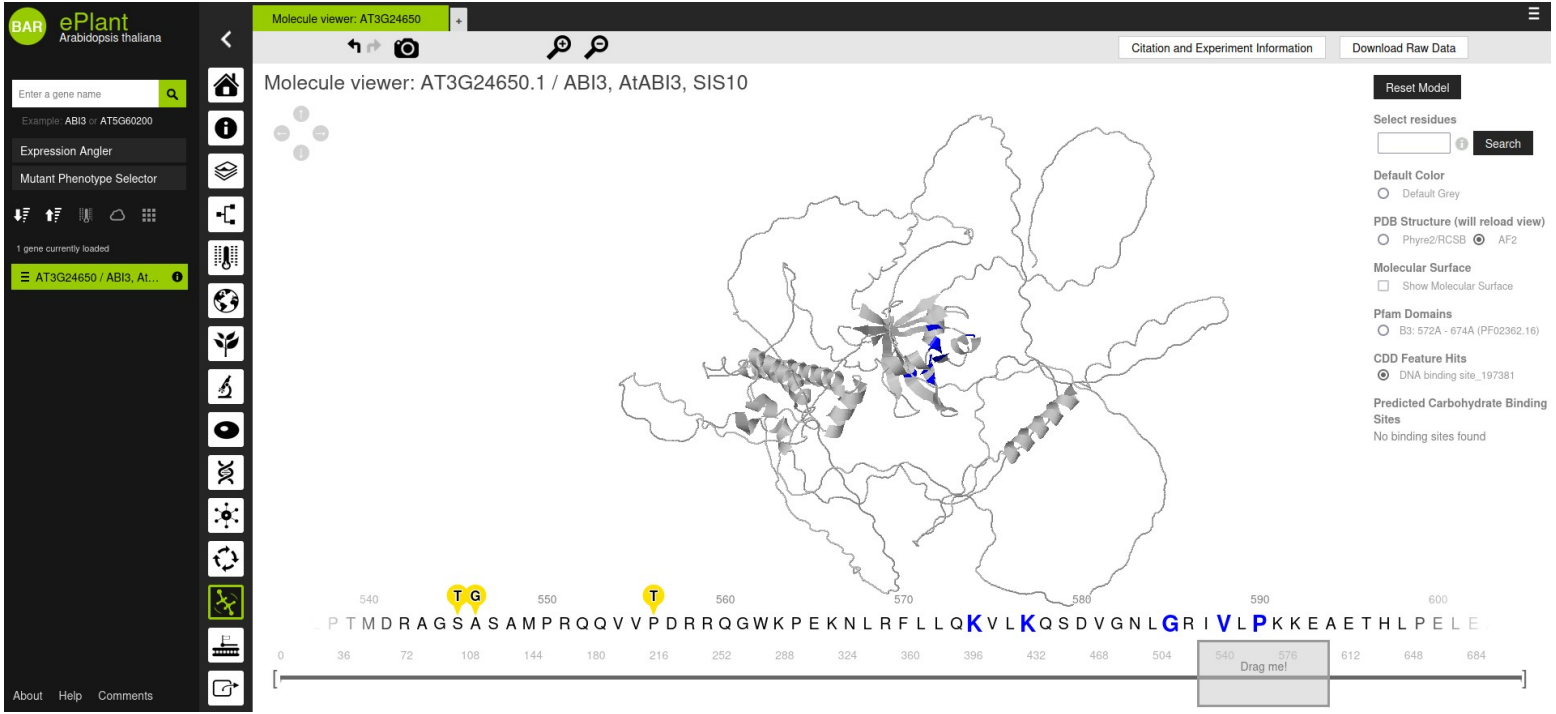
https://bar.utoronto.ca/eplant/ (Waese *et al., 2017, Waese-Perlman et al.* In prep.*)*
https://doi.org/10.1101/2021.04.28.441805

ePlant: Phyre2 and AF2 Structures



https://bar.utoronto.ca/eplant/ (Waese *et al., 2017, Waese-Perlman et al.* In prep.*)*
https://doi.org/10.1101/2021.04.28.441805

ePlant: Phyre2 and AF2 Structures



https://bar.utoronto.ca/eplant/ (Waese *et al., 2017, Waese-Perlman et al.* In prep.*)*
https://doi.org/10.1101/2021.04.28.441805

