活用 TAIR 於植物研究：阿拉伯芥與作物的實用工具
Using TAIR for Plant Research: Practical Tools for Arabidopsis and Crops

tair

The Arabidopsis Information Resource



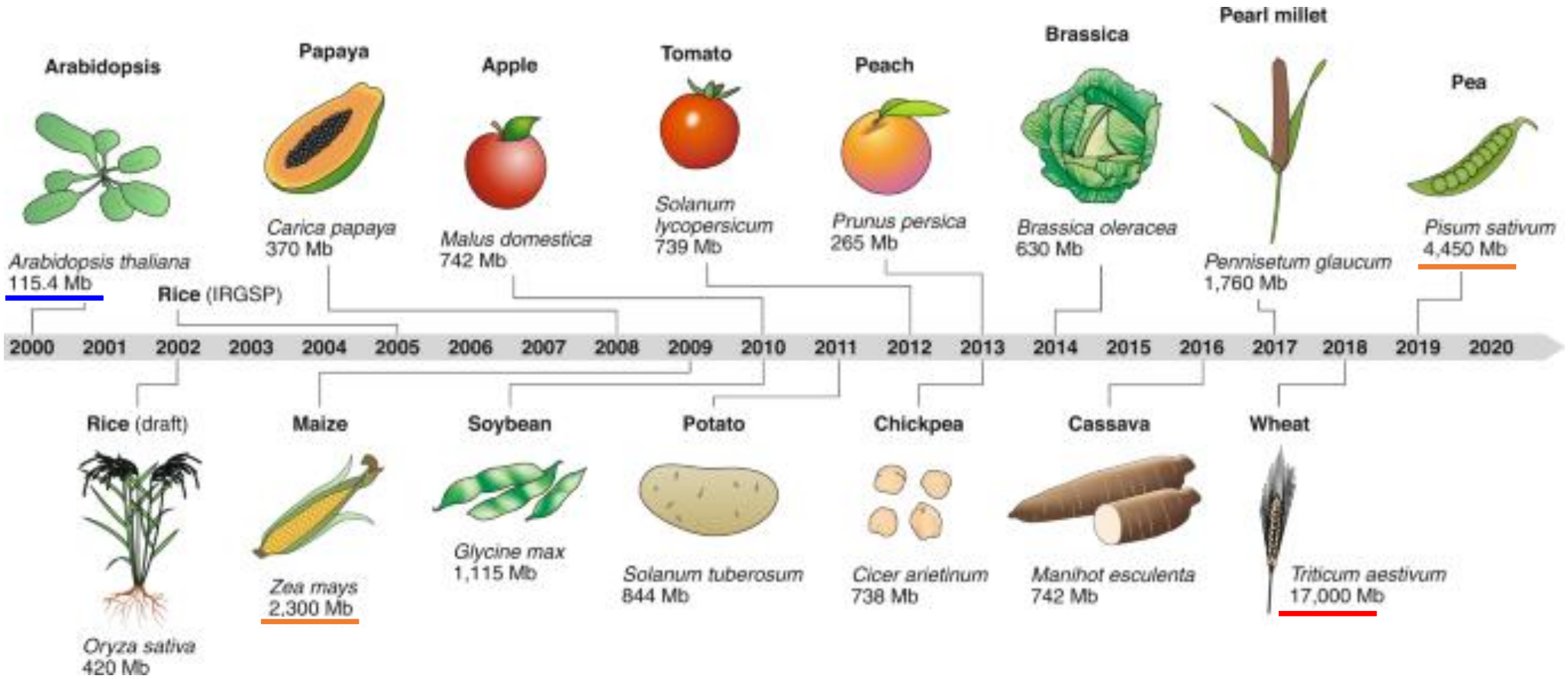
Ho-Ming Chen (陳荷明)

- Associate fellow, Agricultural Biotechnology Research Center (ABRC), Academia Sinica
- 25 years of experience in TAIR and bioinformatics
  - Data: Microarray, Small RNA, RNA degradome, RNA-Seq, Ribo-seq
  - Species: **Arabidopsis (model diploid)**, **banana (crop; triploid)**, **Orchid (crop; tetraploid**), rice, tomato, etc.

Why TAIR database?

- You can do fewer experiments, take fewer wrong turns, and graduate sooner
- You can do biology research from a different perspective, surpass your supervisor, and become a **next-generation** plant researcher.

Timeline of plant genome release

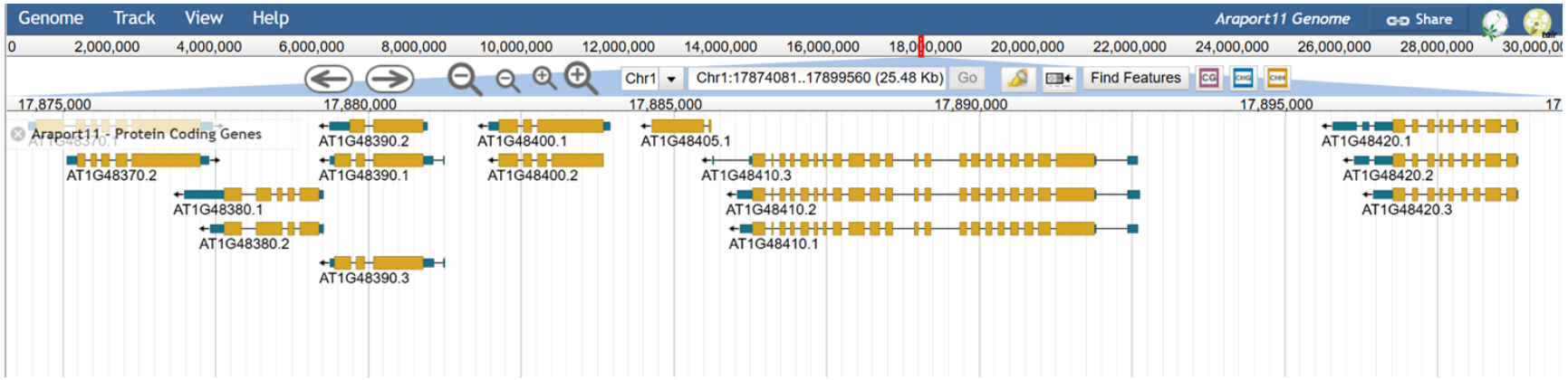


https://www.nature.com/articles/s41588-021-00866-3

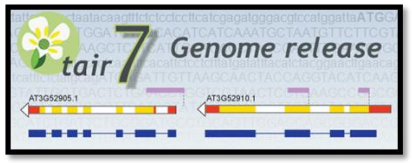
Arabidopsis genome annotation

- High quality sequence with few gaps
- High quality of genome annotation (frequent update)

Genome annotation: The process of identifying the locations of genes and all of the coding regions in a genome and determining what those genes do.











TAIR10

Araport11

TAIR annotations among versions

All Genome Versions Statistics

	Protein Coding Genes	Transposons and pseudogenes	Alternatively spliced genes	Gene density (Kb/gene)	Avg. exons per gene	Avg. exon length	Avg. intron length
TAIR10 (11/10)	27,411	4,827	5,885	4.35	5.89	296	165
TAIR9 (6/09)	27,379	4,827	4,626	4.35	5.67	304	165
TAIR8 (4/08)	27,235	4,759	4,330	4.37	5.62	306	165
TAIR7 (4/07)	26,819	3,889	3,866	4.44	5.79	268	165
TAIR6 (11/05)	26,541	3,818	3,159	4.48	5.64	269	164
TIGR5 (1/04)	26,207	3,786	2,330	4.54	5.42	276	164
TIGR4 (4/03)	27,170	2,218	1,267	4.38	5.31	279	166
TIGR3 (8/02)	27,117	1,967	162	4.32	5.24	266	166
TIGR2 (1/02)	26,156	1,305	28	4.48	5.25	265	167
TIGR1 (8/01)	25,554	1,274	0	4.55	5.23	256	168
Nature (12/00)	25,498	NA	NA	4.50	5.20	250	168

Type	TAIR10	Araport11	Change
(A) Protein-coding genes			
Total number of loci	27 416	27 655[a]	+239
Number of transcript isoforms	35 386	48 359	+12 973
Number of loci with two or more splice variants	5804	10 696	+4892
Number of loci with changes in CDS	–	–	+1158
Number of loci with changes in UTR(s)	–	–	+21 298
Upstream open reading frame (uORF)	58	84	+26

Cheng et al., (2017) Plant Journal

Feb 13, 2026.



https://phoenixbioinfo.org/announcing-the-publication-of-tair12-consult-the-fully-reannotated-genome-on-the-european-nucleotide-archive/

the core features and updates of TAIR 12 include:

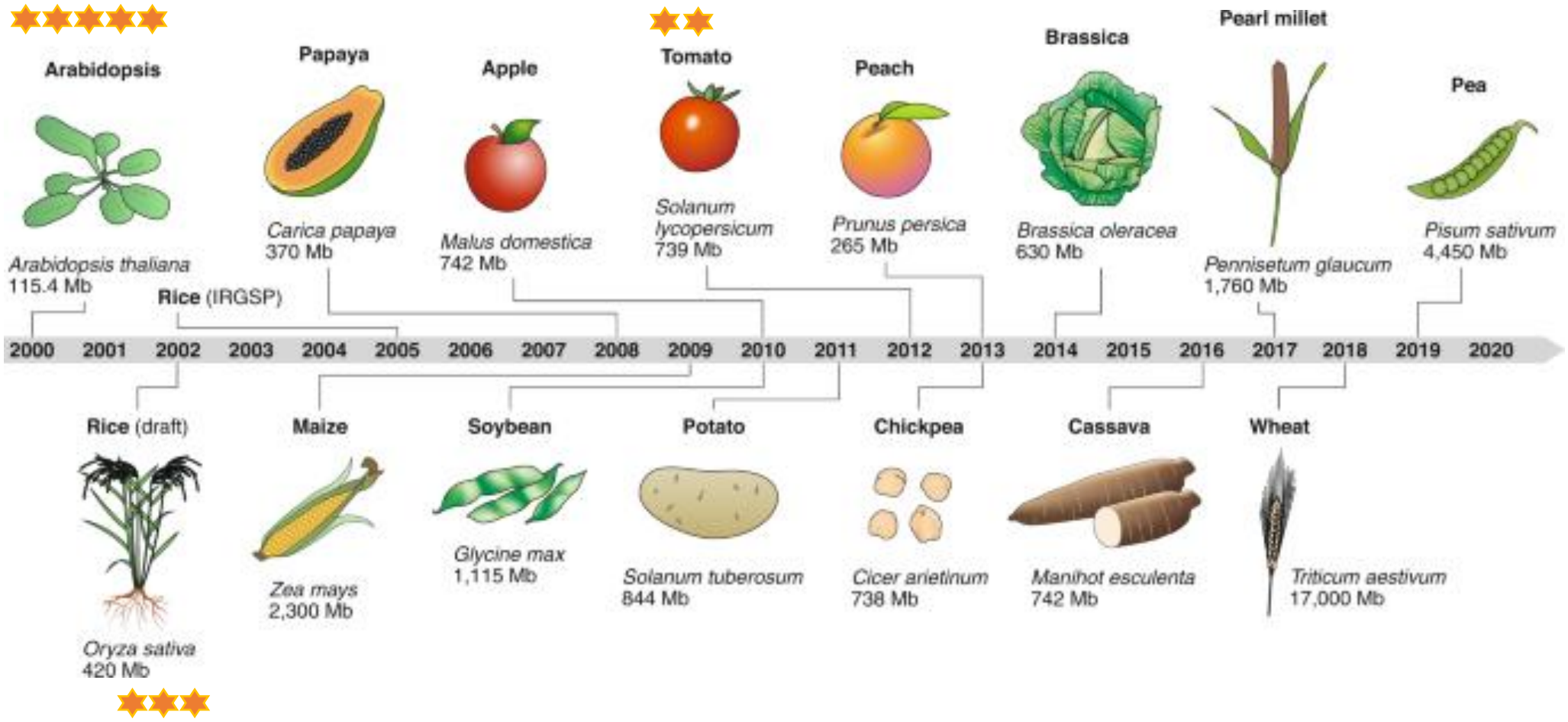
- More Accurate Genome Assembly: near telomere-to-telomere (T2T) reference genome.
- High-Resolution Transcriptome Refinement: long non-coding RNAs (lncRNAs).
- TAIR JBrowse2 genome browser.

Feature type	Araport11	TAIR12
Protein-coding genes	27,562	26,867
mRNAs	48,265	36,573
Pseudogenes (transcribed/other)	926/3917	2,140
Repeat regions (# of repeats)		
centromeric_repeat	not annotated	65,718
satellite_DNA	not annotated	11,123
telomeric_repeat	not annotated	4,218
Non-coding RNA genes		
antisense RNA	92	2,651
lncRNA	3480	2,008
miRNA	428	387
RNase MRP RNA	not annotated	2
rRNA	4	7,528
snoRNA	287	558
snoRNA primary transcript	not annotated	63
snRNA	82	61
snRNA primary transcript	not annotated	7
SRP RNA	not annotated	1
telomerase RNA	not annotated	1
tRNA	625	526
misc_RNA	1826	44
nc_RNA*	not annotated	37

https://phoenixbioinfo.org/announcing-the-publication-of-tair12-consult-the-fully-reannotated-genome-on-the-european-nucleotide-archive/

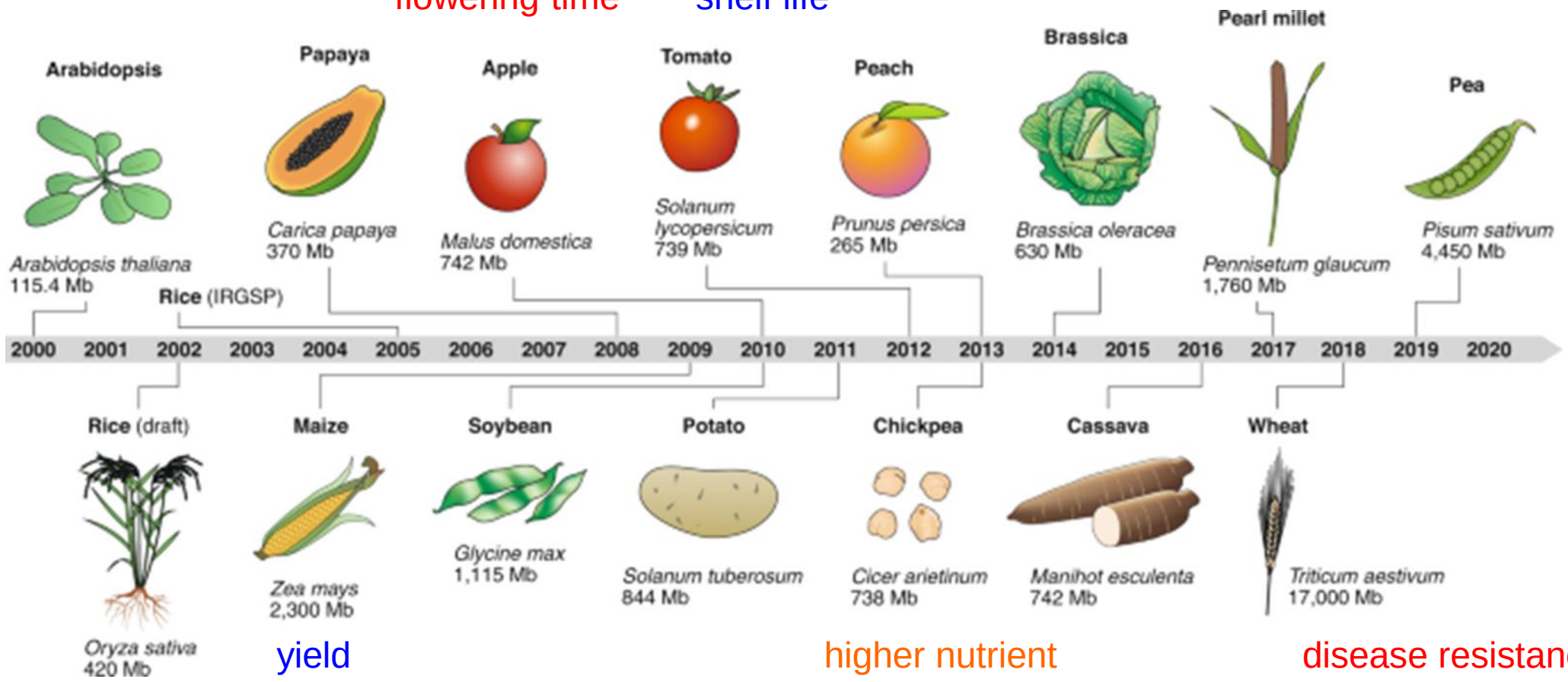
Quality of plant genomes

Genome assemblies and annotations



https://www.nature.com/articles/s41588-021-00866-3

Translational agricultural research



Finding target genes in crops

- Early flowering

ChatGPT

I would like to use genome editing to create an early flowering orange line. Please suggest two targets to edit.

Thought for 12s

Nice — good goal. Below are **two high-confidence gene targets** you can consider editing in orange (Citrus) to promote earlier flowering, with concise rationale, likely phenotypes, and key caveats. I'll keep this at a conceptual, non-actionable level (no editing protocols or sequence/guide design).

2) A citrus TFL1 / CEN homolog (e.g., *CsTFL1* / TFL1-like gene) — *loss-of-function / repression*

Why: TFL1 (and CEN/TFL1-like genes) antagonize FT and maintain meristem indeterminacy; loss of TFL1 function typically releases meristems to flower earlier. TFL1 homologs in woody perennials (including citrus) have been identified and linked to flowering control. PMC +1

Expected outcome: earlier and/or more frequent flowering due to reduced repression of floral meristem identity. Combining reduced TFL1 activity with moderate FT activation can be synergistic. Oxford Academic

Major trade-offs/risks: loss of meristem indeterminacy can change tree form, branching pattern, and possibly long-term productivity. Complete loss-of-function can be more disruptive than controlled down-regulation.

ChatGPT can make mistakes. Check important info.

Read review papers and check TAIR database

TAIR homepage

subscribed

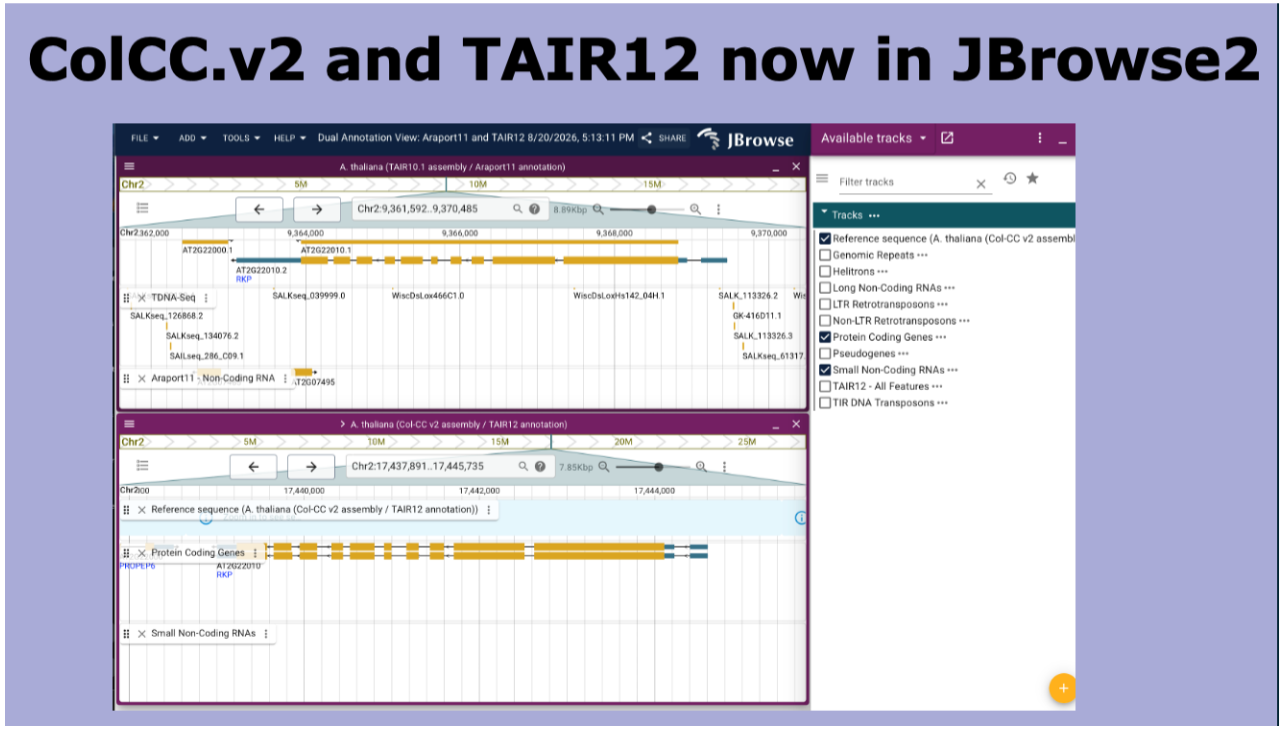
Home | Help | Contact | About Us | Subscribe | Login | Register *Institution: Academia Sinica (subscribed)*

Enter search text Gene

Gene
Clone
Keyword
Protein
Polymorphisms
Ecotype
Person
Germplasm
Marker
Vector
Transposon

Advanced Search | Browse | Tools | Portals | Download | Submit

The Arabidopsis Information Resource



48th p
TAIR@
[Oct 1,

48th p
curate
subscr
model. Files contain new publications, annotations, gene symbols and other data through September 30, 2025.

ICAR2027 Session Proposals due 10/31
[Sep 30, 2026]

The community can submit ICAR 2027 session proposals by October 31.

TAIR subscription

Sinica has institutional subscription

Pay as you go

Package	Amount	Cost
Demo	50 usage units per year *For users who access TAIR 1-2 times a year*	Complimentary when you link your ORCiD in TAIR
Introductory, 60% off	1st bundle of 300 usage units in a 12-month period	$206
Standard	Each subsequent bundle of 300 usage units in a 12-month period *For weekly TAIR users*	$515
Super	Bundle of 900 usage units *For daily TAIR users*	$1,160

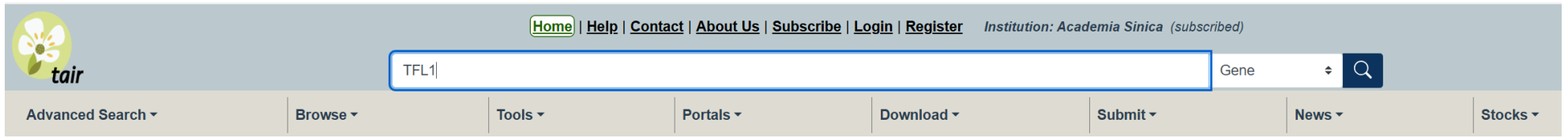
Institutional subscriptions for nonprofits, governments, & universities

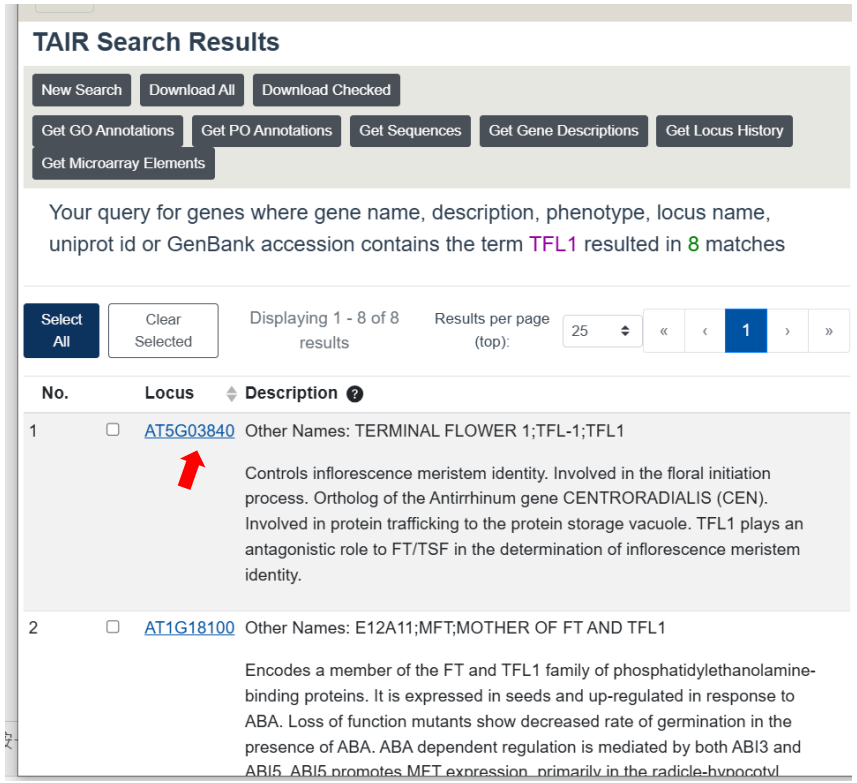
Institutional subscription fees are calculated based on page views from relevant IP addresses from July 1-June 30th of the previous year.

