

TAIR, PMN, SGN and Gramene workshop

Focus on comparative genomics and
new tools

Philippe Lamesch, A. S. Karthikeyan, Aureliano Bombarely Gomez, Pankaj Jaiswal

TAIR

- Genome Browsers
- Synteny Viewer
- Protein interaction Viewer

Tools at TAIR



The screenshot shows the TAIR website interface. At the top left is the TAIR logo, followed by navigation links: Home, Help, Contact, About Us, and Login/Register. A search bar is located at the top right. Below the navigation is a horizontal menu with tabs for Search, Browse, Tools, Stocks, Portals, Download, Submit, and News. The 'Tools' tab is selected, and a dropdown menu is open, listing the following tools: Tools Overview, GBrowse, Synteny Viewer, Seqviewer, Mapviewer, AraCyc Metabolic Pathways, BLAST, WU-BLAST, FASTA, Patmatch, Motif Analysis, Vistasight, Java Tree View, Bulk Data Retrieval, Chromosome Map Tool, Restriction Analysis, Gene Symbol Registry, and Textpresso Full Text. The main content area is divided into three columns. The left column features 'The Arabidopsis Information Resource' and 'The Arabidopsis Biological Resource Project'. The middle column contains a large, semi-transparent dropdown menu. The right column has a 'Breaking News' section with three articles: 'Synteny Viewer At TAIR [July 7, 2009]', 'TAIR9 Genome Release [June 19, 2009]', and 'Tips for searching DNA stocks including vectors and amiRNA clones [May 20, 2009]'. At the bottom, there is a red banner for 'TAIR at the Arabidopsis Conference' with details for two workshops.

Tools

- Tools Overview
- GBrowse
- Synteny Viewer
- Seqviewer
- Mapviewer
- AraCyc Metabolic Pathways
- BLAST
- WU-BLAST
- FASTA
- Patmatch
- Motif Analysis
- Vistasight
- Java Tree View
- Bulk Data Retrieval
- Chromosome Map Tool
- Restriction Analysis
- Gene Symbol Registry
- Textpresso Full Text

The Arabidopsis Information Resource

The Arabidopsis Information Resource (TAIR) is the primary source of genetic and molecular biology data for the model high-throughput genome sequencing project. The complete genome sequence, gene expression, DNA, and sequence annotations, and publications, and information on data is updated every two weeks with new data submissions. Gene structure and manual methods as well as computational methods are provided. Extensive links to other resources are provided.

The Arabidopsis Biological Resource Project

The Arabidopsis Biological Resource Project (ABRC) preserves and distributes seed, DNA, and other biological resources. Stock information and ordering information is available at the ABRC website.

TAIR is located at the University of California, Davis, Department of Plant Biology

Breaking News

Synteny Viewer At TAIR [July 7, 2009]
A synteny viewer, comparing syntenic regions between *A. thaliana* and *A. lyrata*, is now available at TAIR. More genomes will be added soon.

TAIR9 Genome Release [June 19, 2009]
The TAIR9 genome release is now available at TAIR and MCB, with 262 new loci, updates to 1254 gene structures and 739 new splice variants.

Tips for searching DNA stocks including vectors and amiRNA clones [May 20, 2009]
Are you searching for clones and vectors available from ABRC? Here are some tips for finding them using the TAIR searches and ABRC catalog pages. Specific instructions for

TAIR at the Arabidopsis Conference

Workshop I:
How to effectively use the tools and resources at TAIR to enhance your research: July 20 12:20-2pm Room 318 A&B

Workshop II:
TAIR, PMN, Gramene and SGN workshop: focus on comparative genomics and new tools: July 20 7-8:30pm Room 318 A&B

Genome Browsers at TAIR

- Two options (Seqviewer & GBrowse)

SeqViewer Close-up View | SeqViewer Home | Release Note | Print Version

Annotations shown: PERL0087243, PERL0087312, PERL0087308, PERL0087522, PERL0087647, PERL0087846, PERL0087258, PERL0087341, PERL0087432, PERL0087572, PERL0087713, PERL0087953, PERL0087259, PERL0087342, PERL0087434, PERL0087573, PERL0087714, PERL0087957.

Gene models shown: AT1G31319.1, AT1G31340.1, AT1G31355.1, AT1G31370.1, AT1G31320.1, AT1G31350.1, AT1G31358.1, AT1G31380.1, AT1G31390.1, AT1G31330.1, AT1G31360.1.

Transcripts shown: 000086_1061_3294, AY60405TR, RAFL14-58-008 (5'), AYATP75TR, 0005847, GSL7S1L7Z003_206460 (5'), RAFL14-58-008, GSL7S1L362D11, GSL7S1L362D11, 568190, 109156 (3'), RAFL19-70-104 (5'), 11121444, AYATP80TR, GSL7S1L362D11, 119623.

Control Panel (Left):
 Markers
 Polymorphisms
 T-DNA/Tn
 Gene Models
 Transcripts
 Annotation Units

 show 3 data rows
 show all data help

Buttons: show legend, Find, Go, List Genes In Range

Closeup View #1, Chromosome 1 | Remove

Zoom to: 40 Kb

Search by name (e.g. UFD): [] [Find] [Go]

Select range (e.g. 1500-2500): [] [Go]

Annotations shown: PERL0087243, PERL0087312, PERL0087308, PERL0087522, PERL0087647, PERL0087846, PERL0087258, PERL0087341, PERL0087432, PERL0087572, PERL0087713, PERL0087953, PERL0087259, PERL0087342, PERL0087434, PERL0087573, PERL0087714, PERL0087957.

Gene models shown: AT1G31319.1, AT1G31340.1, AT1G31355.1, AT1G31370.1, AT1G31320.1, AT1G31350.1, AT1G31358.1, AT1G31380.1, AT1G31390.1, AT1G31330.1, AT1G31360.1.

Transcripts shown: 000086_1061_3294, AY60405TR, RAFL14-58-008 (5'), AYATP75TR, 0005847, GSL7S1L7Z003_206460 (5'), RAFL14-58-008, GSL7S1L362D11, GSL7S1L362D11, 568190, 109156 (3'), RAFL19-70-104 (5'), 11121444, AYATP80TR, GSL7S1L362D11, 119623.

Details

Annotations shown: YFPBLZ, PERL0087243, PERL0087312, PERL0087308, PERL0087522, PERL0087647, PERL0087846, PERL0087258, PERL0087341, PERL0087432, PERL0087572, PERL0087713, PERL0087953, PERL0087259, PERL0087342, PERL0087434, PERL0087573, PERL0087714, PERL0087957.

Gene models shown: AT1G31319.1, AT1G31340.1, AT1G31355.1, AT1G31370.1, AT1G31320.1, AT1G31350.1, AT1G31358.1, AT1G31380.1, AT1G31390.1, AT1G31330.1, AT1G31360.1.

Transcripts shown: 000086_1061_3294, AY60405TR, RAFL14-58-008 (5'), AYATP75TR, 0005847, GSL7S1L7Z003_206460 (5'), RAFL14-58-008, GSL7S1L362D11, GSL7S1L362D11, 568190, 109156 (3'), RAFL19-70-104 (5'), 11121444, AYATP80TR, GSL7S1L362D11, 119623.

Control Panel (Left):
 Annotations
 Search
 Details
 Reports & Analysis
 Annotate VISTA Plot
 Configure

 Annotate VISTA Plot
 Show 10 Hbp
 2000
 File

Seqviewer

TAIR SeqViewer Whole Genome View

[HELP](#)



Markers

Whole Genome View Options:

Search by name or sequence (15-150 nt)

Paste in name(s) (up to 250) or sequence(s) (up to 4): [help](#)

```
CTACTTCTCTCGGTCGTATTCGTCATTCTGCCTTTTATCTAT
CAACATTTTAGGGTTTTTACTCCTCTAAGGTTTCGCTTCTTC
TCTGTCCTCCGCTTCAGTTGGATCATCGATTTTCTCAATTGC
TCAGCCGGTGAGTTTCTTCCTTTTCACGATTTTGGTGTTTAC
```

or **upload** a file of names or sequences: [help](#)