ePlants

ePlant

Info Go Pub

ePlant Maize

Info Go

ePlant Poplar

Info Go

ePlant Tomato

Info Go

ePlant Camelina

Info Go

ePlant Soybean

Info Go

ePlant Potato

Info Go

ePlant Barley V3

Info Go

ePlant Barley V1

Info Go

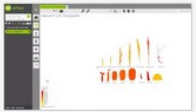
ePlant Medicago

Info Go

ePlant Eucalyptus

Info Go

ePlant Rice



Info Go

ePlant Willow

Info Go

ePlant Sunflower

Info Go

ePlant Cannabis

Info Go

ePlant Wheat

Info Go

ePlant Sugarcane

Info Go

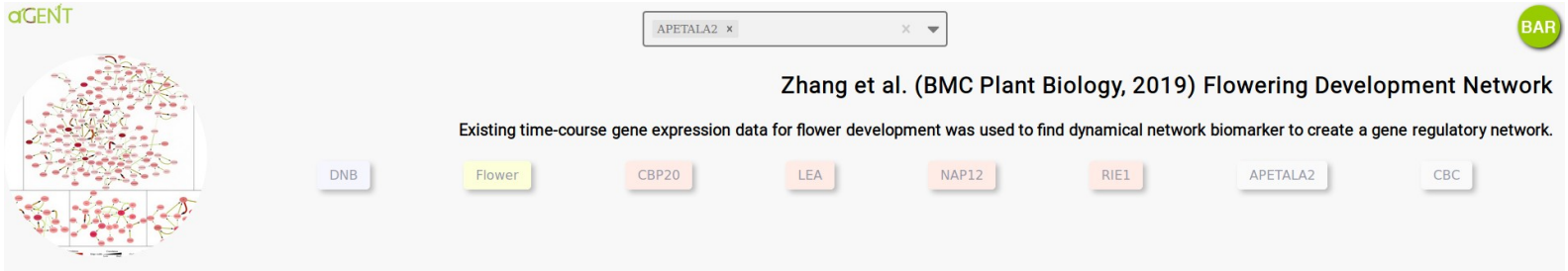
Legacy ePlant (PLOS ONE version)

Info Go Pub

AGENT

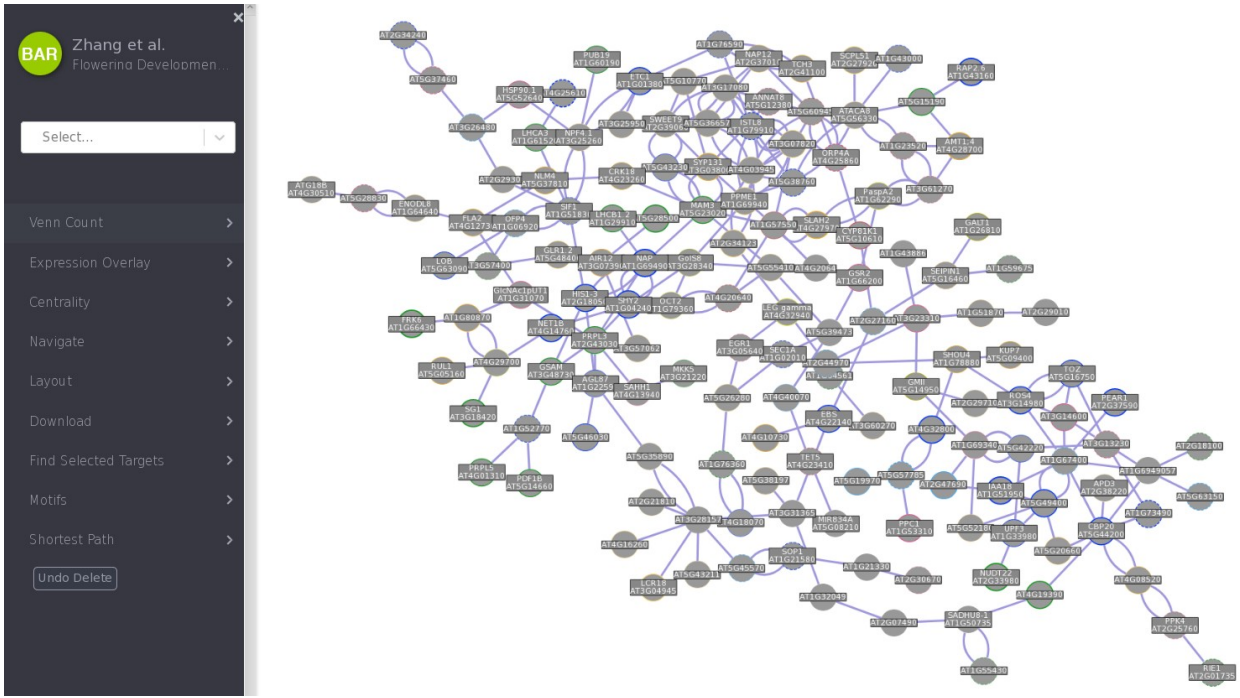
https://bar.utoronto.ca/AGENT (Lau *et al.*, 2021)
https://doi.org/10.1101/2021.04.28.441830