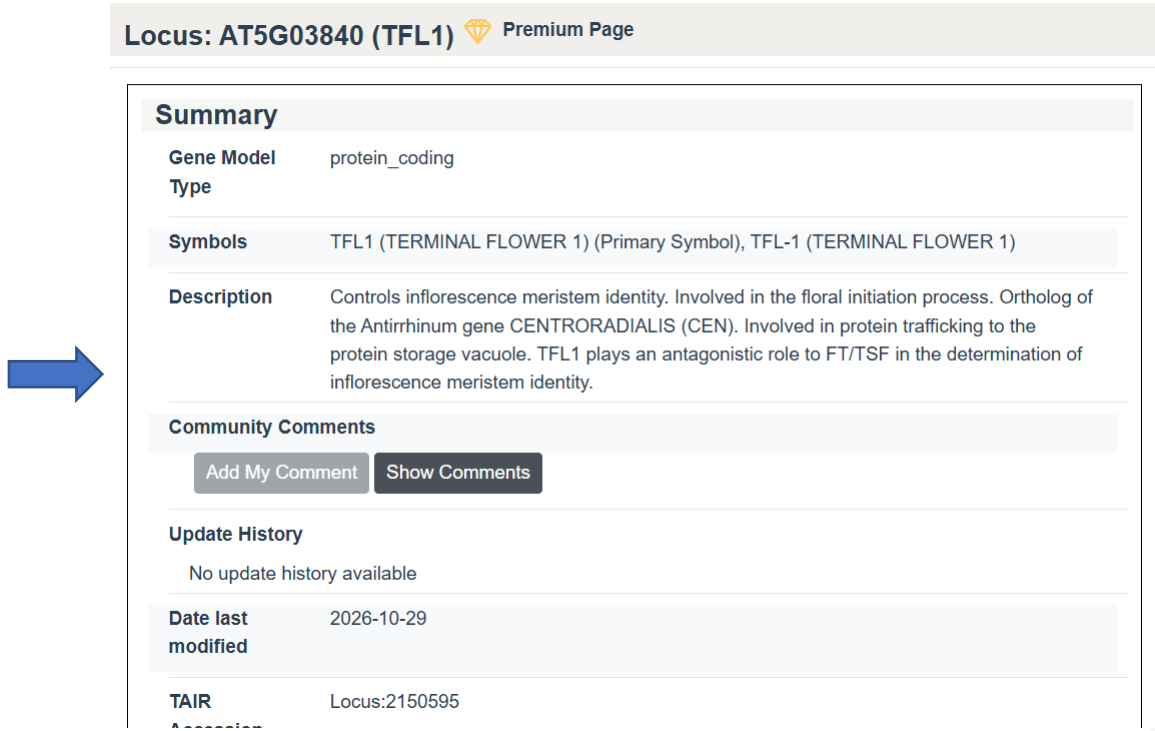
Tier	Paid page views	Cost
Legacy pricing (must renew on time to qualify)	<1,000	$950
Tier 4	<1,000	$1,550
Tier 3	1,000-4,999	$3,950
Tier 2	5,000-9,999	$7,300
Tier 1	10,000+	$10,350

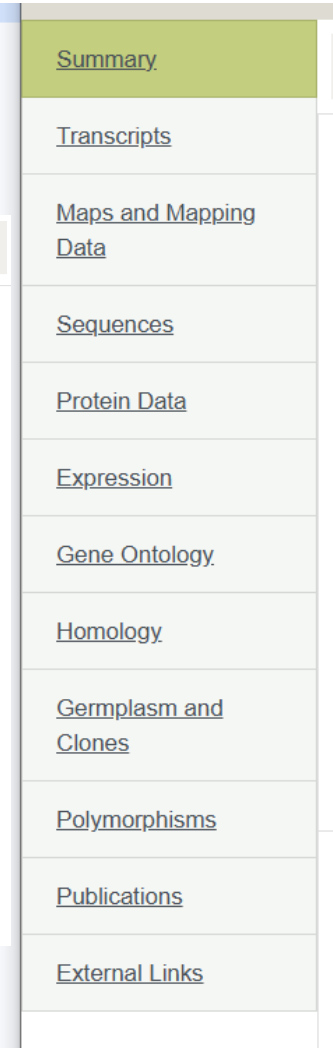
Academic consortia consisting of two or more institutions subscribing at the same time under the same or identical license agreements receive a 10% discount.

Search results and locus page









Transcripts

representative gene model: In TAIR, this is the reference gene model for the locus.
It's likely to be the most common and abundant form.

Locus: AT5G03840 (TFL1) Premium Page

Transcripts

Representative Gene Model	AT5G03840.1

Different models can have different transcription start, end or splicing junction

Associated Transcripts

Type	Transcript Count
EST	(2)
cDNA	(2)

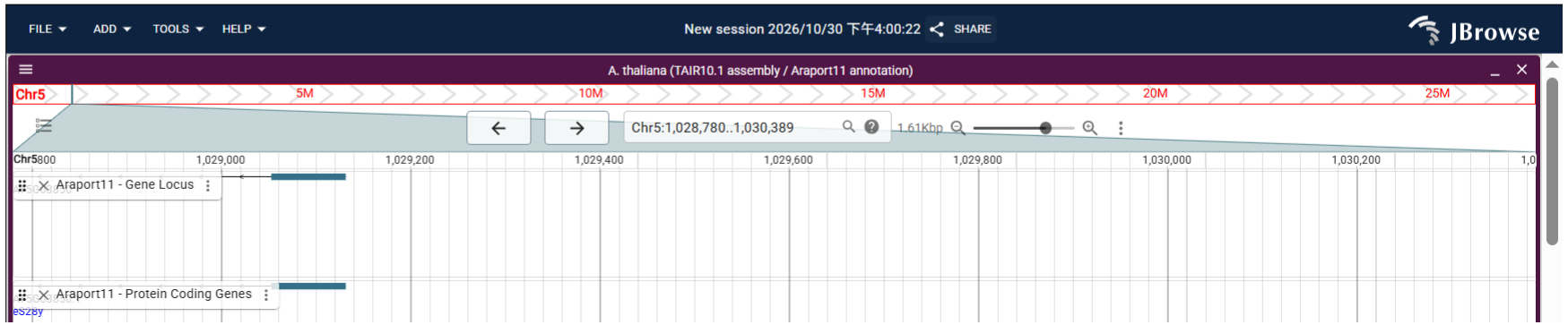
Associated Transposons	No Associated Transposons found

Link to Jbrowse2

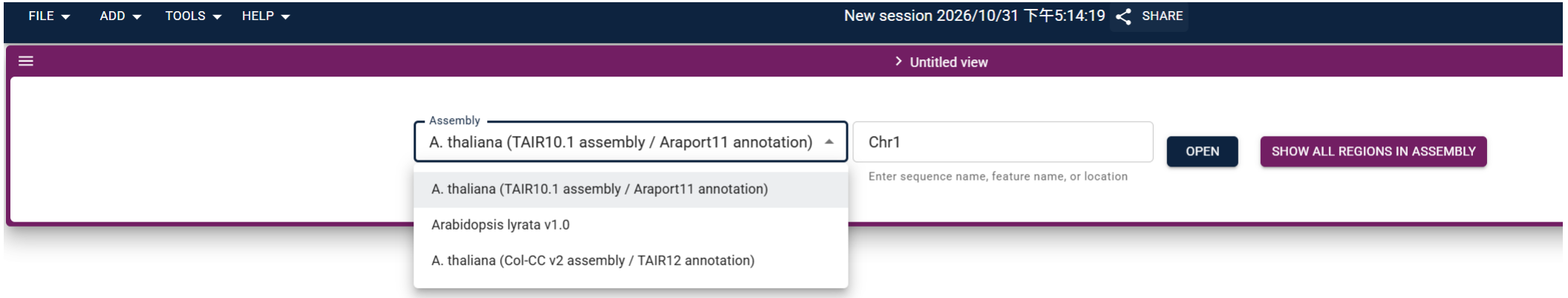
Maps and Mapping Data

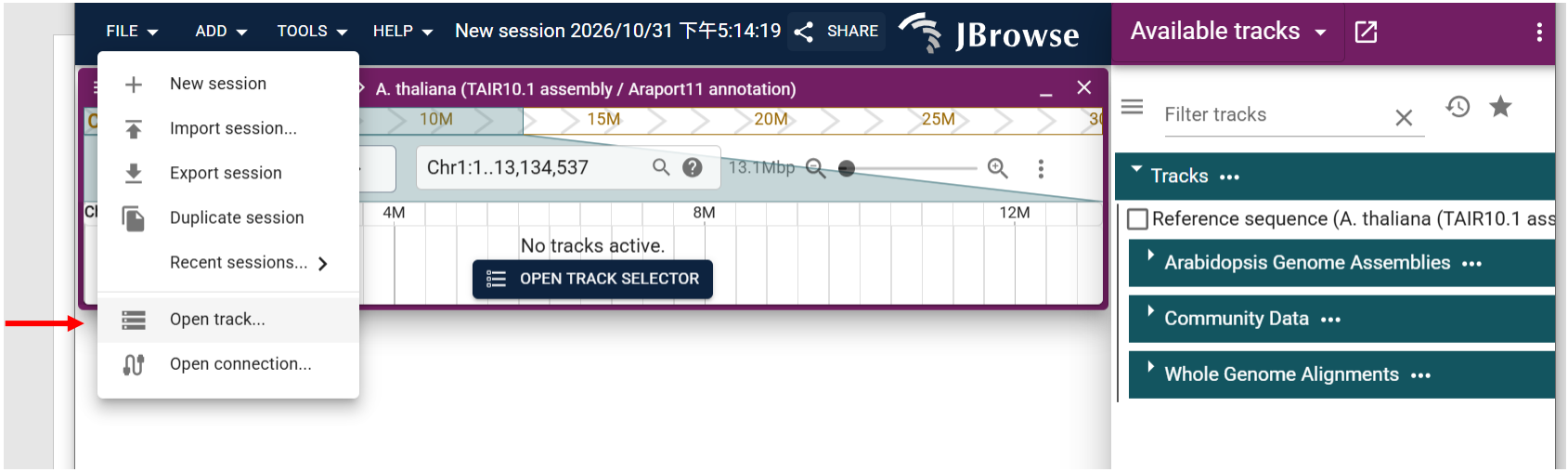
Map Detail Image

Center on AT5G03840 | Full-screen view



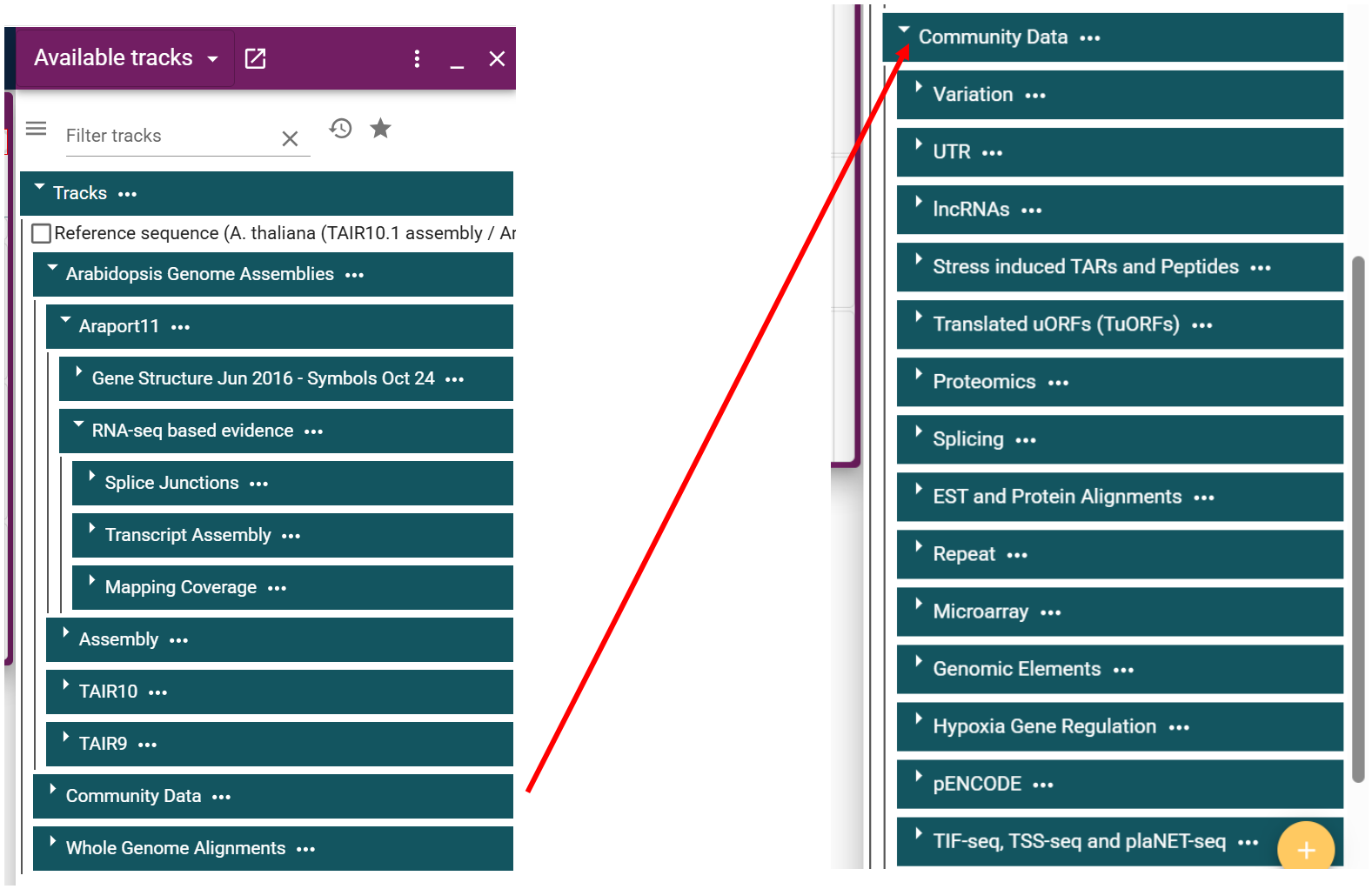
Different assemblies in JBrowse2





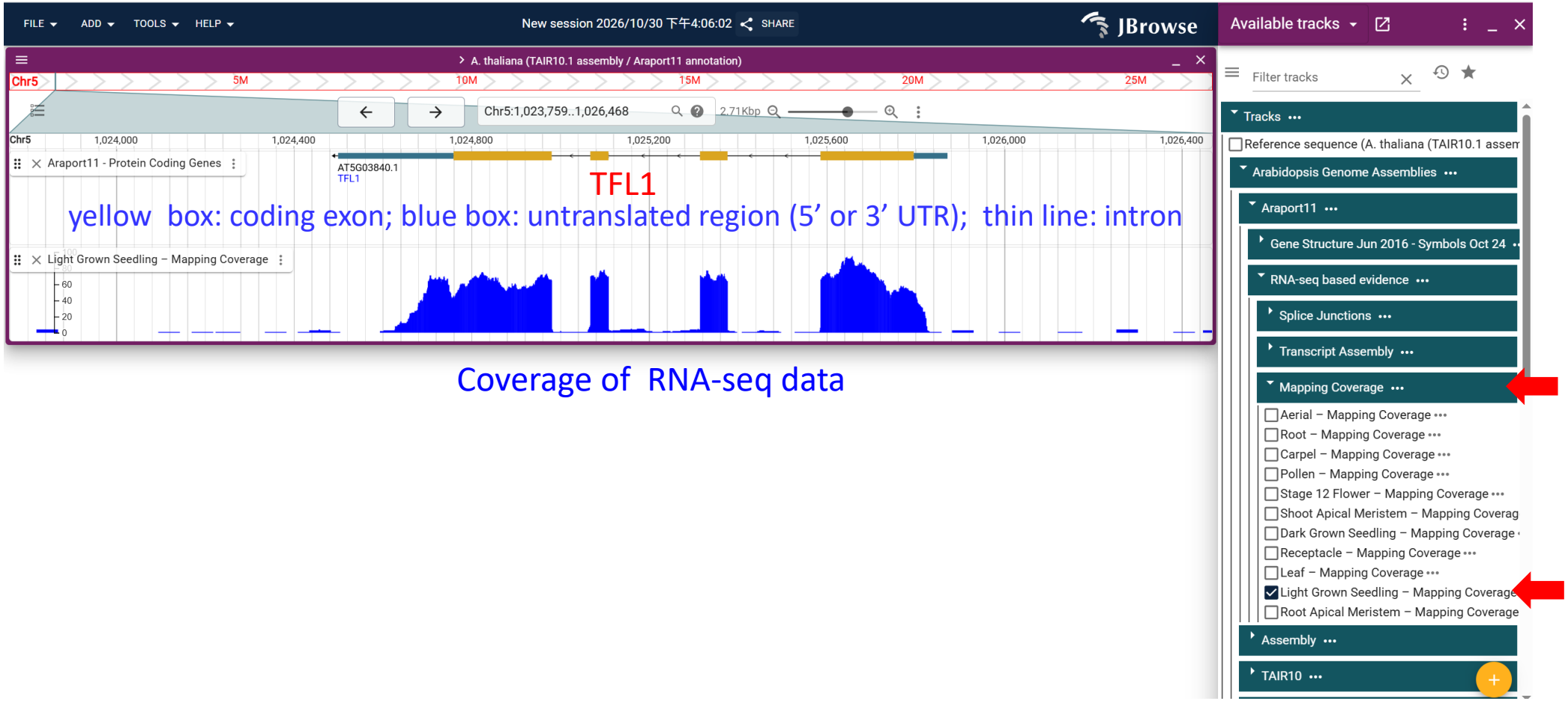
Genome browser

An important tool for users to visualize and browse entire genomes with annotated data or omic data, including gene structure, sequence variation, mutations, RNA-seq data, Ribo-seq data and peptide data etc.



Genome browser: JBrowse2

Check the accuracy of gene model (transcription start, end, introns, and CDS)



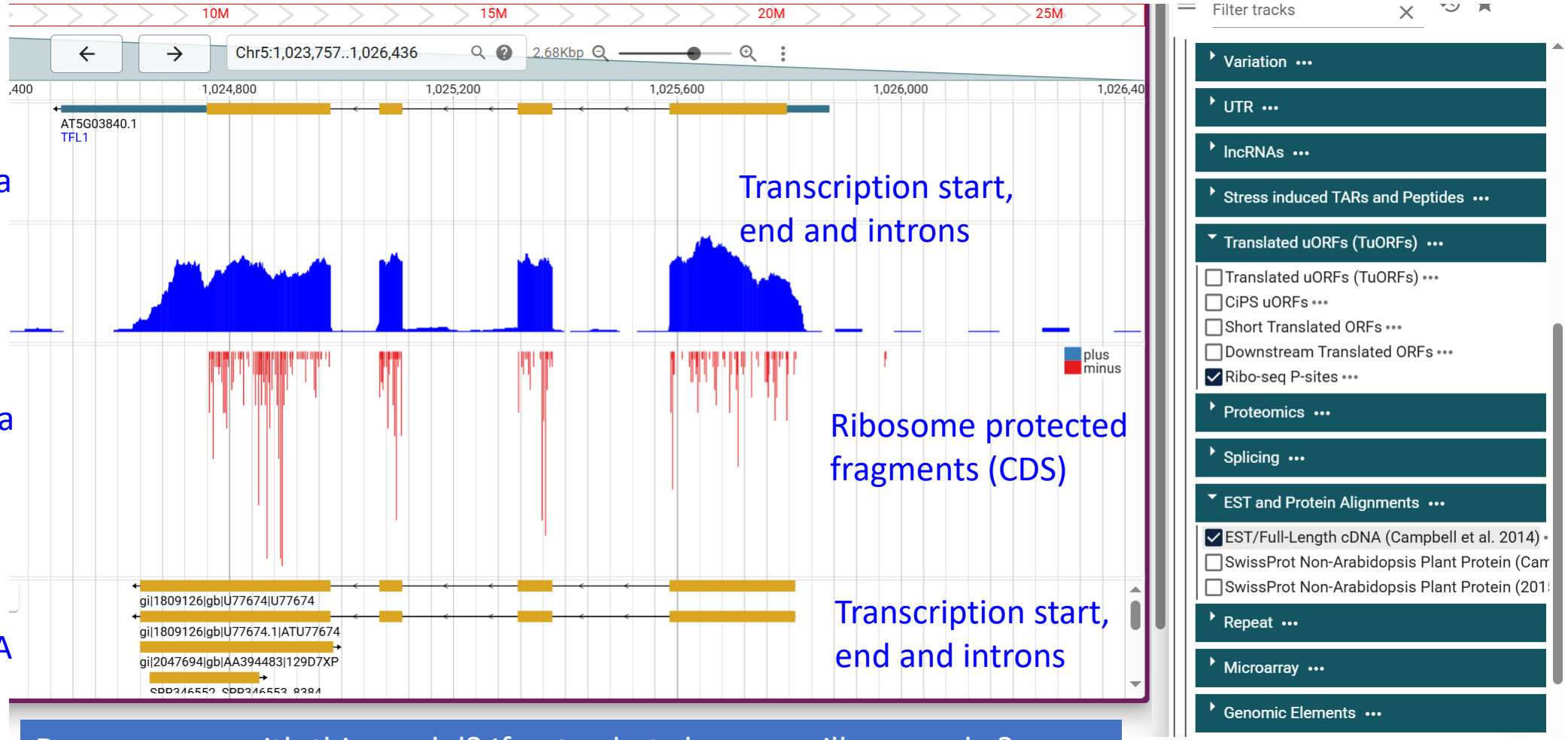
Genome browser: Jbrowse2

Check the accuracy of gene model (transcription start, end, introns, and CDS)

RNA-seq data

Ribo-seq data

EST/cDNA

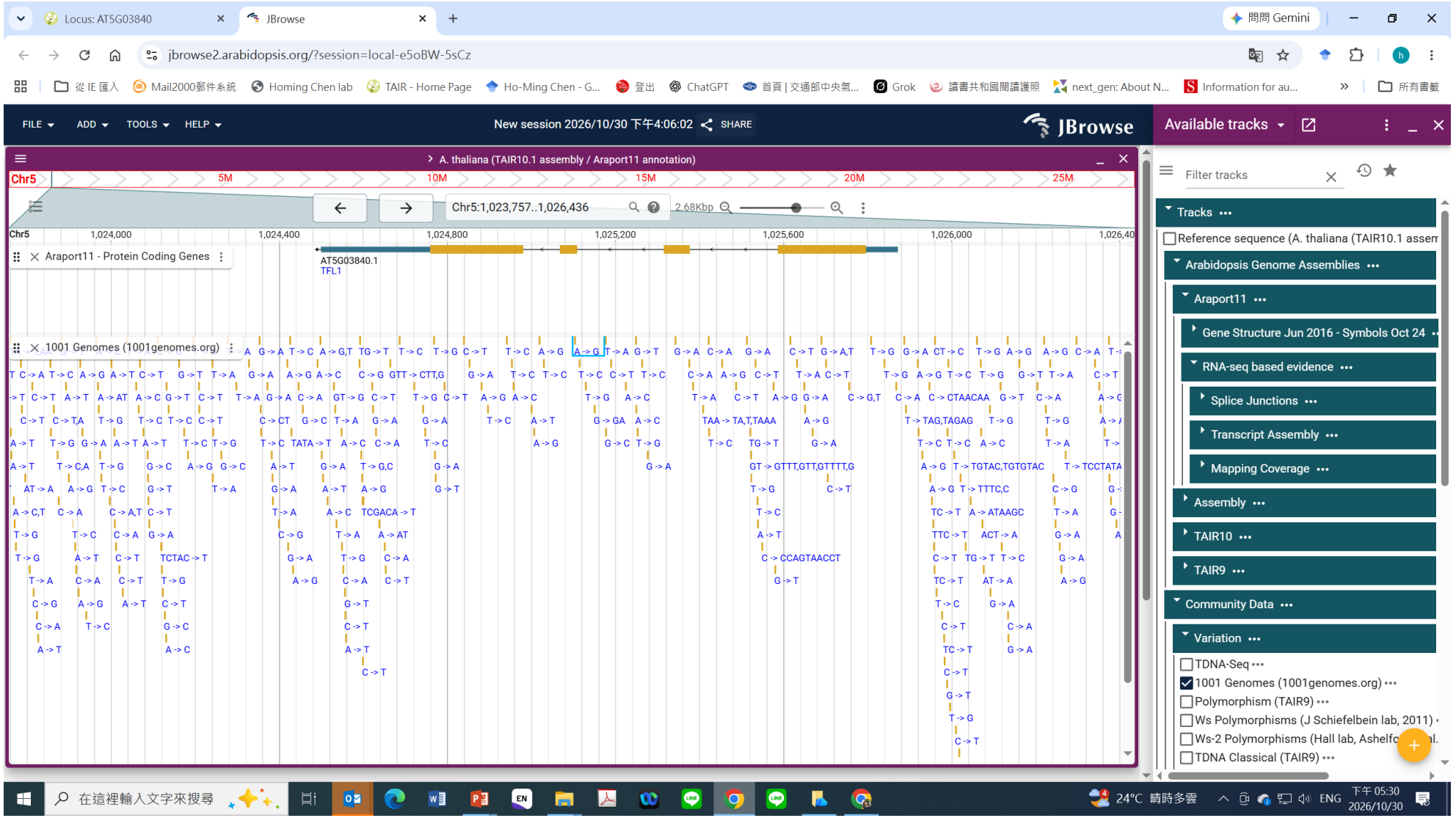


Do you agree with this model? If not, what changes will you make?