Version

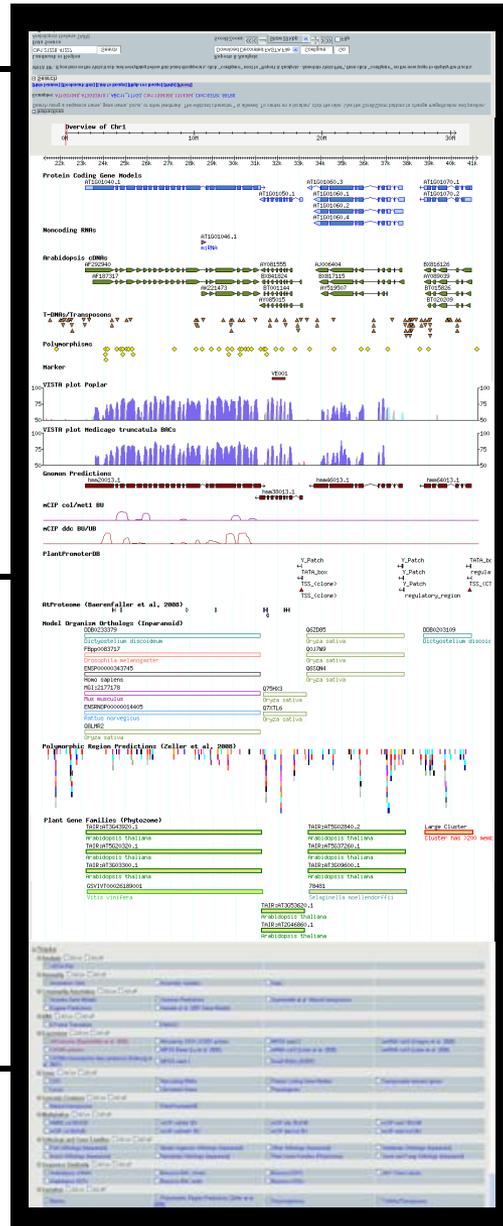


GBrowse

Header

Main Browser Window

Track Menu



GBrowse - Header

Instructions
Searching: Search using a sequence name, gene name, locus, or other landmark. The wildcard character * is allowed.
Navigation: Click one of the rulers to center on a location, or click and drag to select a region. Use the Scroll/Zoom buttons to change magnification and position.
Examples: AT1001040, AT3052910.1, ABC11_17102, Chr1:1504365..1514364, ChrC:63781..60780
[Bookmark this] [Upload your own data] [Hide banner] [Share these tracks] [Link to Image] [High-res Image] [Help] [Reset]

Search

VISTA TIP - If you turn on the Vista track and everything below this band disappears, click "configure" next to "Report & Analysis - Annotate Vista Plot", then click "configure" on the new page to display the tracks.

Landmark or Region:

Data Source: Arabidopsis thaliana TAIR6

Reports & Analysis:

Scroll/Zoom: <<< >>> Flip

- Show 1 kbp
- Show 100 kbp
- Show 200 kbp
- Show 100 kbp
- Show 40 kbp
- Show 20 kbp
- Show 10 kbp
- Show 5 kbp
- Show 2 kbp
- Show 1 kbp
- Show 200 bp
- Show 100 bp
- Show 50 bp

Search by: Vista plot & fasta download

Feature Name AT1G01040
 AY438620

Position Chr2:100500..102500

Track menu

Tracks

<input type="checkbox"/> Analysis <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> VISTA Plot			
<input type="checkbox"/> Assembly <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input checked="" type="checkbox"/> Annotation Units	<input type="checkbox"/> Assembly Updates	<input type="checkbox"/> Gaps	
<input type="checkbox"/> Community Annotation <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> Aceview Gene Models	<input checked="" type="checkbox"/> Genom Prediction	<input type="checkbox"/> Polyadenylation sites (AIPolyA-DB)	
<input type="checkbox"/> Eugene Predictions	<input type="checkbox"/> Hanada et al. 2007 Gene Models	<input type="checkbox"/> Queenville et al. Natural transposons	
<input type="checkbox"/> DNA <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> E Frame Translation <input type="checkbox"/> DNAGC			
<input type="checkbox"/> Expression <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input checked="" type="checkbox"/> Peptide (Castellana et al. 2008)	<input type="checkbox"/> CATMA transcriptome data (analysis) (Aubourg et al. 2007)	<input type="checkbox"/> MPSS seed 1	<input type="checkbox"/> Small RNAs (ASRP)
<input checked="" type="checkbox"/> AtProteome (Baerenfeller et al. 2006)	<input type="checkbox"/> Microarray ATH1-121521 probes	<input type="checkbox"/> MPSS seed 2	<input type="checkbox"/> smRNA col-0 (Gregory et al. 2006)
<input type="checkbox"/> CATMA primers	<input type="checkbox"/> MPSS flower (Lu et al. 2005)	<input checked="" type="checkbox"/> smRNA col-0 (Luster et al. 2008)	<input type="checkbox"/> smRNA col-0 (Luster et al. 2008)
<input type="checkbox"/> Gene <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> CDS	<input type="checkbox"/> Noncoding RNAs	<input checked="" type="checkbox"/> Protein Coding Gene Models	<input type="checkbox"/> Transposable element genes
<input checked="" type="checkbox"/> Locus	<input type="checkbox"/> Obsolete Genes	<input type="checkbox"/> Pseudogenes	
<input type="checkbox"/> Genomic Features <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> Natural transposons <input type="checkbox"/> PlantPromoteDB			
<input type="checkbox"/> Methylation and Phosphorylation <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> HMBD col BUUB	<input type="checkbox"/> mCIP colmet1 BU	<input type="checkbox"/> mCIP met1 BUUB	
<input type="checkbox"/> mCIP col BUUB	<input type="checkbox"/> mCIP ddc BUUB	<input type="checkbox"/> mCIP met1/col BU	
<input type="checkbox"/> mCIP col/ddc BU	<input type="checkbox"/> mCIP ddc/col BU	<input type="checkbox"/> Phosphorylation (PhosPhAt)	
<input type="checkbox"/> Orthologs and Gene Families <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> Fish Orthologs (Inparanoid)	<input type="checkbox"/> Model Organism Orthologs (Inparanoid)	<input type="checkbox"/> Other Orthologs (Inparanoid)	<input type="checkbox"/> Vertebrate Orthologs (Inparanoid)
<input type="checkbox"/> Insect Orthologs (Inparanoid)	<input type="checkbox"/> Nematode Orthologs (Inparanoid)	<input type="checkbox"/> Plant Gene Families (Phytozome)	<input type="checkbox"/> Yeast and Fungi Orthologs (Inparanoid)
<input type="checkbox"/> Sequence Similarity <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input checked="" type="checkbox"/> Arabidopsis cDNAs	<input type="checkbox"/> Brassica cDNAs	<input checked="" type="checkbox"/> Dicot Transcript Assemblies (selected)	<input type="checkbox"/> Radish Clones
<input type="checkbox"/> Arabidopsis ESTs	<input type="checkbox"/> Brassica ESTs	<input type="checkbox"/> JANY Clone Library	
<input type="checkbox"/> Brassica BAC clones	<input type="checkbox"/> Brassica GSSs	<input type="checkbox"/> Monocot Transcript Assemblies (selected)	
<input type="checkbox"/> Variation <input type="checkbox"/> All on <input type="checkbox"/> All off			
<input type="checkbox"/> Bay-0 and Shalbara polymorphisms (JG)	<input type="checkbox"/> Marker	<input type="checkbox"/> Polymorphisms	
<input type="checkbox"/> Bu-0 and Tou-1 polymorphisms (Ossowski et al. 2008)	<input type="checkbox"/> Polymorphic Region Predictions (Zeller et al. 2008)	<input type="checkbox"/> T-CNAs/Transposons	

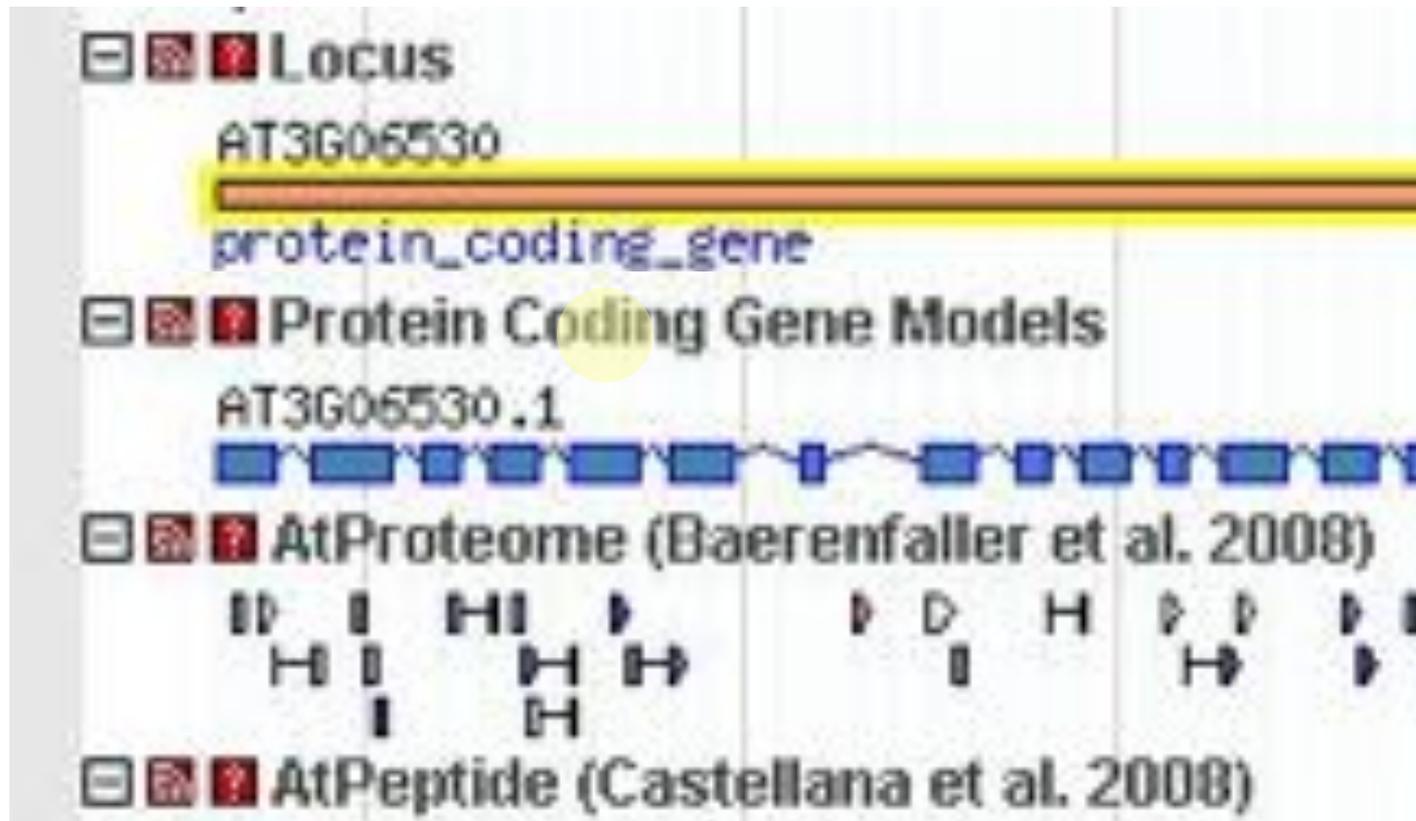
[Configure tracks](#) [Update Image](#)