AGENT



https://bar.utoronto.ca/AGENT (Lau *et al.*, 2021)
https://doi.org/10.1101/2021.04.28.441830

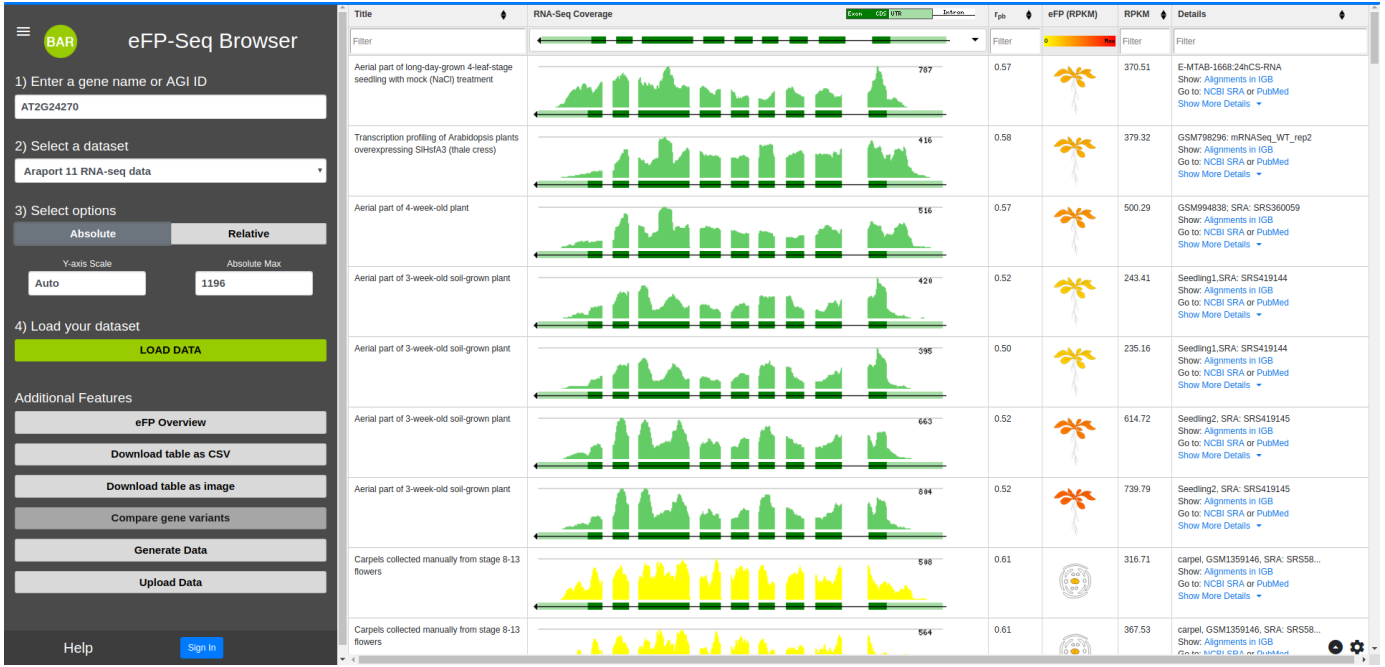
AGENT


BAR
Zhang et al.
Flowering Developmen...
Select...
Venn Count
Expression Overlay
Centrality
Navigate
Layout
Download
Find Selected Targets
Motifs
Shortest Path
Undo Delete