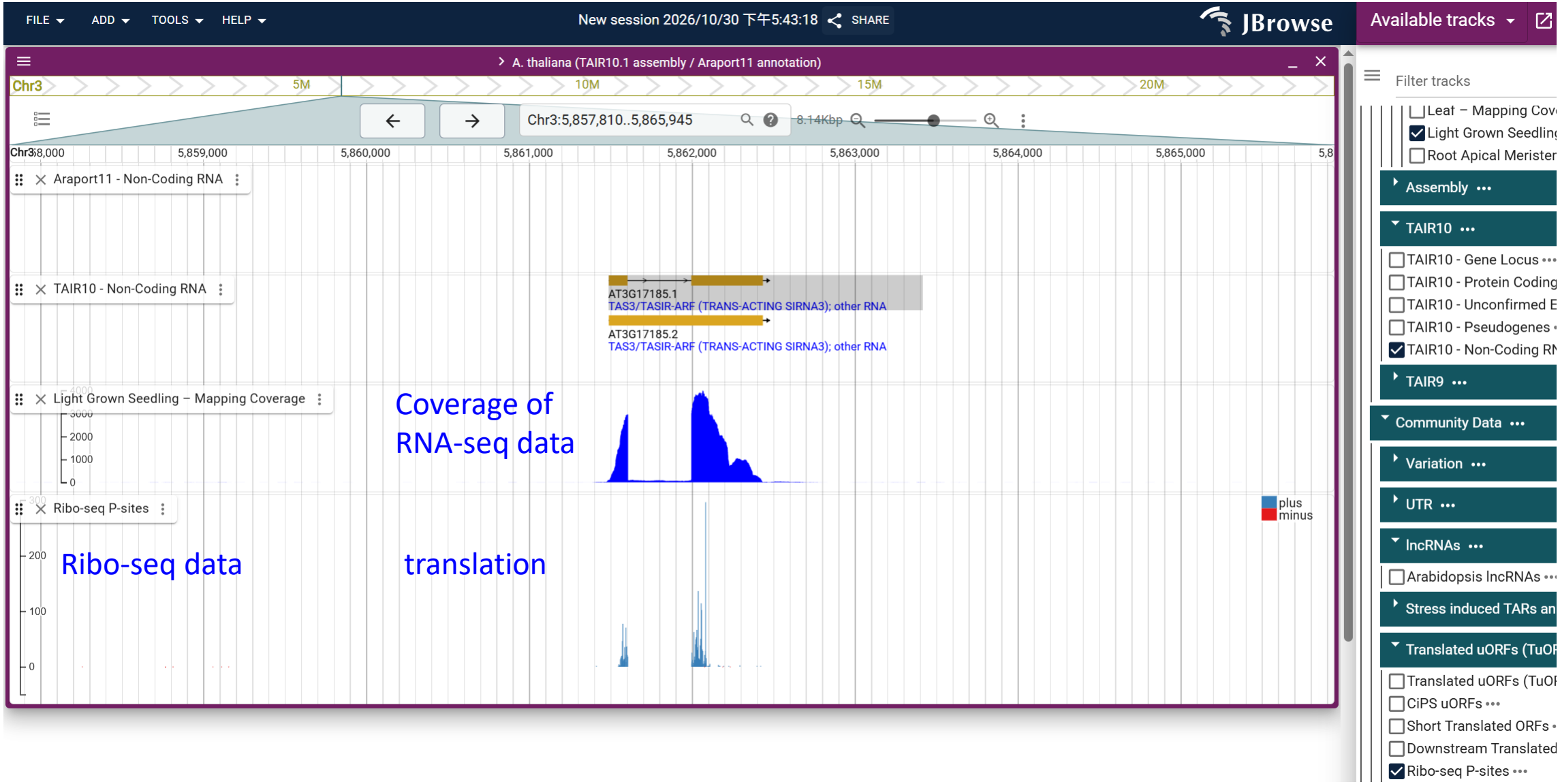
Genome browser: sequence variation among ecotypes

reference genome ≠ your genome

Col-0
Ler
Ws
Cvi
..
..
..



Genome browser: hidden RNA or open reading frame



Genome browser: non-coding RNA in monkey flower

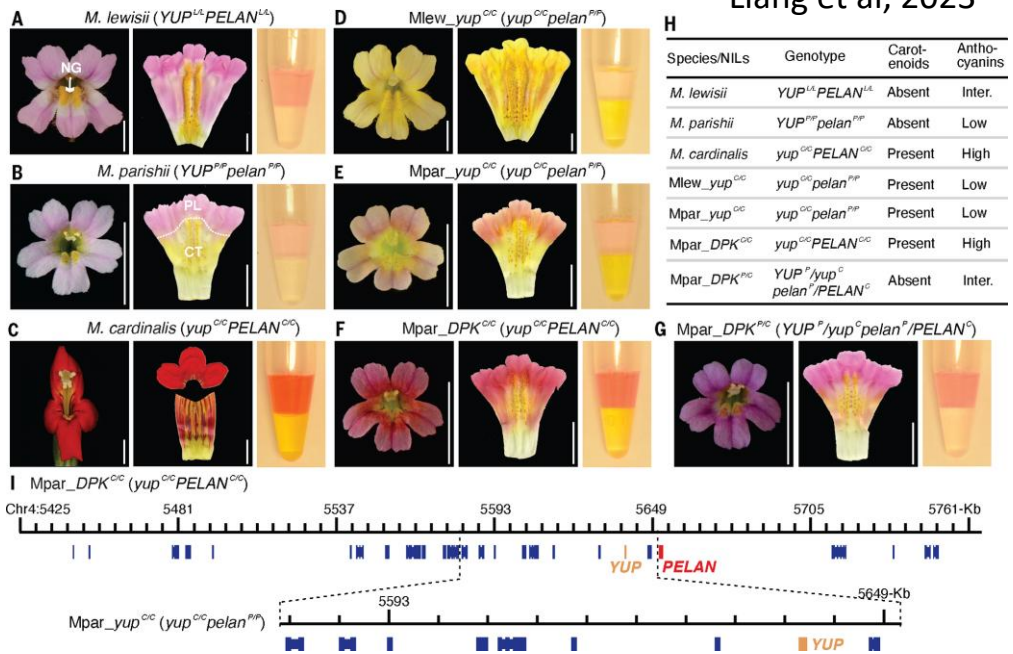
"Unusual suspects in flower color evolution"



Science 2023

YUP: yellow carotenoid pigments in the petal upper epidermis

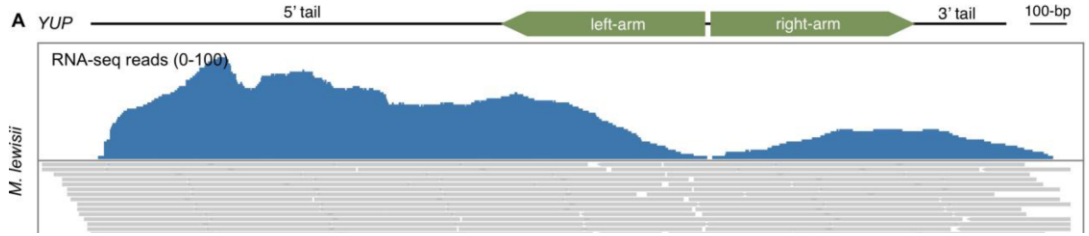
Liang et al, 2023



H

Species/NILs	Genotype	Carot-enoids	Antho-cyanins
M. lewisii	$YUP^{L/L}PELAN^{L/L}$	Absent	Inter.
M. parishii	$YUP^{P/P}pelan^{P/P}$	Absent	Low
M. cardinalis	$yup^{C/C}PELAN^{C/C}$	Present	High
Mlew_$yup^{C/C}$	$yup^{C/C}pelan^{P/P}$	Present	Low
Mpar_$yup^{C/C}$	$yup^{C/C}pelan^{P/P}$	Present	Low
Mpar_$DPK^{C/C}$	$yup^{C/C}PELAN^{C/C}$	Present	High
Mpar_$DPK^{P/C}$	YUP^{P}/yup^{C} $pelan^{P}/PELAN^{C}$	Absent	Inter.

- 8 protein coding gene (RNAi no phenotype)
- 1 non-coding gene (small RNA producing locus)



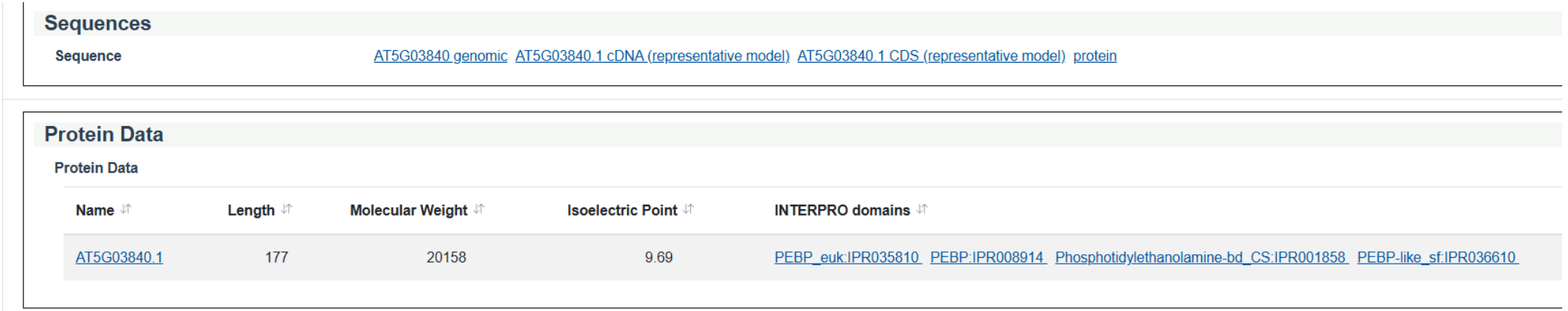
Locus page: DNA and protein sequences

Intron
5’UTR
CDS
3’ UTR

5’UTR
CDS
3’ UTR

CDS

Protein sequence of the representative form



- Some gene models don’t have 5’ UTR or 3’ UTR
- Some gene models have incorrect 5’ UTR or 3 ‘UTR
- If 5’ UTR or 3’ UTR is important to you, compare the annotation with ESTs/full length cDNA or RNA-Seq data in JBrowse

Exercise I

TGA1 (AT5G65210)

1. number of isoforms ? (what are the differences?)
2. the representative form in Araport11? Do you agree (check JBrwose2)? If not, your answer?
3. Are there any upstream open reading frames (small ORFs in 5’UTR)?

Locus: AT5G65210 (TGA1) Premium Page

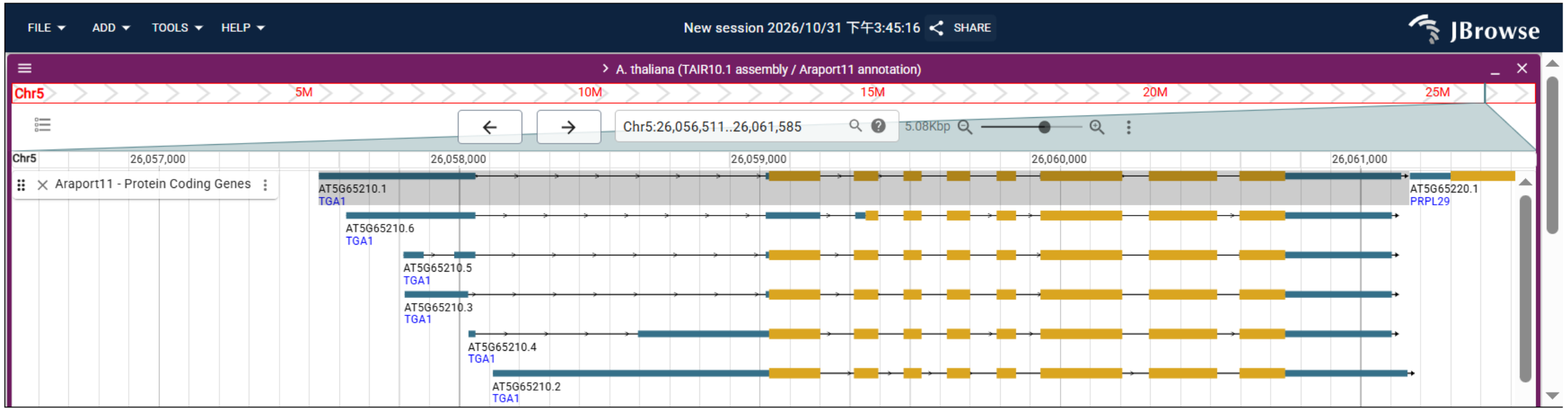
Summary

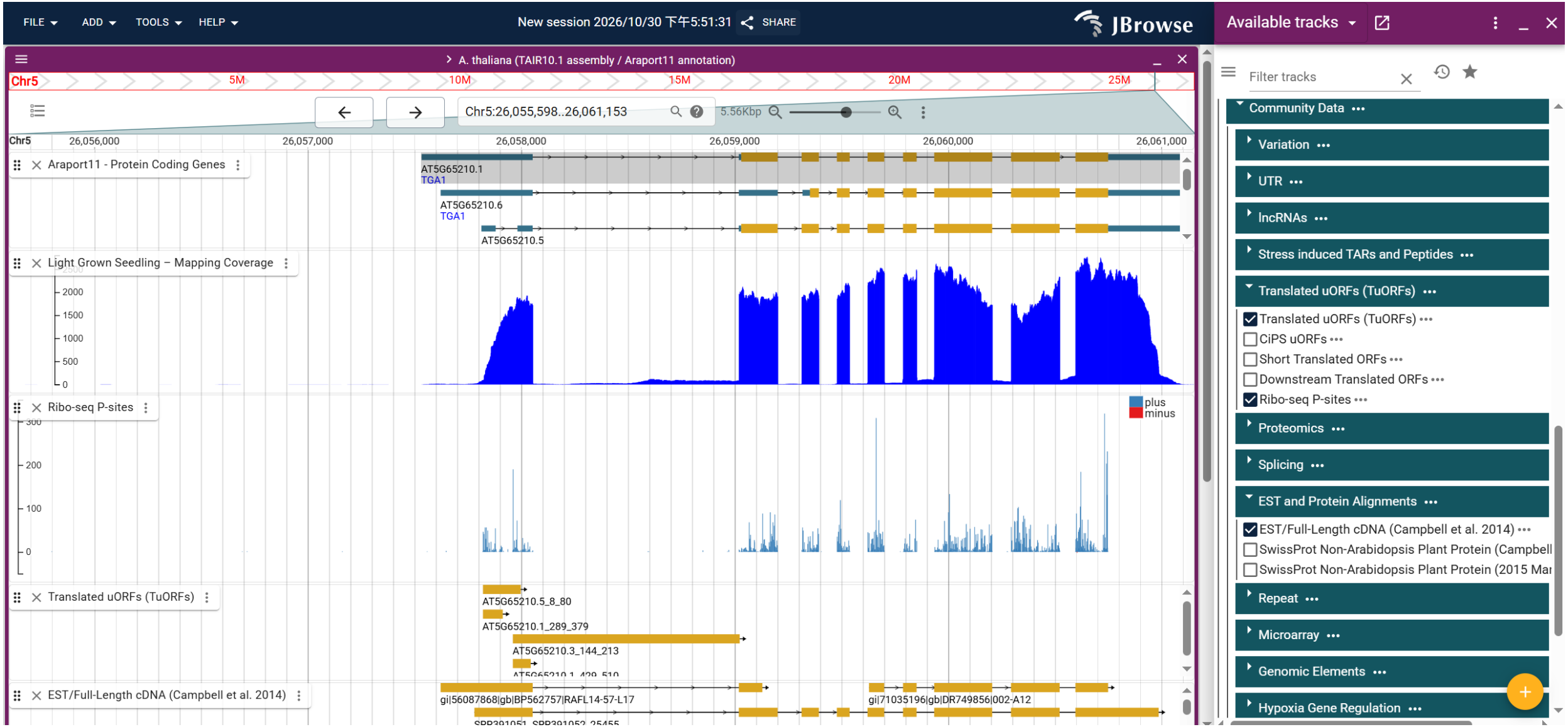
Gene Model Type	protein_coding
Symbols	TGA1 (TGACG sequence-specific binding protein 1) (Primary Symbol)
Description	Encodes TGA1, a redox-controlled regulator of systemic acquired resistance. TGA1 targets the activation sequence-1 (as-1) element of the promoter region of defense proteins. TGA1 are S-nitrosylated.
Community Comments	Add My Comment Show Comments
Update History	No update history available
Date last modified	2026-10-29
TAIR Accession	Locus:2171820

Transcripts

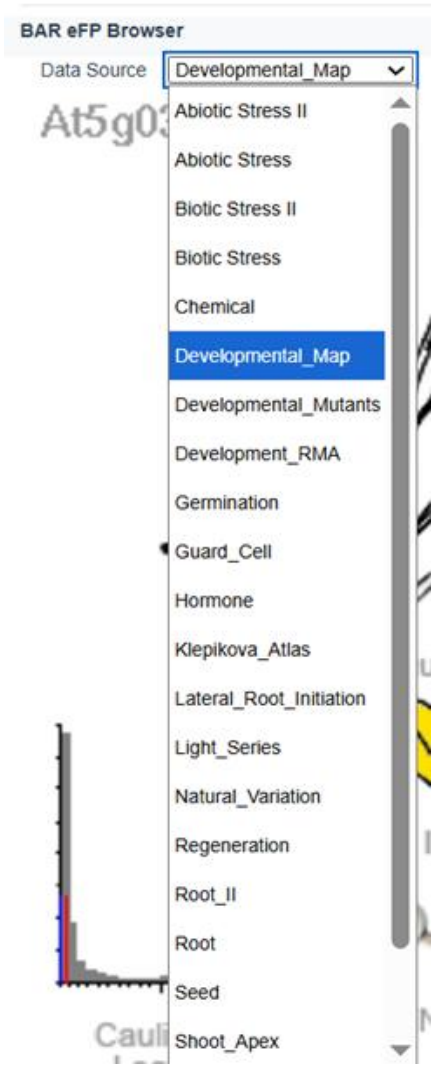
Representative Gene Model	AT5G65210.1
Other Gene Models	**Gene Name**
	AT5G65210.3
	AT5G65210.2
Associated Transcripts	

Center on AT5G65210.3 | Full-screen view




FILE
ADD
TOOLS
HELP
New session 2026/10/30 下午5:51:31
SHARE
JBrowse
Available tracks
A. thaliana (TAIR10.1 assembly / Araport11 annotation)
Chr5
5M
10M
15M
20M
25M
Chr5:26,055,598..26,061,153
5.56Kbp
26,056,000
26,057,000
26,058,000
26,059,000
26,060,000
26,061,000
Araport11 - Protein Coding Genes
AT5G65210.1
TGA1
AT5G65210.6
TGA1
AT5G65210.5
Light Grown Seedling – Mapping Coverage
2000
1500
1000
500
0
Ribo-seq P-sites
300
200
100
0
plus
minus
Translated uORFs (TuORFs)
AT5G65210.5_8_80
AT5G65210.1_289_379
AT5G65210.3_144_213
EST/Full-Length cDNA (Campbell et al. 2014)
gi|56087868|gb|BP562757|RAFL14-57-L17
gi|71035196|gb|DR749856|002-A12
Filter tracks
Community Data
Variation
UTR
lncRNAs
Stress induced TARs and Peptides
Translated uORFs (TuORFs)
Translated uORFs (TuORFs)
CiPS uORFs
Short Translated ORFs
Downstream Translated ORFs
Ribo-seq P-sites
Proteomics
Splicing
EST and Protein Alignments
EST/Full-Length cDNA (Campbell et al. 2014)
SwissProt Non-Arabidopsis Plant Protein (Campbell
SwissProt Non-Arabidopsis Plant Protein (2015 Mar
Repeat
Microarray
Genomic Elements
Hypoxia Gene Regulation

Locus page: Expression (TFL1)



Locus: AT5G03840 (TFL1) Premium Page

Expression

Plant Ontology Annotations 20 unique annotations

View Annotation Details

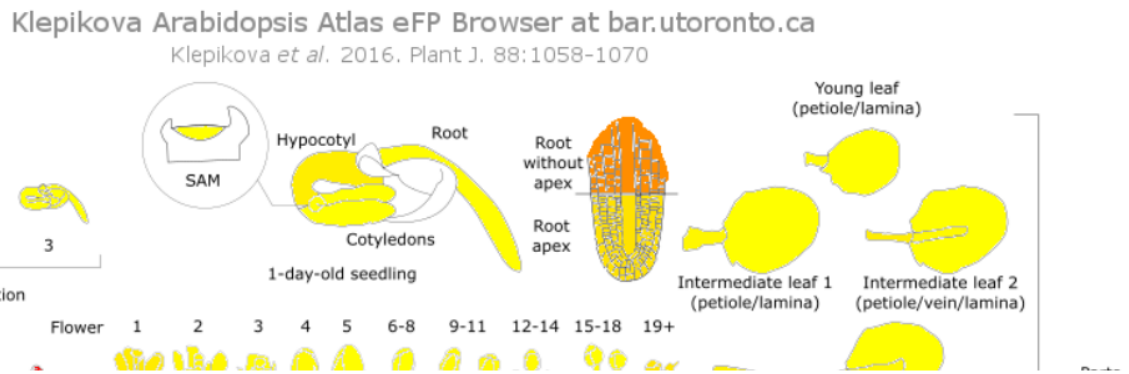
Category	Relationship Type	Keyword
Growth and Developmental Stages	expressed during	mature plant embryo stage, petal differentiation and expansion stage, plant embryo bilateral stage, plant embryo cotyledonary stage, plant embryo globular stage, vascular leaf senescent stage
Plant structure	expressed in	collective leaf structure, flower, flower meristem, hypocotyl, inflorescence meristem, leaf apex, plant embryo, plant sperm cell, root, seed, shoot apex, vascular leaf
Plant structure	located in	flower meristem, shoot system meristem