Proteomics Data

- Baerenfaller 2008
- Castellana et al. 2008



Reposition tracks



Promoter Elements

Genomic Features All on All off

Natural transposons

PlantPromoterDB

plantpromoterdb 20 Arabidopsis thaliana

Top About Promoter Comparisons Index of Genes

PlantPromoterDB
PlantPromoterDB promoter information of AT1G05240.1

Summary of Gene (AT1G05240.1)

Organism	Arabidopsis thaliana
Chromosome	1
Locus	AT1G05240
Gene model	AT1G05240.1
Description	peroxidase, putative, similar to peroxidase, peroxidase ATP 11a (Arabidopsis thaliana) gi1546688 embl CAA67334

Overview

Locus

Protein Coding

Focused view (chromosome 1: 1620201-1621400)

Genome position from initiation codon: 1521101 -100

TSS from cDNA
TSS information
AT1G05240.1
Promoter sequence

5'→3' (+) ATATTCGAAAGCCGCTGGGTGCTATGAAAAGGCTGAGATGTTTCCTGATCTTCCTCCAAAGCTCGTTCACCTTACTTGGCAAG

ATATA252: ATCTATAAG
A-TATA248: ATCTATAAG
A-TATA247: ATATAAAG
A-TATA214: ATATAAAG

ATY019: TCTCGCTC
TCTCGCTC

TSS from cDNA
TSS information
Promoter sequence

3'→5' (-) ATATAAGTTTCGCGAGCCACTAGATAATTTCCGACTCTTACAGAGGAAGTAGAGTAGAGOTTTCGAGCAAGASTGGAAATGAACCCCTT

< 500bp 500bp > Reverse All Sequence for printing

TSS peak TSS TSS clone initiation codon TATA box Y patch REG

Promoter Summary of AT1G05240.1

TSS information

Type	Sequence	TFM score	Genome position	Position from initiation codon
------	----------	-----------	-----------------	--------------------------------

Tracks updated regularly

Tracks

Analysis All on All off

VISTA Plot

Assembly All on All off

Annotation Units Assembly Updates Gaps

Community Annotations All on All off

Accview Gene Models Genomon Predictions Polyadenylation sites (AIPolyA-DB)

Eugene Predictions Hanada et al. 2007 Gene Models Quesneville et al. Natural transposons

DNA All on All off

E Frame Translation DNA/GC

Expression All on All off

AtPeptide (Castellana et al. 2008) CATMA transcriptome data (analysis) (Aubourg et al. 2007) MPSS seed 1 Small RNAs (ASRP)

AtProteome (Baerenfalter et al. 2008) Microarray ATH1-121501 probes MPSS seed 2 smRNA col-0 (Gregory et al. 2008)

CATMA primers MPSS Fower (Lu et al. 2005) mRNA col-0 (Jister et al. 2008) smRNA col-0 (Jister et al. 2008)

Gene All on All off

CDS Noncoding RNAs Protein Coding Gene Models Transposable element genes

Locus Obsoleted Genes Pseudogenes

Genomic Features All on All off

Natural transposons PlantPromoterDB

Methylation and Phosphorylation All on All off

HMBD col BU/B mCIP col/met1 BU mCIP met1 BU/B

mCIP col BU/B mCIP ddc BU/B mCIP met1/col BU

mCIP col/ddc BU mCIP ddc/col BU Phosphorylation (PhosPhAt)

Orthologs and Gene Families All on All off

Fish Orthologs (Inparanoid) Model Organism Orthologs (Inparanoid) Other Orthologs (Inparanoid) Vertebrate Orthologs (Inparanoid)

Insect Orthologs (Inparanoid) Nematode Orthologs (Inparanoid) Plant Gene Families (Phytozome) Yeast and Fung Orthologs (Inparanoid)

Sequence Similarity All on All off

Arabidopsis cDNAs Brassica cDNAs Dicot Transcript Assemblies (selected) Radish Clones

Arabidopsis ESTs Brassica ESTs JAIY Clone Library

Brassica BAC clones Brassica GSSs Monocot Transcript Assemblies (selected)

Variation All on All off

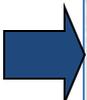
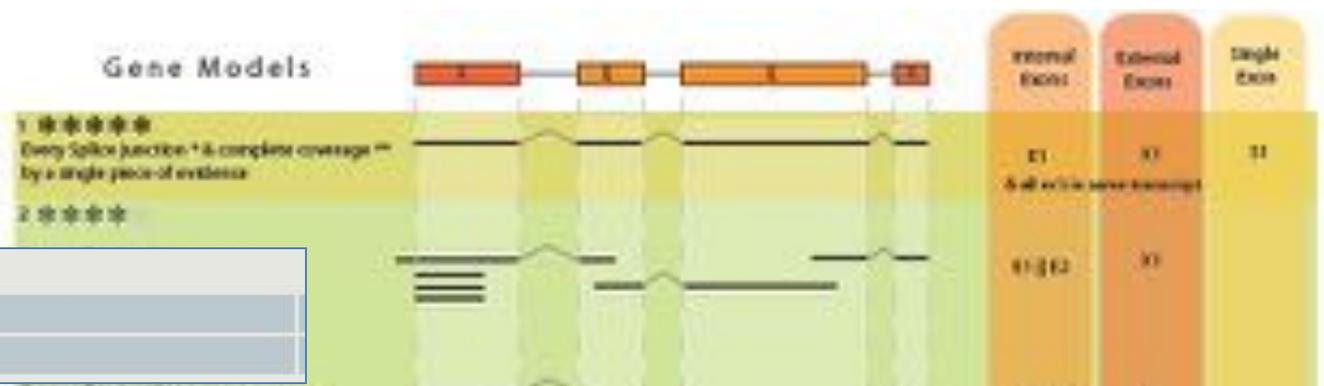
Bay-0 and Shakara polymorphisms (JGI) Marker Polymorphisms

Bur-0 and Tos-1 polymorphisms (Ossowski et al. 2008) Polymorphic Region Predictions (Zeller et al. 2008) T-DNAs/Transposons

[Configure tracks...](#) [Update Image](#)

Full support

Gene All on All off
 CDS
 Gene Confidence Ranking

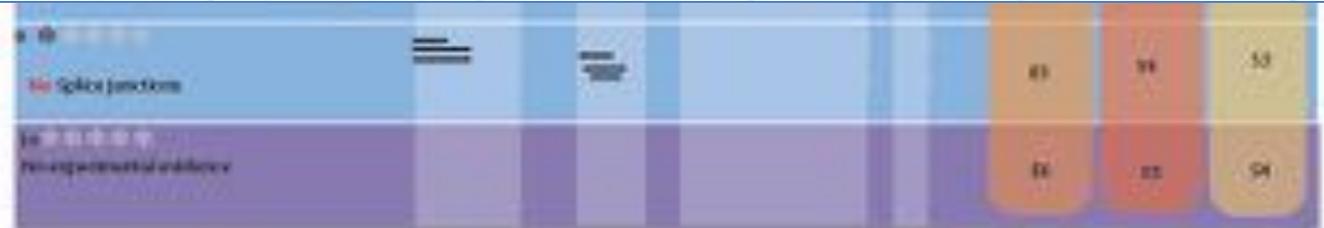


Protein Coding Gene Models
 AT1005330.1 AT1005340.1 AT1005350.1

Gene Confidence Ranking
 AT1005330.1 AT1005340.1 AT1005350.1
 * = (1 stars) **** = (4 stars) **** = (2 stars)