New Tool: Variant Viewer for SNPs

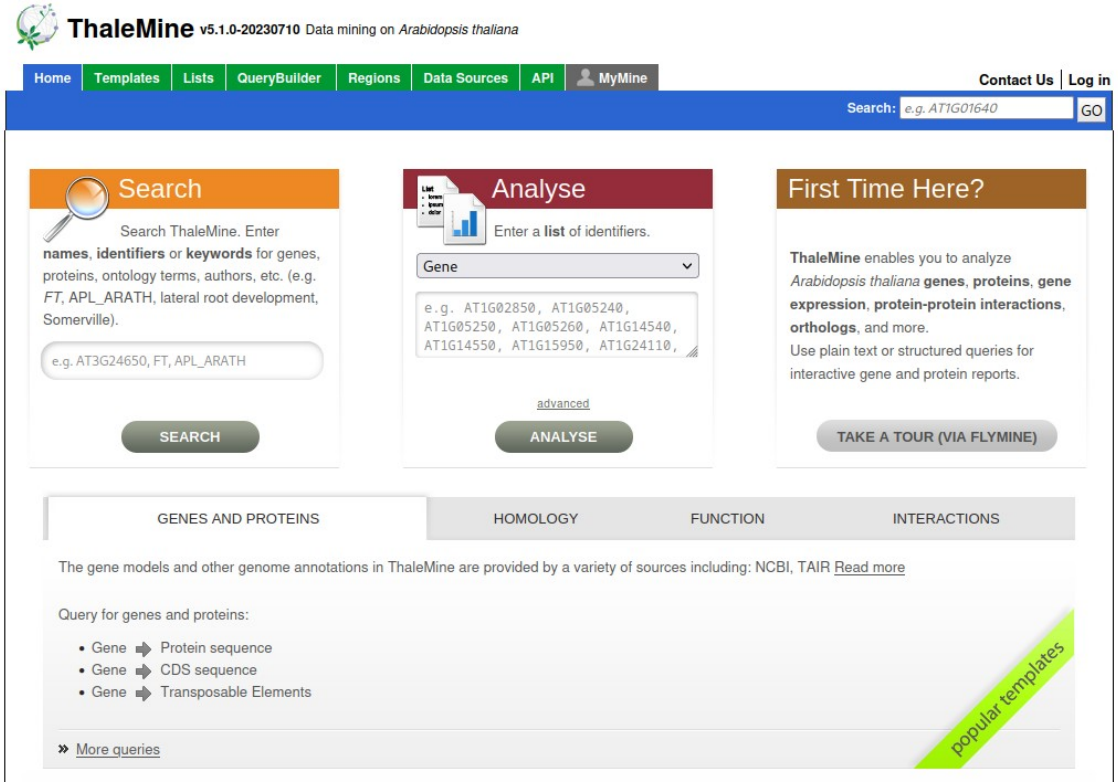
https://bar.utoronto.ca/VariantViewer/ (Cumming *et al.*, 2021)
https://doi.org/10.1101/2021.04.21.440793

eFP-Seq Browser



https://bar.utoronto.ca/eFP-Seq_Browser/ (Sullivan *et al*., 2019)