BAR eFP Browser

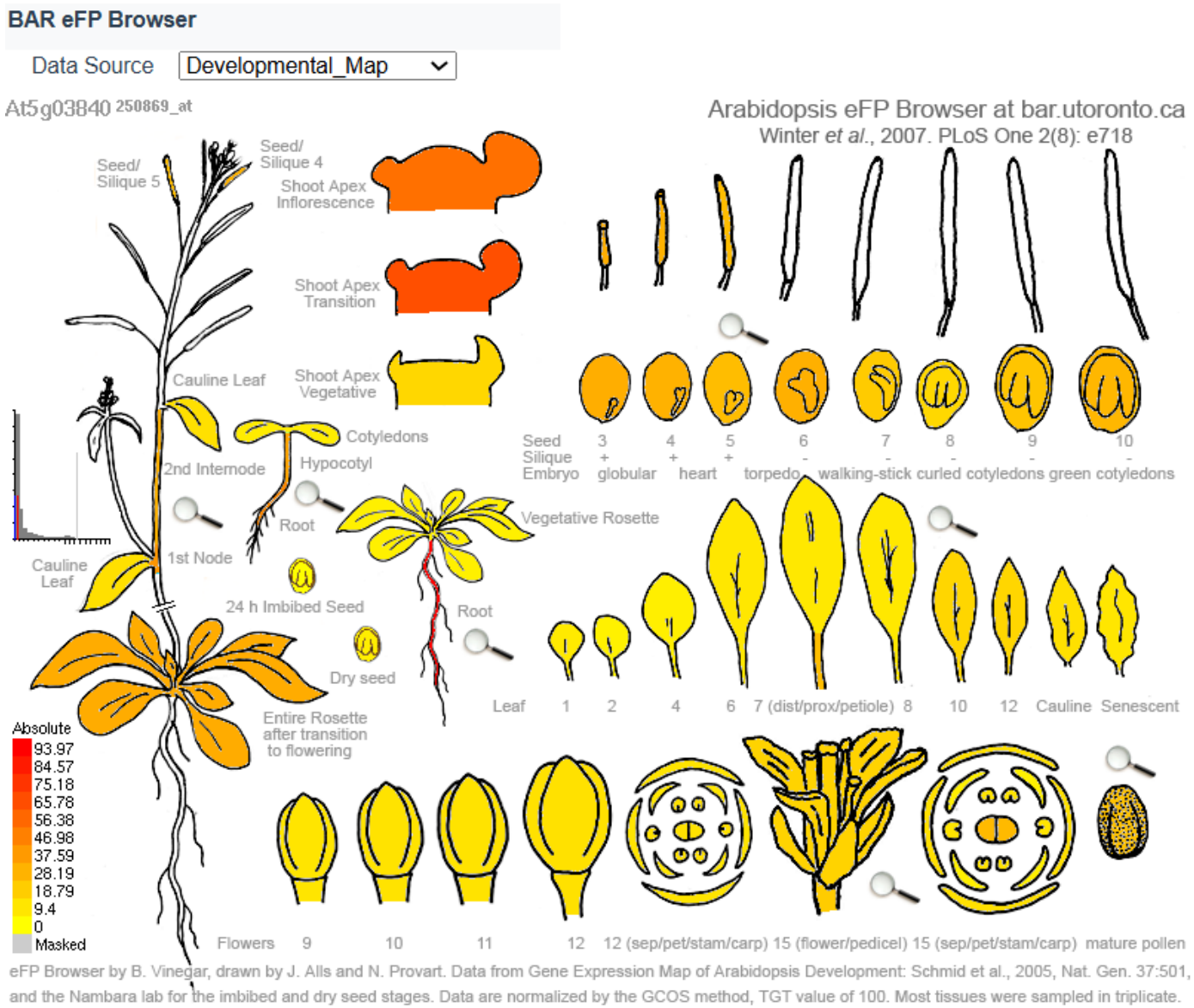
Data Source: Klepikova_Atlas

AT5G03840 AT5G03840

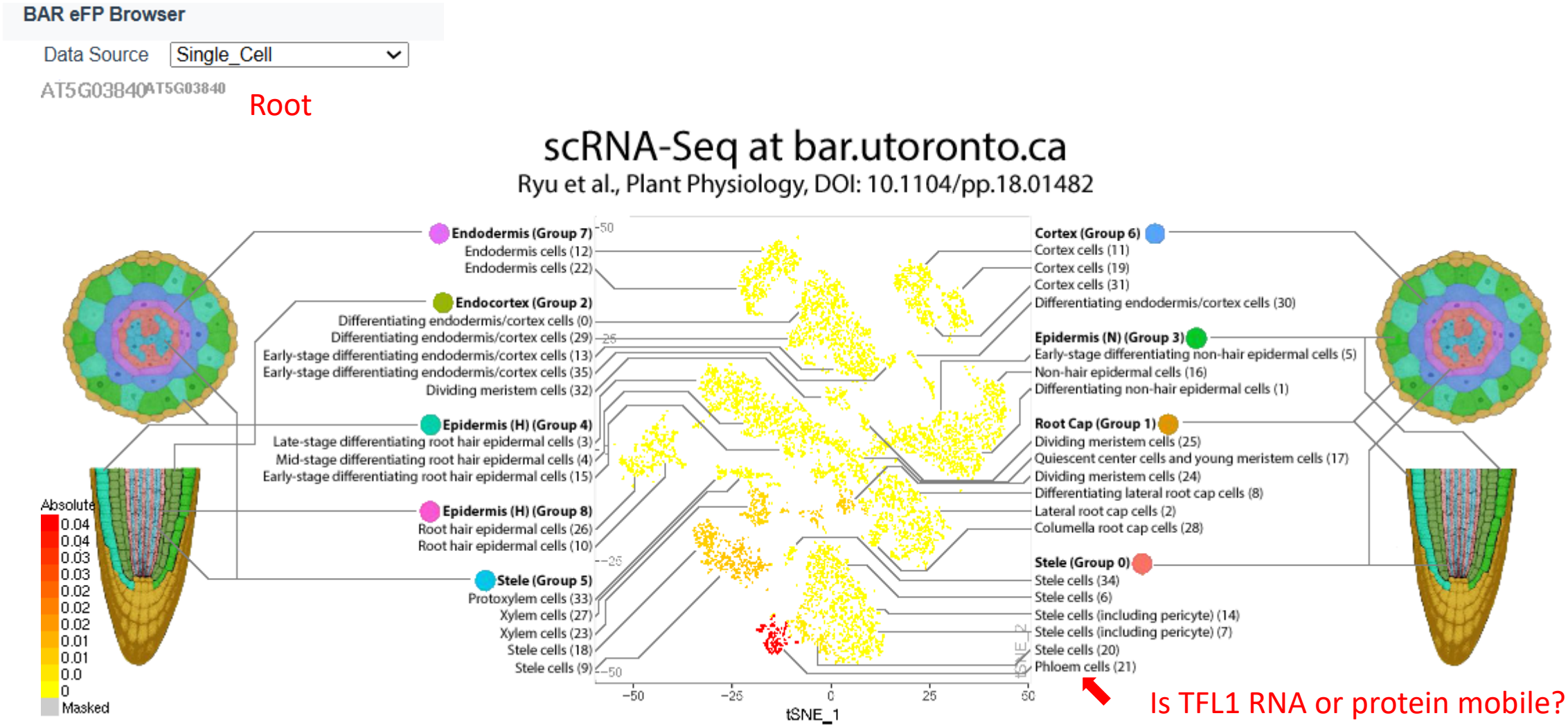
Microarray, RNA-seq, single cell



Development map expression



Single cell expression



Plant Physiology

Plant Physiol. 2020 Jan 29;182(4):2081–2095. doi: 10.1104/pp.19.00867

TERMINAL FLOWER1 Functions as a Mobile Transcriptional Cofactor in the Shoot Apical Meristem[1] [OPEN]

Daniela Goretti [a,2], Marina Silvestre [b,2], Silvio Collani [a,2], Tobias Langenecker [c], Carla Méndez [b], Francisco Madueño [b,3], Markus Schmid [a,d,3,4]

Author information · Article notes · Copyright and License information

PMCID: PMC7140938 PMID: 31996406

TERMINAL FLOWER1 acts in the shoot apical meristem as a mobile cell-non-autonomous transcriptional cofactor that associates with DNA to regulate meristem indeterminacy and flowering.

TFL1 protein is a mobile signal

Abstract

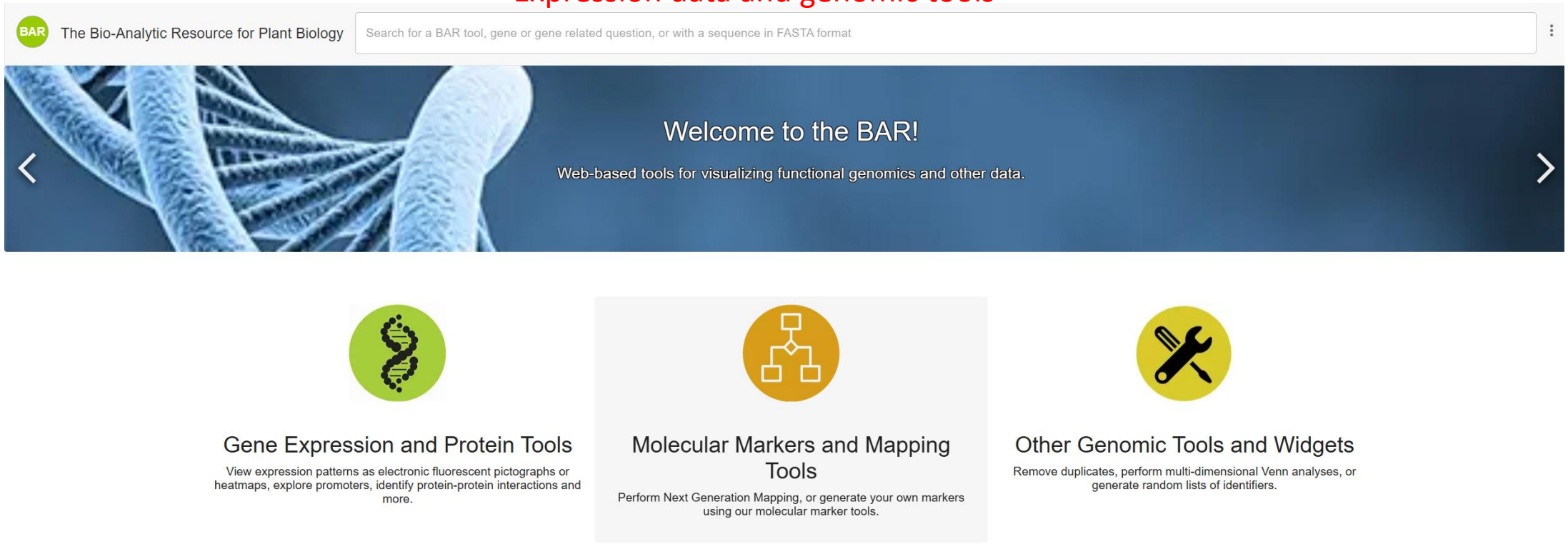
The floral transition is a critical step in the life cycle of flowering plants, and several mechanisms control this finely orchestrated process. TERMINAL FLOWER1 (TFL1) is a floral repressor and close relative of the florigen, FLOWERING LOCUS T (FT). During the floral transition, *TFL1* expression is up-regulated in the inflorescence apex to maintain the

https://pmc.ncbi.nlm.nih.gov/articles/PMC7140938/

BAR: The Bio-Analytic Resource for plant biology

https://bar.utoronto.ca/

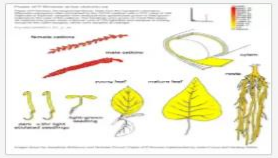
Expression data and genomic tools



BAR: Gene expression in crops and other plants

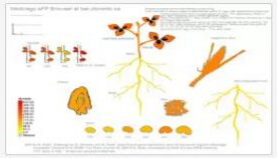
Other Dicot eFP Browsers

Poplar eFP Browser



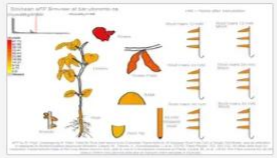
Info Go Pub Pub

Medicago eFP Browser



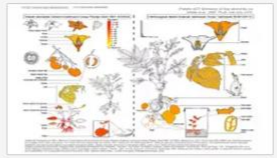
Info Go

Soybean eFP Browser



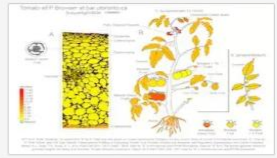
Info Go

Potato eFP Browser



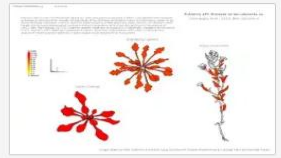
Info Go

Tomato eFP Browser



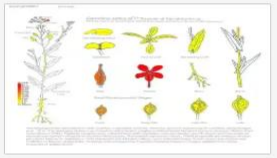
Info Go

E. salsugineum eFP Browser



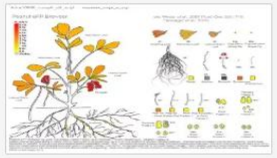
Info Go Pub

C. sativa eFP Browser



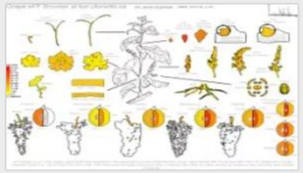
Info Go Pub

Arachis eFP Browser



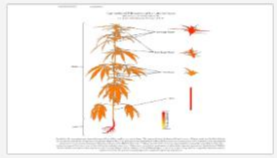
Info Go Pub

Grape eFP Browser



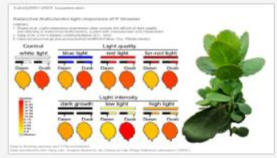
Info Go Pub

Cannabis eFP Browser



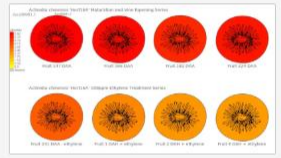
Info Go

Kalanchoe eFP Browser



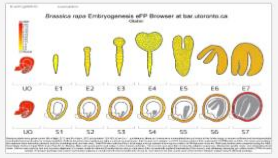
Info Go Pub

Actinidia eFP Browser



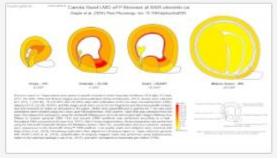
Info Go Pub

B. rapa eFP Browser



Info Go Pub

Canola eFP Browser



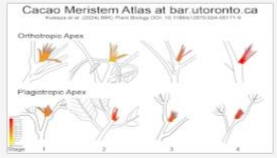
Info Go Pub

Cacao CCN eFP Browser



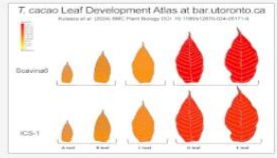
Info Go Pub

Cacao SCA eFP Browser



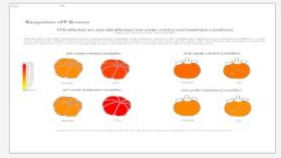
Info Go Pub

Cacao TC eFP Browser



Info Go Pub

Mangosteen eFP Browser



Info Go Pub

Lupin eFP Browser



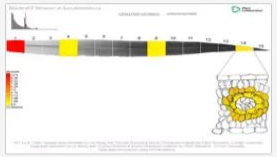
Strawberry eFP Browser



BAR: Gene expression in crops and other plants

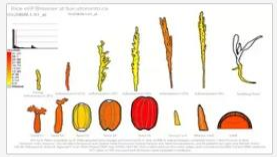
Monocot eFP Browsers

Maize eFP Browser



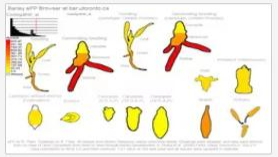
Info Go Pub Pub

Rice eFP Browser



Info Go

Barley eFP Browser



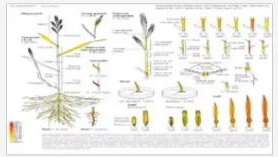
Info Go

Triticale eFP Browser



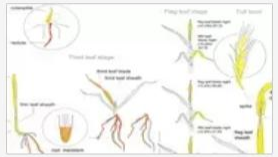
Info Go Pub

Brachypodium eFP Browser



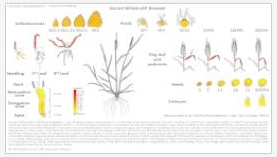
Info Go Pub

Wheat eFP Browser



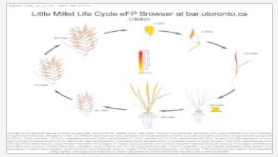
Info Go Pub

Durum Wheat eFP Browser



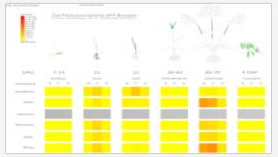
Info Go Pub

Little Millet eFP Browser



Info Go

Oat eFP Browser



Info Go Pub

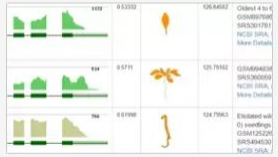
Other Gene Expression & Protein Tools

ThaleMine



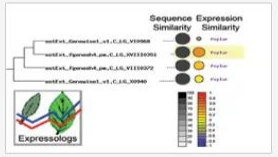
Info Go Pub

eFP-Seq Browser



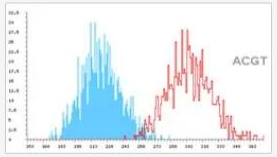
Info Go Pub

Expressolog Tree Viewer



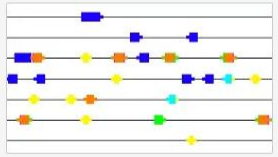
Info Go Pub

Promomer



Info Go Pub

Cistome



Info Go Pub

PMET



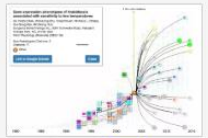
Info Go Pub

BAR: Tools and widgets

Other Genomic Tools and Widgets

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

Arabidopsis Citation Network Viewer



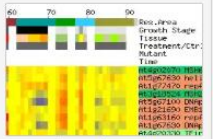
Info Go

ClustalW with MView Output



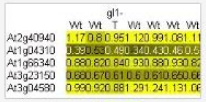
Info Go

DataMetaFormatter



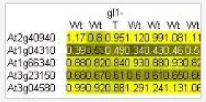
Info Go

Heatmapper



Info Go

Heatmapper Plus



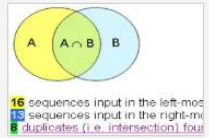
Info Go

Duplicate Remover



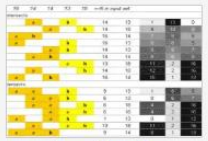
Info Go

Venn Selector



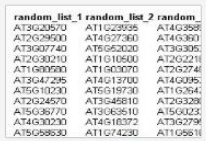
Info Go

Venn SuperSelector



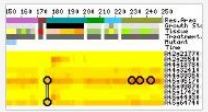
Info Go

Random ID list generator



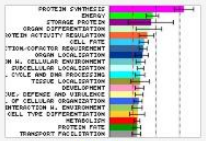
Info Go

AGURR



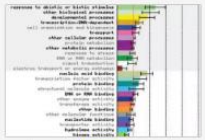
Info Go

Classification Super Viewer



Info Go Pub

Medicago Classification Super Viewer



Info Go Pub

_at to AGI converter



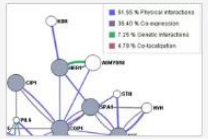
Info Go

MASTA



Info Go

GeneMANIA



Info Go

Topo-phylogeny



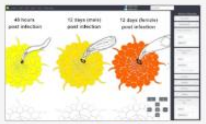
Info Go Pub

BAR API



Info Go

Custom eFP Creator



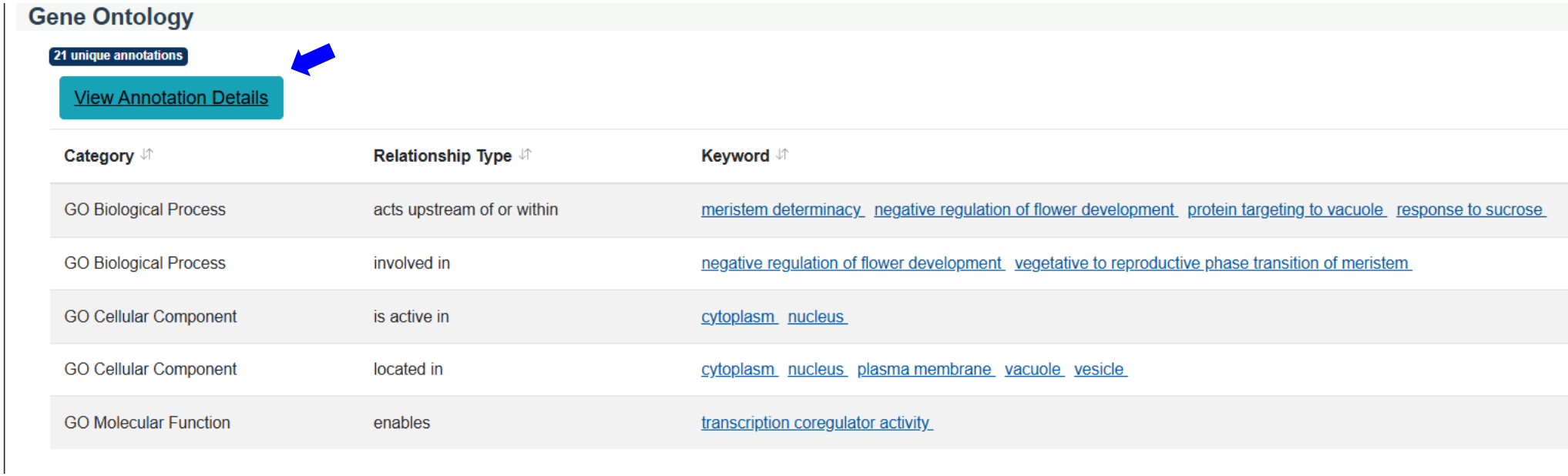
Info Go

Downloads



Info Go

Gene ontology (GO)



Gene Ontology

21 unique annotations

View Annotation Details

Category	Relationship Type	Keyword
GO Biological Process	acts upstream of or within	meristem determinacy, negative regulation of flower development, protein targeting to vacuole, response to sucrose
GO Biological Process	involved in	negative regulation of flower development, vegetative to reproductive phase transition of meristem
GO Cellular Component	is active in	cytoplasm, nucleus
GO Cellular Component	located in	cytoplasm, nucleus, plasma membrane, vacuole, vesicle
GO Molecular Function	enables	transcription coregulator activity

Annotations with **controlled vocabulary** terms :

Gene Ontology (GO)

Genes function and **localization** of the gene product.

Most gene annotations are made by curators using the published literature as sources. Some annotations are based on computational methods. Detailed information including references and supporting evidence can be obtained by clicking on the link to view Annotation Detail.

