PeanutBase

The peanut community database

Ethy Cannon

Iowa State University

APRES 2017

Contributors to PeanutBase

Iowa State University/PeanutBase

- **Ethy Cannon** – lead
- Wei Huang – computational biologist
- Paul Otyama – PhD student

Iowa State University/Legume Federation

- Jacqueline Campbell – data curator and outreach

USDA-ARS at Ames, IA

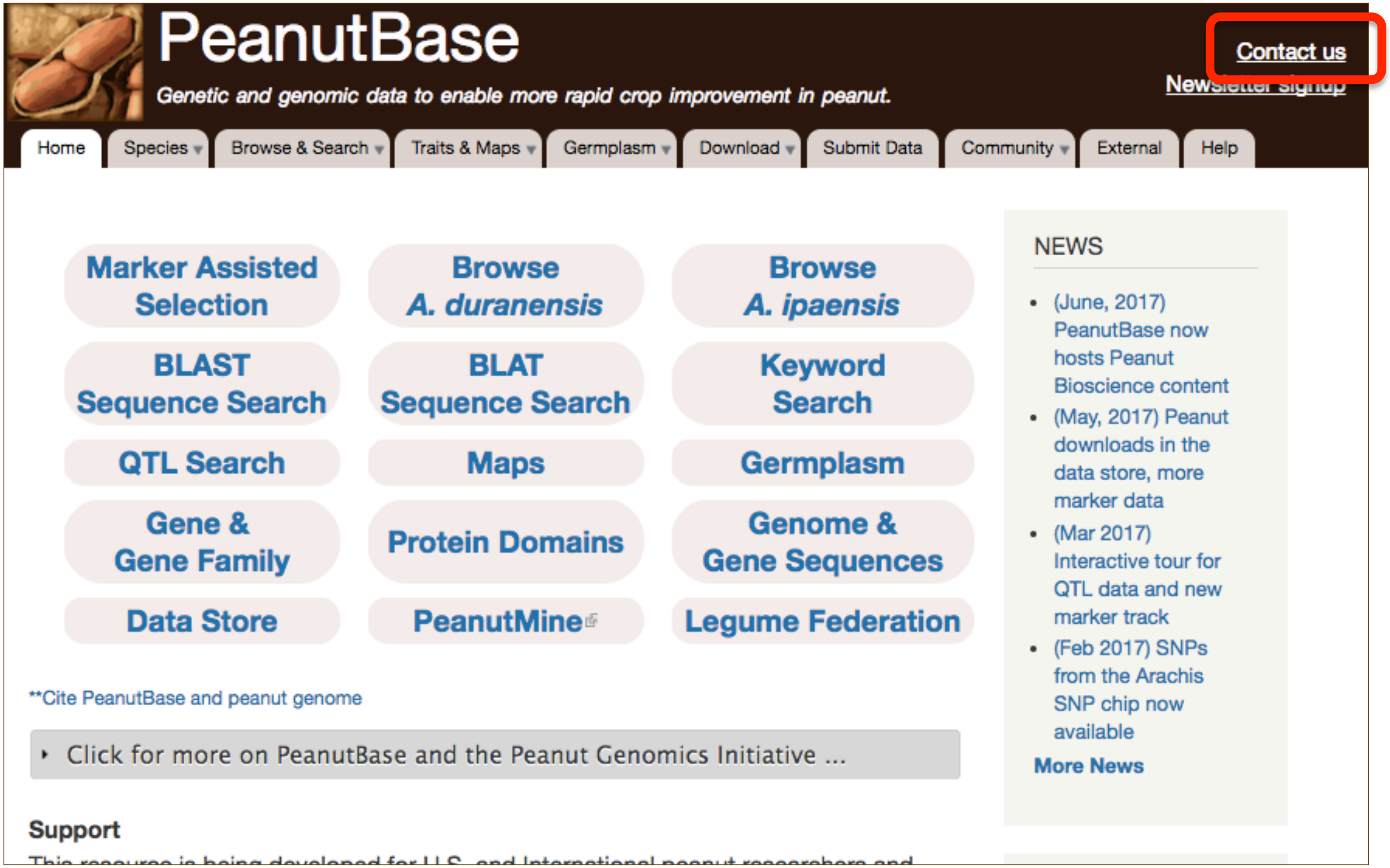
- **Steven Cannon** – lead scientist, ARS legume database project
- Nathan Weeks – IT specialist and computational biologist
- Scott Kalberer – data curator
- Andrew Wilkey – programmer

National Center for Genomic Resources

- **Andrew Farmer** – lead, Legume Information System
- Sudhansu Dash - geneticist and computational biologist
- Alex Rice - programmer

PeanutBase is a community resource

Let us know what you need



PeanutBase

Genetic and genomic data to enable more rapid crop improvement in peanut.

**Cite PeanutBase and peanut genome

‣ Click for more on PeanutBase and the Peanut Genomics Initiative ...

Support

This resource is being developed for U.S. and International peanut researchers and

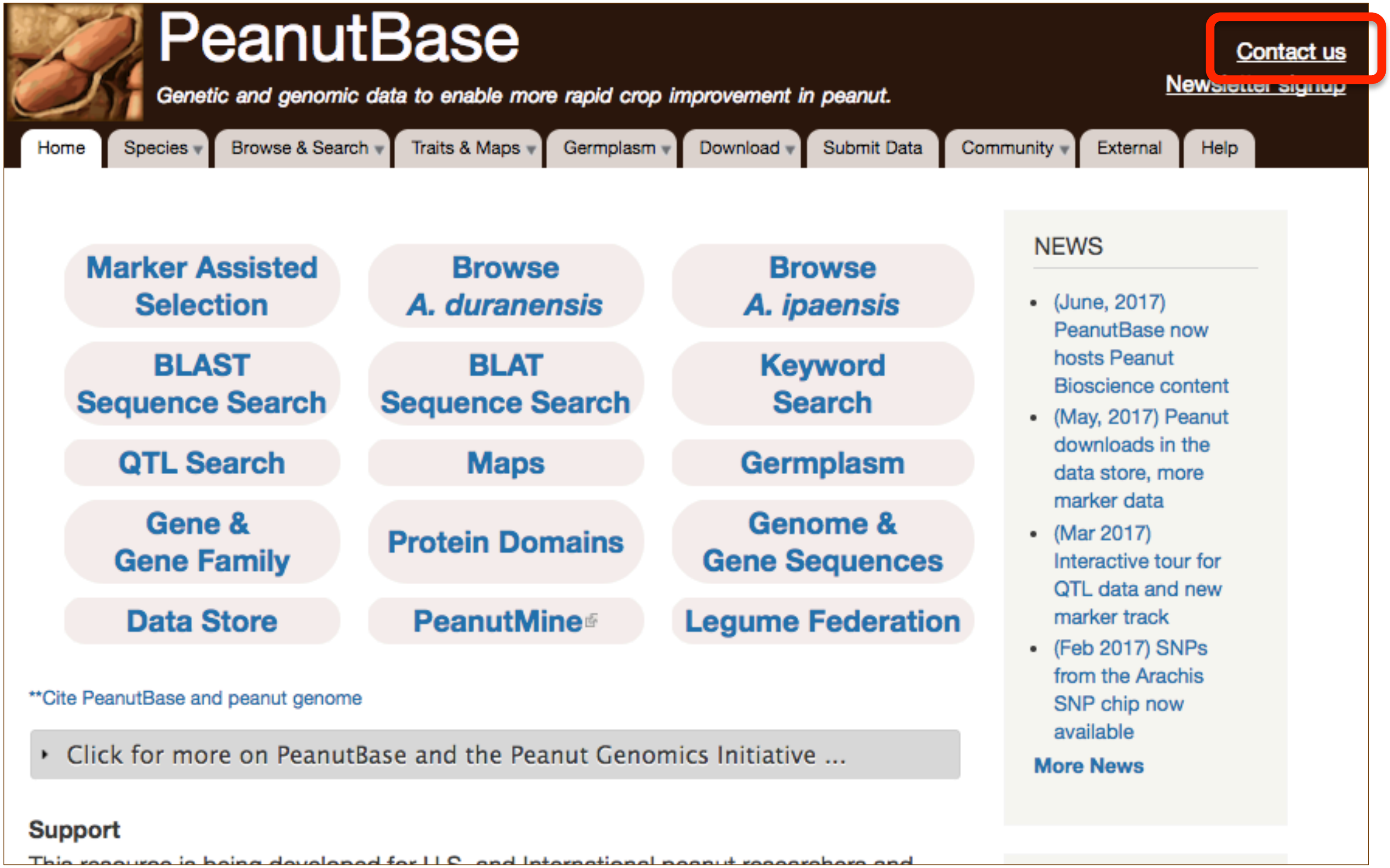
NEWS

- (June, 2017) PeanutBase now hosts Peanut Bioscience content
- (May, 2017) Peanut downloads in the data store, more marker data
- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis SNP chip now available

More News

PeanutBase is a community resource

Report errors



Genetic and genomic data to enable more rapid crop improvement in peanut.

NEWS

- (June, 2017) PeanutBase now hosts Peanut Bioscience content
- (May, 2017) Peanut downloads in the data store, more marker data
- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis SNP chip now available

**Cite PeanutBase and peanut genome

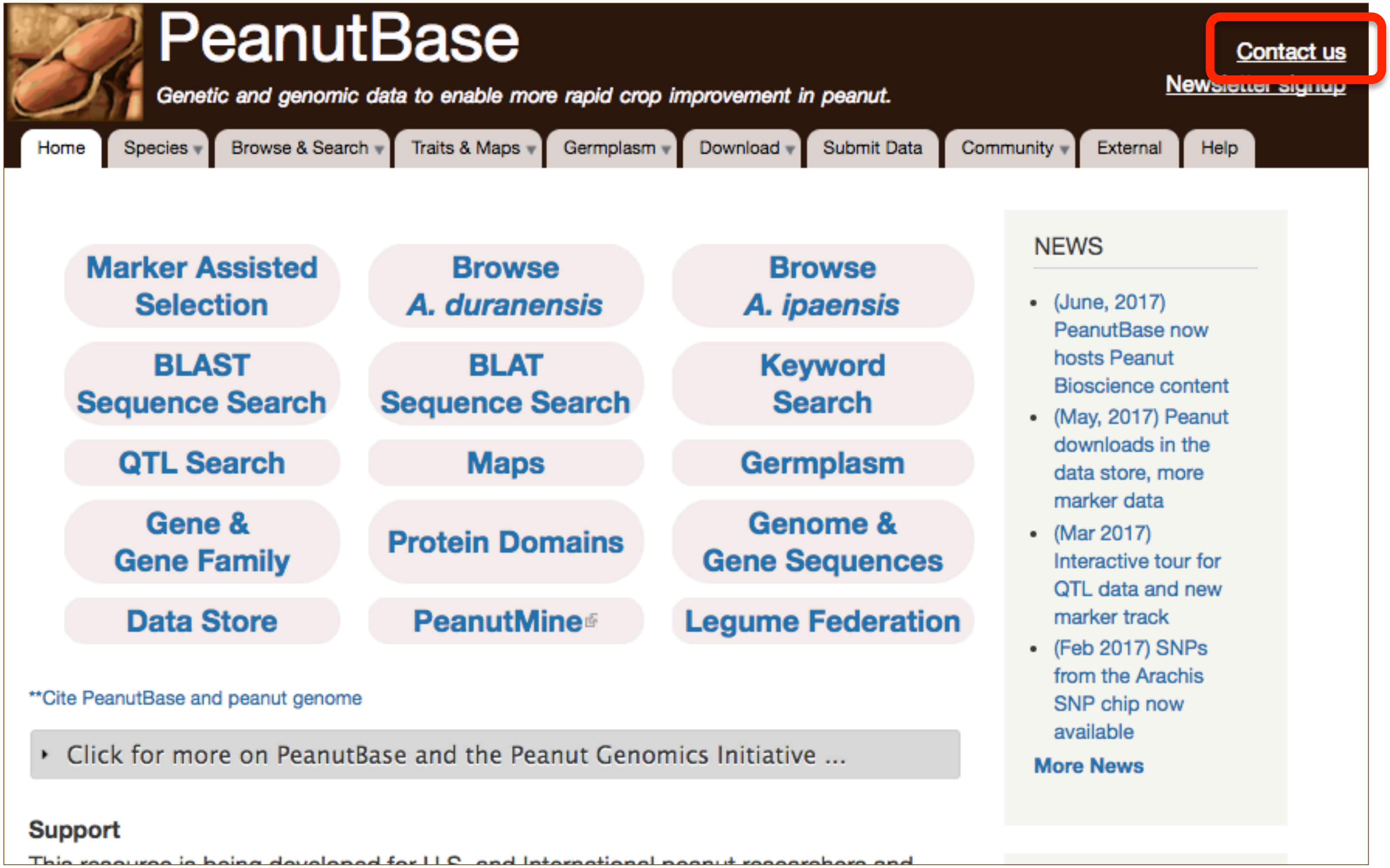
‣ Click for more on PeanutBase and the Peanut Genomics Initiative ...

Support

This resource is being developed for U.S. and International peanut researchers and

PeanutBase is a community resource

Work with us to host your public data



Genetic and genomic data to enable more rapid crop improvement in peanut.

NEWS

- (June, 2017) PeanutBase now hosts Peanut Bioscience content
- (May, 2017) Peanut downloads in the data store, more marker data
- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis SNP chip now available

More News

**Cite PeanutBase and peanut genome

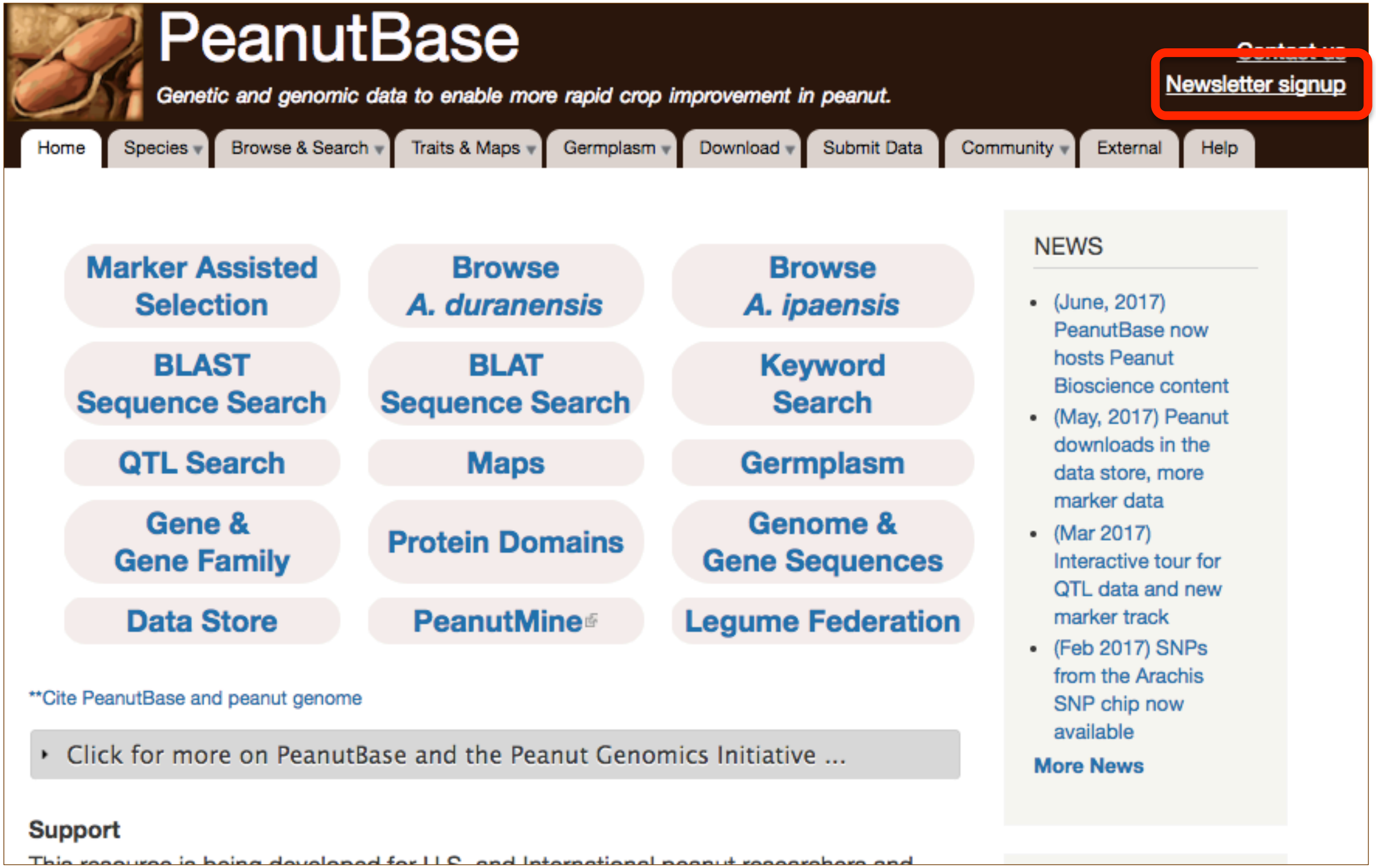
‣ Click for more on PeanutBase and the Peanut Genomics Initiative ...

Support

This resource is being developed for U.S. and International peanut researchers and

PeanutBase is a community resource

Sign up for the ~quarterly newsletter



PeanutBase

Genetic and genomic data to enable more rapid crop improvement in peanut.

NEWS

- (June, 2017) PeanutBase now hosts Peanut Bioscience content
- (May, 2017) Peanut downloads in the data store, more marker data
- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis SNP chip now available

More News

**Cite PeanutBase and peanut genome

Click for more on PeanutBase and the Peanut Genomics Initiative ...