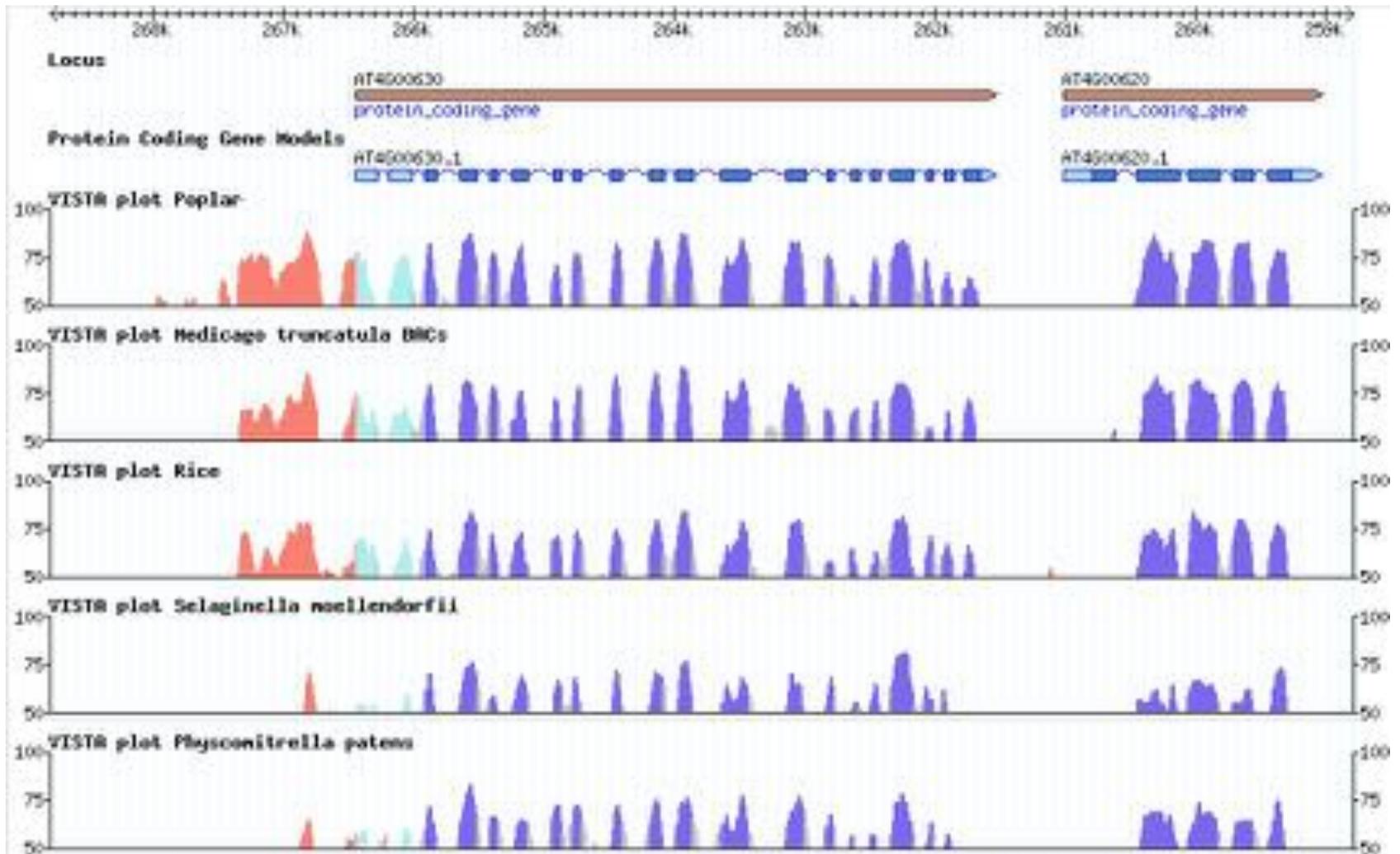
Arabidopsis cDNAs
 E0041888
 E0047567
 AY065980
 AF360144
 AF226423
 E0010526
 E0017674
 AY066369