ThaleMine

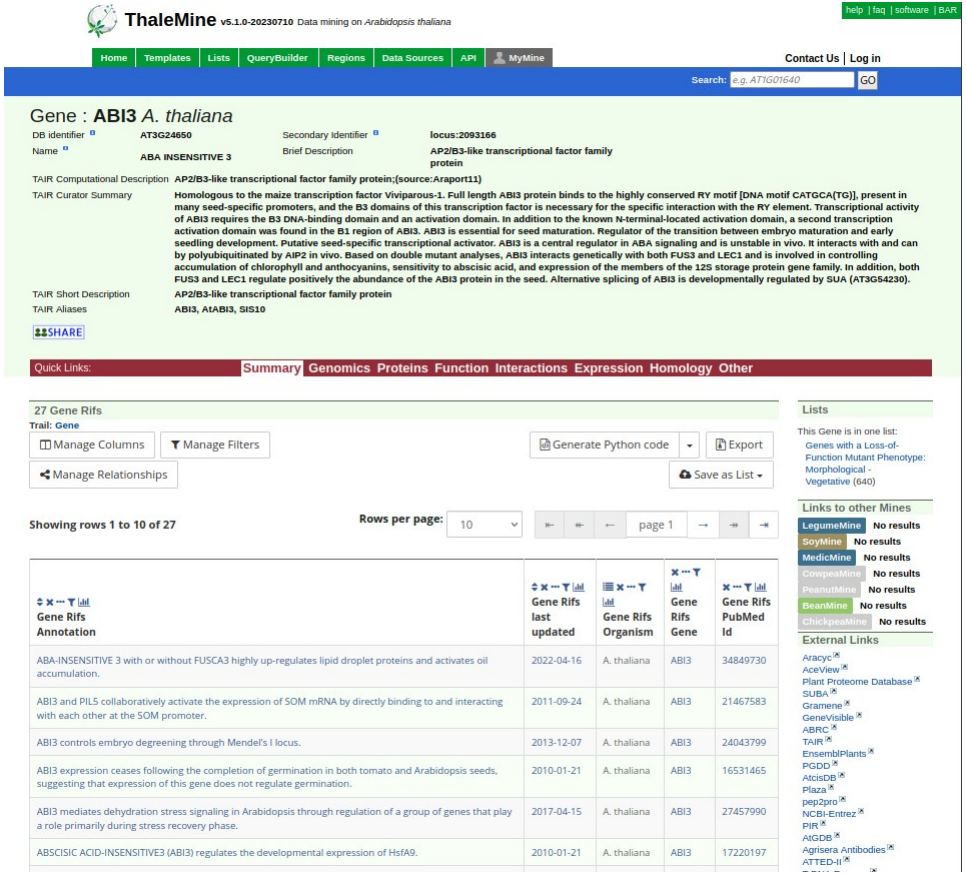


(Pasha *et al.*, 2020)