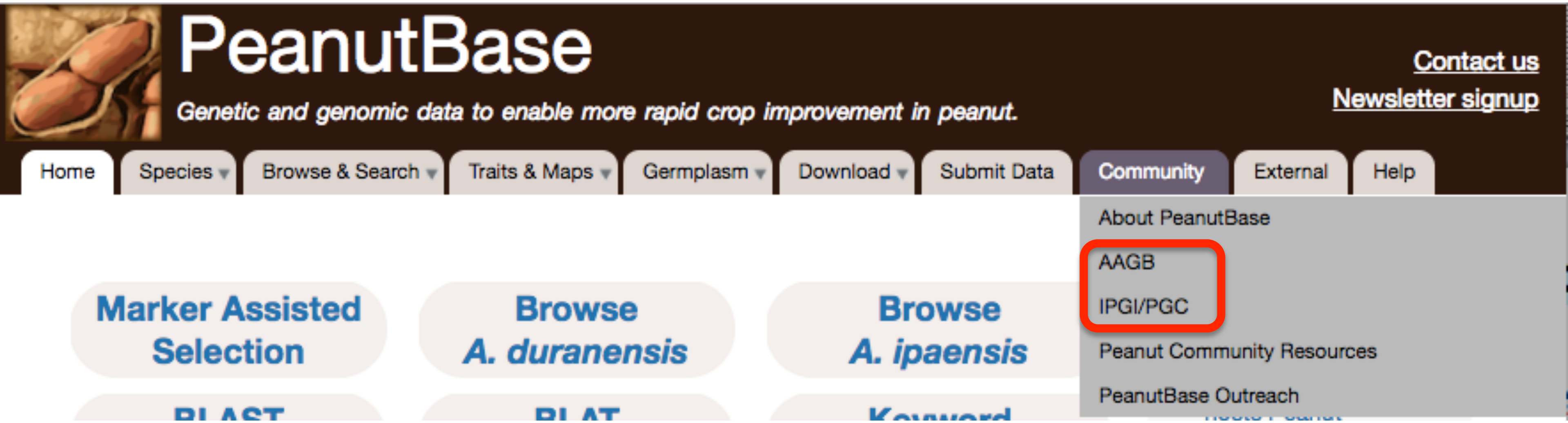
Support

This resource is being developed for U.S. and International peanut researchers and

Hot off the presses

PeanutBase now hosts PeanutBioscience.com

- AAGB archives and information
- IPGC and PGP meeting minutes and information

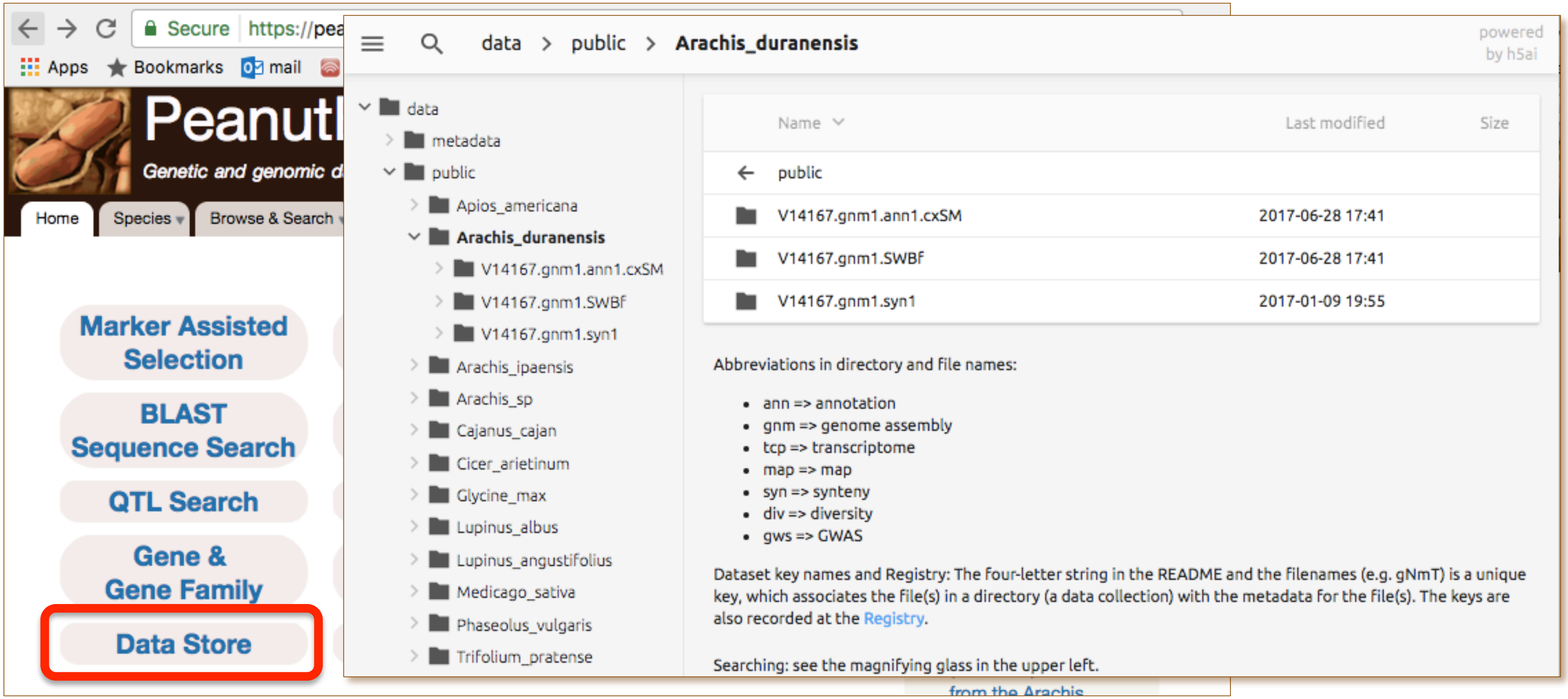


PeanutBase

Genetic and genomic data to enable more rapid crop improvement in peanut.

Hot off the presses

- Data store for downloading datasets for peanut and related species



Name	Last modified	Size
public		
V14167.gnm1.ann1.cxSM	2017-06-28 17:41	
V14167.gnm1.SWBF	2017-06-28 17:41	
V14167.gnm1.syn1	2017-01-09 19:55	

Abbreviations in directory and file names:

- ann => annotation
- gnm => genome assembly
- tcp => transcriptome
- map => map
- syn => synteny
- div => diversity
- gws => GWAS

Dataset key names and Registry: The four-letter string in the README and the filenames (e.g. gNmT) is a unique key, which associates the file(s) in a directory (a data collection) with the metadata for the file(s). The keys are also recorded at the Registry.

Searching: see the magnifying glass in the upper left.

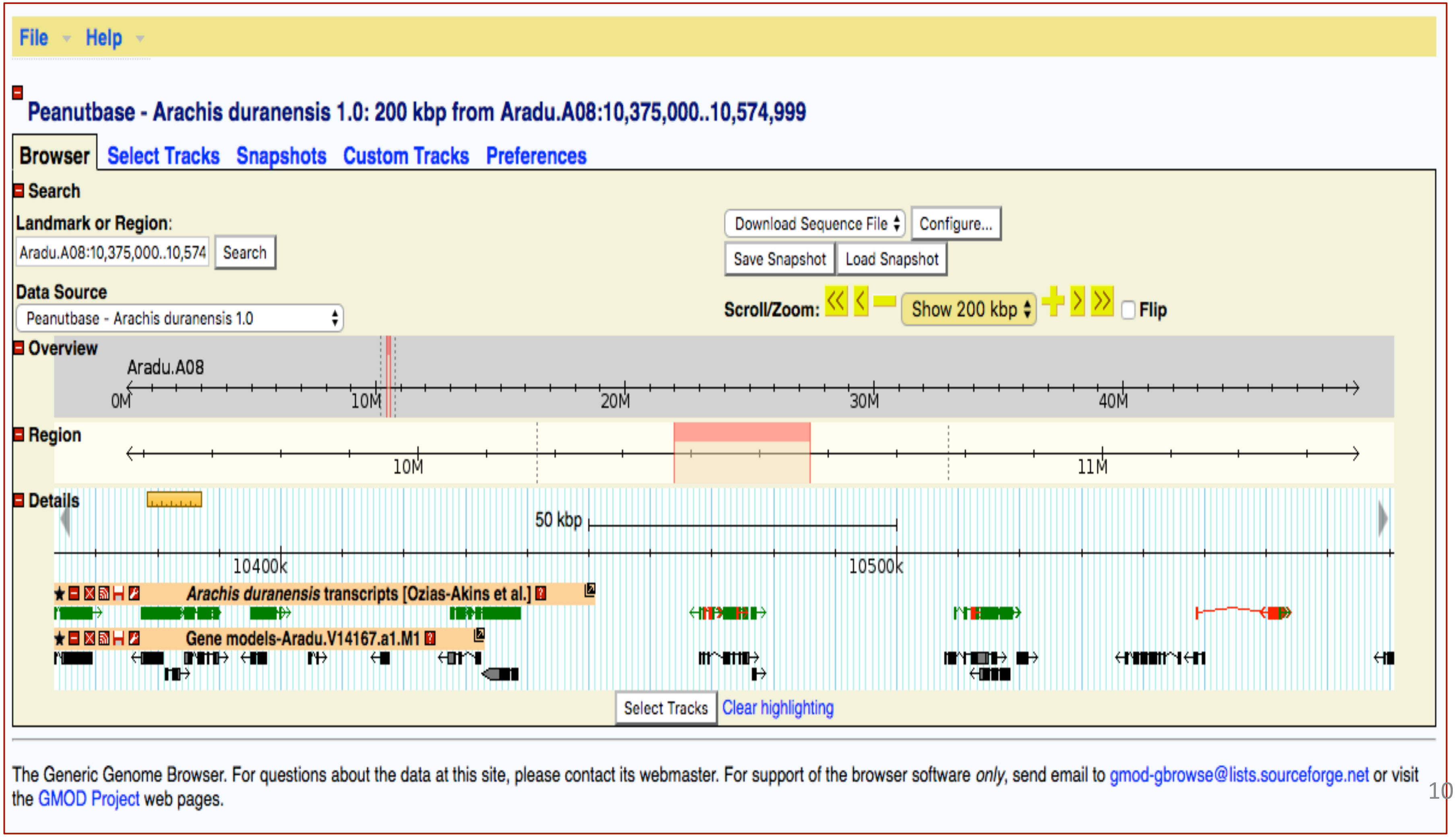
from the Arachis

What is in PeanutBase?

1. Genomes
2. Genes
3. Maps & Markers
4. Germplasm & Traits
5. Germplasm geographic viewer
6. Sequence search tools
7. Gene families & synteny
8. Website tours (and other help)

1. Genome sequences for *A. duranensis* and *A. ipaensis*

Genome browsers maintained by Wei Huang



Peanutbase - Arachis duranensis 1.0: 200 kbp from Aradu.A08:10,375,000..10,574,999

Landmark or Region: Aradu.A08:10,375,000..10,574

Data Source: Peanutbase - Arachis duranensis 1.0

Arachis duranensis transcripts [Ozias-Akins et al.]

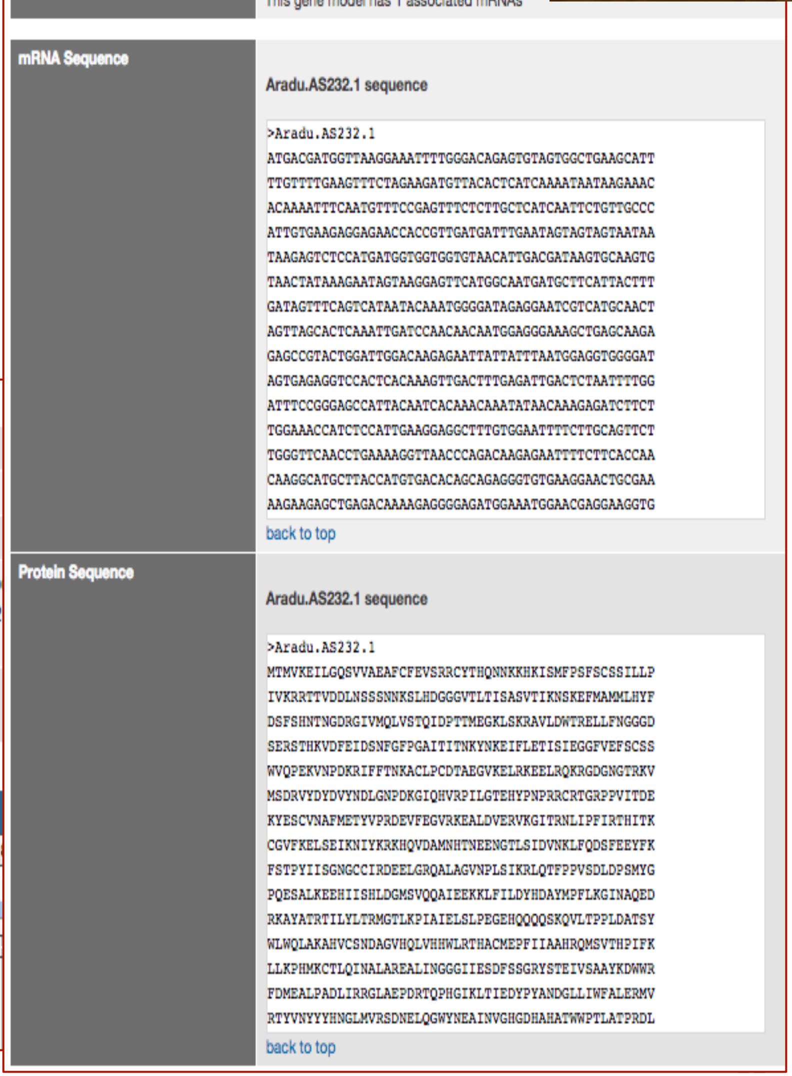
Gene models-Aradu.V14167.a1.M1

The Generic Genome Browser. For questions about the data at this site, please contact its webmaster. For support of the browser software *only*, send email to gmod-gbrowse@lists.sourceforge.net or visit the GMOD Project web pages.

2. Predicted genes, with sequences, structure, functional descriptions, and expression data.

Sequences

Gene Model Name	Aradu.AS232
	This gene model has 1 associated mRNAs



mRNA Sequence

Aradu.AS232.1 sequence

```
>Aradu.AS232.1
ATGACGATGGTTAAGGAAATTTTGGGACAGAGTGTAGTGGCTGAAGCATT
TTGTTTTGAAGTTTCTAGAAGATGTTACACTCATCAAAATAATAAGAAAC
ACAAAATTTCAATGTTTCCGAGTTTCTCTTGCTCATCAATTCTGTTGCCC
ATTGTGAAGAGGAGAACCACCGTTGATGATTTGAATAGTAGTAGTAATAA
TAAGAGTCTCCATGATGGTGGTGGTGTAACATTGACGATAAGTGCAAGTG
TAACTATAAAGAATAGTAAGGAGTTCATGGCAATGATGCTTCATTACTTT
GATAGTTTCAGTCATAATACAAATGGGGATAGAGGAATCGTCATGCAACT
AGTTAGCACTCAAATTGATCCAACAACAATGGAGGGAAAGCTGAGCAAGA
GAGCCGTACTGGATTGGACAAGAGAATTATTATTTAATGGAGGTGGGGAT
AGTGAGAGGTCCACTCACAAAGTTGACTTTGAGATTGACTCTAATTTTGG
ATTTCCGGGAGCCATTACAATCACAAACAAATATAACAAAGAGATCTTCT
TGGAAACCATCTCCATTGAAGGAGGCTTTGTGGAATTTTCTTGCAGTTCT
TGGGTTCAACCTGAAAAGGTTAACCCAGACAAGAGAATTTTCTTCACCAA
CAAGGCATGCTTACCATGTGACACAGCAGAGGGTGTGAAGGAACTGCGAA
AAGAAGAGCTGAGACAAAAGAGGGGAGATGGAAATGGAACGAGGAAGGTG
```

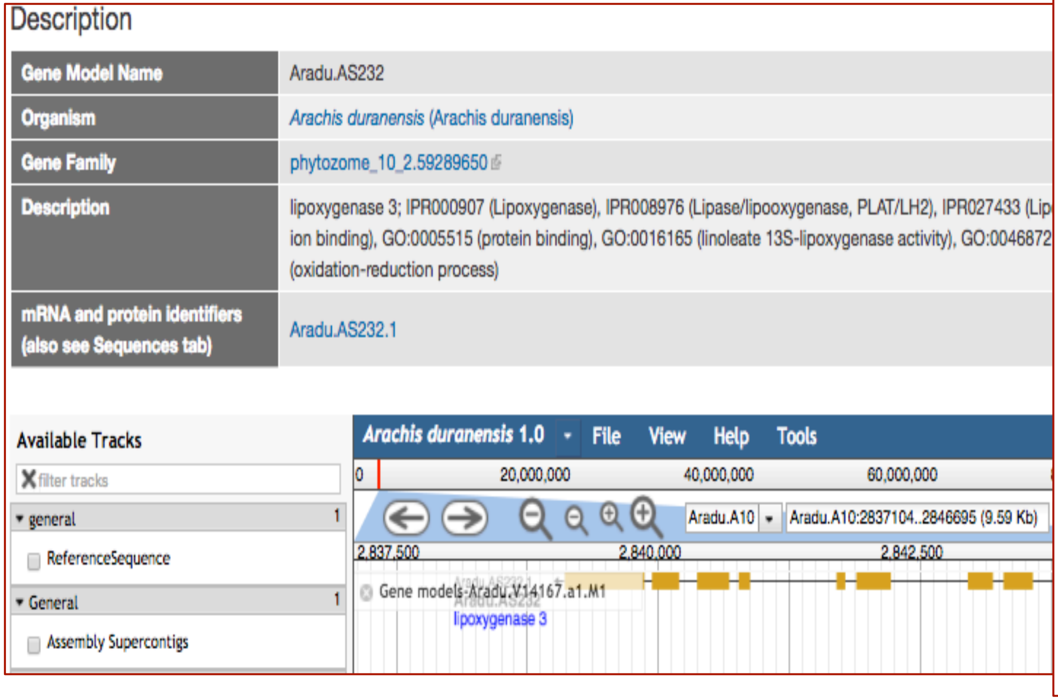
back to top

Protein Sequence

Aradu.AS232.1 sequence

```
>Aradu.AS232.1
MTMVKEILGQSVVAEAFCFEVSRRCYTHQNNKKHKISMFPSFSCSSILLP
IVKRRTTVDDLNSSSNNKSLHDGGGVTLTISASVTIKNSKEFMAMMLHYF
DSFSHNTNGDRGIVMQLVSTQIDPTTMEGKLSKRAVLDWTRELLFNGGGD
SERSTHKVDFEIDSNFGFPGAITITNKYNKEIFLETISIEGGFVEFSCSS
WVQPEKVNPDKRIFFTNKACLPCDTAEGVKELRKEELRQKRGDGNGTRKV
MSDRVYDYDVYNDLGNPDKGIQHVRPILGTEHYPNPRRCRTGRPPVITDE
KYESCVNAFMETYVPRDEVFEGVRKEALDVERVKGITRNLIPFIRTHITK
CGVFKELSEIKNIYKRKHQVDAMNHTNEENGTLSIDVNKLFQDSFEEYFK
FSTPYIISGNGCCIRDEELGRQALAGVNPLSIKRLQTFPPVSDLDPSMYG
PQESALKEEHIISHLDGMSVQQAIEEKKLFILDYHDAYMPFLKGINAQED
RKAYATRTILYLTRMGTLKPIAIELSLPEGEHQQQQSKQVLTPPLDATSY
WLWQLAKAHVCSNDAGVHQLVHHWLRTHACMEPFIIAAHRQMSVTHPIFK
LLKPHMKCTLQINALAREALINGGGIIESDFSSGRYSTEIVSAAYKDWWR
FDMEALPADLIRRGLAEPDRTQPHGIKLTIEDYPYANDGLLIWFALERMV
RTYVNYYYHNGLMVRSDNELQGWYNEAINVGHGDHAHATWWPTLATPRDL
```