Annotation details

Select All | Clear Selected | Displaying 1 - 23 of 23 results | Results per page (top): 25

No.		Locus	Gene Model	Gene Symbol/Full Name	Relationship Type	Keyword	Keyword Category	Evidence:Description:With	Reference
1	☐		AT5G03840.1		located in	cytoplasm	cellular component	*inferred from sequence model:* predicted protein features:	AtSubP analysis prediction
2	☐	AT5G03840		TFL1/TERMINAL FLOWER 1	acts upstream of or within	response to sucrose	biological process	*inferred from mutant phenotype:* mutant growth experiment with supplementation of substrates: NONE	Ohto et al (2001)
3	☐	AT5G03840		TFL1/TERMINAL FLOWER 1	acts upstream of or within	negative regulation of flower development	biological process	*inferred from mutant phenotype:* analysis of physiological response: NONE	Shannon et al (1991)
4	☐	AT5G03840		TFL1/TERMINAL FLOWER 1	acts upstream of or within	protein targeting to vacuole	biological process	*inferred from mutant phenotype:* biochemical/chemical analysis: NONE	Sohn et al (2007)
5	☐	AT5G03840		TFL1/TERMINAL FLOWER 1	involved in	vegetative to reproductive phase transition of meristem	biological process	*Inferred from Biological aspect of Ancestor:* none: PANTHER:PTN002260847|TAIR:locus:2038643	Annotation inferences using phylogenetic trees prediction

Using GO to identify genes with similar functions

Summary

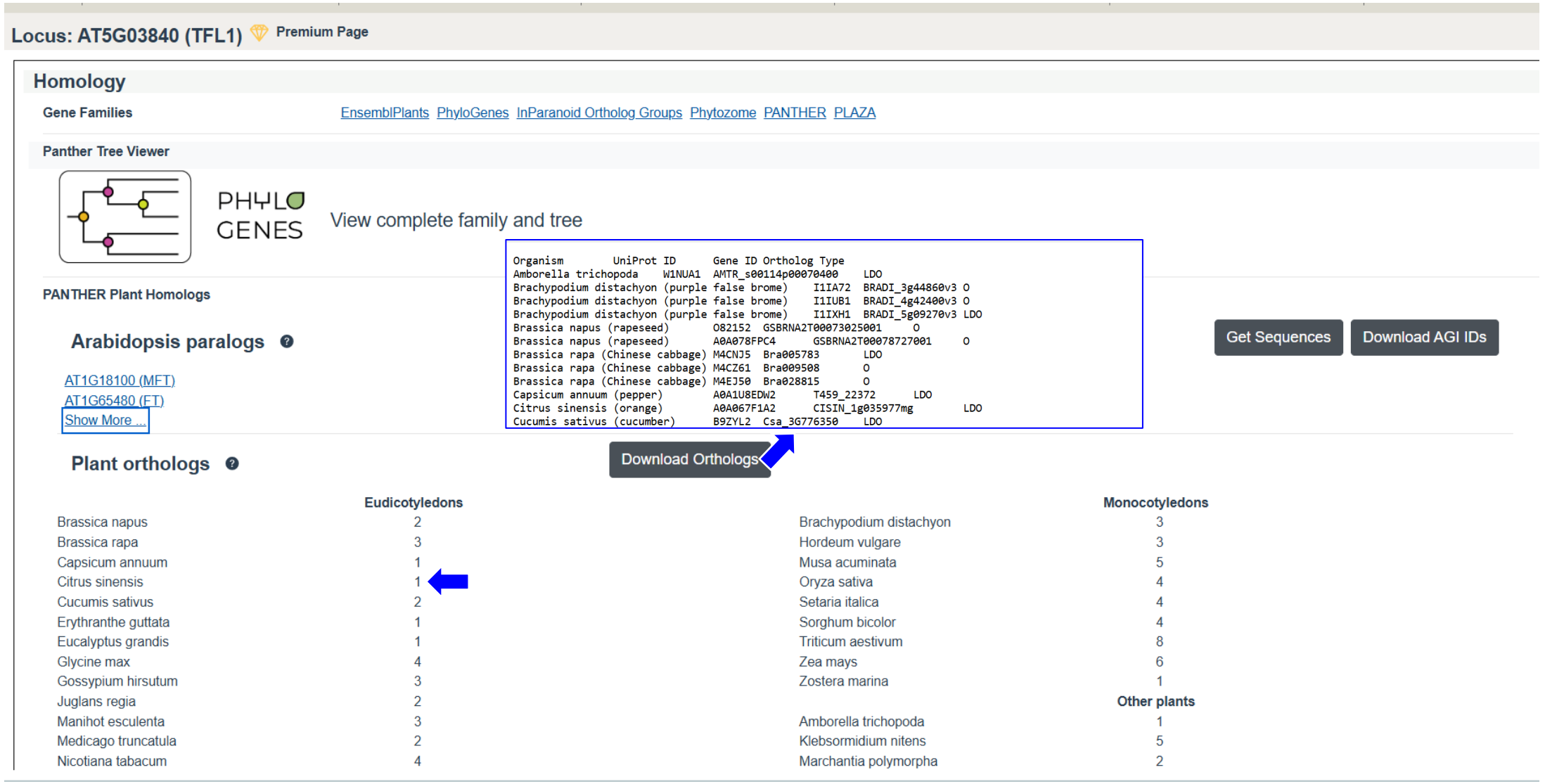
Name	negative regulation of flower development
Synonyms	inhibition of flower development,down-regulation of flower development,down regulation of flower development,downregulation of flower development
Date Last Modified	2010-11-18
Definition	Any process that stops, prevents, or reduces the frequency, rate or extent of flower development.
External ID	GO:0009910
Keyword Category	biological process
Data Associated To This Term	

Publications	Annotations	Loci
44	71	57

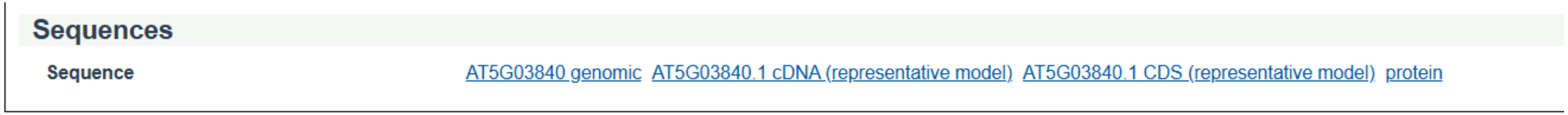
Data Associated To Children Terms	No data associated to this term.
TreeView	Keyword TreeViewer

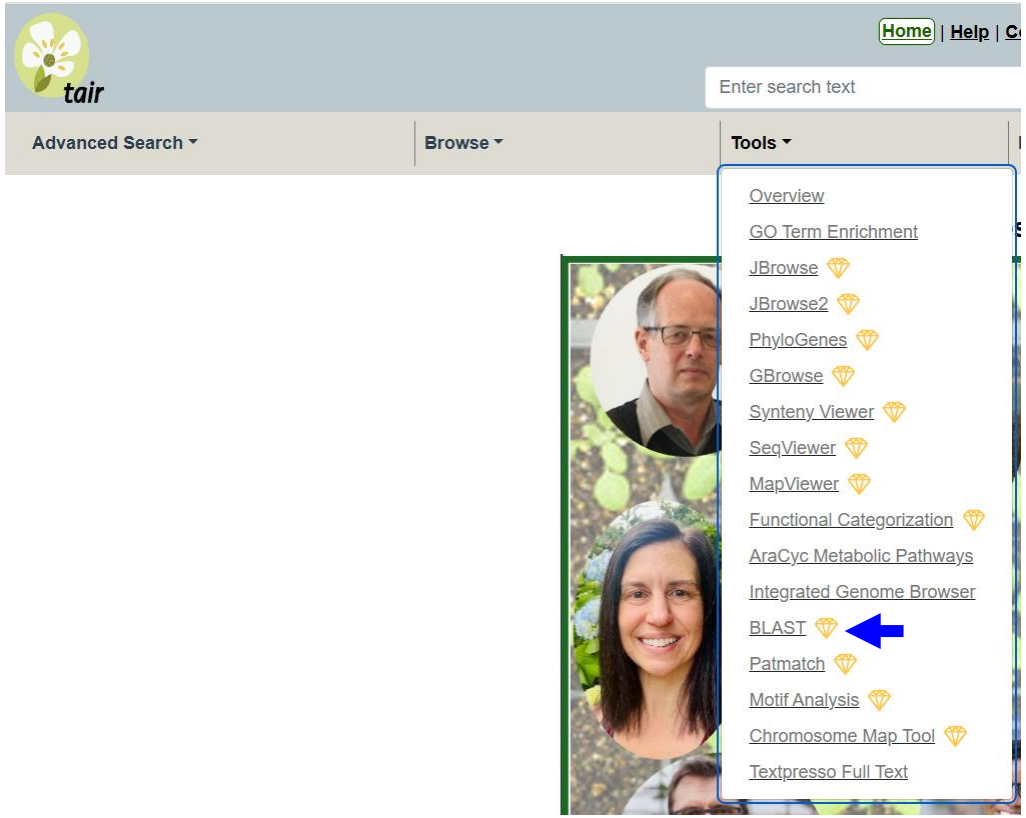
No.	Locus	Description
1	AT5G13790	Other Names: AGAMOUS-LIKE 15;AGL15 AGL15 (AGAMOUS-Like 15) is a member of the MADS domain family of regulatory factors. Although AGL15 is preferentially expressed during embryogenesis, AGL15 is also expressed in leaf primordia, shoot apical meristems and young floral buds, suggesting that AGL15 may play a role during post-germinative development. Transgenic plants that ectopically express AGL15 show delays in the transition to flowering, perianth abscission and senescence and fruit and seed maturation. Role in embryogenesis and gibberellic acid catabolism. Targets B3 domain transcription factors that are key regulators of embryogenesis.AGL15 binds the HAE promoter in floral receptacles and represses HAE expression. AGL15 is phosphorylated in a MKK4/5 dependent manner in floral receptacles. Serines 231 and 257 are phosphorylated in floral receptacles. AGL15 also directly regulates the expression of the peroxidase PRX17, linking it to lignified tissue expression.
2	AT4G20400	Other Names: JMJ14;JUMONJI 14;PKDM7B Encodes a histone H3K4 demethylase repressing floral transition.Coordinates and structure features have been deposited in the RCSB Protein Data Bank with the accession codes 5YKN (JMJ14CD) and 5YKO(JMJ14CD-NOG-H3K4em3 complex). Glu285 and Glu516 contribute to the recognition of H3R2. Asp312 contributes to the recognition of H3Q5. Phe171 and Val363 contribute to the recognition of H3A7. Ser290, Trp296, and Tyr298 contribute to the recognition of H3K4me3 (PMID:29233856).
3	AT5G48890	Other Names: LATE;LATE FLOWERING Encodes a C(2) H(2) -type zinc-finger transcriptional regulator and is expressed in the leaf vasculature and the vegetative shoot apical meristem and controls the transition to flowering.
4	AT5G11530	Other Names: EMF1;embryonic flower 1 Involved in regulating reproductive development . Part of PcG like complex mediating floral gene expression repression including AG.

Finding crop orthologs

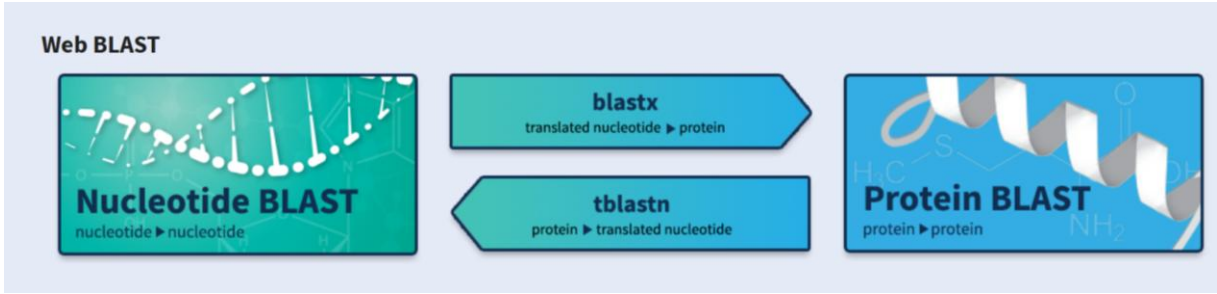


Finding orthologs with Blast search





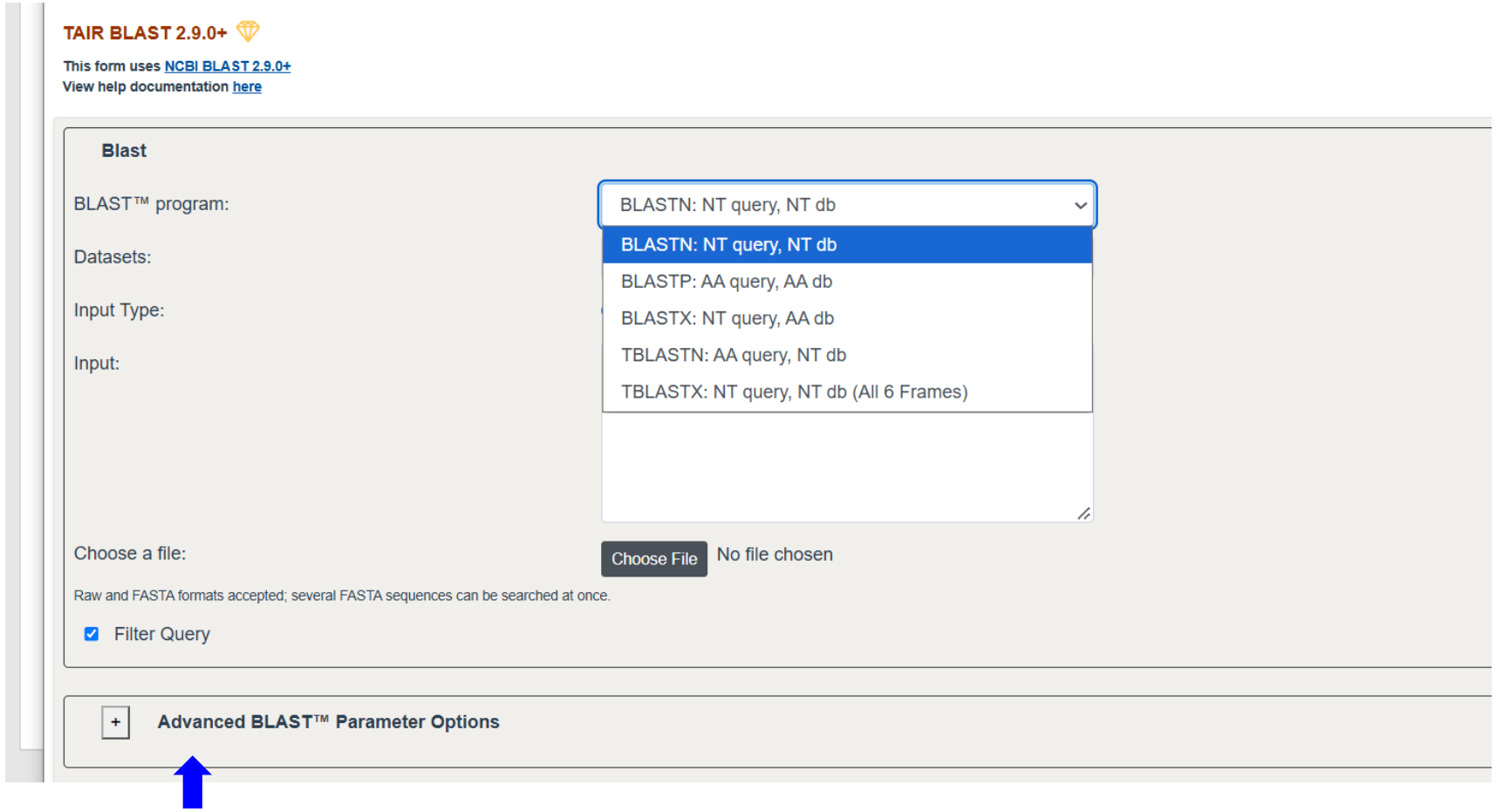
NCBI https://blast.ncbi.nlm.nih.gov/Blast.cgi



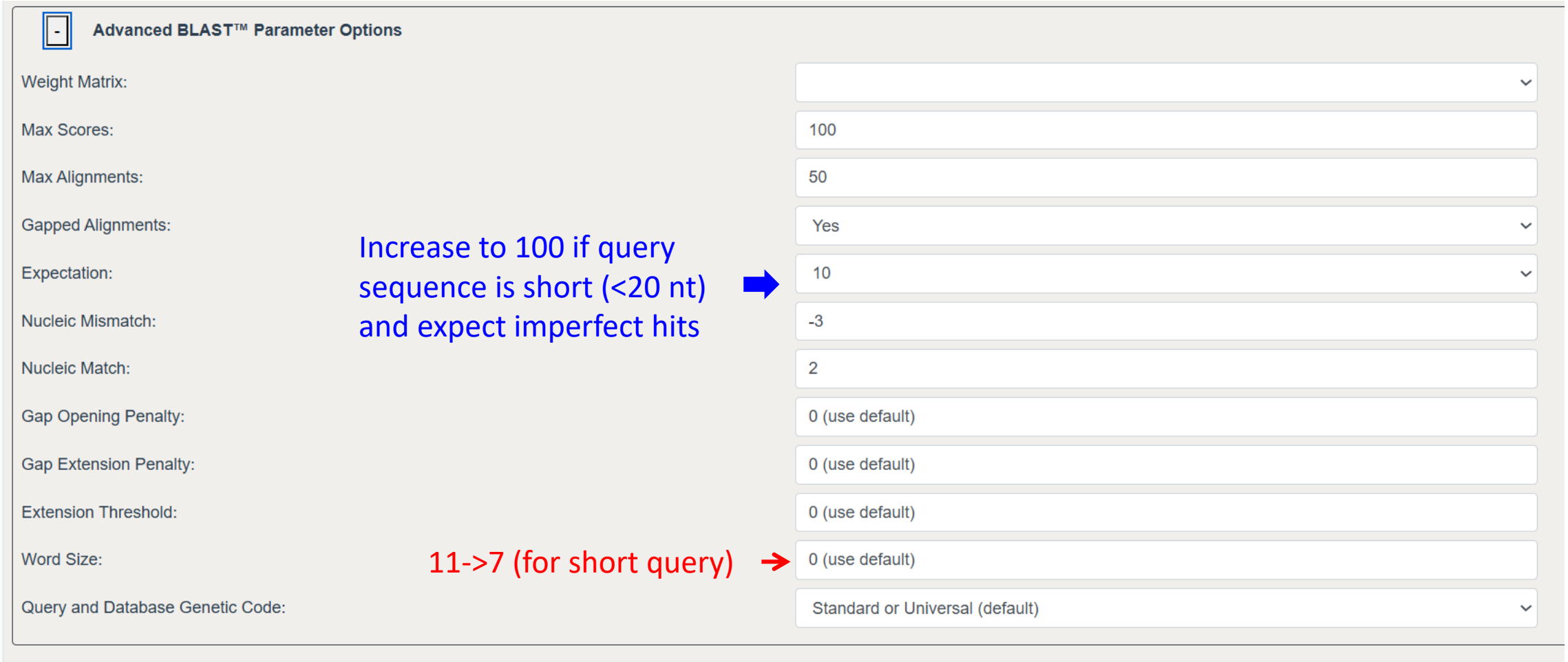
- search homologs
- map transcripts
- examine the specificity of primers
- …..

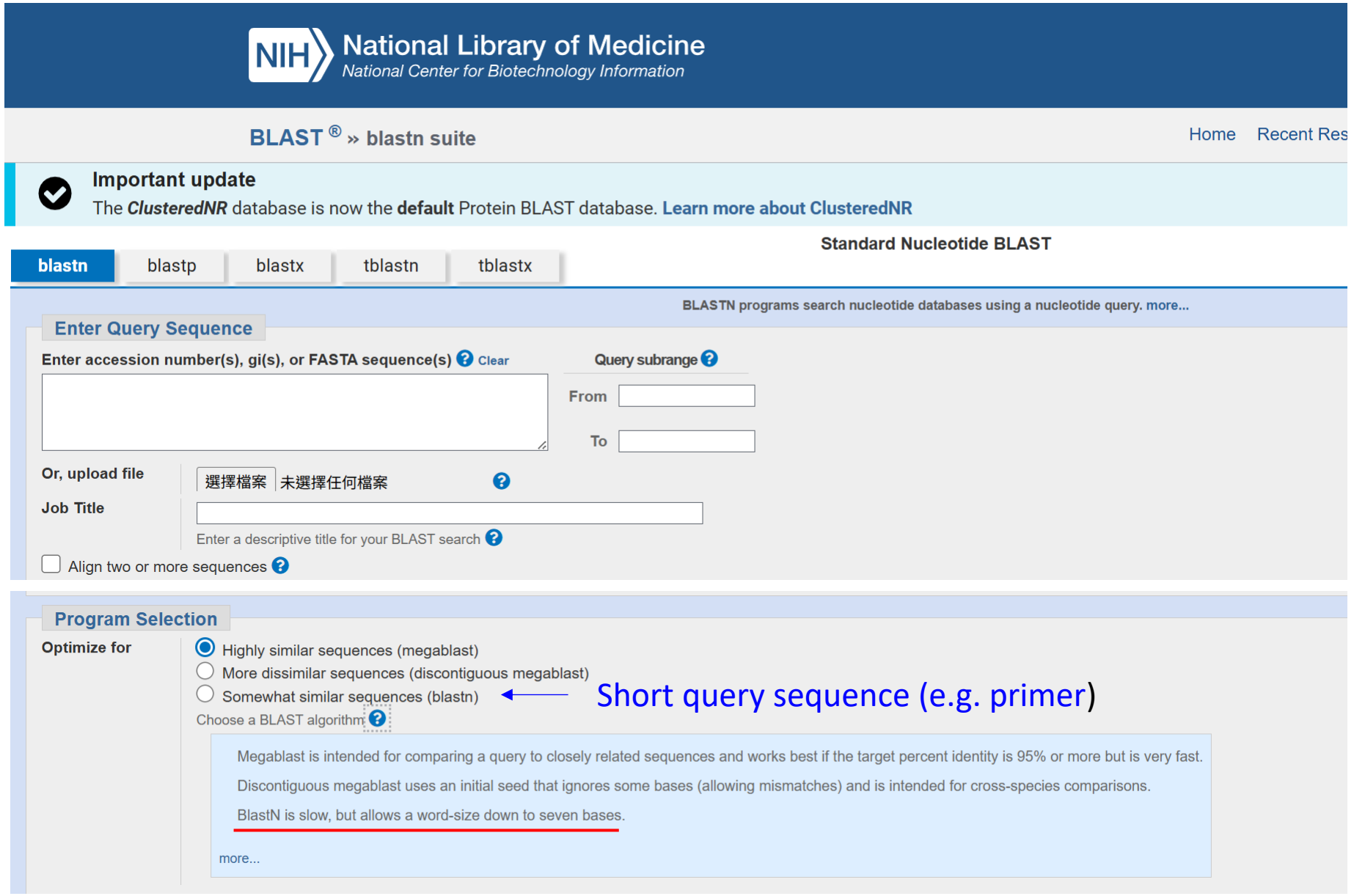
Blast : homolog search

Protein search is more sensitive than nucleotide search
Use BLASTP or BLASTX for searching homologs



Blast search: short sequences (e.g. primers)




National Library of Medicine
National Center for Biotechnology Information
BLAST® » blastn suite
Home
Recent Res
Important update
The ClusteredNR database is now the default Protein BLAST database. Learn more about ClusteredNR
Standard Nucleotide BLAST
blastn
blastp
blastx
tblastn
tblastx
BLASTN programs search nucleotide databases using a nucleotide query. more...
Enter Query Sequence
Enter accession number(s), gi(s), or FASTA sequence(s)
Clear
Query subrange
From
To
Or, upload file
選擇檔案
未選擇任何檔案
Job Title
Enter a descriptive title for your BLAST search
Align two or more sequences
Program Selection
Optimize for
Highly similar sequences (megablast)
More dissimilar sequences (discontiguous megablast)
Somewhat similar sequences (blastn)
Short query sequence (e.g. primer)
Choose a BLAST algorithm
Megablast is intended for comparing a query to closely related sequences and works best if the target percent identity is 95% or more but is very fast.
Discontiguous megablast uses an initial seed that ignores some bases (allowing mismatches) and is intended for cross-species comparisons.
BlastN is slow, but allows a word-size down to seven bases.
more...

Mutants and phenotypes

Locus: AT5G03840 (TFL1) Premium Page

Germplasms 15 / 20 entries

View Complete List | Search ABRC | Search RIKEN | Search NASC

Name	Polymorphisms	Background	Images
CS6167	gl1-3 ; tfl1-1		

Phenotypes

Phenotype curated by ABRC.