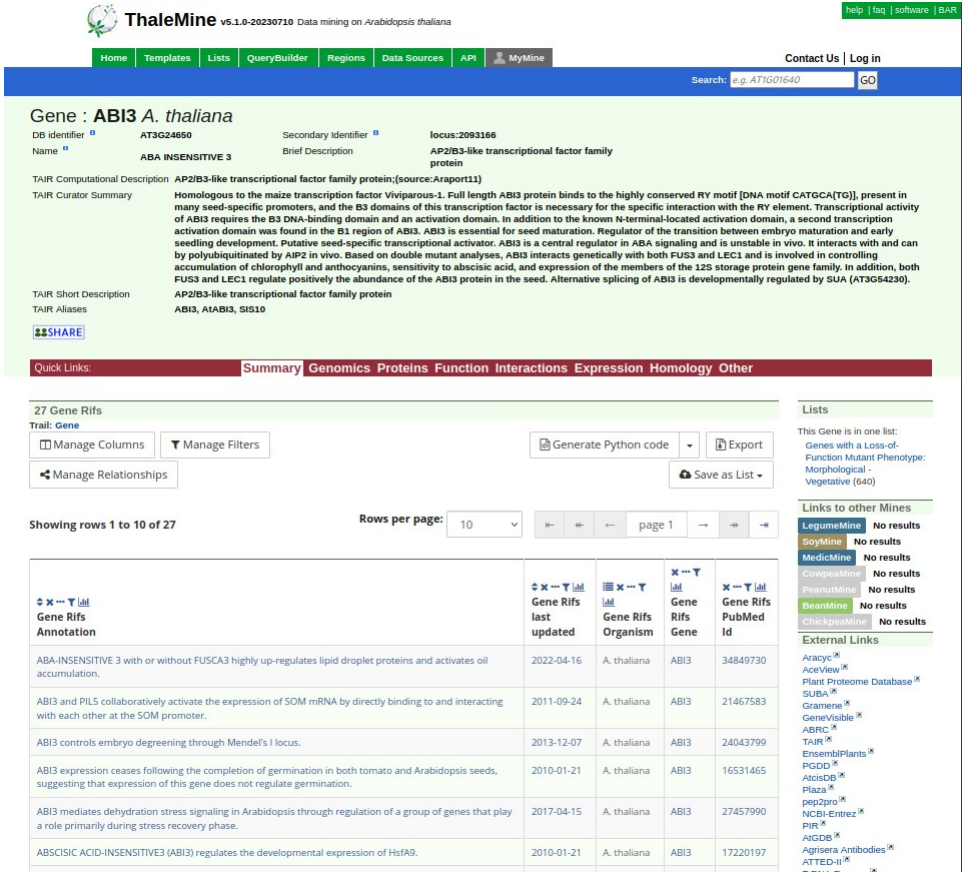
ThaleMine

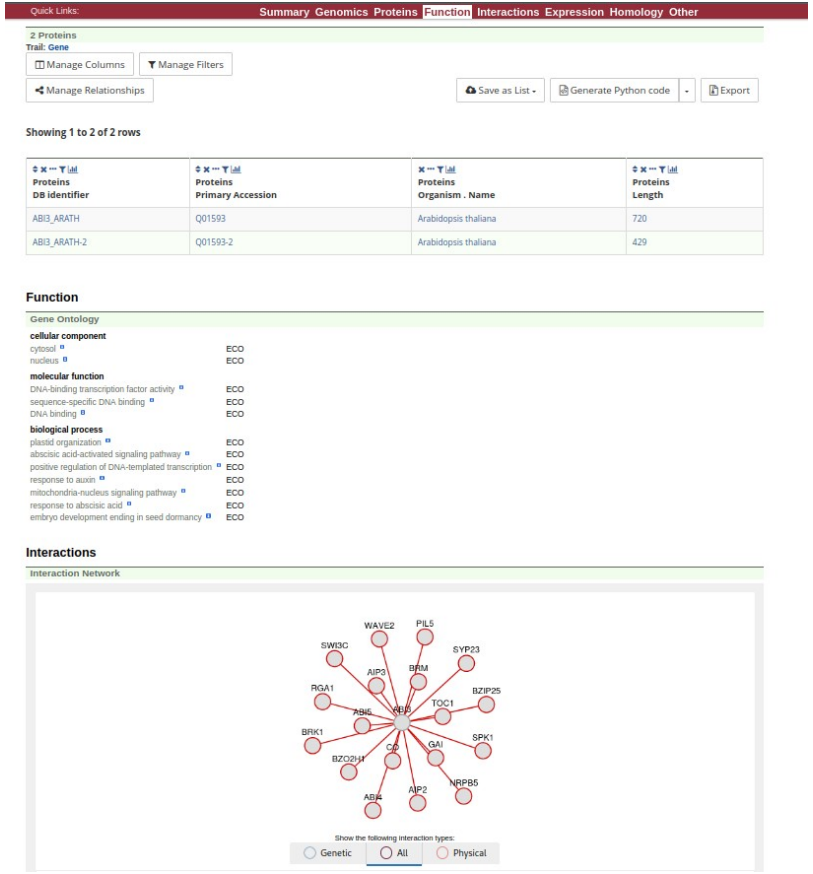
- ThaleMine v5.1.0-20230710 was release with InterMine 5.1.0 and data as of July 10, 2023.
- TAIR datasets upgraded to public release 20220630
- Panther dataset is upgraded to 17.0
- BioGrid data are upgraded to 4.4.233
- BAR eFP images are now loaded using a new Python 3 based web service with cached images.
- We continue to support InterMine staff.