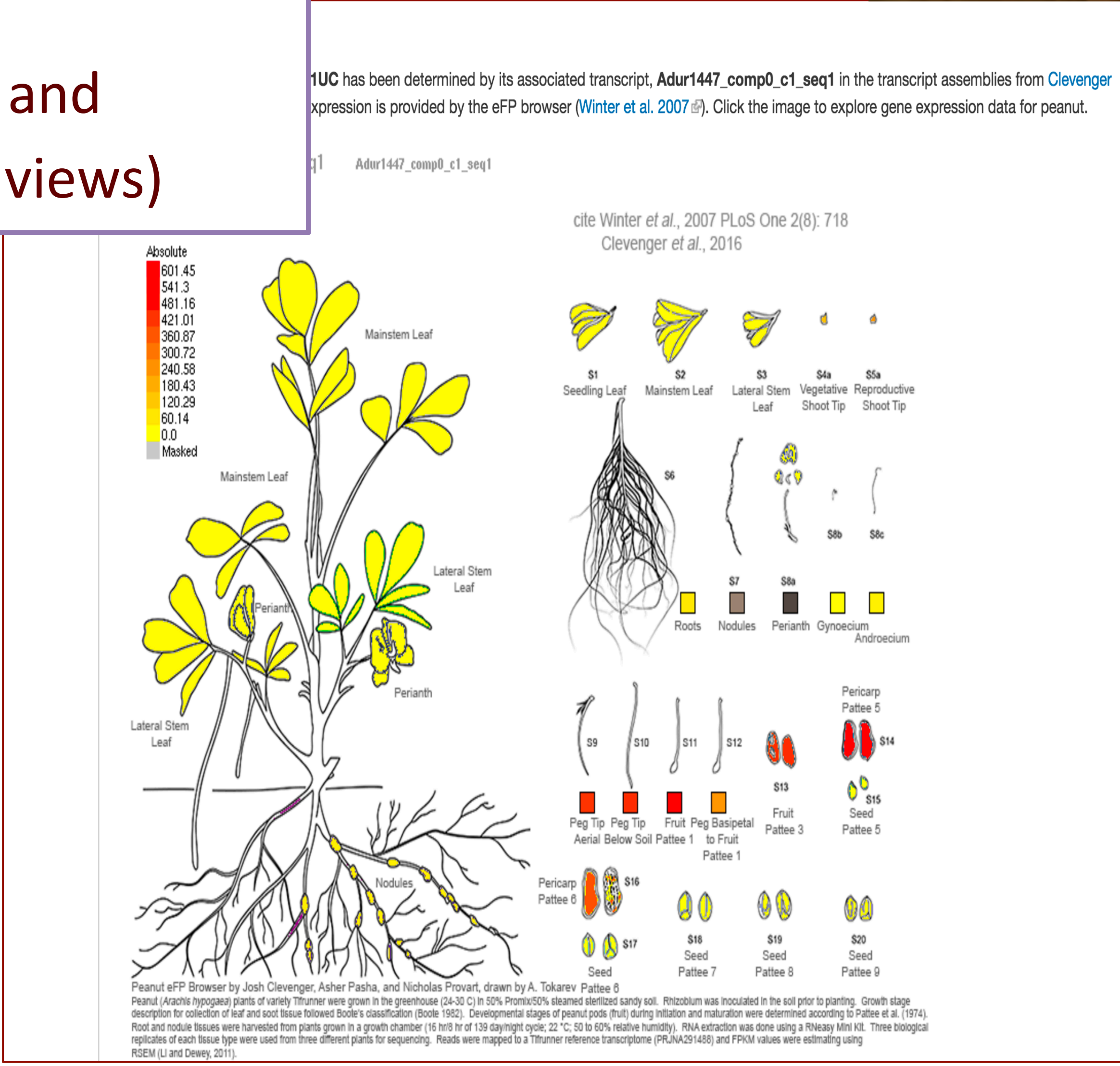
back to top



Description

Gene Model Name	Aradu.AS232
Organism	*Arachis duranensis* (Arachis duranensis)
Gene Family	phytozome_10_2.59289650
Description	lipoxygenase 3; IPR000907 (Lipoxygenase), IPR008976 (Lipase/lipooxygenase, PLAT/LH2), IPR027433 (Lip ion binding), GO:0005515 (protein binding), GO:0016165 (linoleate 13S-lipoxygenase activity), GO:0046872 (oxidation-reduction process)
mRNA and protein identifiers (also see Sequences tab)	Aradu.AS232.1

2. Predicted genes, with sequences, structure, functional descriptions, and expression data.
(thanks Josh, Peggy et al., and Nick Provart lab for these views)



1UC has been determined by its associated transcript, **Adur1447_comp0_c1_seq1** in the transcript assemblies from Clevenger
xpression is provided by the eFP browser (Winter et al. 2007). Click the image to explore gene expression data for peanut.

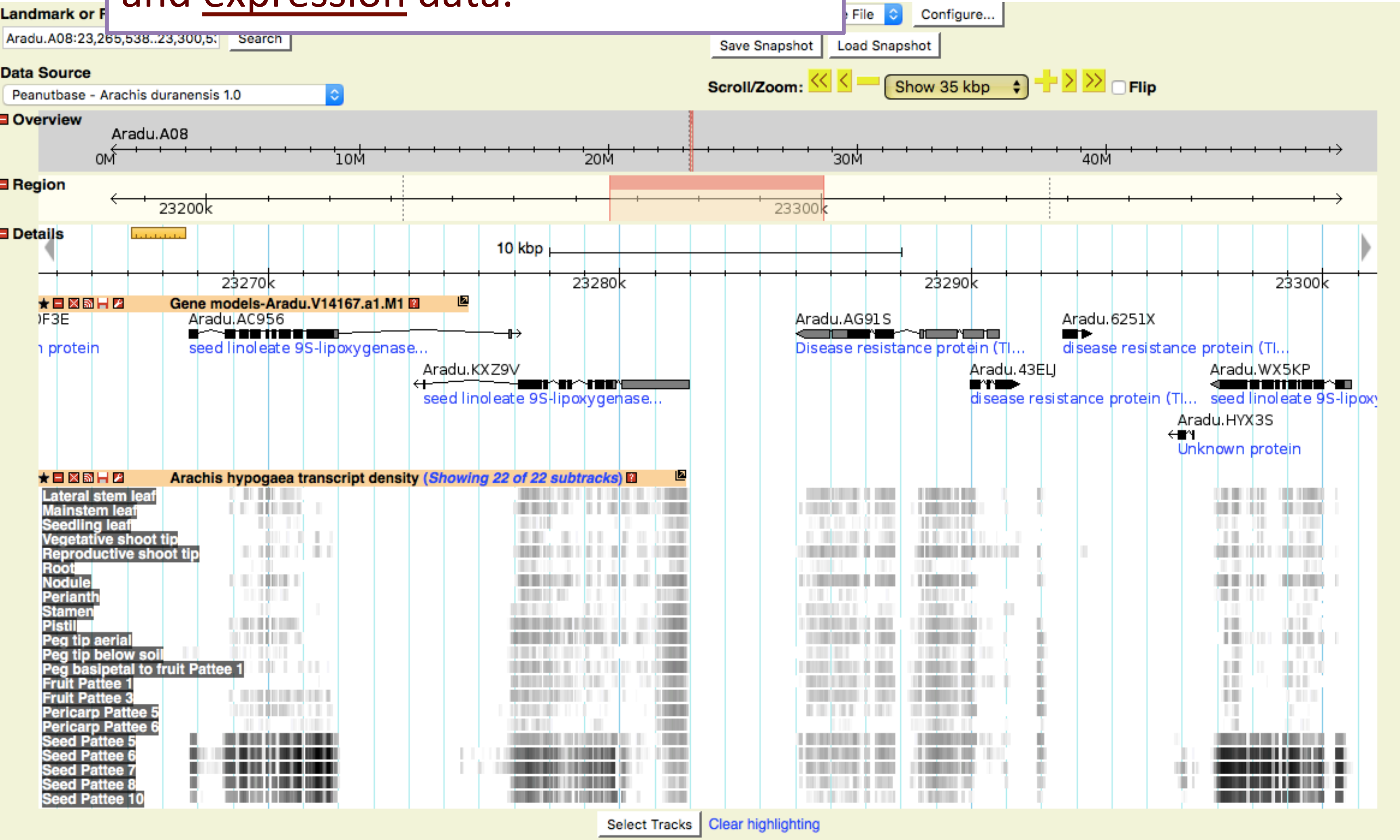
q1 Adur1447_comp0_c1_seq1

cite Winter *et al.*, 2007 PLoS One 2(8): 718
Clevenger *et al.*, 2016

Peanut eFP Browser by Josh Clevenger, Asher Pasha, and Nicholas Provart, drawn by A. Tokarev

Peanut (*Arachis hypogaea*) plants of variety Tifrunner were grown in the greenhouse (24-30 C) in 50% Promix/50% steamed sterilized sandy soil. Rhizobium was inoculated in the soil prior to planting. Growth stage description for collection of leaf and soot tissue followed Boote's classification (Boote 1982). Developmental stages of peanut pods (fruit) during initiation and maturation were determined according to Pattee et al. (1974). Root and nodule tissues were harvested from plants grown in a growth chamber (16 hr/8 hr of 139 day/night cycle; 22 °C; 50 to 60% relative humidity). RNA extraction was done using a RNeasy Mini Kit. Three biological replicates of each tissue type were used from three different plants for sequencing. Reads were mapped to a Tifrunner reference transcriptome (PRJNA291488) and FPKM values were estimating using RSEM (Li and Dewey, 2011).

2. Predicted <u>genes</u>, with sequences, structure, functional descriptions, and <u>expression</u> data.

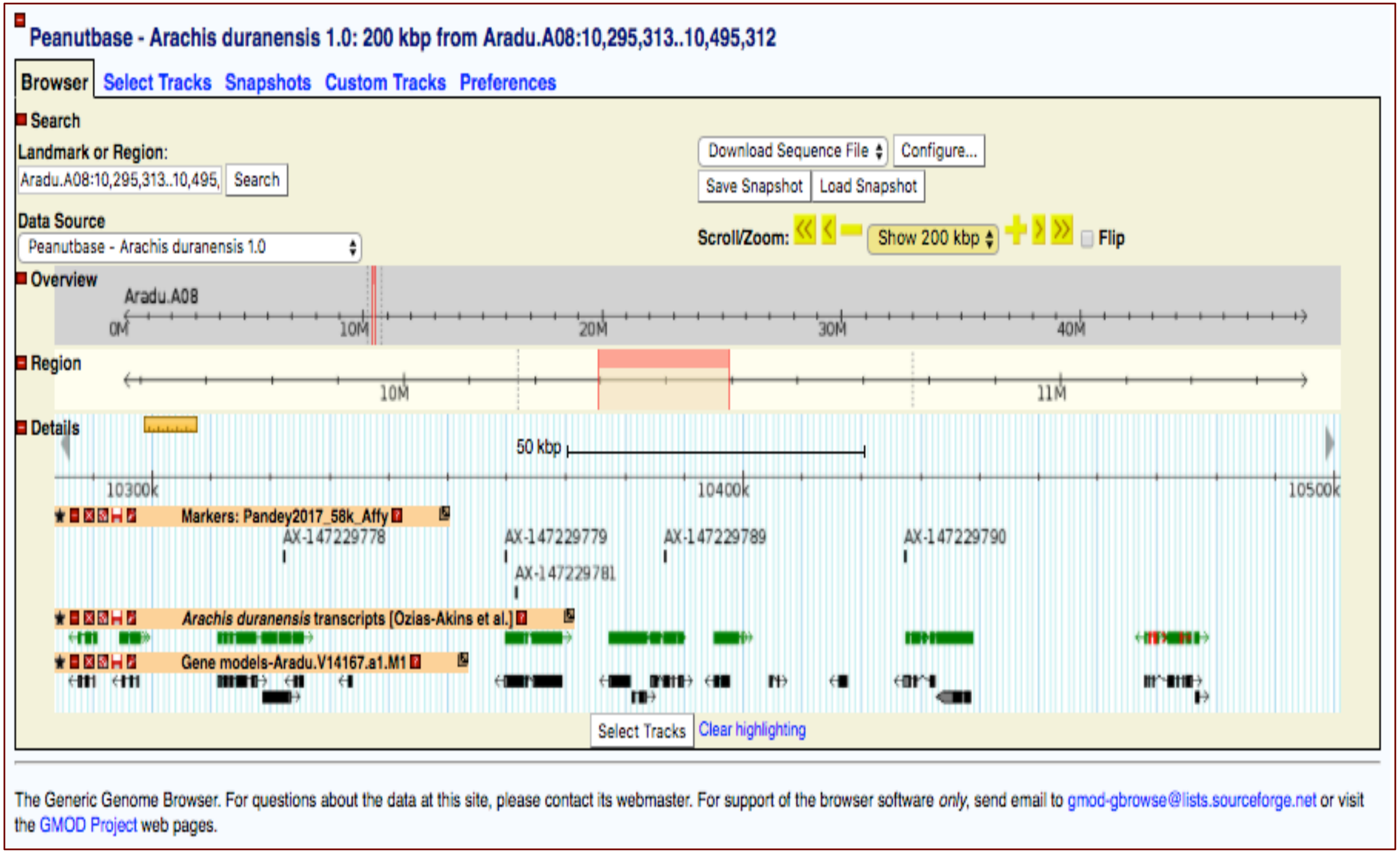


Landmark or R
Aradu.A08:23,265,538..23,300,5

Data Source
Peanutbase - Arachis duranensis 1.0

Scroll/Zoom: Show 35 kbp ☐ Flip

3. Various genetic map and marker sets, including the Affymetrix SNP array. (Pandey et al., 2017, Clevenger et al., 2017)



Peanutbase - Arachis duranensis 1.0: 200 kbp from Aradu.A08:10,295,313..10,495,312

Landmark or Region: Aradu.A08:10,295,313..10,495,

Data Source: Peanutbase - Arachis duranensis 1.0

Scroll/Zoom: Show 200 kbp

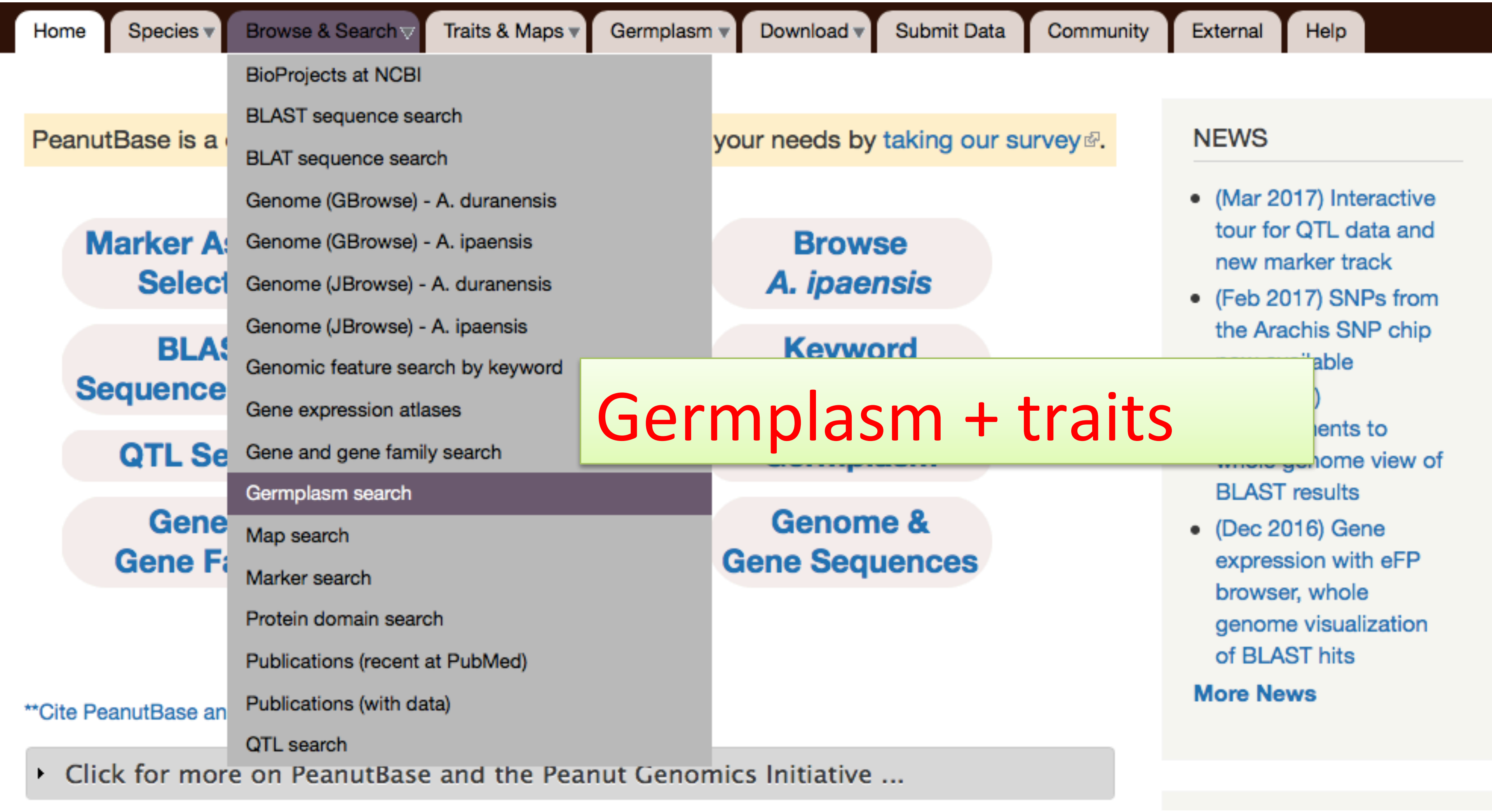
Markers: Pandey2017_58k_Affy

Arachis duranensis transcripts [Ozias-Akins et al.]

Gene models-Aradu.V14167.a1.M1

The Generic Genome Browser. For questions about the data at this site, please contact its webmaster. For support of the browser software *only*, send email to gmod-gbrowse@lists.sourceforge.net or visit the GMOD Project web pages.

4. germplasm + traits



PeanutBase is a ... your needs by taking our survey.