No support

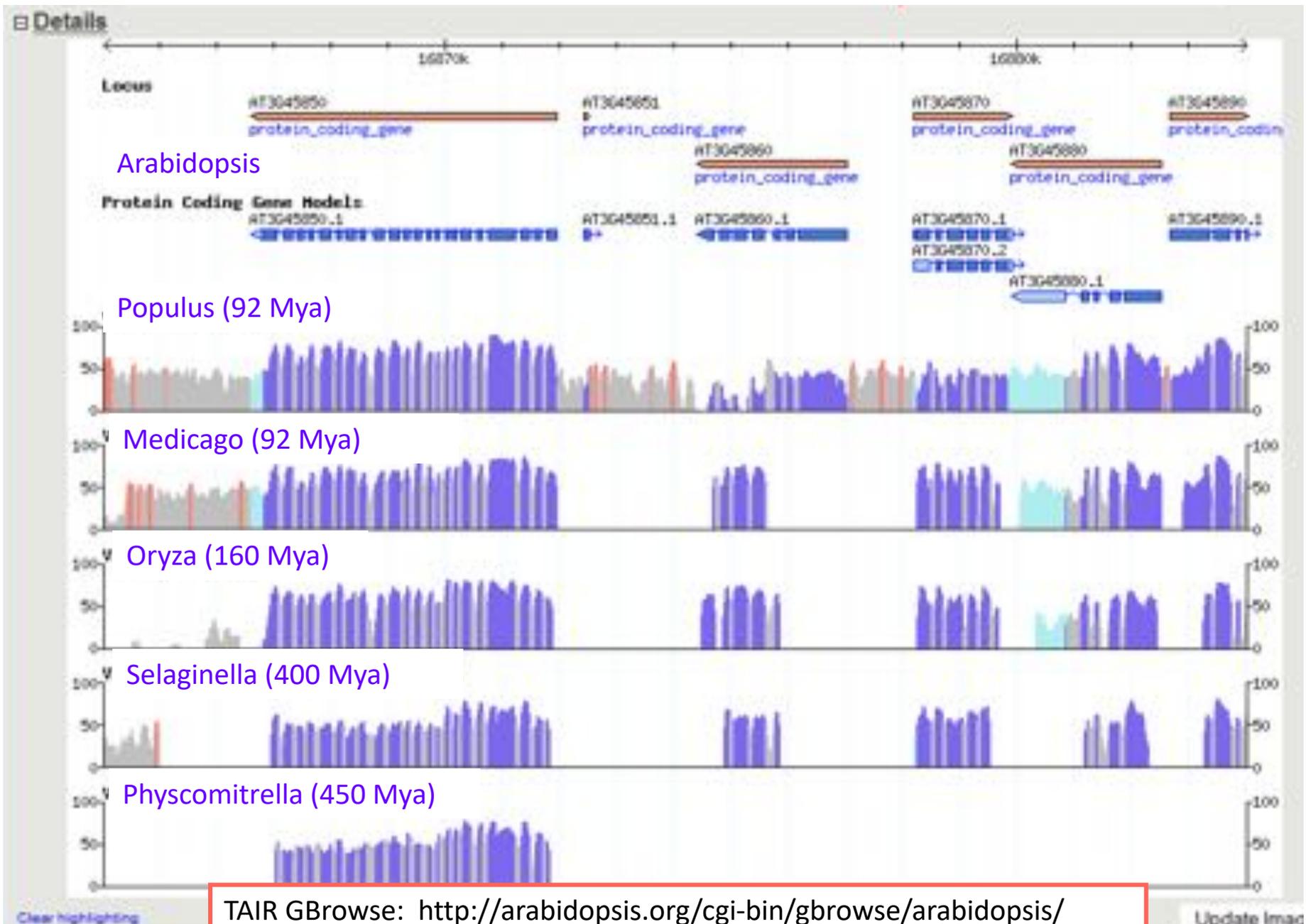


*except for single exon genes
 ** Transcript may extend beyond end of gene model

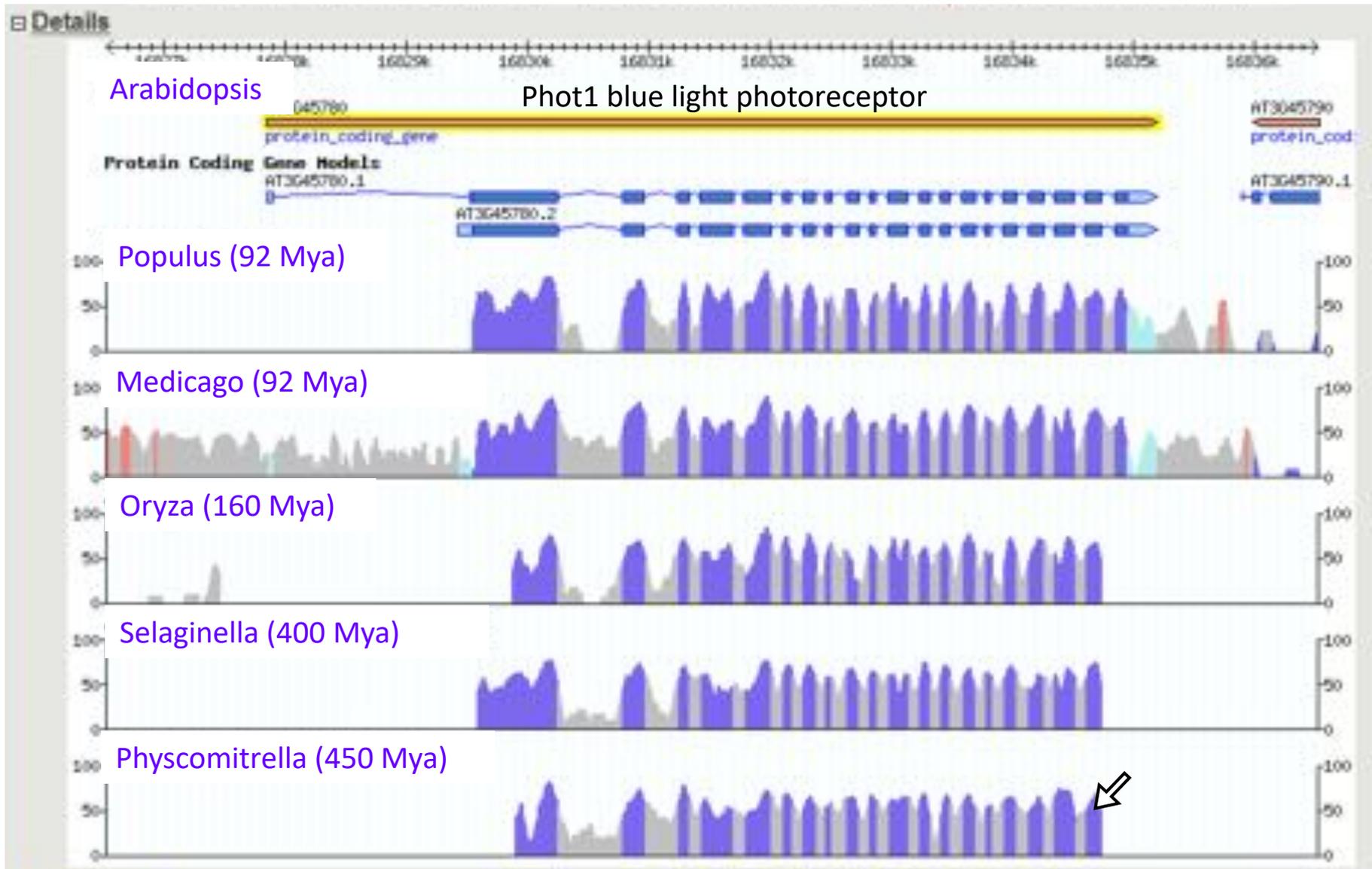
VISTA plot Gbrowse track



Nucleotide conservation across large evolutionary time spans



Nucleotide conservation across large evolutionary time spans



TAIR GBrowse: <http://arabidopsis.org/cgi-bin/gbrowse/arabidopsis/>

Arabidopsis chr3:16,826,523-16,836,522

<< >>

VISTA Browser

Change Annotation:

[Download JGI -- Filtered Models](#) [Get CNS: Arabidopsis-P. patens](#)

Location on Arabidopsis	Location on P. patens	Alignment
<p>chr3:16,829,890-16,834,733 (+)</p>  <p>Sequence (softmasked) length: 4844bp</p>	<p>scaffold_509:51,139-55,737 (-)</p> <p>Sequence (softmasked) length: 4599bp</p> <p>VISTA Browser</p> <p>View in: <input type="text" value="Select Browser"/></p>	<p>rankVISTA</p> <p>Alignment: Arabidopsis-P. patens</p> <p>MFA: Arabidopsis-P. patens</p> <p>CNS: Arabidopsis-P. patens</p> <p>rVISTA: Arabidopsis-P. patens</p> <p>PDF: Arabidopsis-P. patens</p>
<p>chr3:16,830,091-16,834,744 (+)</p>  <p>Sequence (softmasked) length: 4654bp</p> <p>Overlap=4643bp</p>	<p>scaffold_179:442,743-447,338 (+)</p> <p>Sequence (softmasked) length: 4596bp</p> <p>VISTA Browser</p> <p>View in: <input type="text" value="Select Browser"/></p>	<p>rankVISTA</p> <p>Alignment: Arabidopsis-P. patens</p> <p>MFA: Arabidopsis-P. patens</p> <p>CNS: Arabidopsis-P. patens</p> <p>rVISTA: Arabidopsis-P. patens</p> <p>PDF: Arabidopsis-P. patens</p>
<p>chr3:16,830,127-16,834,512 (+)</p>  <p>Sequence (softmasked) length: 4385bp</p> <p>Overlap=4385bp</p>	<p>scaffold_75:1,260,093-1,264,602 (-)</p> <p>Sequence (softmasked) length: 4510bp</p> <p>VISTA Browser</p> <p>View in: <input type="text" value="Select Browser"/></p>	<p>rankVISTA</p> <p>Alignment: Arabidopsis-P. patens</p> <p>MFA: Arabidopsis-P. patens</p> <p>CNS: Arabidopsis-P. patens</p> <p>rVISTA: Arabidopsis-P. patens</p> <p>PDF: Arabidopsis-P. patens</p>



```

>Arabidopsis Mar. 2004 chr3:16830091-16834744 (+)
>P. patens v.1.1 scaffold_179:442743-447338 (+)

016830091 GTATCGGAGATCTAAAAGATGCGTTGTCGACGTTTCAACAACGTTTGTGGTCTCAGAT 016830150
>>>>>>>> || ||||| ||| || || || ||||| | || | ||||| ||||| ||||| <<<<<<<<<
000442743 GTTTCGGAGGATTTACTTGGACGCATTGCTTCCCTTCAAGCAAACCTTTGTGGTATCAGAT 000442802

016830151 GCTACAAAACCTGATTATCCGATTATGTATGCAAGTGCTGGTTTTCATCATGACTGGT 016830210
>>>>>>>> || || ||||| || || || || ||||| || ||||| || ||||| ||||| <<<<<<<<<
000442803 GCAACGAAACCGGACTACCCCATCATGTACGCCGAGTGCGGGATTCTTCAGCATGACTGGG 000442862

016830211 TACACTTCCAAGAGTCTGTCGGCAGAAACTGGTAAAAAGTTATTCAAAAATTTTAGAG 016830270
>>>>>>>> |||| | || || || ||||| <<<<<<<<<
000442863 TACAGCCC GAAGGAGGTGATCGC----- 000442886

016830271 CATTTTTTTAATATATAAAATACGTTTCGTTTTCTGTAATAATTACAAAATTGCAAAGT 016830330
>>>>>>>> ----- <<<<<<<<<<
000000000 ----- 000000000

016830331 TTTTCTAATTAATATCATAGTTTTGTTTTTAGGTTTATGGTTACTATAAATGTATAAA 016830390
>>>>>>>> ----- <<<<<<<<<<
000000000 ----- 000000000

```

```

>Arabidopsis Mar. 2004 chr3:16830091-16834744 (+)
>P. patens v.1.1 scaffold_179:442743-447338 (+)

```

```

016830091 GTATCGGAGATCTAAAAGATGCGTTGTCGACGTTTCAACAAC
>>>>>>>> || ||||| ||| || || || ||||| | || | |||||
000442743 GTTTCGGAGGATTTACTTGGACGCATTGCTTCCCTTCAAGCAAAC

016830151 GCTACAAAACCTGATTATCCGATTATGTATGCAAGTGCTGGTTT
>>>>>>>> || || ||||| || || || || ||||| || ||||| || || ||
000442803 GCAACGAAACCGGACTACCCCATCATGTACGCCGAGTGCGGGATT

```

```

TTATT 016830450
----- 000000000

AGGTA 016830510
----- 000000000

ACCTA 016830570
----- 000000000

AATGA 016830630
----- 000000000

TTACA 016830690
----- 000000000

GAAAA 016830750
----- 000442891

```

```

000442887 -----TACAA-----
016830751 GAACATAATTTGGGTTTTTGTATTGTTAGCCGATTTTACAGGATCAGGTACAGATG 016830810
>>>>>>>> | ||||| || ||||| | | || ||| <<<<<<<<<
000442892 -----TTGCCGATTCCTTCAAGGCCCTGACACGGATC 000442923

016830811 CAGATGAGTTAGCAAAGATAAGAGAGACATTAGCTGCTGGTAACAATTATTGTGGGCGTA 016830870
>>>>>>>> | || | ||||| || || | ||||| || || ||||| || <<<<<<<<<
000442924 CCATGGAAAGTTGAAAAGATTCCGCCAGGC TGTGAGGACTGGGAAGCC TTTC TGTGGCCGGT 000442983

016830871 TATTGAATTACAGAAGATGGGACCTCTTTTGGAACTTCTCAGATTGCTCCCATTA 016830930
>>>>>>>> || ||||| |||| || ||||| || ||||| || || || | || || | <<<<<<<<<
000442984 TACTGAAC TACAGAAGGATGGGACGCAGTTCTGGAAATTTGCTAACCATCACGCCAATCA 000443043