- early flowering, determinate inflorescence with terminal compound floral structure consisting of 2 or 3, sometimes incomplete, flowers; more severe phenotype at higher temperatures and longer photoperiods; lacks trichomes on stems and leaves.



Name	Polymorphisms	Background	Images
CS6235	tfl1-11		

Phenotypes

Phenotype curated by ABRC.

- strong allele, early flowering, reduced plant height and increased number of rosette inflorescences; determinate inflorescences; terminal region of primary inflorescence may consist of a single flower or of 2-3 clustered flowers separated by short internodes, in which organs of the 3 outer whorls are missing or are mosaic; secondary inflorescence apices are determinate and usually converted into a single normal flower; 2-5 normal flowers/inflorescence; photoperiod affects phenotype.



Name	Polymorphisms	Background	Images
CS6237	tfl1-13		

Phenotypes

Phenotype curated by ABRC.

- strong allele; early flowering, reduced plant height, increased number of rosette inflorescences, determinate inflorescences; terminal region of primary inflorescence may consist of a single flower or of 2-3 clustered flowers separated by short internodes, in which the organs in the 3 outer whorls are missing or are mosaic; secondary inflorescence apices are determinate and usually converted into a single normal flower; 2-5 normal flowers/inflorescence; photoperiod affects phenotype.

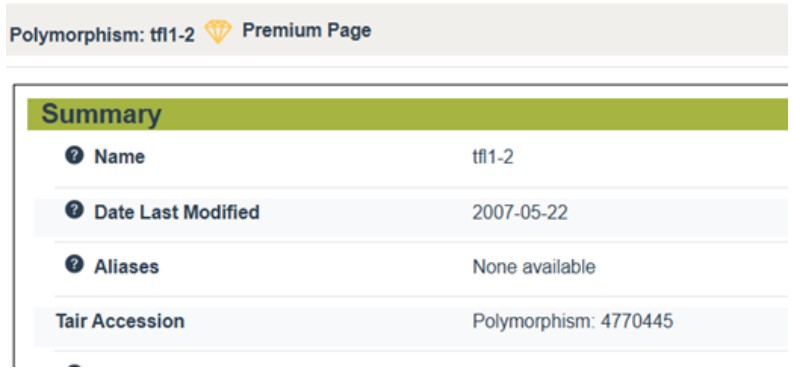


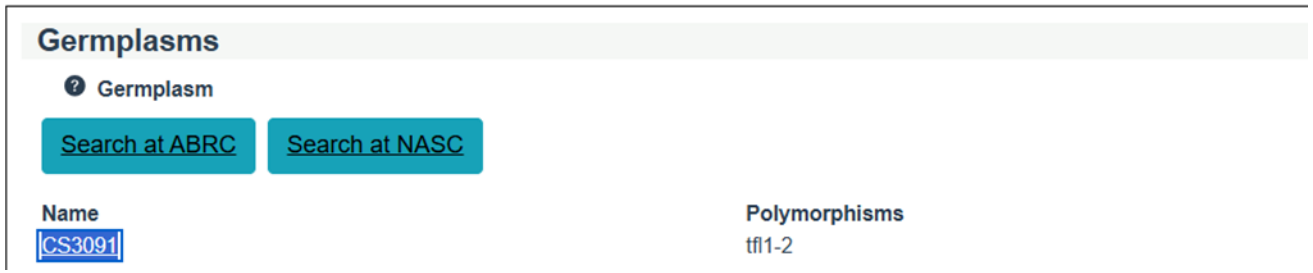
Test your hypothesis using Arabidopsis mutants

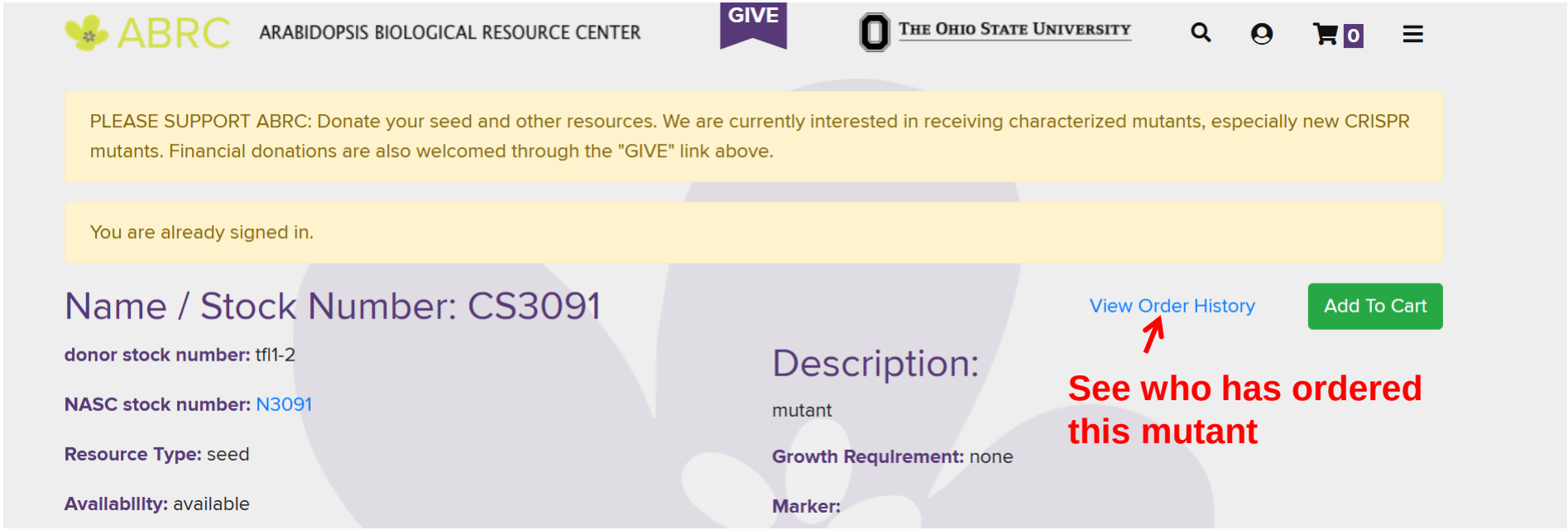


Mutant order through ABRC (not our ABRC)

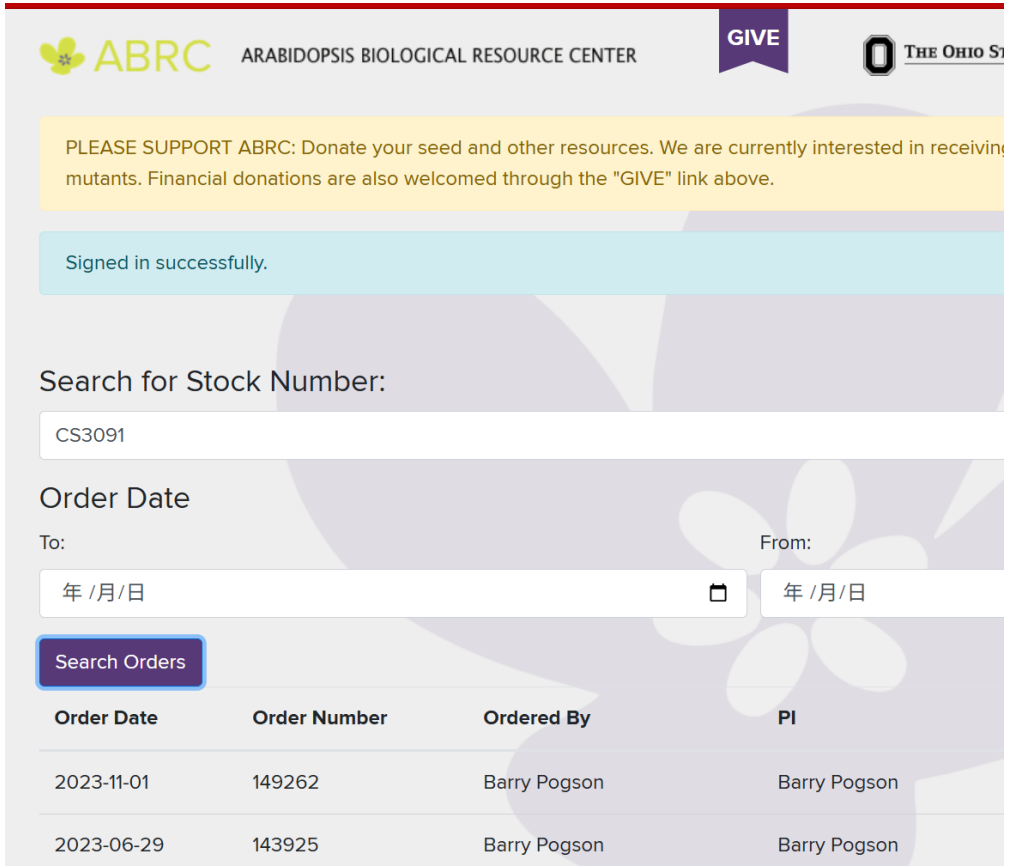
Need to apply an account for making order







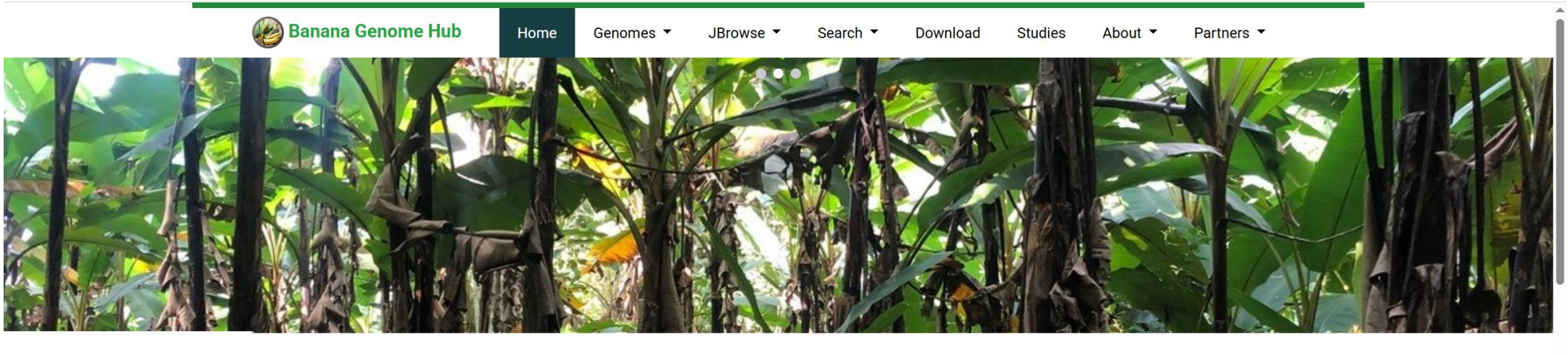
Stock order histories



2013-11-05	T104218	Chung-Mo Park	Chung-Mo Park
2013-08-16	T102573	Ariel Orellana	Ariel Orellana
2013-04-05	T99460	Vi Nguyen	Yuhai Cui
2013-03-21	T99143	Zheng Wang	Zhixi Tian
2012-12-19	T97305	Ying Wang	Paula Pijut
2012-11-14	T96529	Huazhong Ren	Huazhong Ren
2012-09-27	T95467	Hidetoshi Ikegami	Hidetoshi Ikegami
2012-05-27	T92681	Bin Liu	Bin Liu
2012-05-15	T92386	Tien-Shin Yu	Tien-Shin Yu
2012-04-11	T91510	Byeong-ha Lee	Byeong-ha Lee

Banana genome

- Banana: AA, AAA, BB, AAB, AAAB genome
- Triploid banana: sterile
- A reference A genome (DH Pahang): 2012 (v1); 2021 Telomere-to-telomere (v4)
- AAA Cavendish: 2023



Mapping genes in crops

Identify candidate genes responsible for a trait

- e.g. Disease resistance (Fusarium wilt of banana/Panama disease)

Hypothesis:
A negative regulator of plant immunity

PC

susceptible to Fusarium wilt



mutations

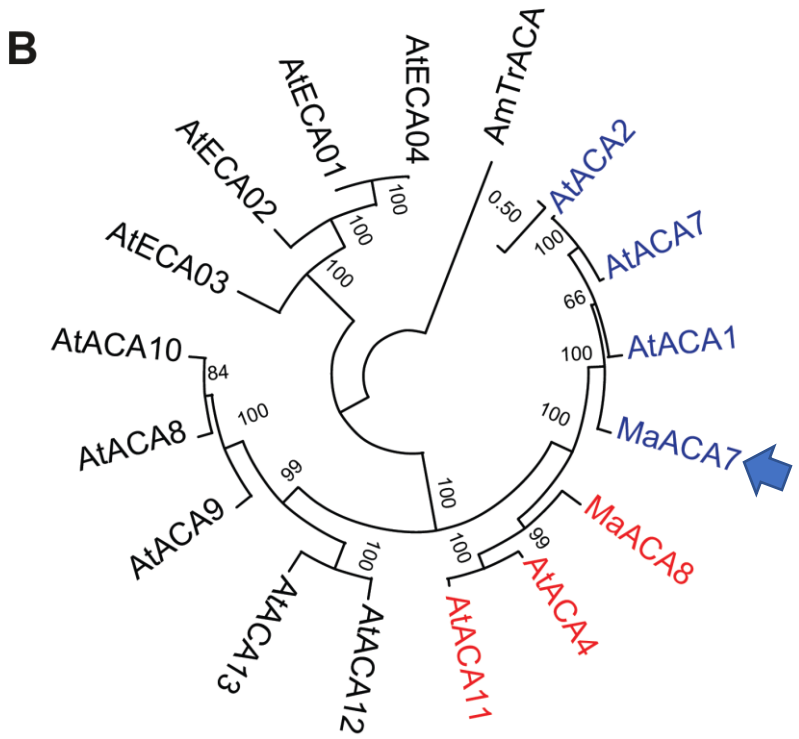
?

TC5

resistant to Fusarium wilt



Finding Arabidopsis orthologs



1. BLAST banana protein against TAIR protein database
2. Identify orthologs
3. Check the functions Arabidopsis orthologs

Locus: AT1G27770 (ACA1) Premium Page

Summary

Gene Model Type	protein_coding
Symbols	ACA1 (autoinhibited Ca2+-ATPase 1) (Primary Symbol), PEA1 (PLASTID ENVELOPE ATPASE 1)
Other Names	AT1G27770_Araport11; AT1G27770_TAIR12
Description	Encodes a chloroplast envelope Ca2+-ATPase with an N-terminal autoinhibitor.

Gene Ontology

15 unique annotations

View Annotation Details

Category	Relationship Type	Keyword
GO Biological Process	involved in	biological_process
GO Cellular Component	is active in	intracellular membrane-bounded organelle, plasma membrane
GO Cellular Component	located in	chloroplast inner membrane, endoplasmic reticulum, nucleus
GO Molecular Function	enables	P-type calcium transporter activity, calcium channel activity, calmodulin binding

Functions of Arabidopsis orthologs

- Annotations could be incomplete or incorrect
- Read key papers

Publications

1-10 of 16 entries

View Complete List

Results per page: 10

Author/Title	Source	Associated Loci	Date
Wang, S., Zhang, C., Chen, R. et al. H2S is involved in drought-mediated stomatal closure through PLD?1 in Arabidopsis	PLANTA	AT1G12480 ; AT1G27770 ; AT2G13610 ; AT2G26770 ; AT2G38170 ; AT3G15730 ; AT3G52720 ; AT3G62130 ; AT4G33950	2024
Kramer, M. C., Kim, H. J., Palos, K. R. et al. A Conserved Long Intergenic Non-coding RNA Containing snoRNA Sequences, lncCOBRA1, Affects Arabidopsis Germination and Development	FRONT PLANT SCI	AT1G05907 ; AT1G05913 ; AT1G05917 ; AT1G18080 ; AT1G27770 ; AT3G03415 ; AT3G03425 ; AT3G03445 ; AT3G05655	2022
Rahmati Ishka, M., Brown, E., Rosenberg, A. et al. Arabidopsis Ca2+-ATPases 1, 2, and 7 in the endoplasmic reticulum contribute to growth and pollen fitness	PLANT PHYSIOLOGY	AT1G27770 ; AT2G14610 ; AT2G41560 ; AT3G57260	2021
Ascencio-Ibanez, Jose Trinidad, Sozzani, Rosangela, Lee, Tae-Jin et al. Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection.	PLANT PHYSIOLOGY	AT1G01010 ; AT1G01030 ; AT1G01040 ; AT1G01060 ; AT1G01080 ; AT1G01090 ; AT1G01120 ; AT1G01140 ; AT1G01190 ; AT1G01290 ; AT1G01300 ; AT1G01320 ; AT1G01340 ; AT1G01350 ; AT1G01430 ; AT1G01470 ; AT1G01550 ; AT1G01560 ; AT1G01620 ; AT1G01725 ; ... more	2008
Mahé, Laetitia, Combes, Marie-Christine, Lashermes, Philippe et al. Comparison between a coffee single copy chromosomal region and Arabidopsis duplicated counterparts evidenced high level synteny between the coffee genome and the ancestral Arabidopsis genome.	PLANT MOLECULAR BIOLOGY	AT1G01110 ; AT1G01140 ; AT1G03800 ; AT1G04080 ; AT1G04150 ; AT1G04160 ; AT1G04190 ; AT1G04210 ; AT1G04220 ; AT1G04440 ; AT1G05460 ; AT1G08840 ; AT1G16800 ; AT1G19440 ; AT1G22610 ; AT1G25450 ; AT1G27770 ; AT1G30270 ; AT1G34320 ; AT1G51570 ; ... more	2007

Functions of Arabidopsis orthologs

Plant Physiology®

Issues | Advance Articles | More Content ▾ | Submit ▾ | Purchase | About ▾ | Plant Physiology

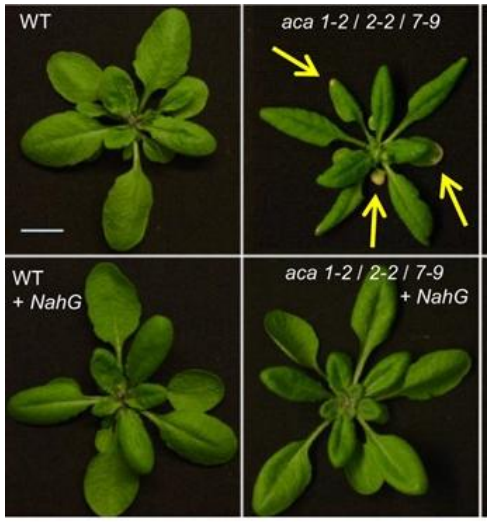


Volume 185, Issue 4

JOURNAL ARTICLE

Arabidopsis Ca^{2+}-ATPases 1, 2, and 7 in the endoplasmic reticulum contribute to growth and pollen fitness FREE

Maryam Rahmati Ishka ✉, Elizabeth Brown, Alexa Rosenberg, Shawn Romanowsky, James A Davis, Won-Gyu Choi, Jeffrey F Harper ✉ Author Notes



SA-dependent rosette size reductions and leaf lesions

↓

A negative regulator of plant immunity

Abstract

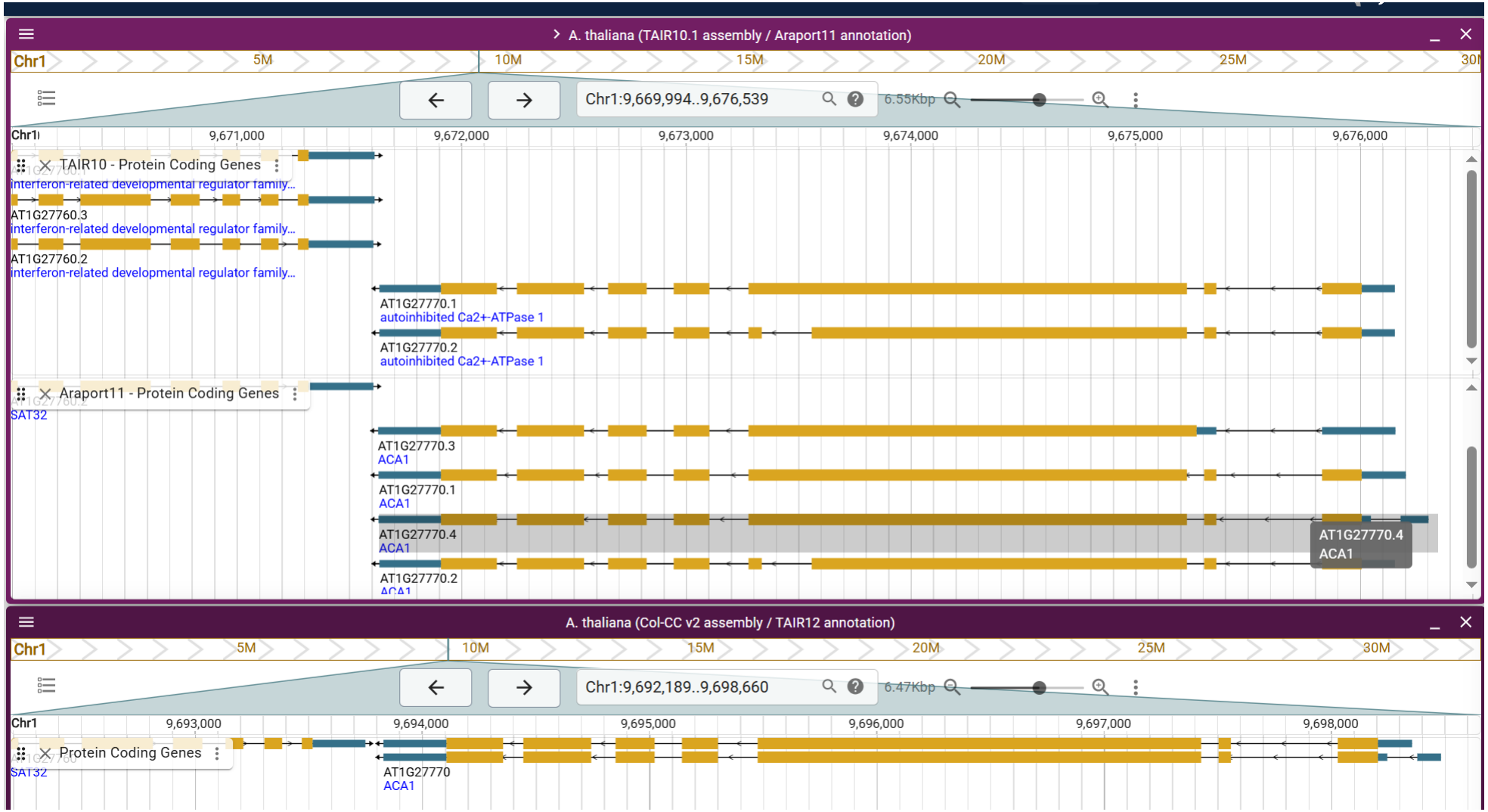
Generating cellular Ca^{2+} signals requires coordinated transport activities from both Ca^{2+} influx and efflux pathways. In Arabidopsis (*Arabidopsis thaliana*), multiple efflux pathways exist, some of which involve Ca^{2+}-pumps belonging to the Autoinhibited Ca^{2+}-ATPase (ACA) family. Here, we show that ACA1, 2, and 7 localize to the endoplasmic reticulum (ER) and are important for plant growth and pollen fertility. While phenotypes for plants harboring single-gene knockouts (KOs) were weak or undetected, a triple KO of *aca1/2/7* displayed a 2.6-fold decrease in pollen transmission efficiency, whereas inheritance through female gametes was normal. The triple KO also resulted in smaller rosettes showing a high frequency of lesions. Both vegetative and reproductive phenotypes were rescued by transgenes encoding either ACA1, 2, or 7, suggesting that all three isoforms are biochemically redundant. Lesions were suppressed by expression of a transgene encoding NahG, an enzyme that degrades salicylic acid (SA). Triple KO mutants showed elevated mRNA expression for two SA-inducible marker genes, *Pathogenesis-related1 (PR1)* and *PR2*. The *aca1/2/7* lesion phenotype was similar but less severe than SA-dependent lesions associated with a double KO of vacuolar pumps *aca4* and *11*. Imaging of Ca^{2+} dynamics triggered by blue light or the pathogen elicitor flg22 revealed that *aca1/2/7* mutants display Ca^{2+} transients with increased magnitudes and durations. Together, these results indicate that ER-localized ACAs play important roles in regulating Ca^{2+} signals, and that the loss of these pumps results in male fertility and vegetative growth deficiencies.