NEWS

- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis SNP chip ...able
- ...ents to ...ome view of BLAST results
- (Dec 2016) Gene expression with eFP browser, whole genome visualization of BLAST hits

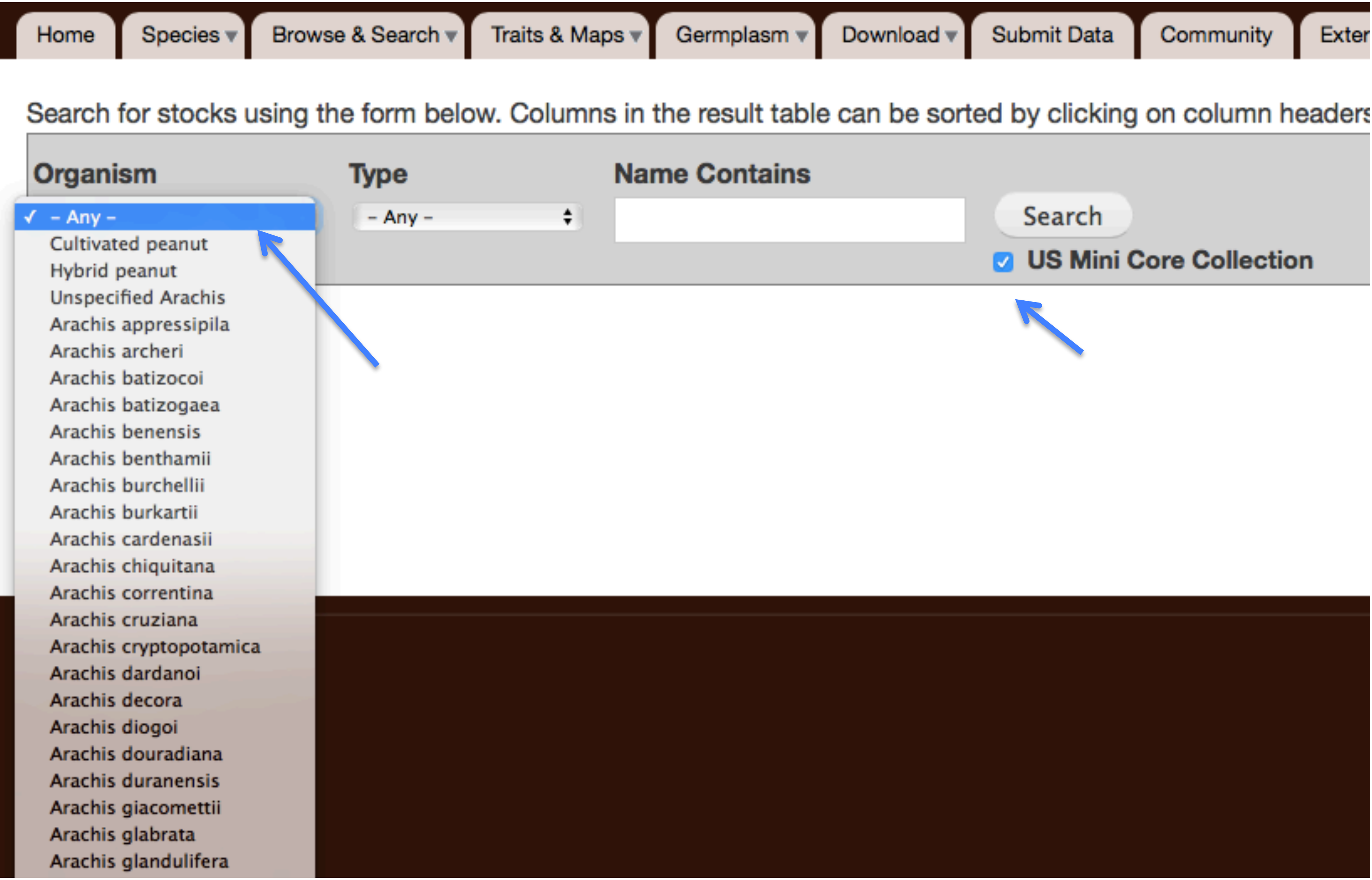
More News

**Cite PeanutBase an

Click for more on PeanutBase and the Peanut Genomics Initiative ...

Germplasm + traits

4. germplasm + traits



Search for stocks using the form below. Columns in the result table can be sorted by clicking on column headers

Organism

✓ – Any –

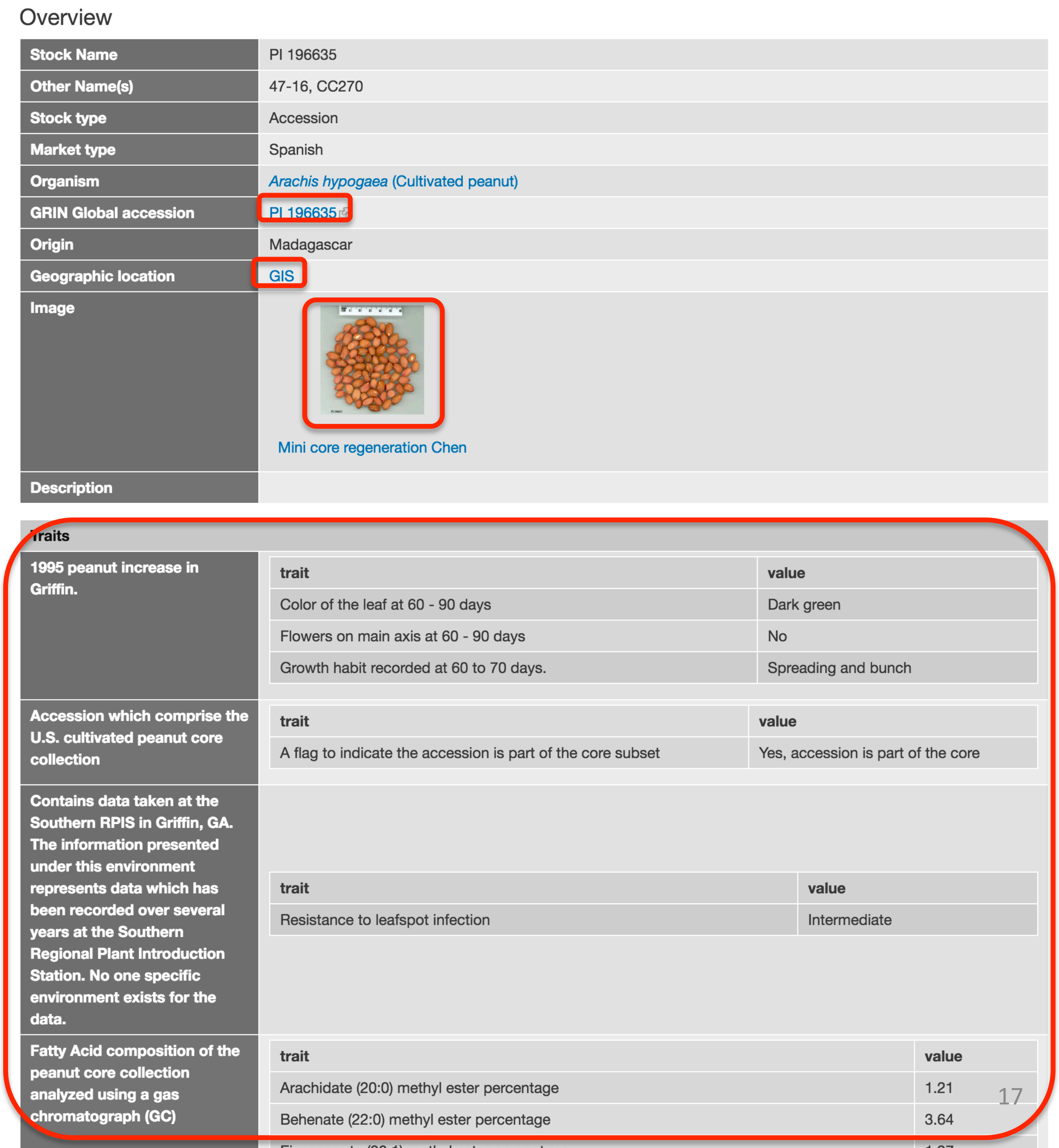
Type

– Any –

Name Contains

☑ **US Mini Core Collection**

Germplasm record



Overview

Stock Name	PI 196635
Other Name(s)	47-16, CC270
Stock type	Accession
Market type	Spanish
Organism	*Arachis hypogaea* (Cultivated peanut)
GRIN Global accession	PI 196635
Origin	Madagascar
Geographic location	GIS
Image	Mini core regeneration Chen
Description	

Traits

1995 peanut increase in Griffin.

trait	value
Color of the leaf at 60 - 90 days	Dark green
Flowers on main axis at 60 - 90 days	No
Growth habit recorded at 60 to 70 days.	Spreading and bunch

Accession which comprise the U.S. cultivated peanut core collection

trait	value
A flag to indicate the accession is part of the core subset	Yes, accession is part of the core

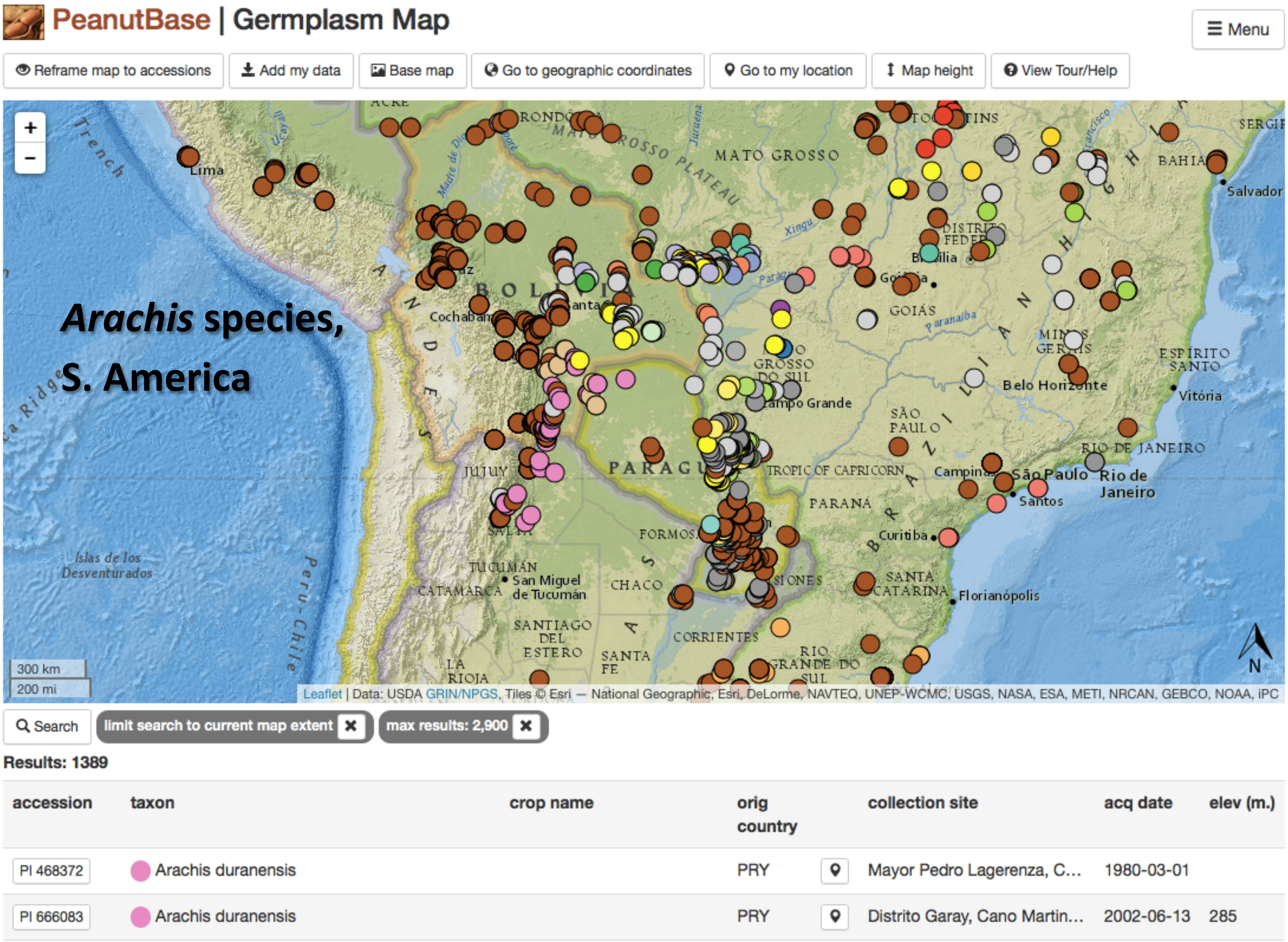
Contains data taken at the Southern RPIS in Griffin, GA. The information presented under this environment represents data which has been recorded over several years at the Southern Regional Plant Introduction Station. No one specific environment exists for the data.

trait	value
Resistance to leafspot infection	Intermediate

Fatty Acid composition of the peanut core collection analyzed using a gas chromatograph (GC)

trait	value
Arachidate (20:0) methyl ester percentage	1.21
Behenate (22:0) methyl ester percentage	3.64

Arachis species, S. America



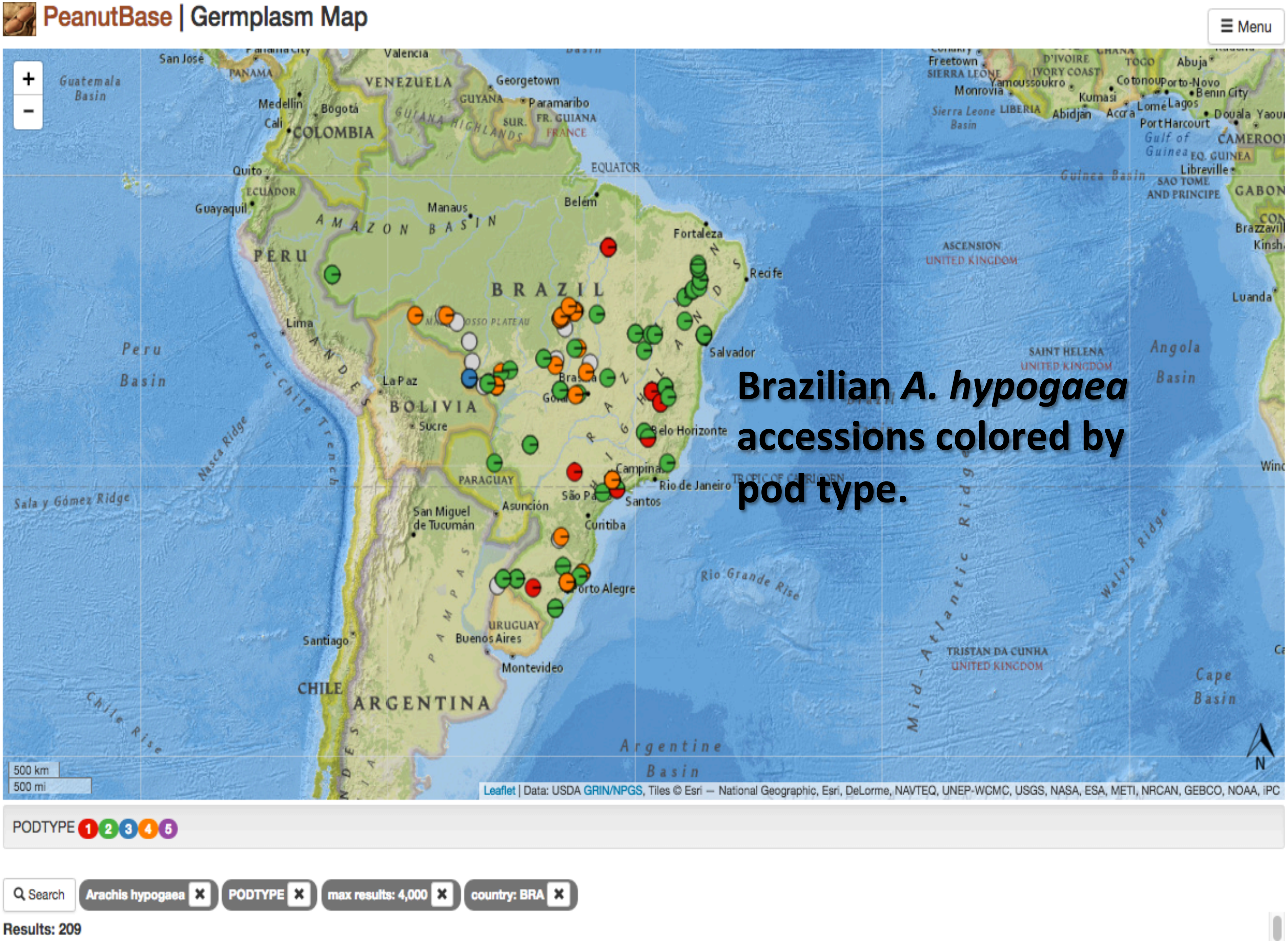
Results: 1389

accession	taxon	crop name	orig country		collection site	acq date	elev (m.)
PI 468372	● Arachis duranensis		PRY	📍	Mayor Pedro Lagerenza, C...	1980-03-01	
PI 666083	● Arachis duranensis		PRY	📍	Distrito Garay, Cano Martin...	2002-06-13	285

PeanutBase | Germplasm Map

Mini-core locations

Leaflet | Data: USDA GRIN/NPGS, Tiles © Esri — National Geographic, Esri, DeLorme, NAVTEQ, UNEP-WCMC, USGS, NASA, ESA, METI, NRCAN, GEBCO, NOAA, iPC



PeanutBase | Germplasm Map

☰ Menu

Leaflet | Data: USDA GRIN/NPGS, Tiles © Esri — National Geographic, Esri, DeLorme, NAVTEQ, UNEP-WCMC, USGS, NASA, ESA, METI, NRCAN, GEBCO, NOAA, iPC