ThaleMine Gene Search

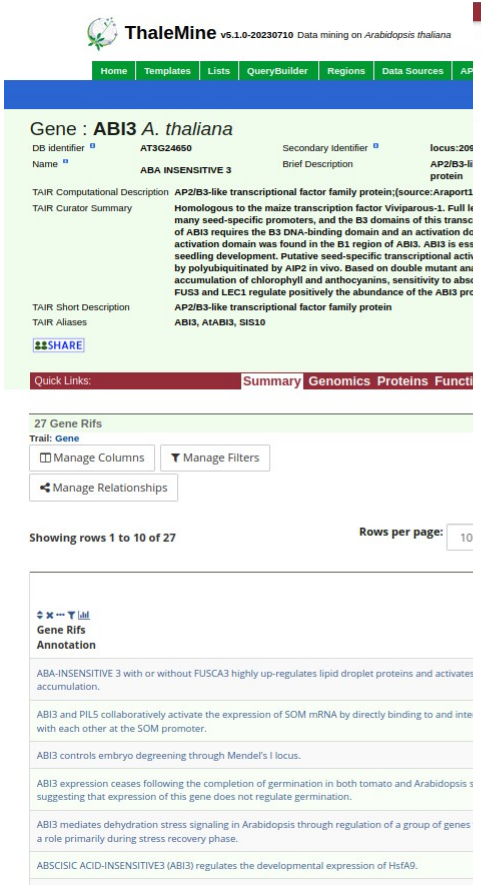


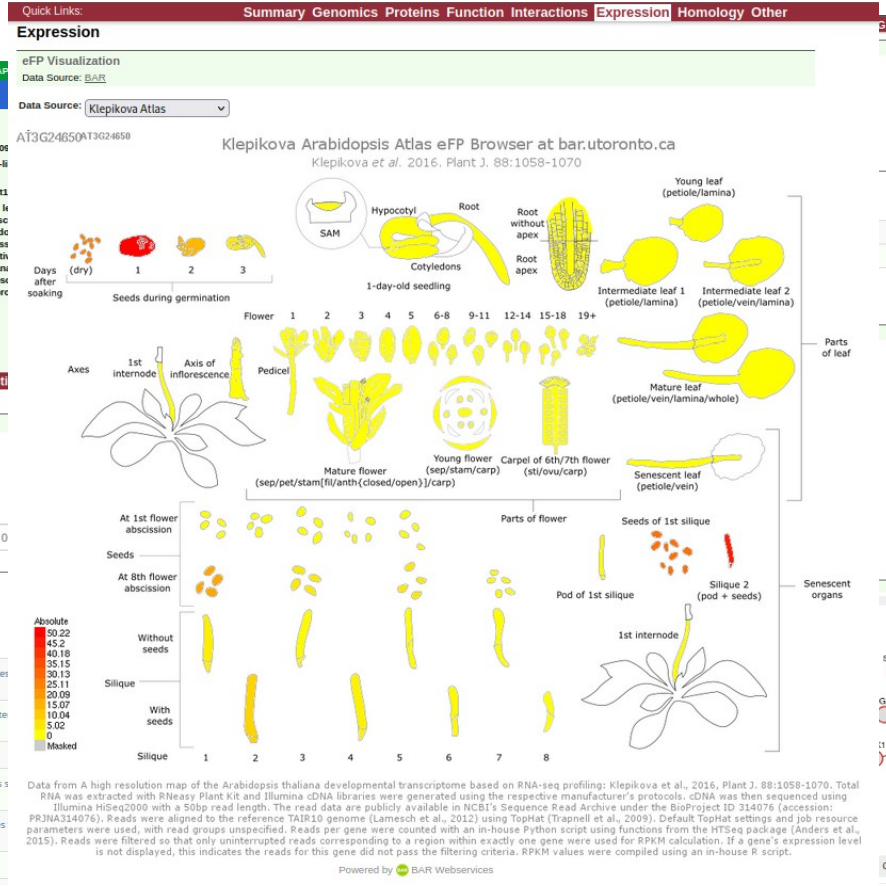
ThaleMine Gene Search

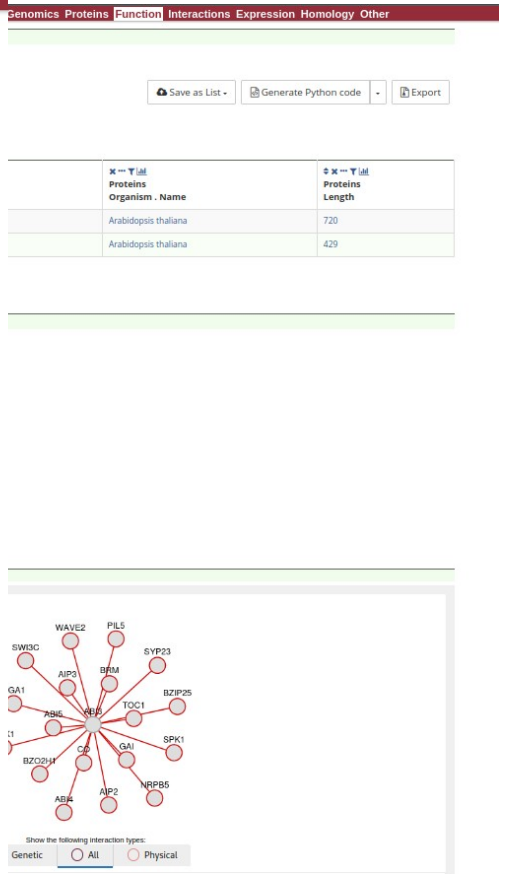




ThaleMine Gene Search







ThaleMine Data Sources

Data Category	Data	Source	PubMed
Genome	TAIR10 Genome assembly (5 chromosomes plus chloroplast and mitochondrial assemblies)	NCBI - Release TAIR10 (2018/04/06)	Arabidopsis Genome Initiative - PubMed: 11130711
	Araport11 GFF3 data from TAIR	TAIR - Release Araport11 (2016/06/17)	Cheng et al., 2016 - PubMed: 27862469
Proteins	High-quality, manually annotated, non-redundant protein sequence database.	Swiss-Prot - Release 2023_03	UniProt Consortium - PubMed: 17142230
	Computationally analysed records, enriched with automatic annotation	TrEMBL - Release 2023_03	
	Protein family and domain assignments to proteins	InterPro - Release v95.0	Mitchell et al., 2019 - PubMed: 30398656
Homology	Orthologue relationships based on the inferred speciation and gene duplication events in the phylogenetic tree.	Panther - Release 17.0	Mi et al - PubMed: 23193289
	Paralogue relationships based on the inferred speciation and gene duplication events in the phylogenetic tree.	Panther - Release 17.0	Mi et al - PubMed: 23193289
	Phytozome Homologs generated with InParanoid	Phytozome - realtime	Goodstein et al - PubMed: 22110026
Curation	Manually curated TAIR functional descriptions	TAIR - Release 20220630	Huala et al - PubMed: 11125061
	Manually curated TAIR gene aliases	TAIR - Release 20220630	Huala et al - PubMed: 11125061
Gene Ontology	GO annotations from Gene Ontology	Gene Ontology - Release 2023-06-11	Berardini et al., 2004 - PubMed: 15173566 Gene Ontology Consortium - PubMed:10802651
	Several electronic and manual GO annotation methods utilized by UniProt	UniProt - Release 2023_03	UniProt Consortium - PubMed: 17142230
Interactions	Curated set of genetic and physical interactions for Arabidopsis thaliana	BioGRID - Release 4.4.233	Chatr-Aryamontri et al., 2014 - PubMed: 25428363