Gene models differ between different versions

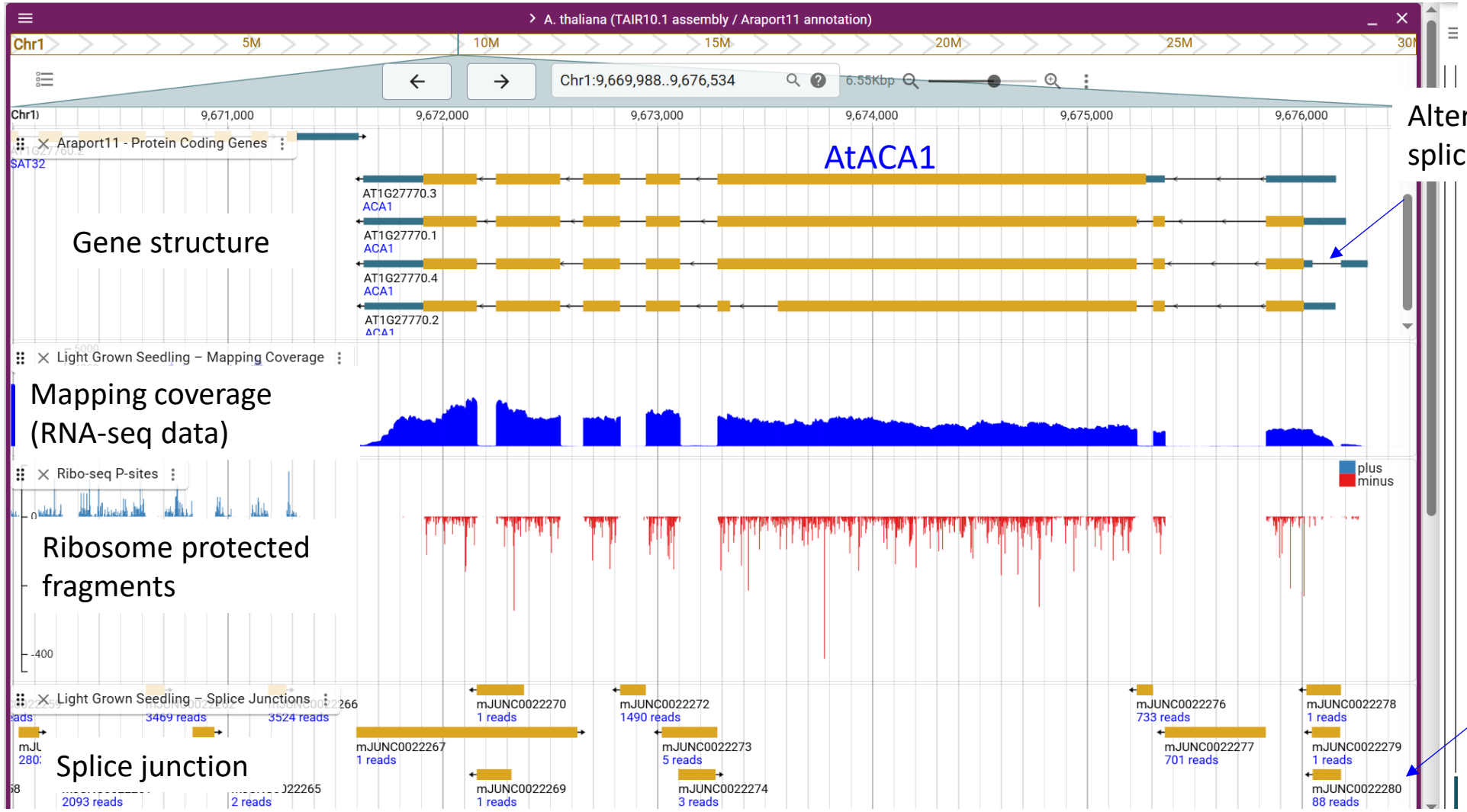
TAIR 10
4 isoforms

Araport 11
2 isoforms

TAIR 12
2 isoforms

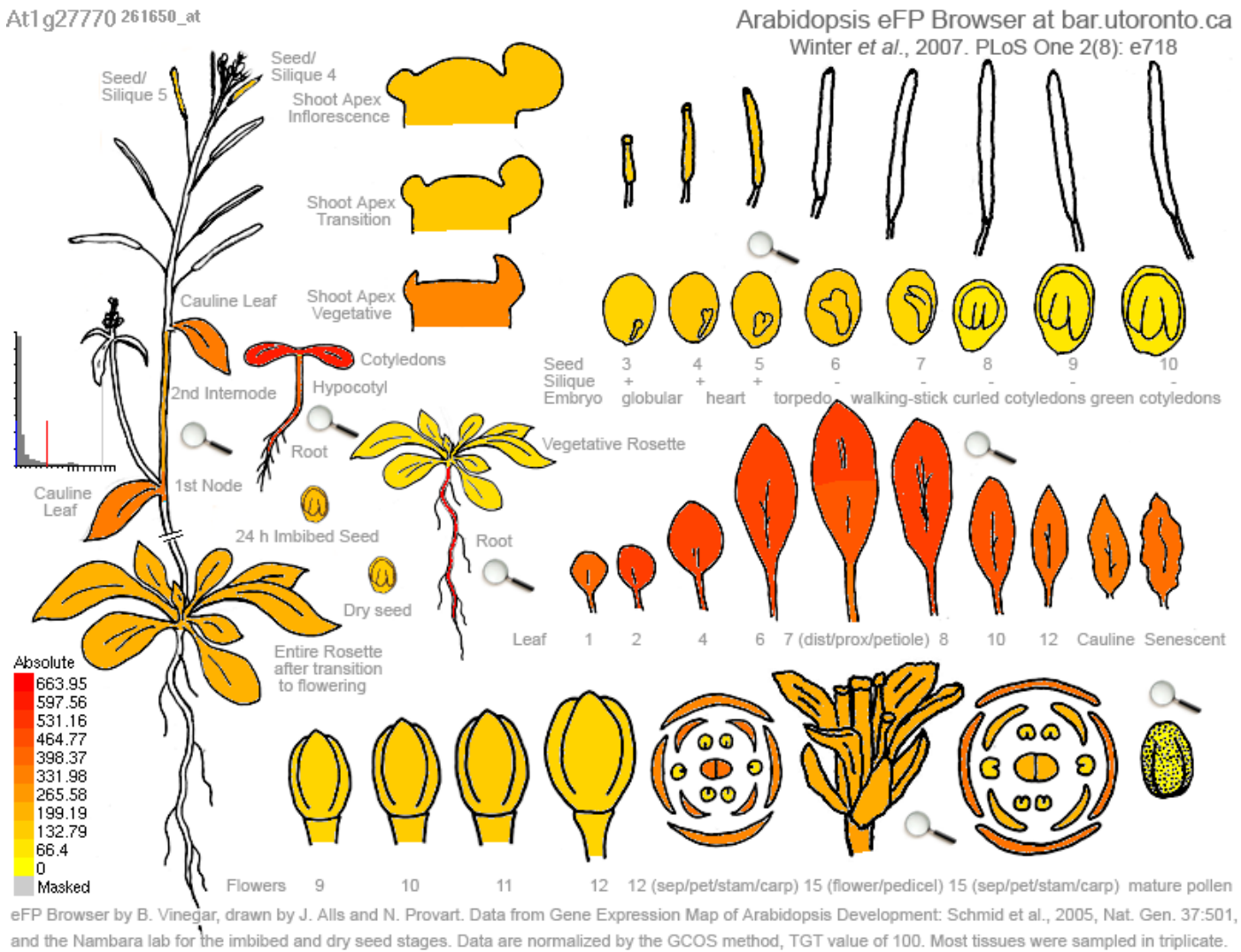


Gene structure and alternative splicing

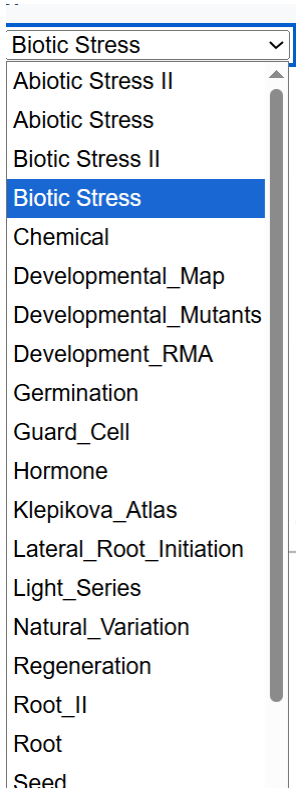


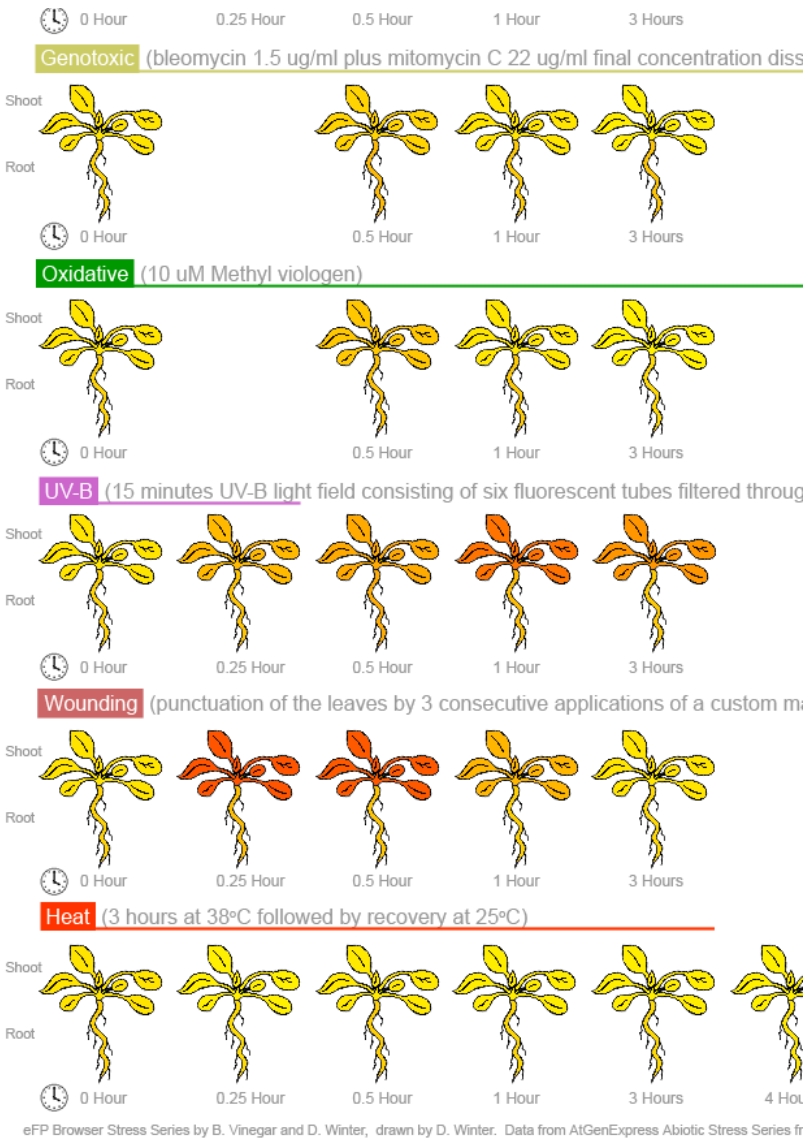
Wrong gene models can cause problems when you design primers for mRNA measurement.

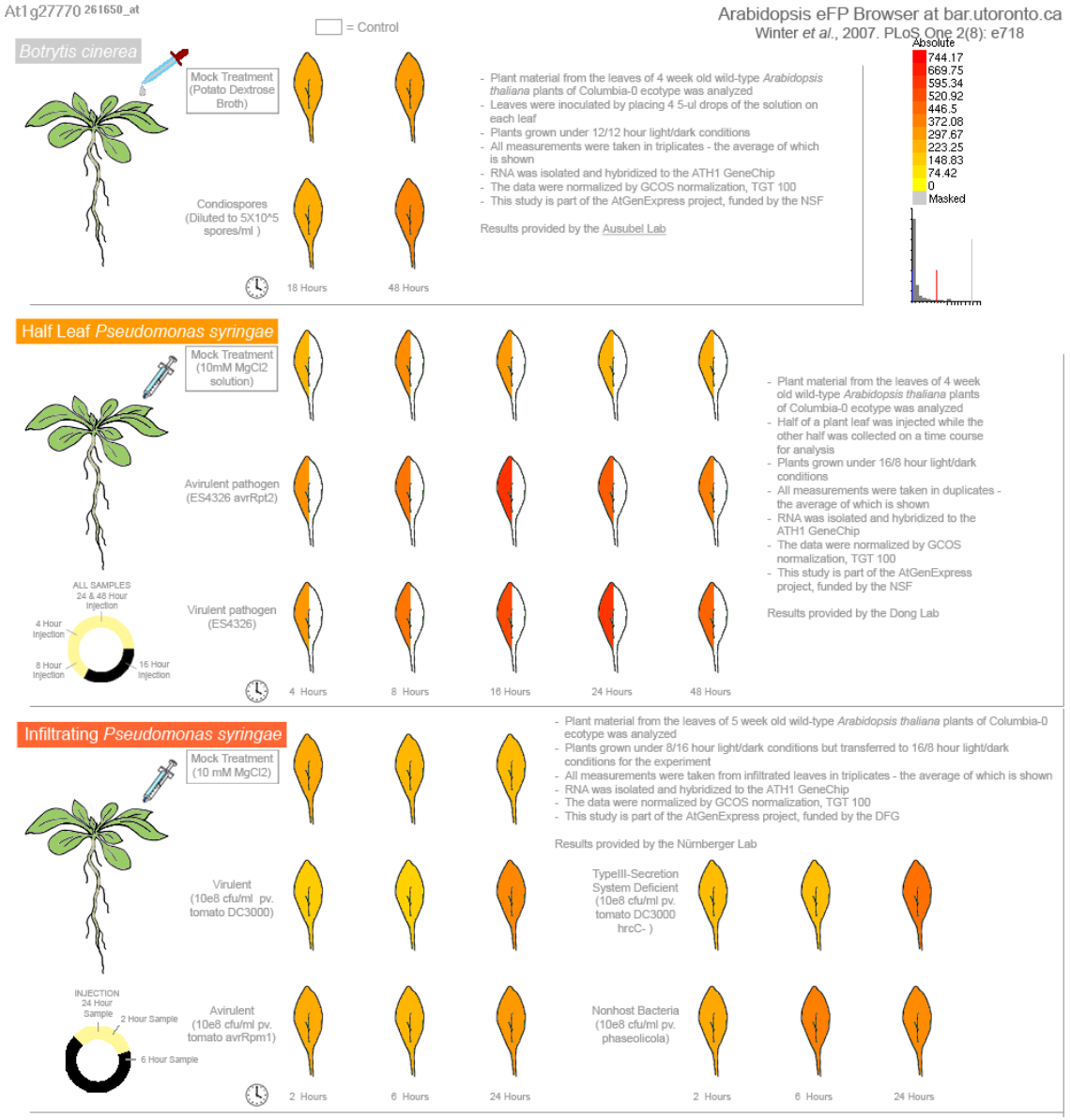
Locus page: Development map expression



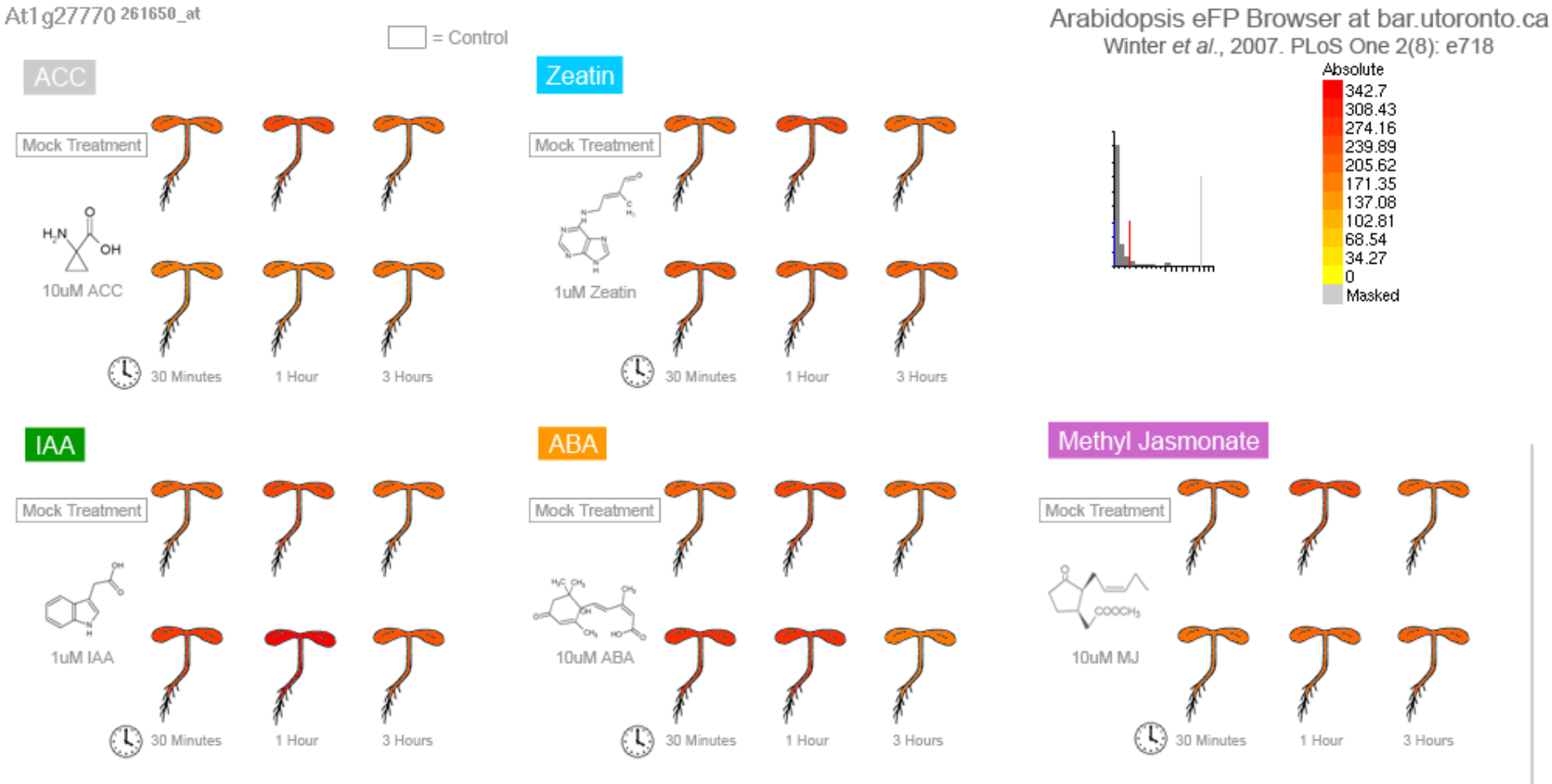
Abiotic and biotic stresses



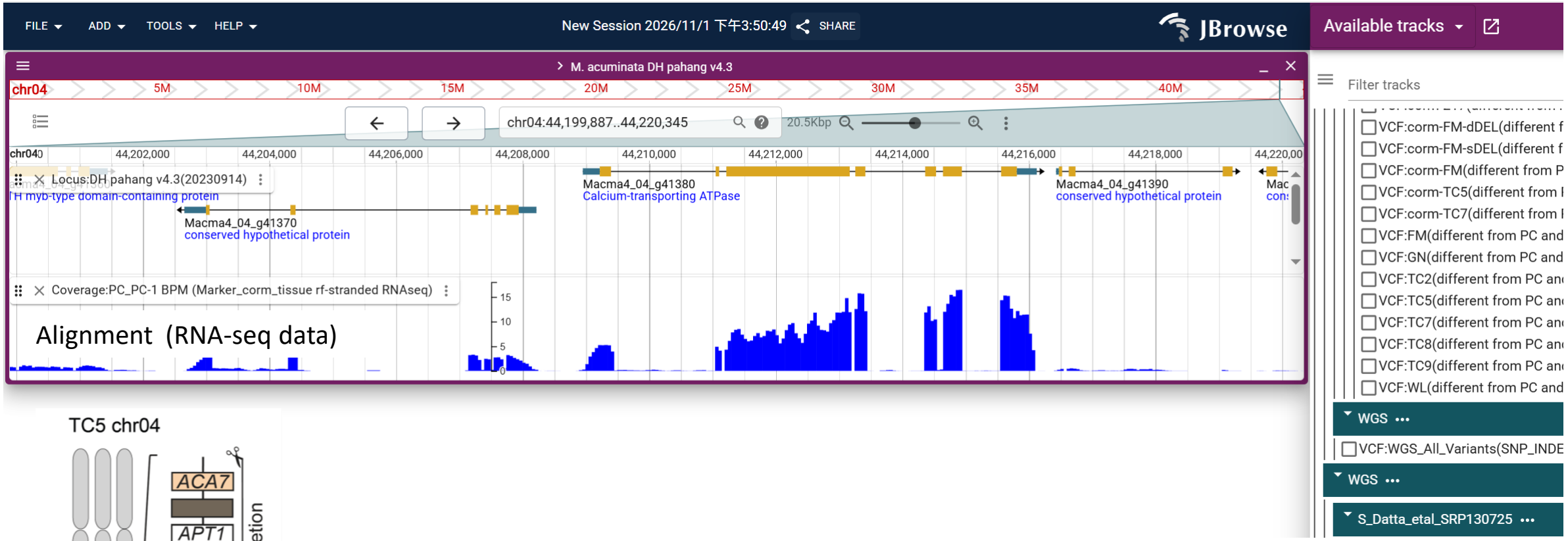


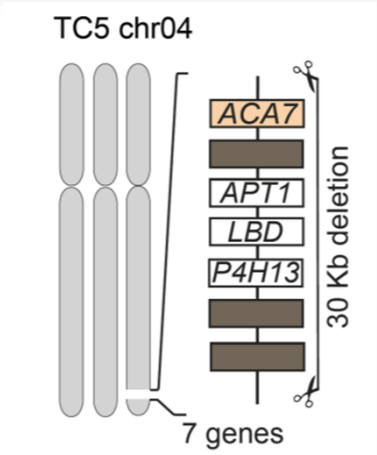


Hormone response



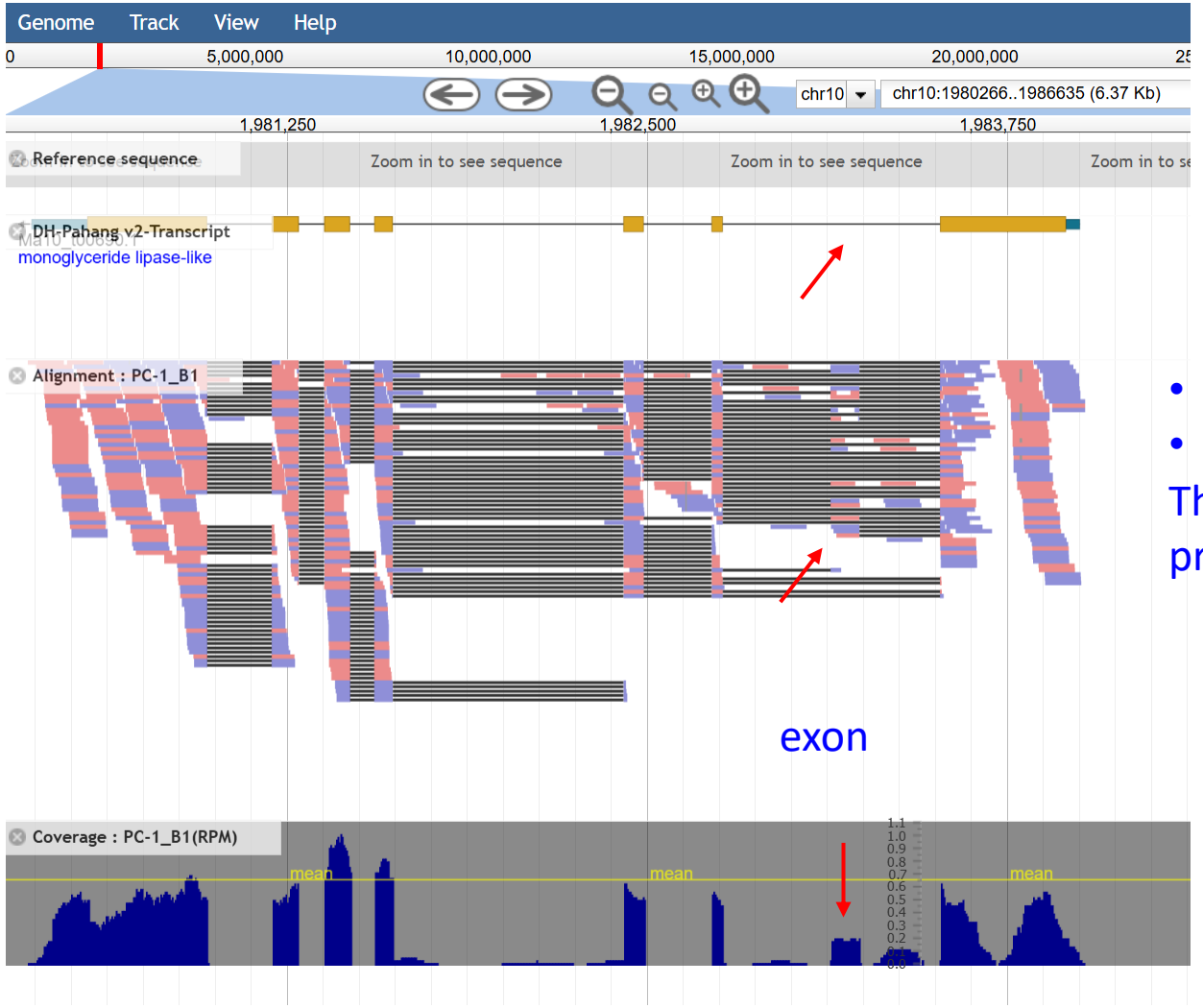
In-house Jbrowse of banana: gene model





In-house Jbrowse of banana: gene model

JBrowse not JBrowse2



Alignment
(RNA-seq data)