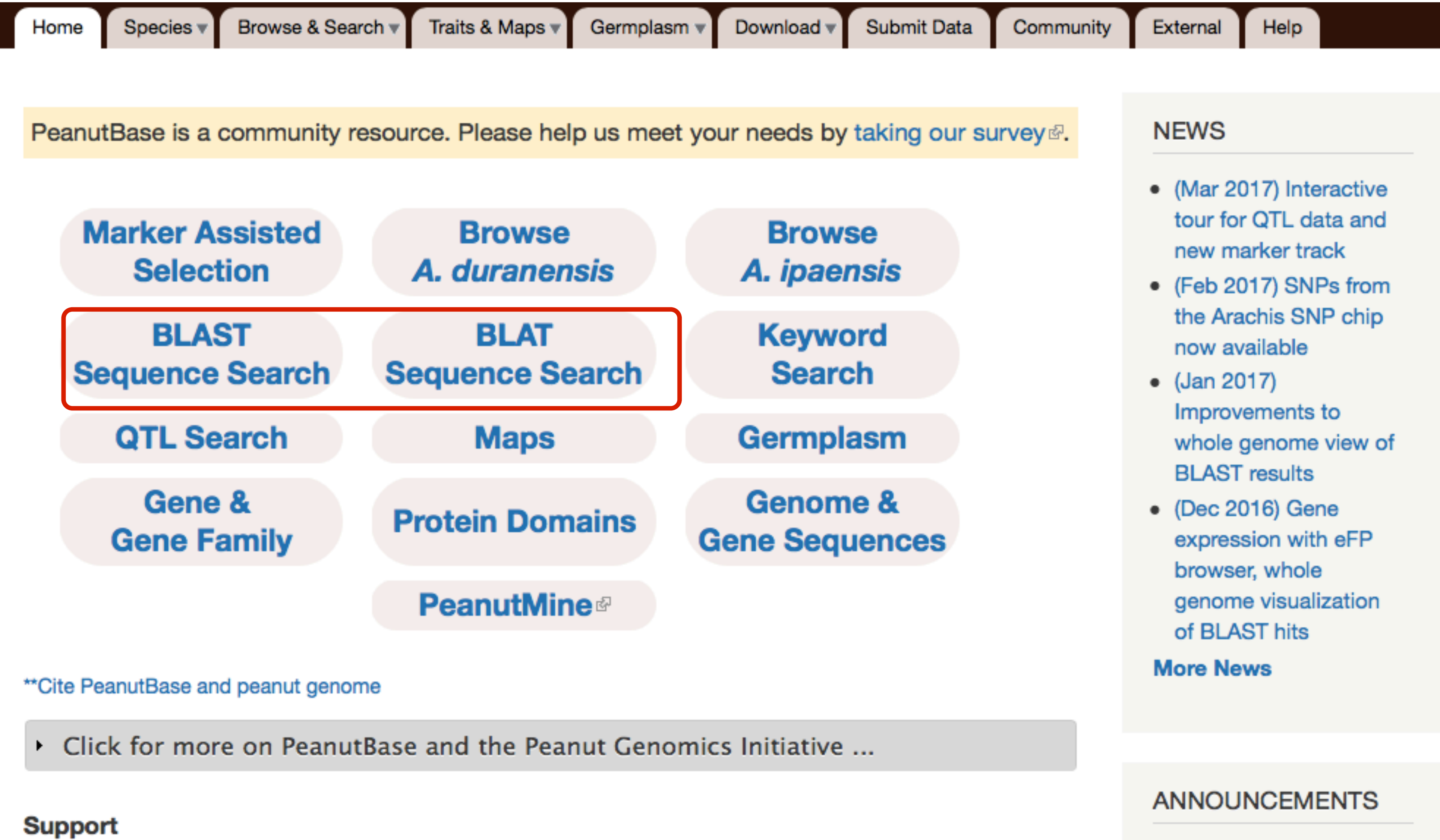
PODTYPE 1 2 3 4 5

Search | Arachis hypogaea ✖ | PODTYPE ✖ | max results: 4,000 ✖ | country: BRA ✖

Results: 209

Brazilian *A. hypogaea* accessions colored by pod type.

6. Sequence search tools



PeanutBase is a community resource. Please help us meet your needs by taking our survey.

**Cite PeanutBase and peanut genome

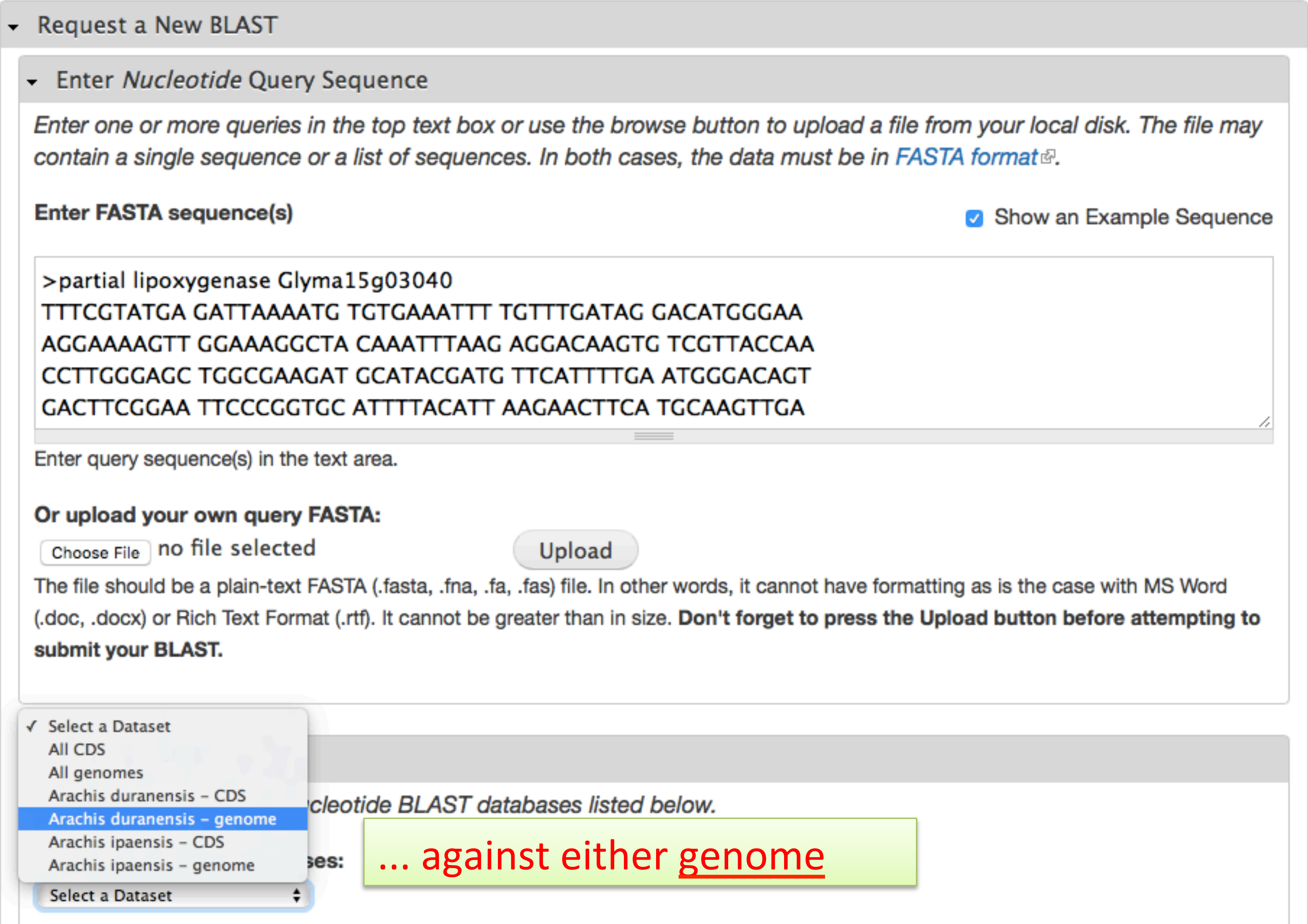
▸ Click for more on PeanutBase and the Peanut Genomics Initiative ...

Support

NEWS

- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis SNP chip now available
- (Jan 2017) Improvements to whole genome view of BLAST results
- (Dec 2016) Gene expression with eFP browser, whole genome visualization of BLAST hits

ANNOUNCEMENTS



Enter one or more queries in the top text box or use the browse button to upload a file from your local disk. The file may contain a single sequence or a list of sequences. In both cases, the data must be in FASTA format.

Enter FASTA sequence(s)

☑ Show an Example Sequence

```
>partial lipoxygenase Glyma15g03040
TTTCGTATGA GATTAAAATG TGTGAAATTT TGTTTGATAG GACATGGGAA
AGGAAAAGTT GGAAAGGCTA CAAATTTAAG AGGACAAGTG TCGTTACCAA
CCTTGGGAGC TGGCGAAGAT GCATACGATG TTCATTTTGA ATGGGACAGT
GACTTCGGAA TTCCCGGTGC ATTTTACATT AAGAACTTCA TGCAAGTTGA
```

Enter query sequence(s) in the text area.

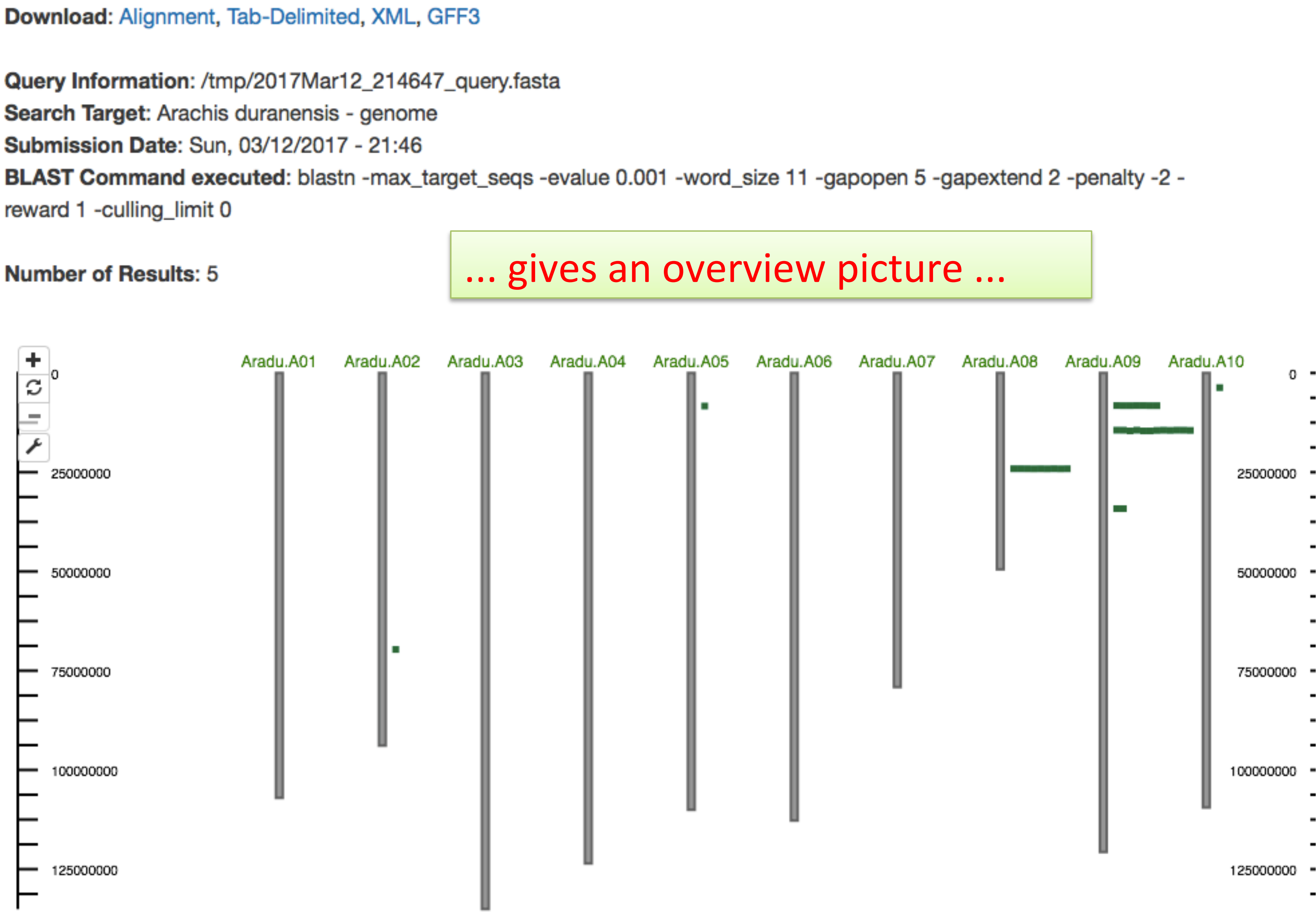
Or upload your own query FASTA:

The file should be a plain-text FASTA (.fasta, .fna, .fa, .fas) file. In other words, it cannot have formatting as is the case with MS Word (.doc, .docx) or Rich Text Format (.rtf). It cannot be greater than in size. **Don't forget to press the Upload button before attempting to submit your BLAST.**

cleotide BLAST databases listed below.

ses:

... against either <u>genome</u>

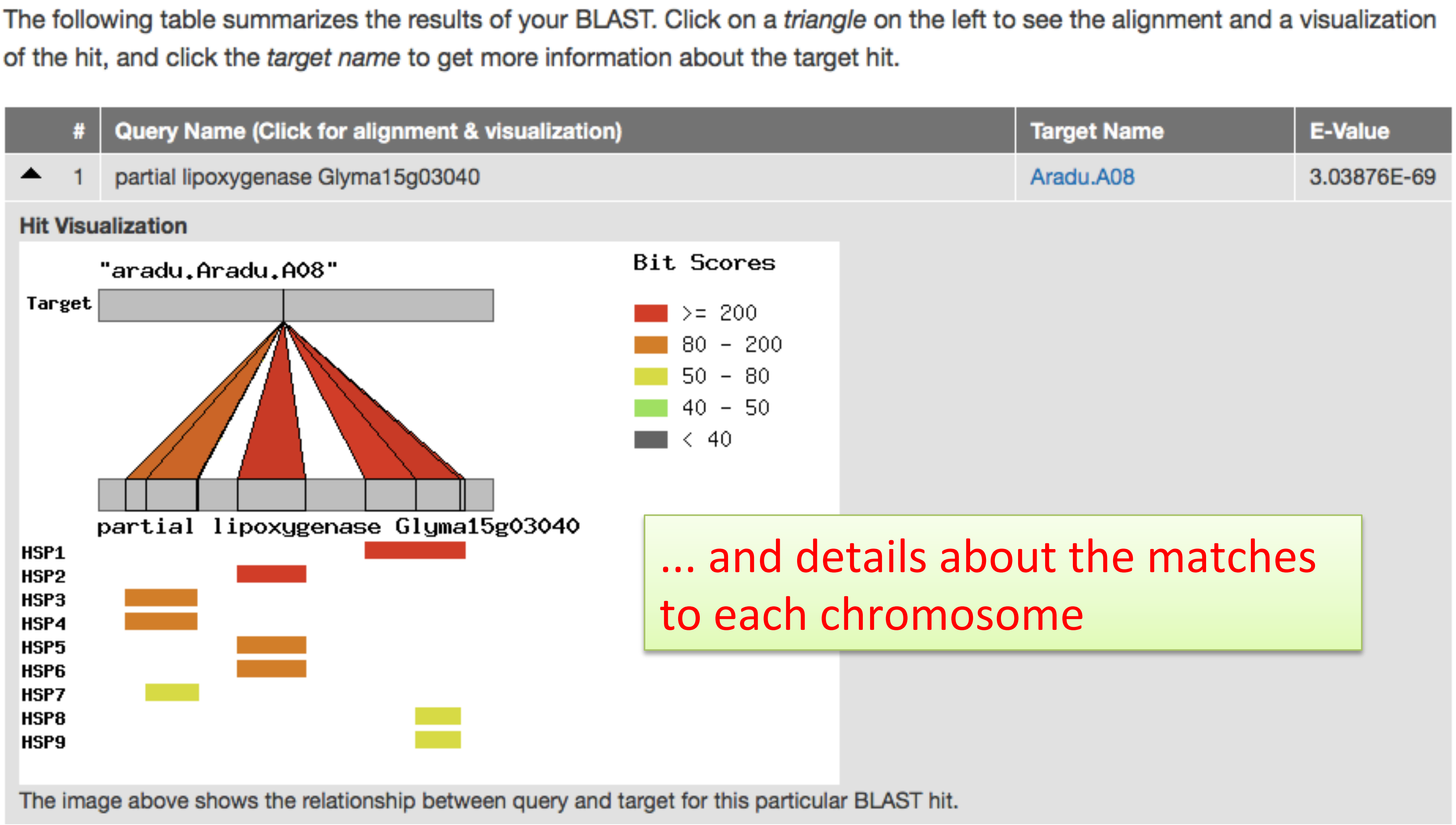


Download: Alignment, Tab-Delimited, XML, GFF3

Query Information: /tmp/2017Mar12_214647_query.fasta
Search Target: Arachis duranensis - genome
Submission Date: Sun, 03/12/2017 - 21:46
BLAST Command executed: blastn -max_target_seqs -evalue 0.001 -word_size 11 -gapopen 5 -gapextend 2 -penalty -2 -reward 1 -culling_limit 0

Number of Results: 5

... gives an overview picture ...



The following table summarizes the results of your BLAST. Click on a *triangle* on the left to see the alignment and a visualization of the hit, and click the *target name* to get more information about the target hit.