```

Orthologs and Gene Families

INSECTS

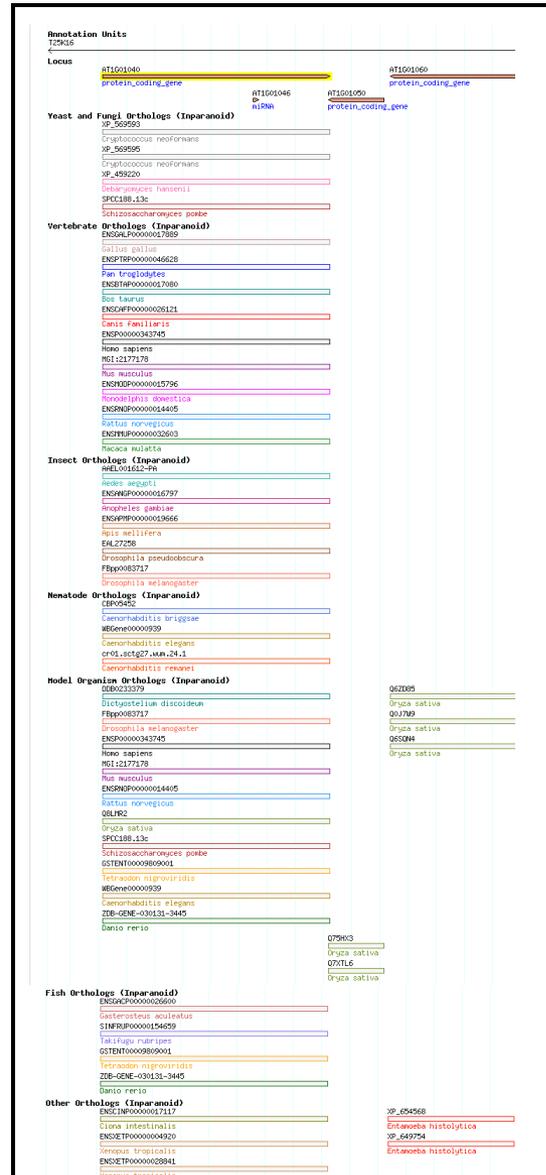
- Aedes aegypti (Yellow fever mosquito)
- Anopheles gambiae (Malaria mosquito)
- Apis mellifera (Western honeybee)
- Drosophila melanogaster (Fly)
- Drosophila pseudoobscura (Fly)

MAMMALS

- Bos taurus (Cow)
- Canis familiaris (Dog)
- Homo sapiens (human)
- Mus Musculus (mouse)
- Gallus gallus (Chicken)
- Macaca mulatta (Rhesus Macaque)
- Monodelphis domestica (Opossum)
- Pan troglodytes (Chimpanzee)
- Rattus norvegicus (Rat)

YEAST

- Candida glabrata (Haploid yeast)
- Cryptococcus neoformans (Yeast-like fungus)
- Debaryomyces hansenii (Yeast)
- Kluyveromyces lactis (Yeast)
- Yarrowia lipolytica (yeast)
- Schizosaccharomyces pombe (Fission yeast)
- Saccharomyces cerevisiae (Budding yeast)



FISH

- Takifugu rubripes (Pufferfish/Fugu)
- Tetraodon nigroviridis (Pufferfish/Green-spotted)
- Gasterosteus aculeatus (Stickleback)
- Danio rerio (Zebrafish)

NEMATODES

- Caenorhabditis elegans (Nematode)
- Caenorhabditis briggsae (Nematode)
- Caenorhabditis remanei (Nematode)

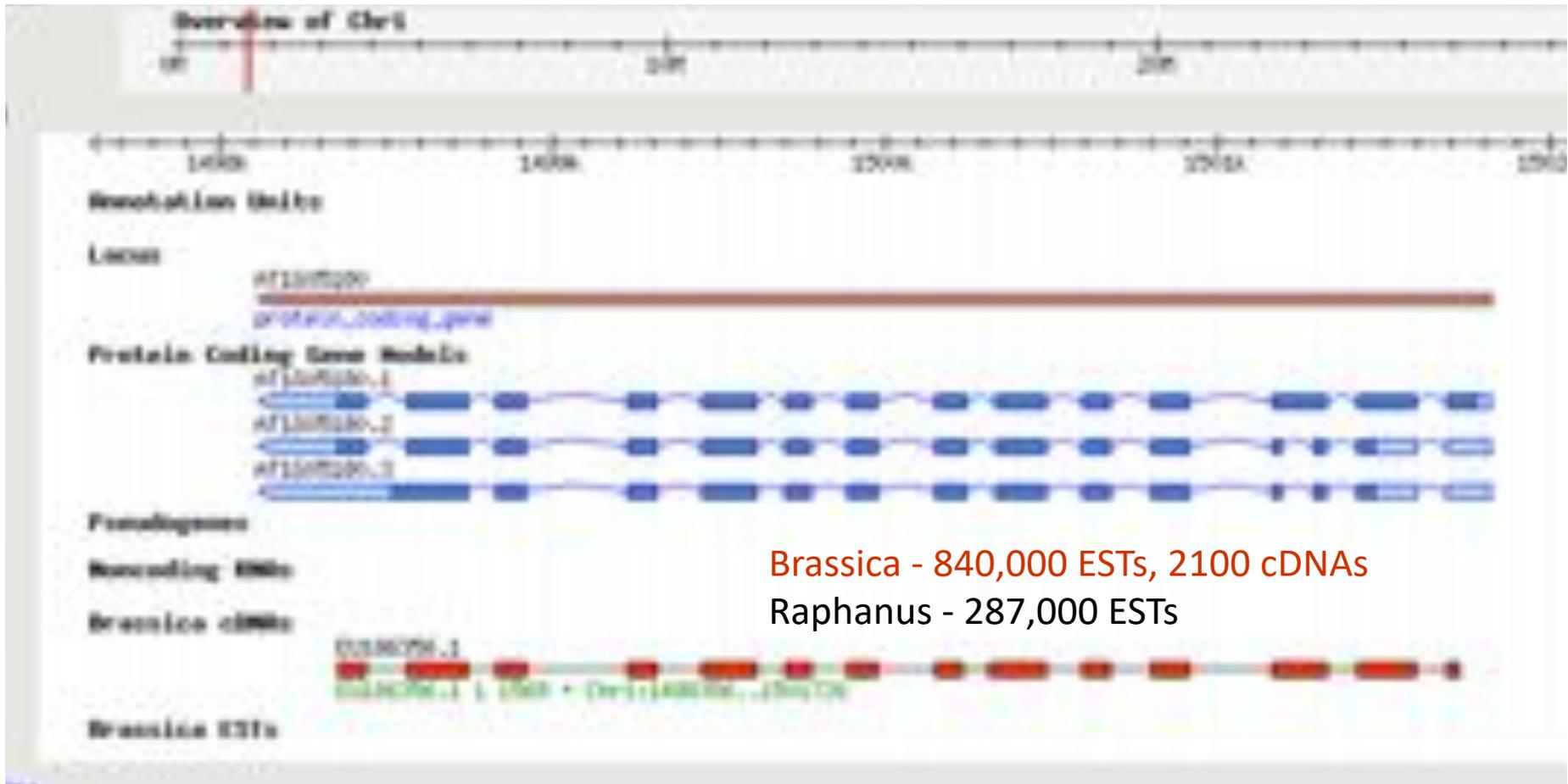
PLANTS

- Oryza sativa (rice)

OTHER

- Xenopus tropicalis (Frog)
- Ciona intestinalis (Sea squirt)
- Entamoeba histolytica (Amoebozoan)
- Dictyostelium discoideum (Slime mold)

Other ways to navigate between genomes: Brassica and radish sequences



Nucleotide alignments to Arabidopsis using CAT
(Cross-species Alignment Tool, Li et al 2007)

TAIR survey April 2008



TAIR survey April 2008



What tools should we add to TAIR?

TAIR survey April 2008



What tools should we add to TAIR?