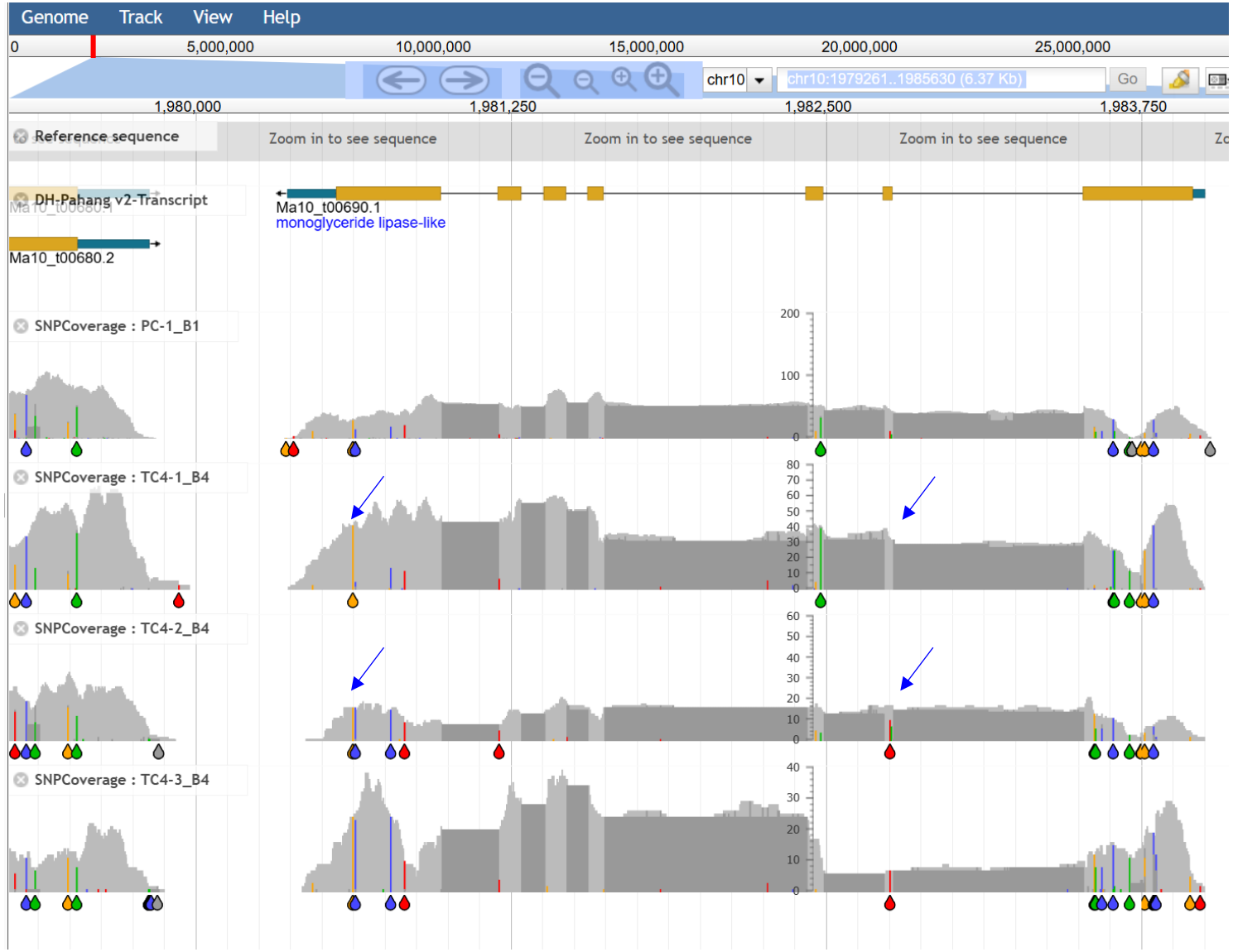
Mapping coverage
(RNA-seq data)

- Alternative splicing
- AAA genome

Three copies may not produce identical transcripts

In-house Jbrowse of banana: sequence polymorphism

Different colors indicate different bases



aca-mediated TR4 resistance in Cavendish banana

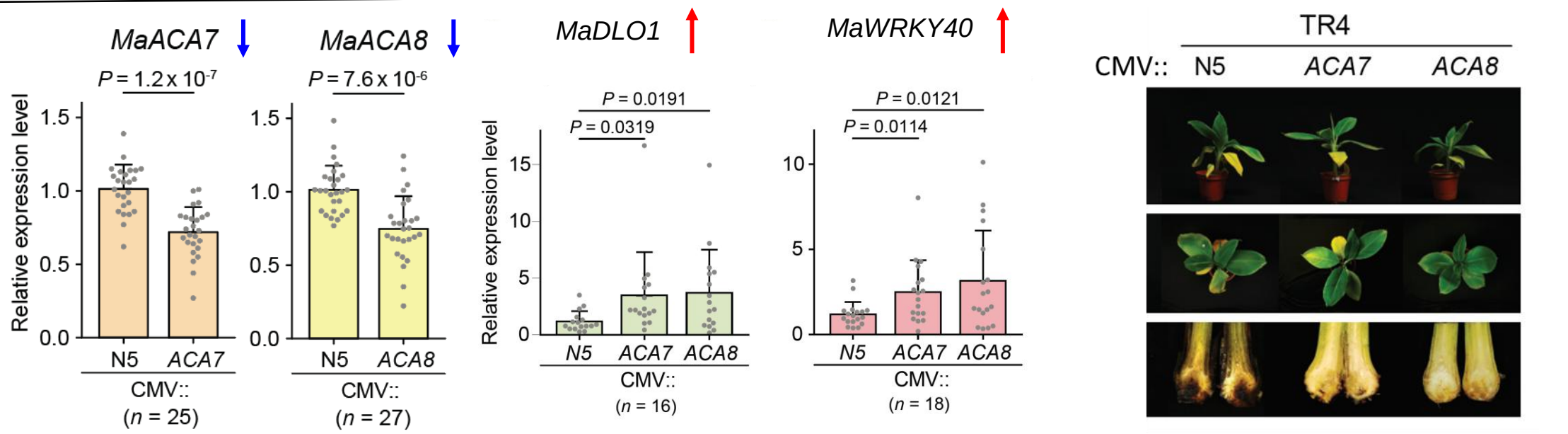
Virus-induced gene silencing (VIGS) in banana

(Dr. Hsin-Hung Yeh's group)

Banana *ACA* genes

Banana SA marker genes

TR4 disease symptom



Singh et al. (2025) PNAS

in the susceptible cultivar

3 suggestions for next-generation researchers

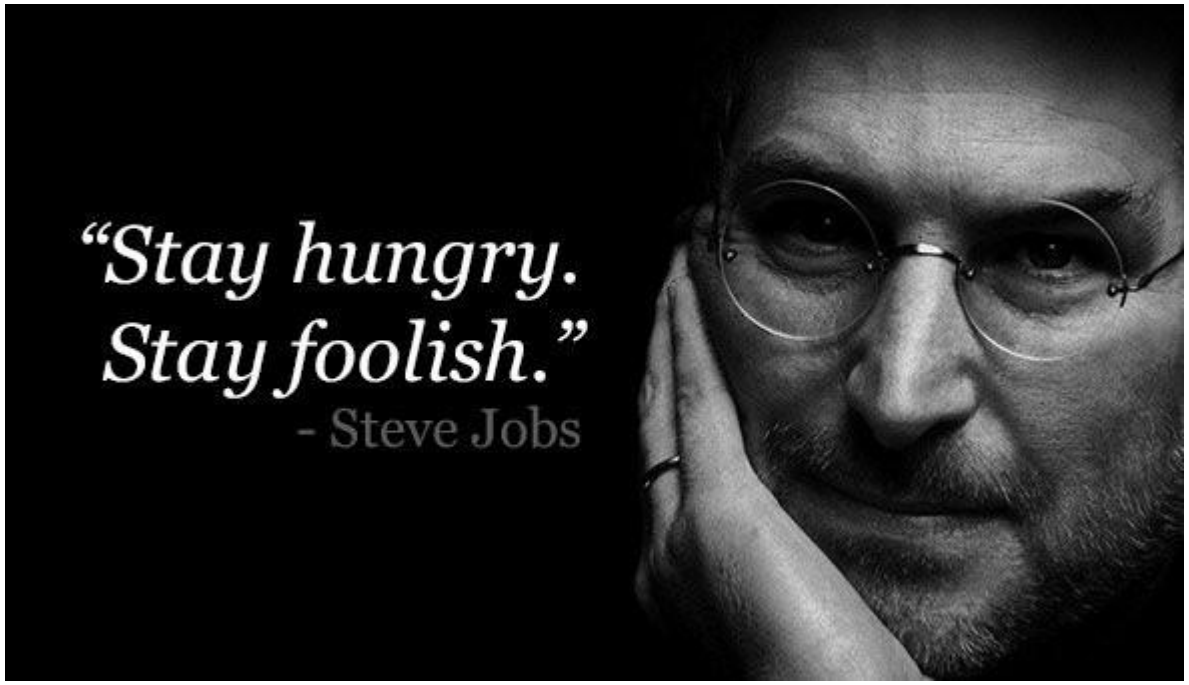
Play like a kindergarten child (no need to read the whole manual)

Warning: For molecular biology experiments, DO read the whole protocol carefully



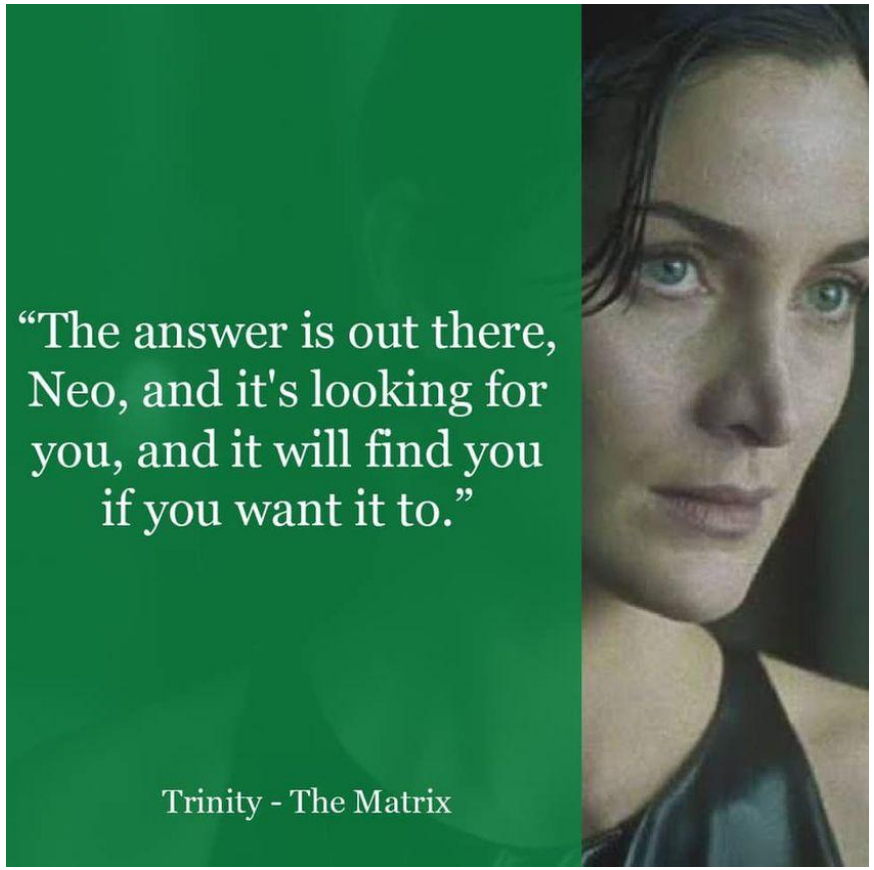
https://tw.123rf.com/

Suggestions for next-generation researchers



(ask as many questions as possible; it is free)

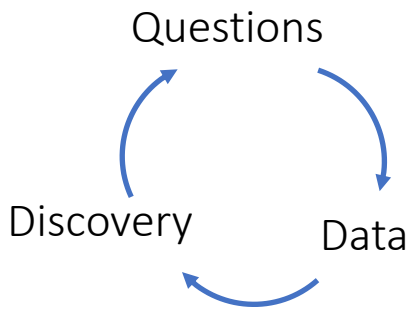
Suggestions for next-generation researchers



https://www.facebook.com/thematrixquotes/

Questions for TAIR

- **To what extent is known for the gene of your interest**?
  - DNA, RNA, protein sequences, isoforms
  - Functions and cellular localization (predicted, supported by experimental evidence)
  - Gene expression (spatial, temporal, response to stress, hormone ...)
  - Others: mutants, small RNAs, RNA degradation, ribosome footprints and DNA methylation
- **Are there other genes having similar features as the gene of your interest?**
  - Similar DNA, RNA, protein sequences
  - Similar functions
  - Similar expression pattern
  - ....

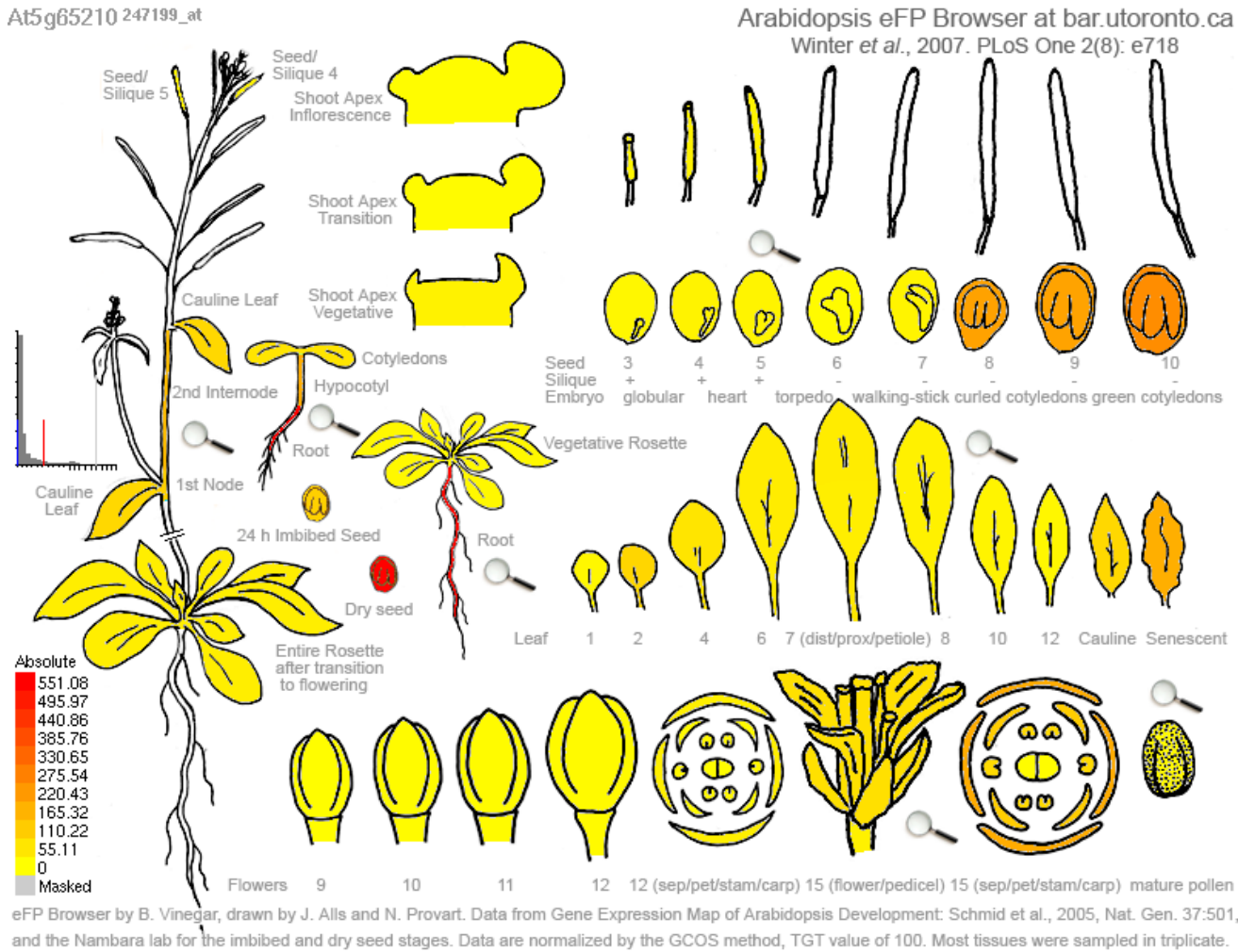




Exercise II

TGA1 (AT5G65210)

1. GO biological process: number of keywords?; number of loci in each keyword?
2. In annotation, what functions are supported by experimental results?
3. RNA expression: highest in which tissue, at what stage, in which ecotype? diurnal gene? response to any stress or hormone?

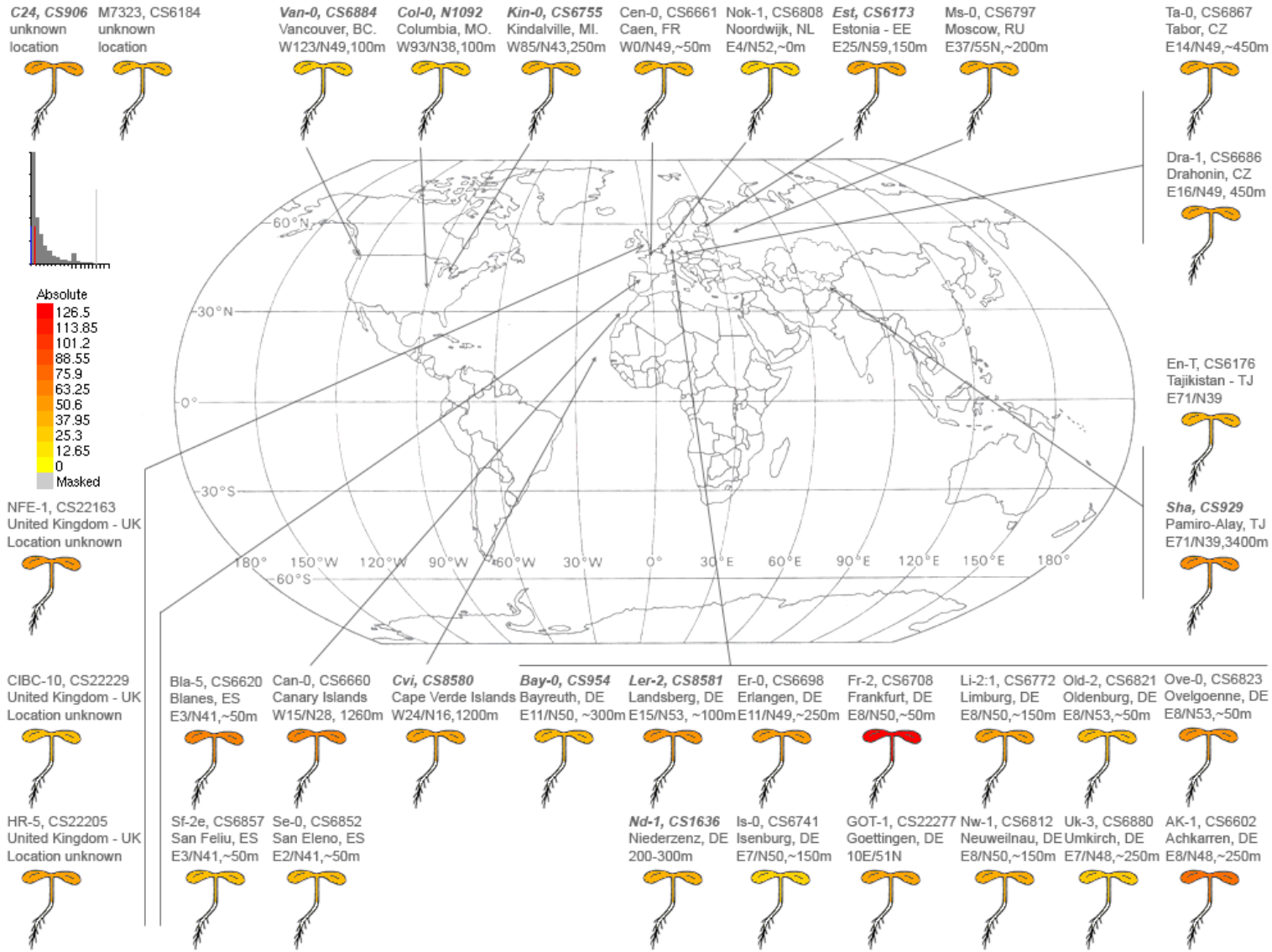


eFP Browser by B. Vinegar, drawn by J. Alls and N. Provart. Data from Gene Expression Map of Arabidopsis Development: Schmid et al., 2005, Nat. Gen. 37:501, and the Nambara lab for the imbibed and dry seed stages. Data are normalized by the GCOS method, TGT value of 100. Most tissues were sampled in triplicate.

At5g65210 247199_at

Arabidopsis eFP Browser at bar.utoronto.ca
Winter *et al*., 2007. PLoS One 2(8): e718

Natural Variation eFP Browser. Data from the Weigel Lab (Lempe *et al*., 2005, PLoS Genetics 1:e6). Aerial parts of 4 day old seedlings greenhouse-grown in soil at 23C under continuous light were sampled. Data normalized by the GCOS method, TGT value of 100. Plant material sampled in triplicate where ***indicated by italics***, otherwise once.





Thank you !

For questions, you can contact

Ho-Ming Chen 陳荷明

Office: Agricultural Technology Building A638
Email: homing@as.edu.tw
Phone: 2787-2073