	Curated binary and complex protein-protein interactions for Arabidopsis thaliana	IntAct - Downloaded 20230706	Kerrien et al., 2012 - PubMed: 22121220
Expression	Electronic Fluorescent Pictograph (eFP) Visualization paints gene expression information from one of the AtGenExpress data sets or other compendia for a desired gene onto a diagrammatic representation of *Arabidopsis thaliana* plants.	BAR eFP Webservice - realtime	Winter et al., 2007 - PubMed: 17684564 Brady et al., 2009 - PubMed: 19401381
Co-Expression	Co-regulated gene relationships deduced from microarray and RNA-seq data via ATTED-II web services	ATTED-II Co-expression - realtime	Obayashi et al., 2014 - PubMed: 24334350
Publications	Curated associations between publications and genes from UniProt	UniProt - Release 2023_03	UniProt Consortium - PubMed: 17142230
	Publications from InterPro	InterPro - Release v95.0	Mitchell et al., 2019 - PubMed: 30398656
	Publications from NCBI	NCBI - Downloaded 20230706	Maglott et al., 2007 - PubMed: 17148475
GeneRIF	Concise phrase describing gene function and publication associated with NCBI Gene records	NCBI - Downloaded 20230706	Maglott et al., 2007 - PubMed: 17148475

Acknowledgments

Kiana Toufighi
Ryan Austin
Siobhan Brady
Nina Wang
Hardeep Nahal
Joe Wu
Matthew Ierullo
Ron Ammar
Vincent Lau
Matt Cumming
Michael Dong
Matthew Ierullo
Anna van Weringh
David Di Biase
Geoff Fucile
Jamie Waese
Bruno Pereira
Ben Waese-Perlman
Alex Waese-Perlman

Ben Vinegar
Debbie Winter
Richard Song
Anjali Silva
Rohan Patel
Sakshi Shinghal
Rachel Woo
Ida Liu
Wen Kai Cao
Bruno Pereira
YueXin Yu
Hans Yu
Jim Fan
Fausto de Oliveira
Eddi Esteban
Guilherme Hosaka
Sylva Donaldson
Isaiah Hazelwood

NSERC CRSNG
GenomeCanada
OntarioGenomicsInstitute
UNIVERSITY OF TORONTO
Canada Foundation for Innovation
Fondation canadienne pour l'innovation
CAGEF University of Toronto Centre for the Analysis of Genome Evolution & Function

@BAR_PlantBio

@BAR_PlantBio@genomic.social

https://www.facebook.com/BioAnalyticResource

http://www.BAR.utoronto.ca