Syntenry Viewer
Protein-Protein Interaction Viewer

Synteny Viewer

- GBrowse_syn: GBrowse-based synteny browser developed by Sheldon McKay
- Helps to study and analyze syntenic regions, homologous genes and other conserved elements between sequences
- By comparing less studied genomes to the well annotated Arabidopsis genome in Gbrowse_syn, scientists can identify novel genes and putative regulatory elements.
- First version of TAIR synteny viewer contains A.thaliana to A. lyrata genome alignments (more will be added soon)

Synteny Viewer

Instructions

Select a Region to Browse and a Reference species:

Examples: [thaliana Chr1:1504365..1514364](#), [lyrata scaffold_1:18300..29099](#)

To change from the current reference genome to another within the same region, please click on the ruler within the colored box for that genome

Search

Landmark

Arabidopsis thaliana TAIR9 (syn)

Genome to Search

▼

▼

Synteny Viewer

Instructions

Select a Region to Browse and a Reference species:

Examples: [thaliana Chr1:1504365..1514364](#), [lyrata scaffold_1:18300..29099](#)

To change from the current reference genome to another within the same region, please click on the ruler within the colored box for that genome

Search

Landmark

Chr1:1504365..1514364

Search

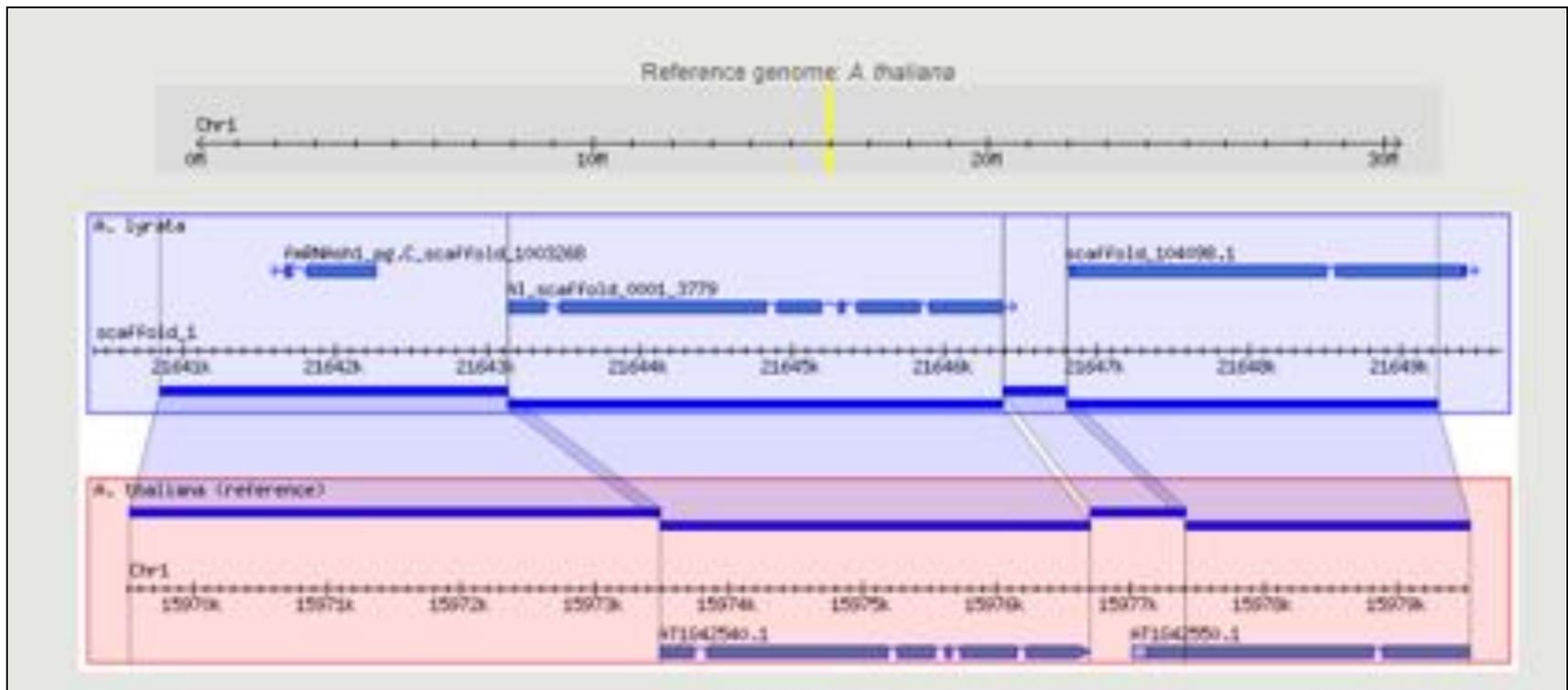
Reset

Genome to Search

A. thaliana

<<< Show 10 kbp >>>

Arabidopsis thaliana TAIR9 (syn)

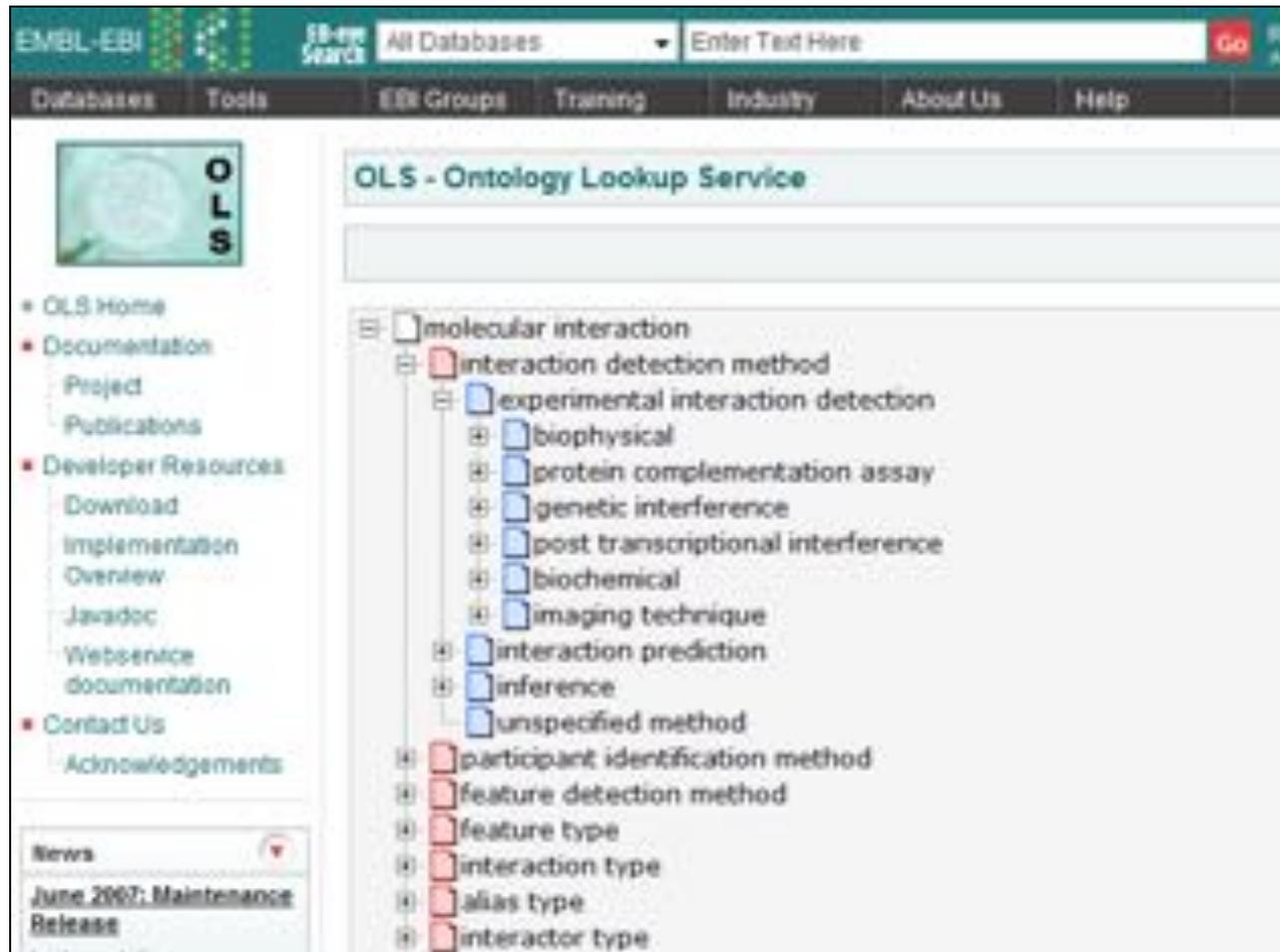


TAIR Protein-Protein interaction viewer using N-browse

- Protein-protein interactions curated by Intact, Biogrid and TAIR
- Using the generic network browser N-browse developed by the Gunsalus Lab (Curr Protoc Bioinformatics. 2008 Sep; Chapter 9:Unit 9.11)
- First version contains interactions determined experimentally
- Interactions can be filtered by type of experiment and biological modules
- User can overlay own interaction set on curated set