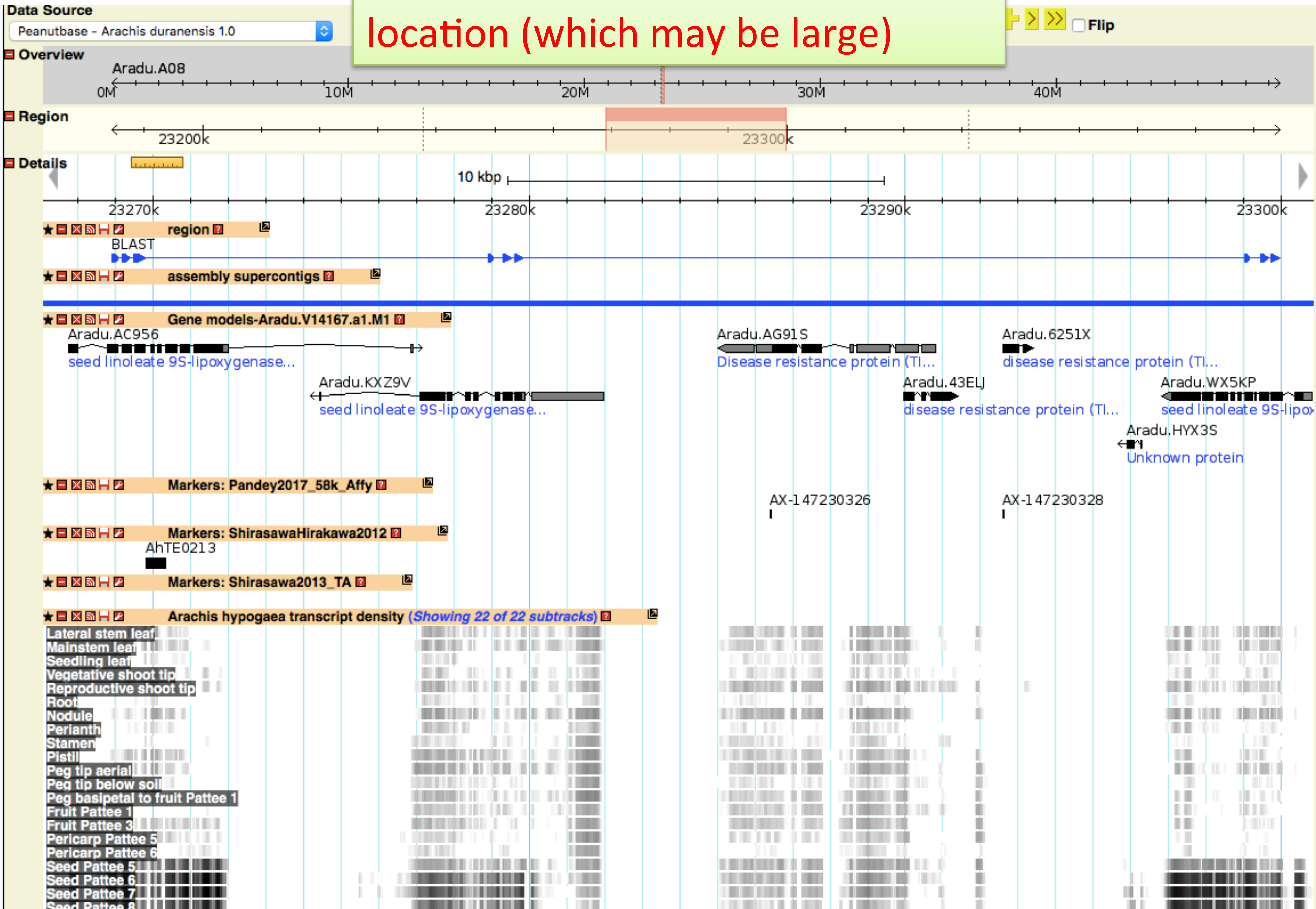
	#	Query Name (Click for alignment & visualization)	Target Name	E-Value
▲	1	partial lipoxygenase Glyma15g03040	Aradu.A08	3.03876E-69

Hit Visualization

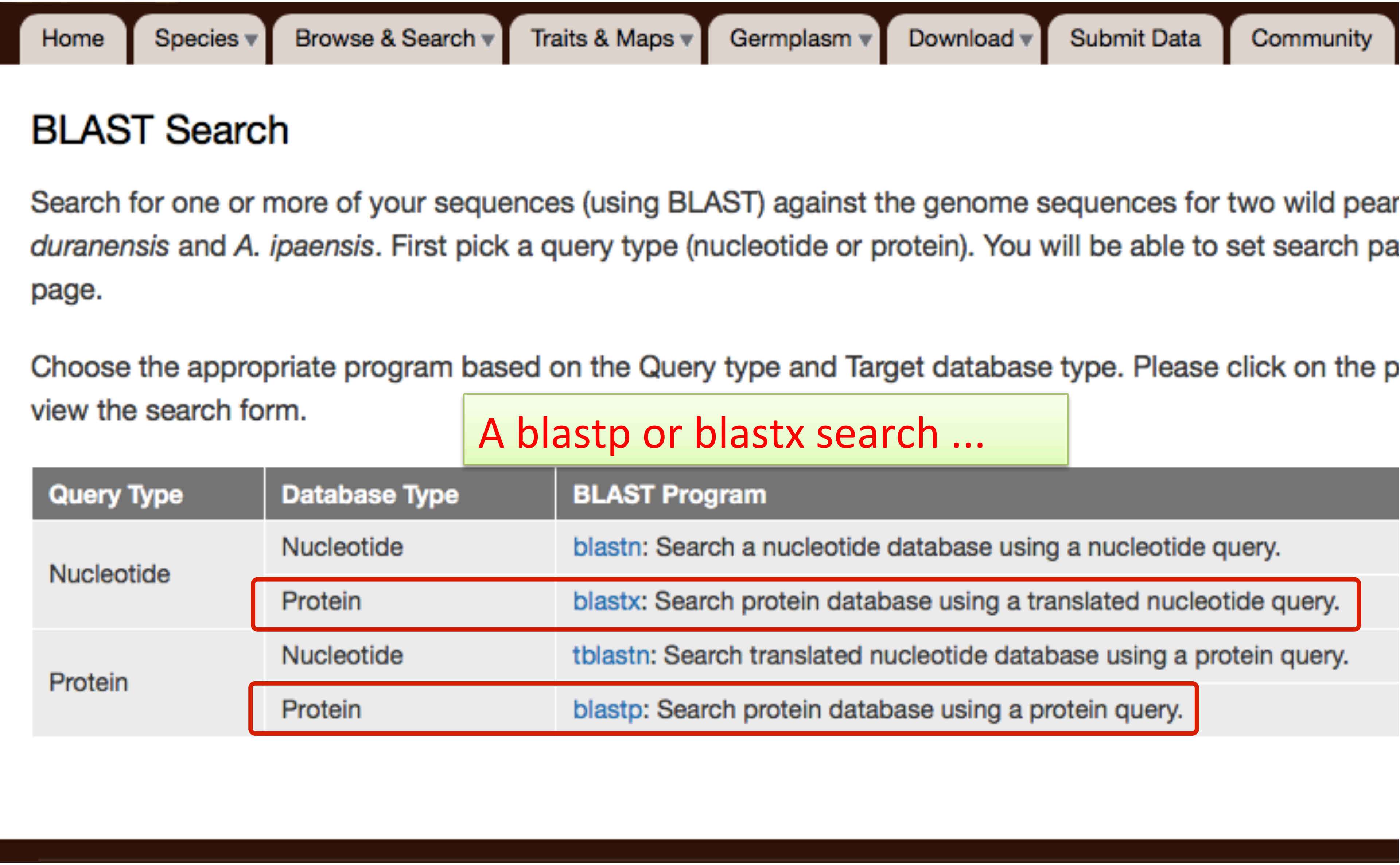
The image above shows the relationship between query and target for this particular BLAST hit.

... and details about the matches to each chromosome

… which link to GBrowse at that location (which may be large)



7. Gene families and synteny with other legumes



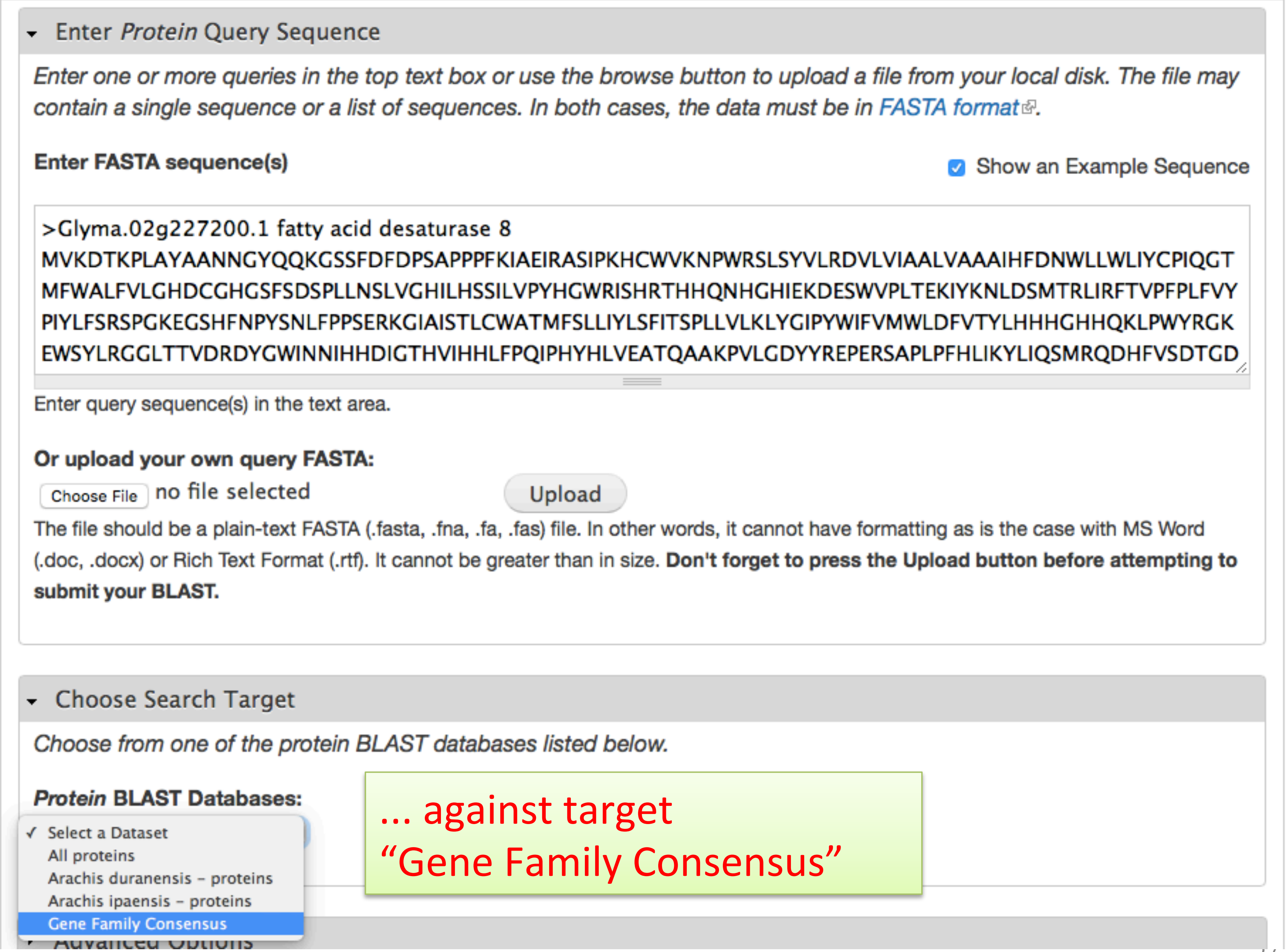
BLAST Search

Search for one or more of your sequences (using BLAST) against the genome sequences for two wild pear *duranensis* and *A. ipaensis*. First pick a query type (nucleotide or protein). You will be able to set search pa page.

Choose the appropriate program based on the Query type and Target database type. Please click on the p view the search form.

A blastp or blastx search ...

Query Type	Database Type	BLAST Program
Nucleotide	Nucleotide	blastn: Search a nucleotide database using a nucleotide query.
	Protein	blastx: Search protein database using a translated nucleotide query.
Protein	Nucleotide	tblastn: Search translated nucleotide database using a protein query.
	Protein	blastp: Search protein database using a protein query.



Enter *Protein* Query Sequence

Enter one or more queries in the top text box or use the browse button to upload a file from your local disk. The file may contain a single sequence or a list of sequences. In both cases, the data must be in FASTA format.

Enter FASTA sequence(s)

☑ Show an Example Sequence

```
>Glyma.02g227200.1 fatty acid desaturase 8
MVKDTKPLAYAANNGYQQKGSSFDFDPSAPPPFKIAEIRASIPKHCWVKNPWRSLSYVLRDVLVIAALVAAAIHFDNWLLWLIYCPIQGT
MFWALFVLGHDCGHGSFSDSPLLNSLVGHILHSSILVPYHGWRISHRTHHQNHGHIEKDESWVPLTEKIYKNLDSMTRLIRFTVPFPLFVY
PIYLFSRSPGKEGSHFNPYSNLFPPSERKGIAISTLCWATMFSLLIYLSFITSPLLVLKLYGIPYWIFVMWLDFVTYLHHHGHHQKLPWYRGK
EWSYLRGGLTTVDRDYGWINNIHHDIGTHVIHHLFPQIPHYHLVEATQAAKPVLGDYYREPERSAPLPFHLIKYLIQSMRQDHFVSDTGD
```

Enter query sequence(s) in the text area.

Or upload your own query FASTA:

no file selected

The file should be a plain-text FASTA (.fasta, .fna, .fa, .fas) file. In other words, it cannot have formatting as is the case with MS Word (.doc, .docx) or Rich Text Format (.rtf). It cannot be greater than in size. **Don't forget to press the Upload button before attempting to submit your BLAST.**

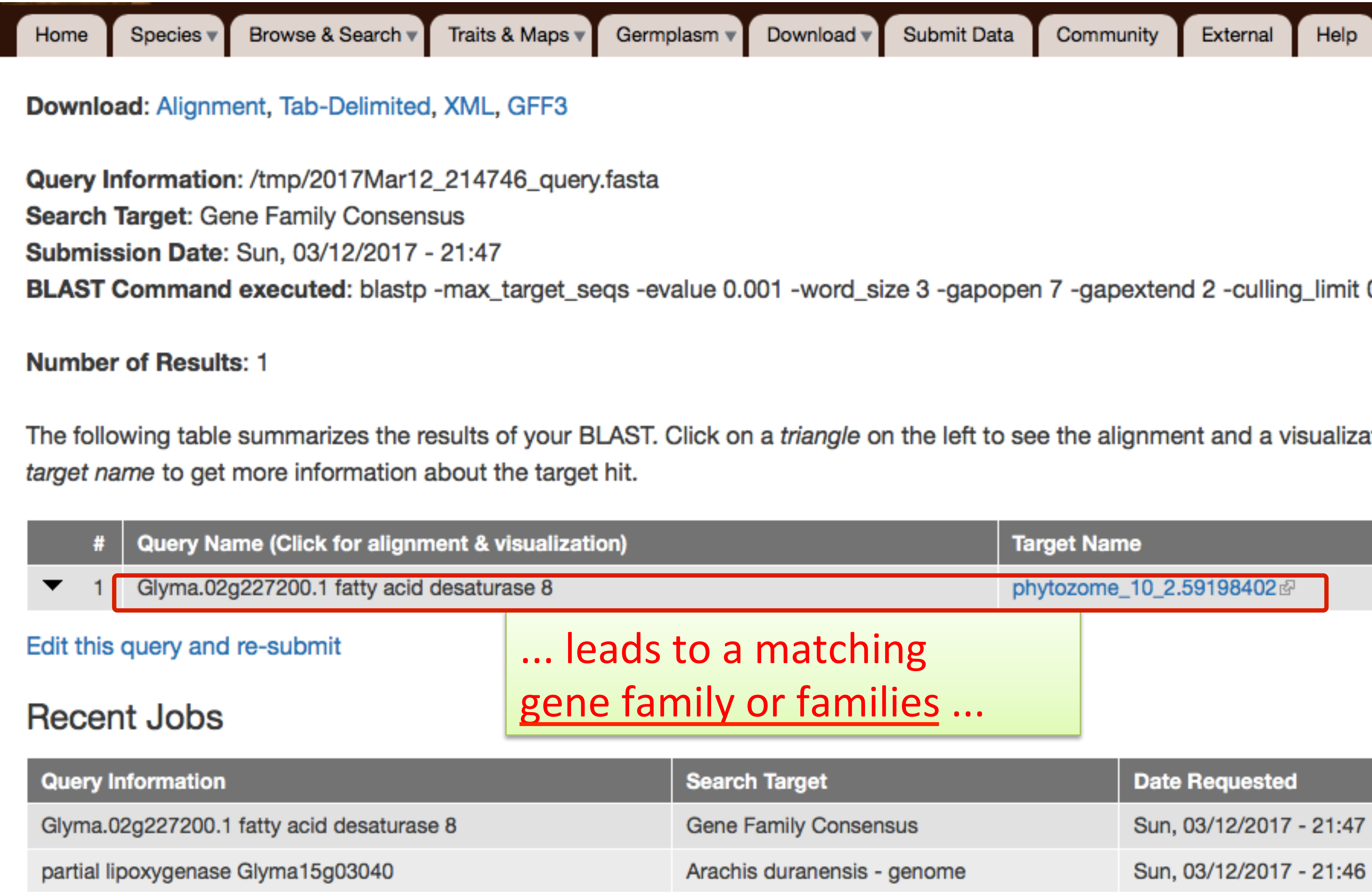
Choose Search Target

Choose from one of the protein BLAST databases listed below.

***Protein* BLAST Databases:** Gene Family Consensus

... against target
“Gene Family Consensus”

7. Gene families and synteny with other legumes



Download: Alignment, Tab-Delimited, XML, GFF3

Query Information: /tmp/2017Mar12_214746_query.fasta
Search Target: Gene Family Consensus
Submission Date: Sun, 03/12/2017 - 21:47
BLAST Command executed: blastp -max_target_seqs -evalue 0.001 -word_size 3 -gapopen 7 -gapextend 2 -culling_limit

Number of Results: 1

The following table summarizes the results of your BLAST. Click on a *triangle* on the left to see the alignment and a visualiza
target name to get more information about the target hit.

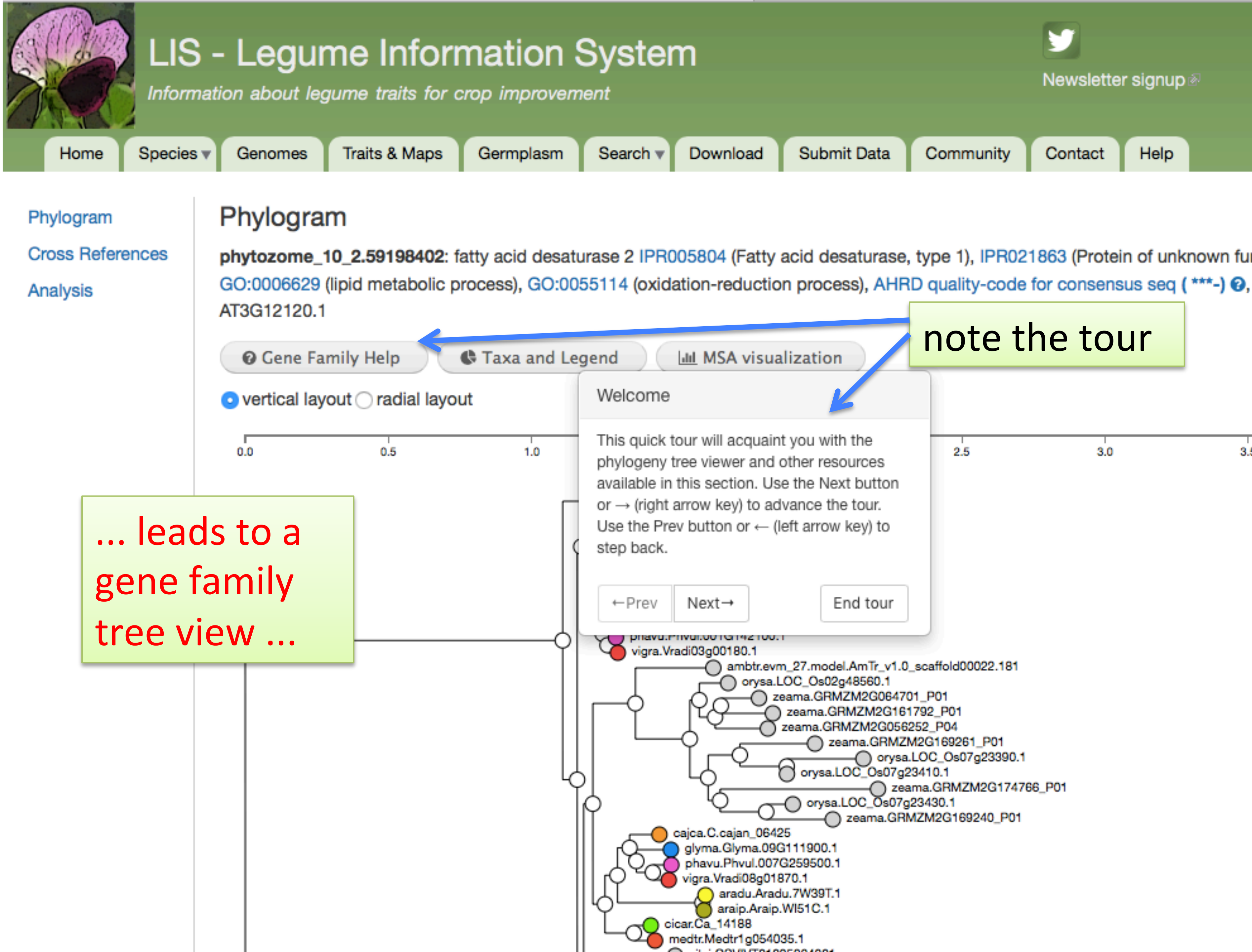
	#	Query Name (Click for alignment & visualization)	Target Name
▾	1	Glyma.02g227200.1 fatty acid desaturase 8	phytozome_10_2.59198402

Edit this query and re-submit

... leads to a matching
<u>gene family or families</u> ...