Minimal information about an Interaction

<http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName=MI>



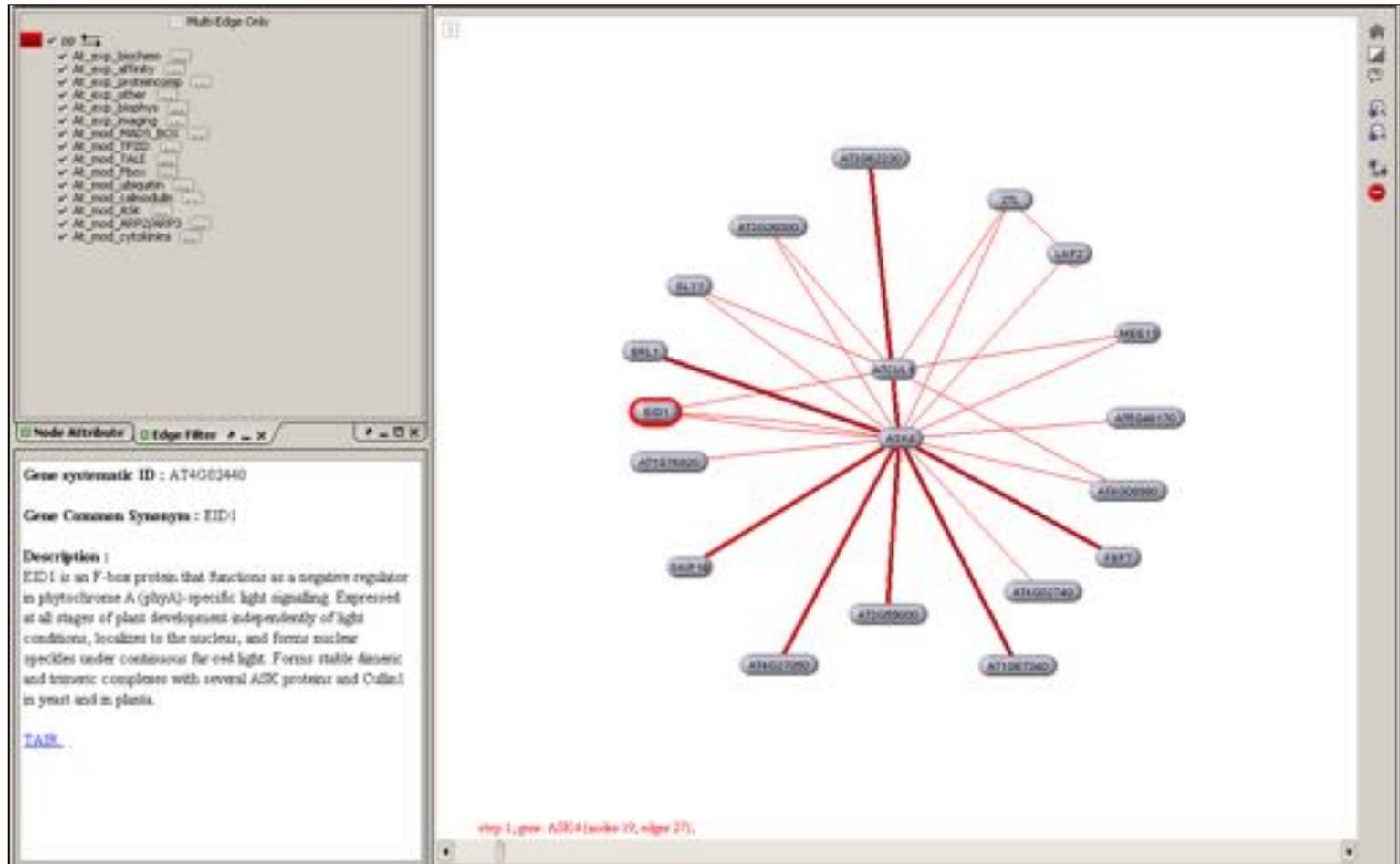
The screenshot displays the EMBL-EBI Ontology Lookup Service (OLS) interface. At the top, there is a search bar with the text "All Databases" and "Enter Text Here", and a "Go" button. Below the search bar is a navigation menu with links for "Databases", "Tools", "EBI Groups", "Training", "Industry", "About Us", and "Help".

The main content area is titled "OLS - Ontology Lookup Service". On the left side, there is a sidebar with a logo and a list of links: "OLS Home", "Documentation" (with sub-links for "Project" and "Publications"), "Developer Resources" (with sub-links for "Download", "Implementation", "Overview", and "Javadoc"), "Web service documentation", and "Contact Us" (with a sub-link for "Acknowledgements").

The main content area shows a tree view of the "molecular interaction" ontology. The tree is expanded to show the following structure:

- molecular interaction
 - interaction detection method
 - experimental interaction detection
 - biophysical
 - protein complementation assay
 - genetic interference
 - post transcriptional interference
 - biochemical
 - imaging technique
 - interaction prediction
 - inference
 - unspecified method
 - participant identification method
 - feature detection method
 - feature type
 - interaction type
 - alias type
 - interactor type

Arabidopsis Protein-Protein interaction viewer



> 1,500 interactions in TAIR Nbrowse Interaction Viewer

Acknowledgments

A long-exposure photograph of the Golden Gate Bridge at night, showing light trails from traffic and the bridge's illuminated towers and cables against a dark sky and city lights in the background.

PIs

Eva Huala
Sue Rhee

Curators

David Swarbreck
Donghui Li
Tanya Berardini
Kate Dreher
Peifen Zhang

TAIR Tech Team:

Vanessa Kirkuo
Chris Wilks
Tom Meyer
Cindy Lee
Raymond Chetty
Bob Muller

N-Browse

Kris Gunsalus (NYU)
Mark Gibson

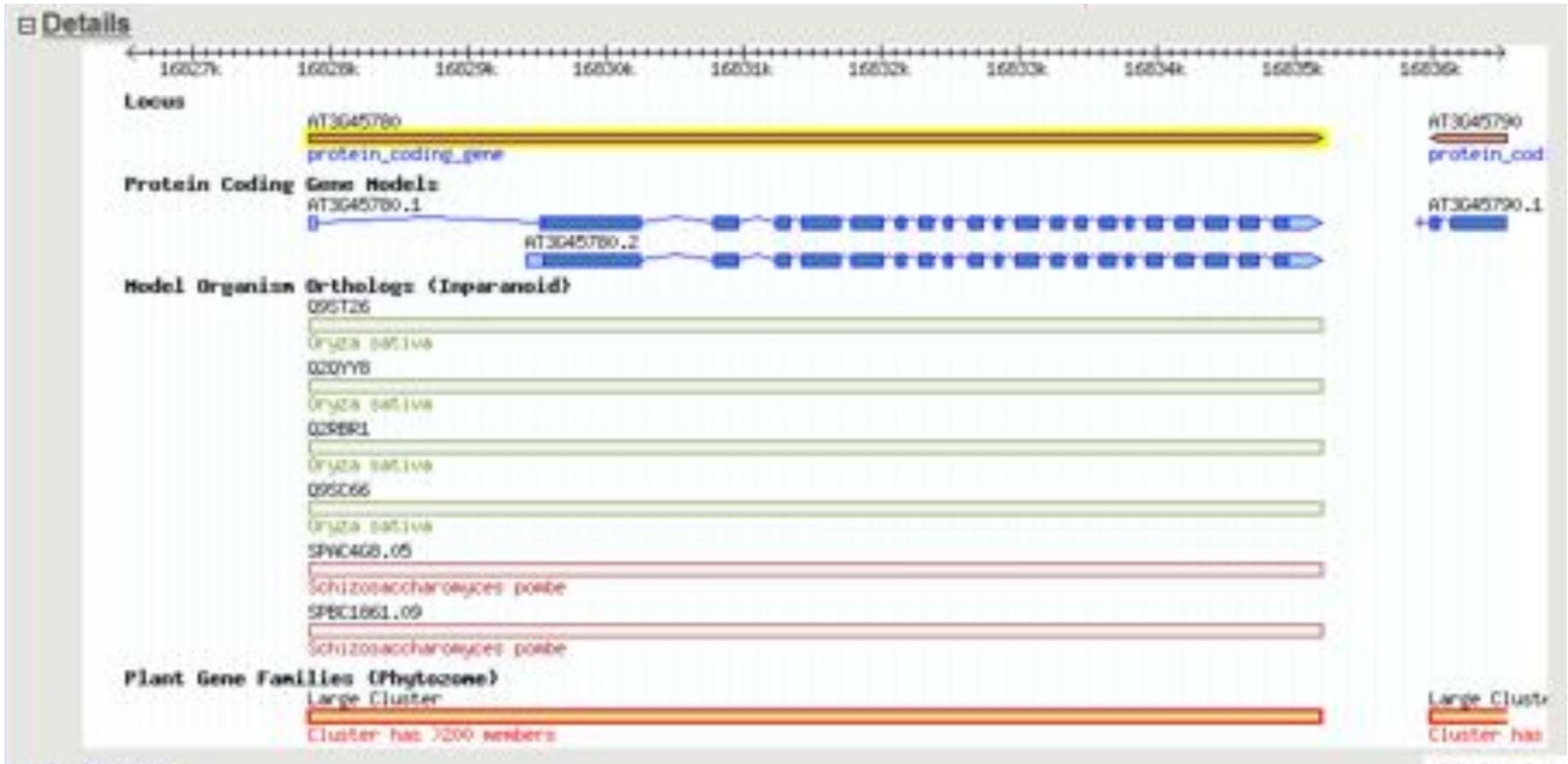
Gbrowse_syn

Sheldon McKay (CSHL)

Vista Browser

Ina Dubchak (LBNL)

Other ways to navigate between genomes: Orthologs and gene families



TAIR GBrowse: <http://arabidopsis.org/cgi-bin/gbrowse/arabidopsis/>