Recent Jobs

Query Information	Search Target	Date Requested
Glyma.02g227200.1 fatty acid desaturase 8	Gene Family Consensus	Sun, 03/12/2017 - 21:47
partial lipoxygenase Glyma15g03040	Arachis duranensis - genome	Sun, 03/12/2017 - 21:46



Phylogram

phytozome_10_2.59198402: fatty acid desaturase 2 IPR005804 (Fatty acid desaturase, type 1), IPR021863 (Protein of unknown fu GO:0006629 (lipid metabolic process), GO:0055114 (oxidation-reduction process), AHRD quality-code for consensus seq (***-) , AT3G12120.1

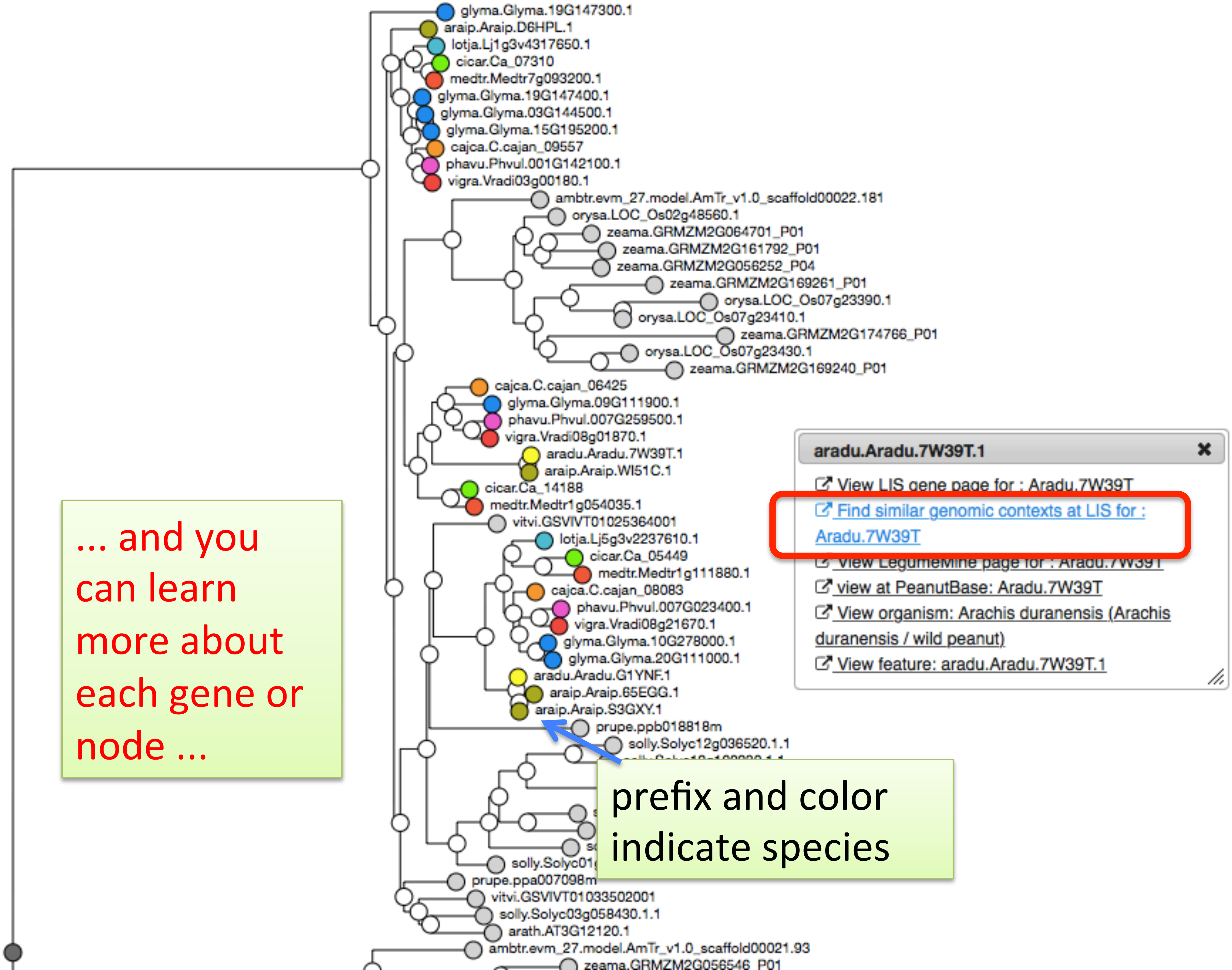
Welcome

This quick tour will acquaint you with the phylogeny tree viewer and other resources available in this section. Use the Next button or → (right arrow key) to advance the tour. Use the Prev button or ← (left arrow key) to step back.

note the tour

... leads to a gene family tree view ...

... and you can learn more about each gene or node ...

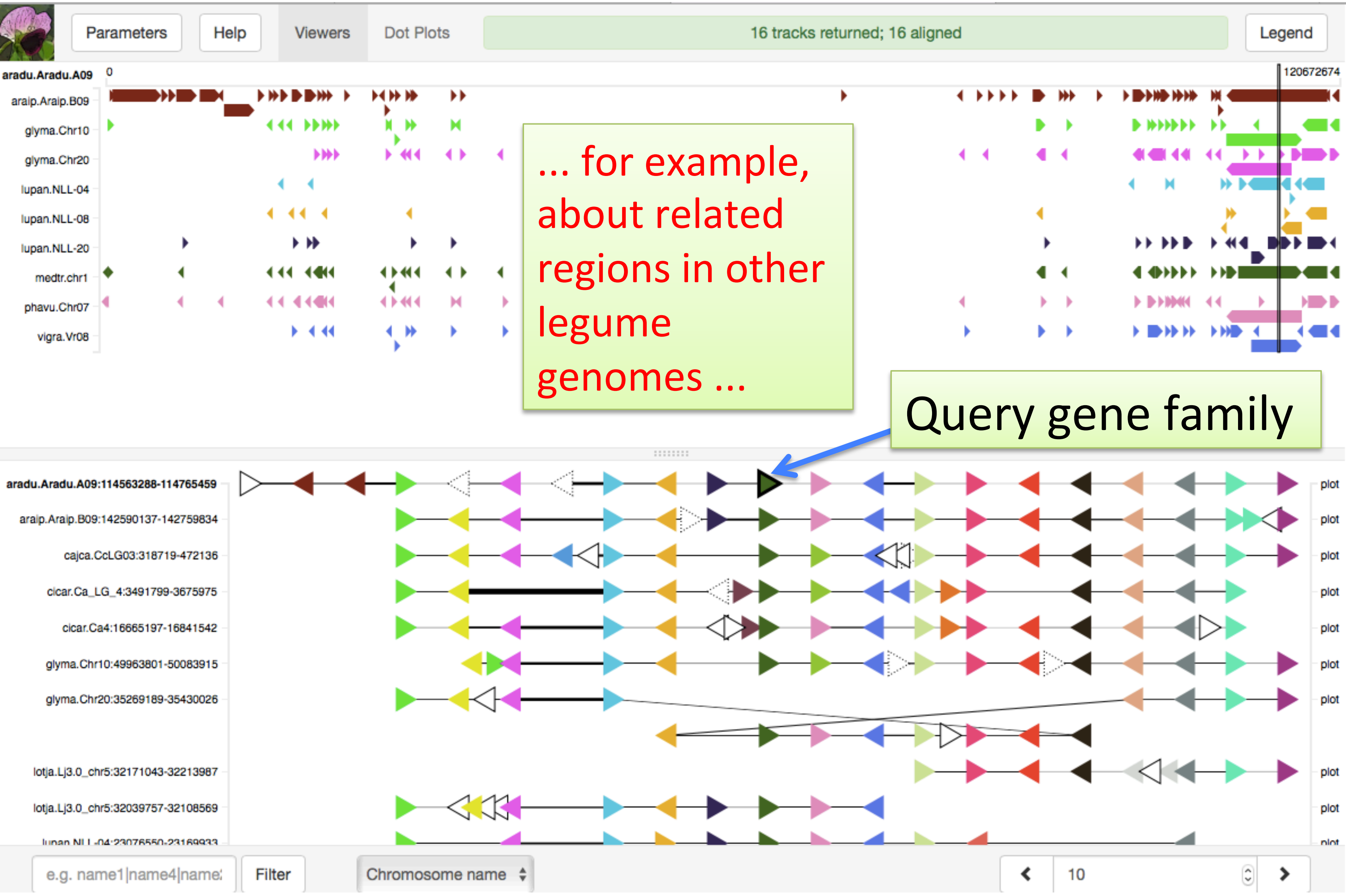


aradu.Aradu.7W39T.1 ✖

- View LIS gene page for : Aradu.7W39T
- Find similar genomic contexts at LIS for : Aradu.7W39T
- View LegumeMine page for : Aradu.7W39T
- view at PeanutBase: Aradu.7W39T
- View organism: Arachis duranensis (Arachis duranensis / wild peanut)
- View feature: aradu.Aradu.7W39T.1

prefix and color indicate species

7. Gene families and synteny with other legumes

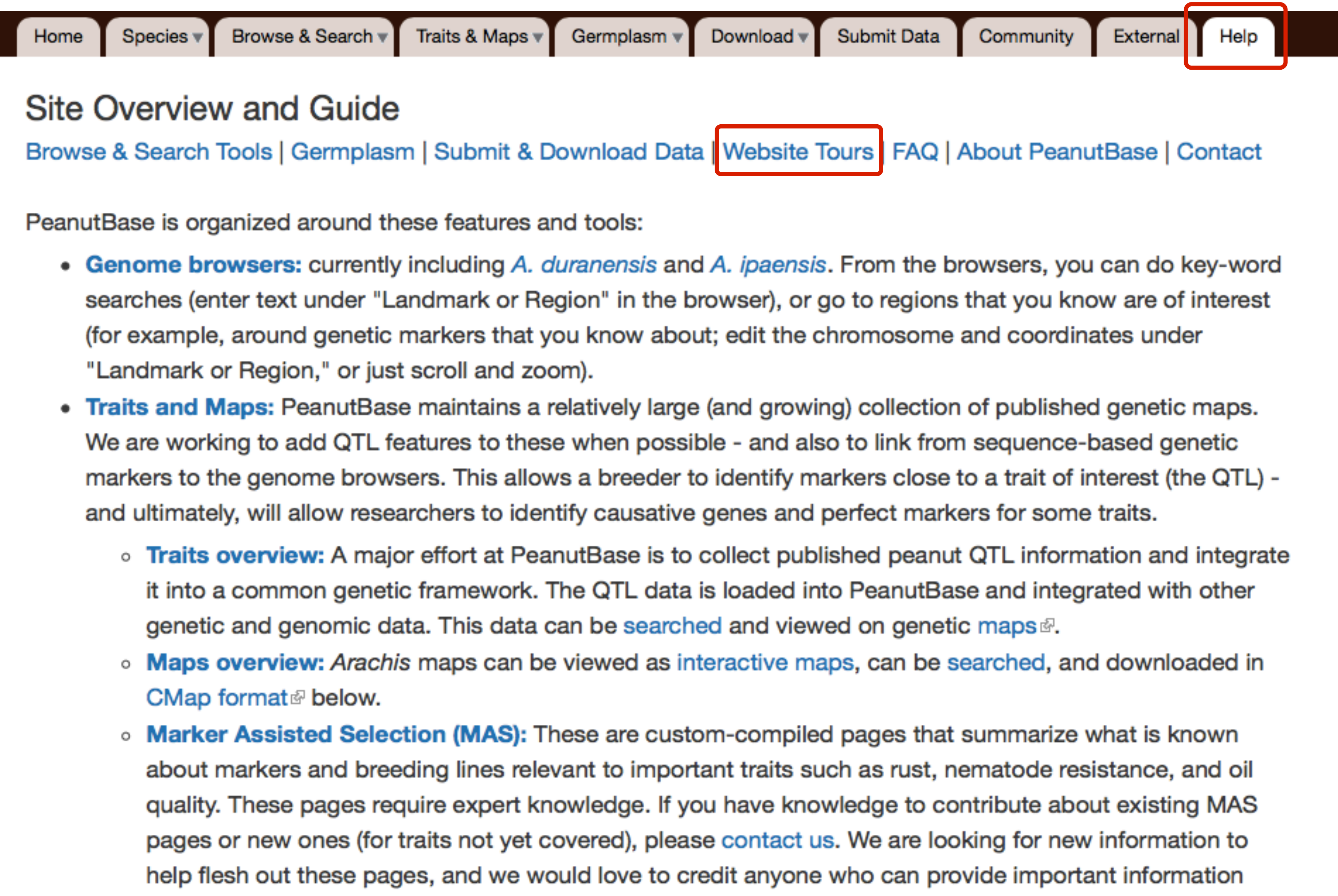


16 tracks returned; 16 aligned

... for example, about related regions in other legume genomes ...

Query gene family

8. Website tours (and other help)



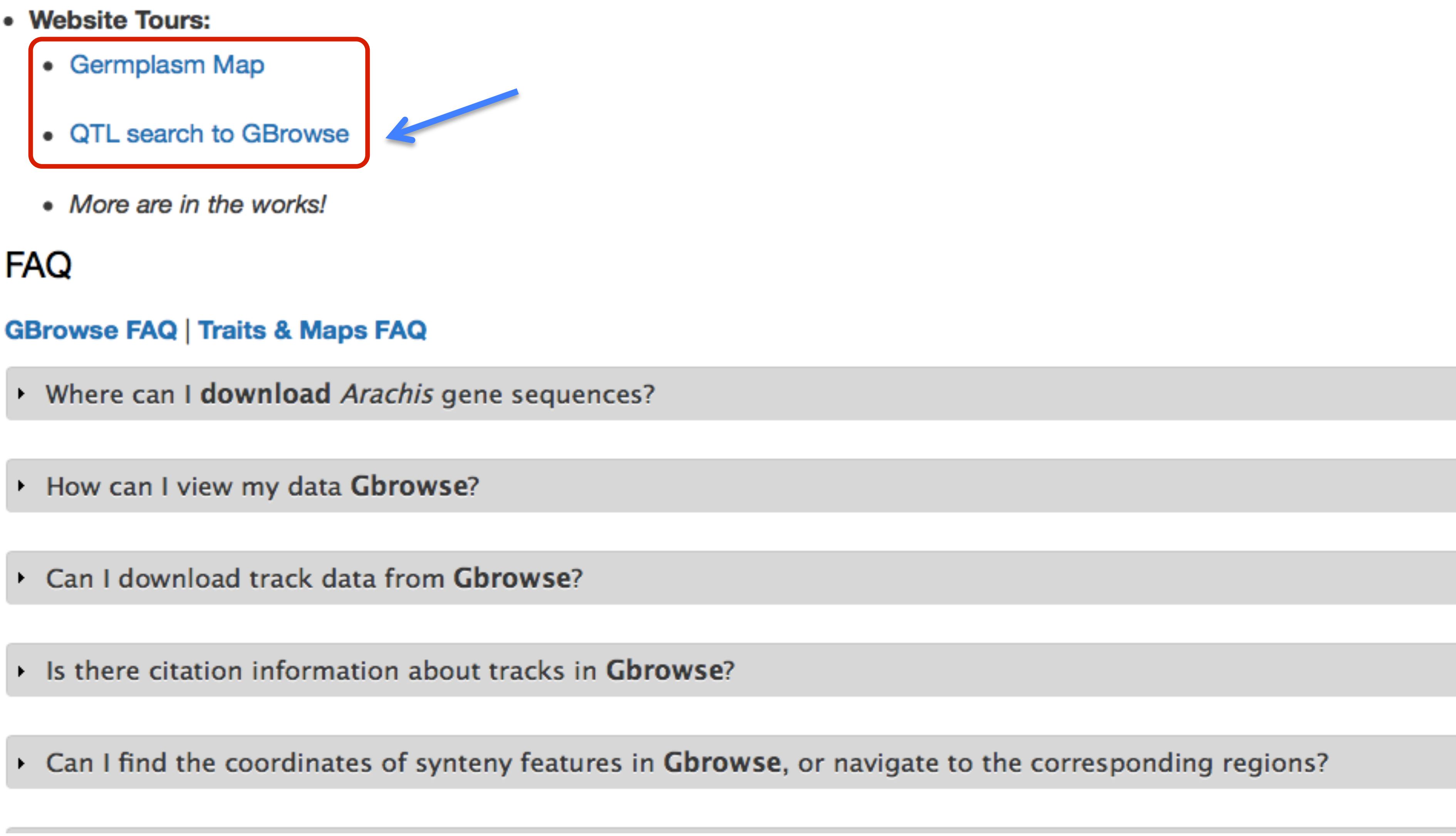
Site Overview and Guide

Browse & Search Tools | Germplasm | Submit & Download Data | Website Tours | FAQ | About PeanutBase | Contact

PeanutBase is organized around these features and tools:

- **Genome browsers:** currently including *A. duranensis* and *A. ipaensis*. From the browsers, you can do key-word searches (enter text under "Landmark or Region" in the browser), or go to regions that you know are of interest (for example, around genetic markers that you know about; edit the chromosome and coordinates under "Landmark or Region," or just scroll and zoom).
- **Traits and Maps:** PeanutBase maintains a relatively large (and growing) collection of published genetic maps. We are working to add QTL features to these when possible - and also to link from sequence-based genetic markers to the genome browsers. This allows a breeder to identify markers close to a trait of interest (the QTL) - and ultimately, will allow researchers to identify causative genes and perfect markers for some traits.
    - **Traits overview:** A major effort at PeanutBase is to collect published peanut QTL information and integrate it into a common genetic framework. The QTL data is loaded into PeanutBase and integrated with other genetic and genomic data. This data can be searched and viewed on genetic maps.
    - **Maps overview:** *Arachis* maps can be viewed as interactive maps, can be searched, and downloaded in CMap format below.
    - **Marker Assisted Selection (MAS):** These are custom-compiled pages that summarize what is known about markers and breeding lines relevant to important traits such as rust, nematode resistance, and oil quality. These pages require expert knowledge. If you have knowledge to contribute about existing MAS pages or new ones (for traits not yet covered), please contact us. We are looking for new information to help flesh out these pages, and we would love to credit anyone who can provide important information

8. Website tours (and other help)

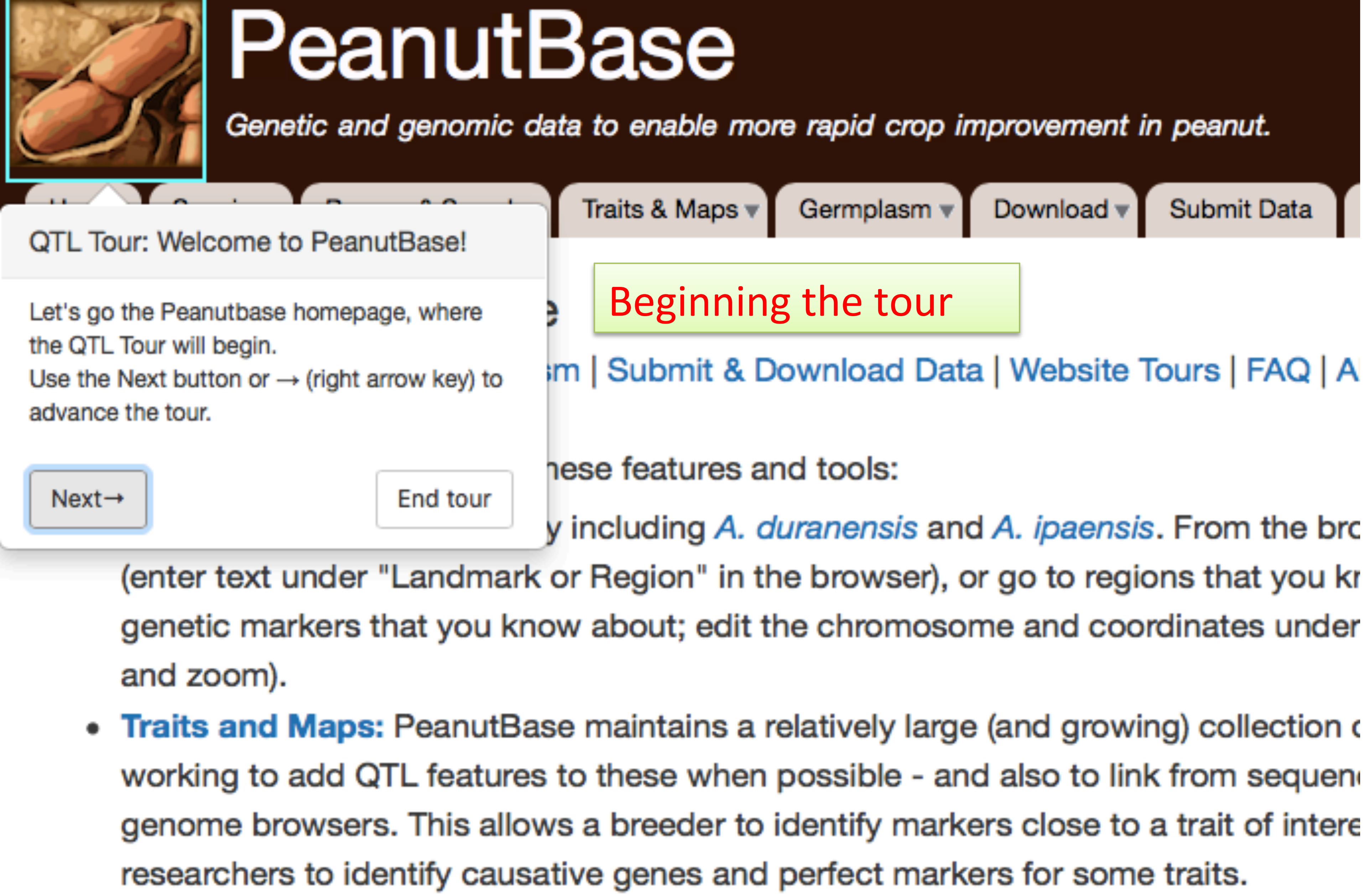


- **Website Tours:**
  - Germplasm Map
  - QTL search to GBrowse
  - *More are in the works!*

FAQ

GBrowse FAQ | **Traits & Maps FAQ**

- Where can I **download** *Arachis* gene sequences?
- How can I view my data **Gbrowse**?
- Can I download track data from **Gbrowse**?
- Is there citation information about tracks in **Gbrowse**?
- Can I find the coordinates of synteny features in **Gbrowse**, or navigate to the corresponding regions?



PeanutBase

Genetic and genomic data to enable more rapid crop improvement in peanut.

QTL Tour: Welcome to PeanutBase!

Let's go the Peanutbase homepage, where the QTL Tour will begin.
Use the Next button or → (right arrow key) to advance the tour.

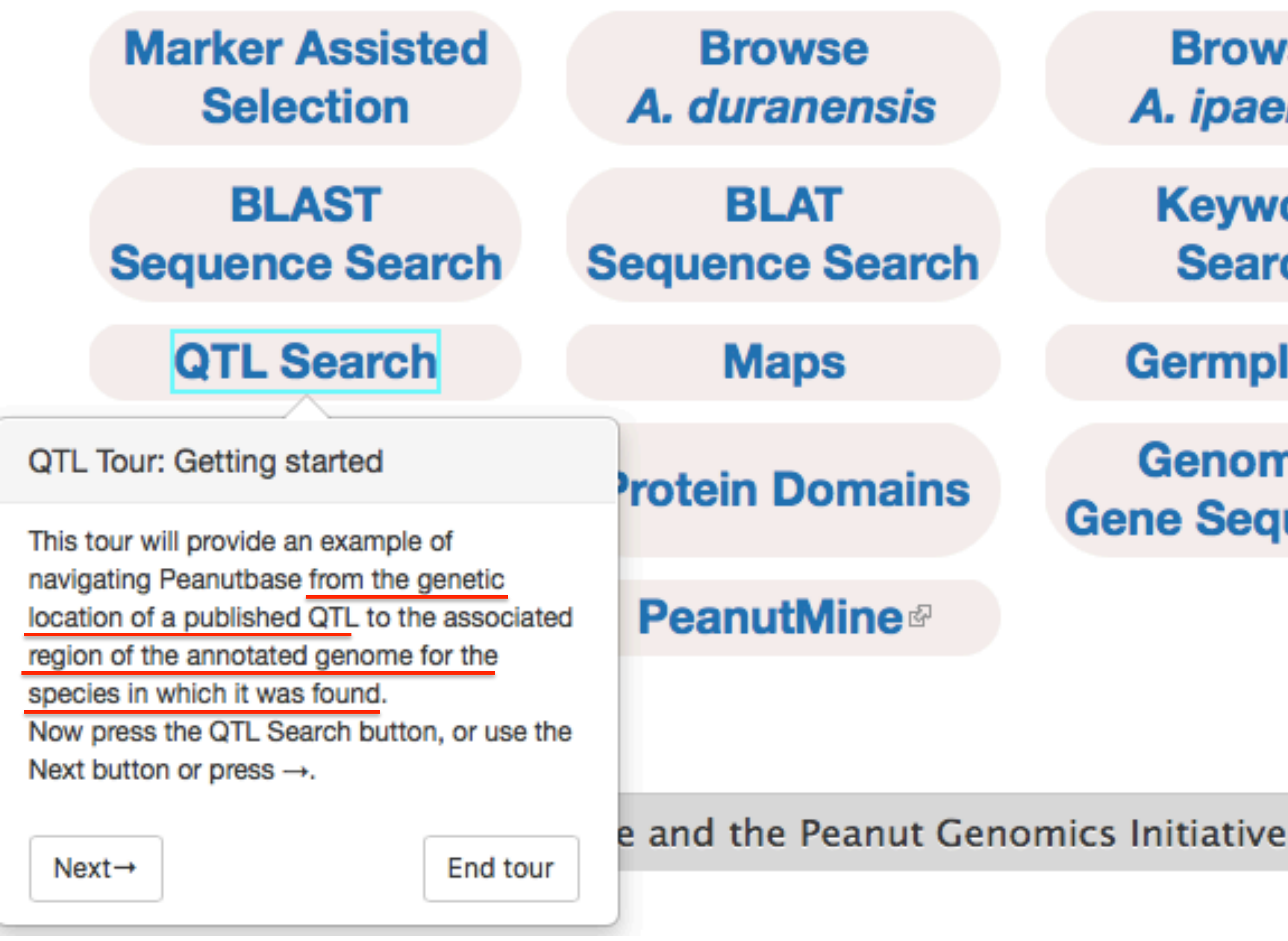
Beginning the tour

sm | Submit & Download Data | Website Tours | FAQ | A

nese features and tools:

y including *A. duranensis* and *A. ipaensis*. From the bro (enter text under "Landmark or Region" in the browser), or go to regions that you kr genetic markers that you know about; edit the chromosome and coordinates under and zoom).

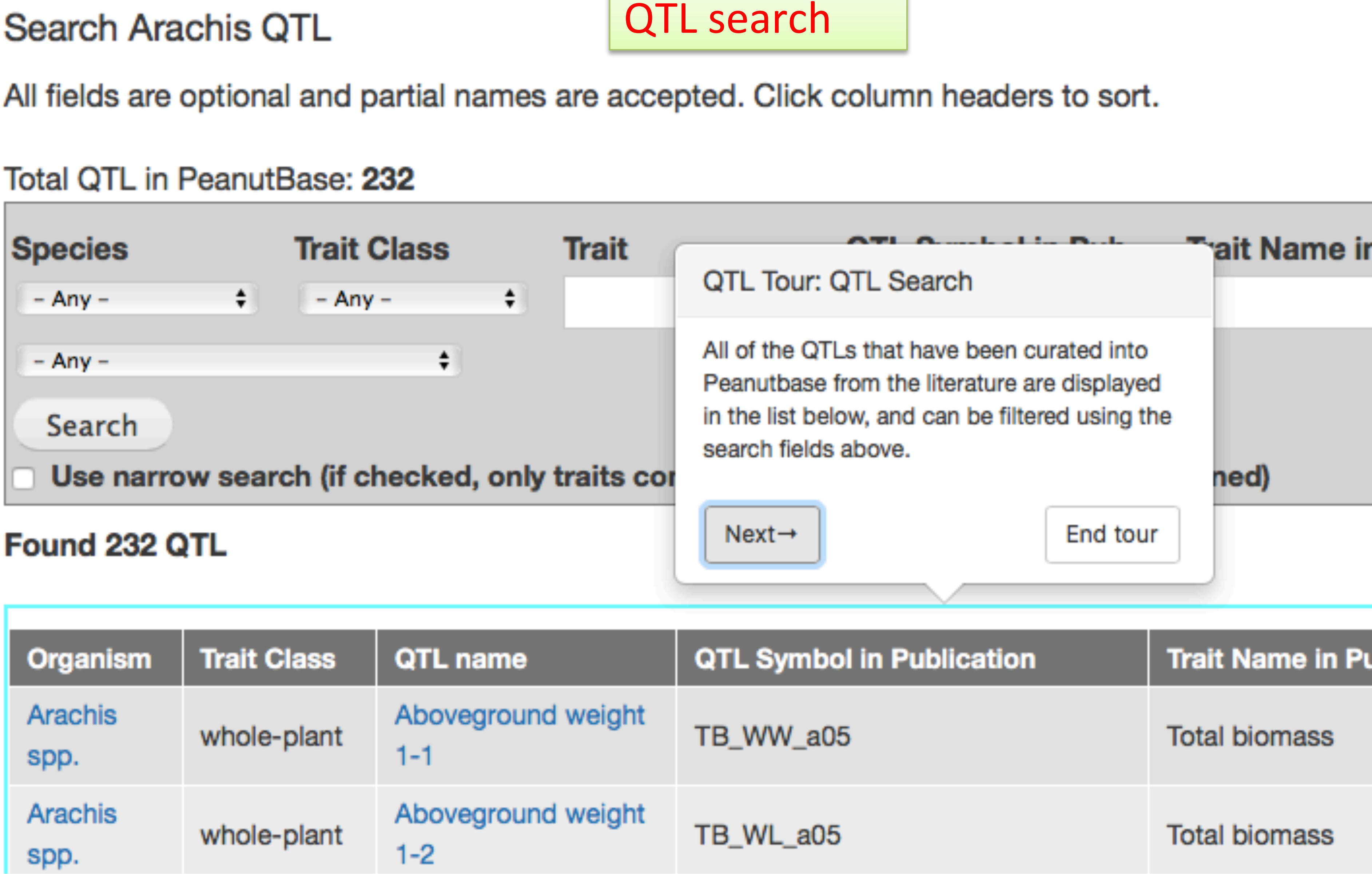
- **Traits and Maps:** PeanutBase maintains a relatively large (and growing) collection working to add QTL features to these when possible - and also to link from sequen genome browsers. This allows a breeder to identify markers close to a trait of intere researchers to identify causative genes and perfect markers for some traits.



QTL Tour: Getting started

This tour will provide an example of navigating Peanutbase from the genetic location of a published QTL to the associated region of the annotated genome for the species in which it was found.
Now press the QTL Search button, or use the Next button or press →.

QTL search



Search Arachis QTL

All fields are optional and partial names are accepted. Click column headers to sort.

Total QTL in PeanutBase: **232**

☐ **Use narrow search (if checked, only traits co** **ned)**

QTL Tour: QTL Search

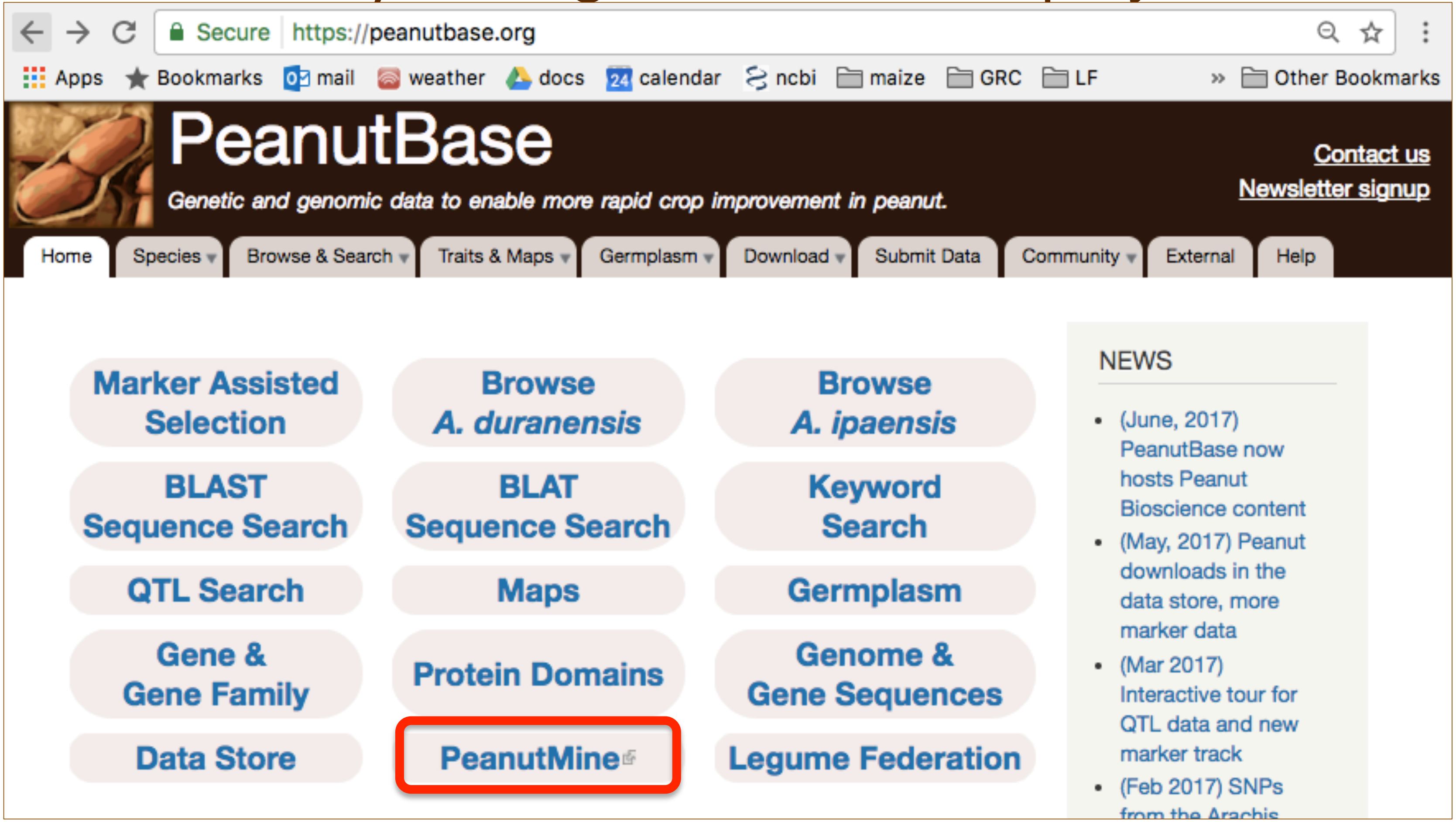
All of the QTLs that have been curated into Peanutbase from the literature are displayed in the list below, and can be filtered using the search fields above.

Found 232 QTL

Organism	Trait Class	QTL name	QTL Symbol in Publication	Trait Name in Pu
Arachis spp.	whole-plant	Aboveground weight 1-1	TB_WW_a05	Total biomass
Arachis spp.	whole-plant	Aboveground weight 1-2	TB_WL_a05	Total biomass

Related resource - PeanutMine

- Provided by the Legume Federation project



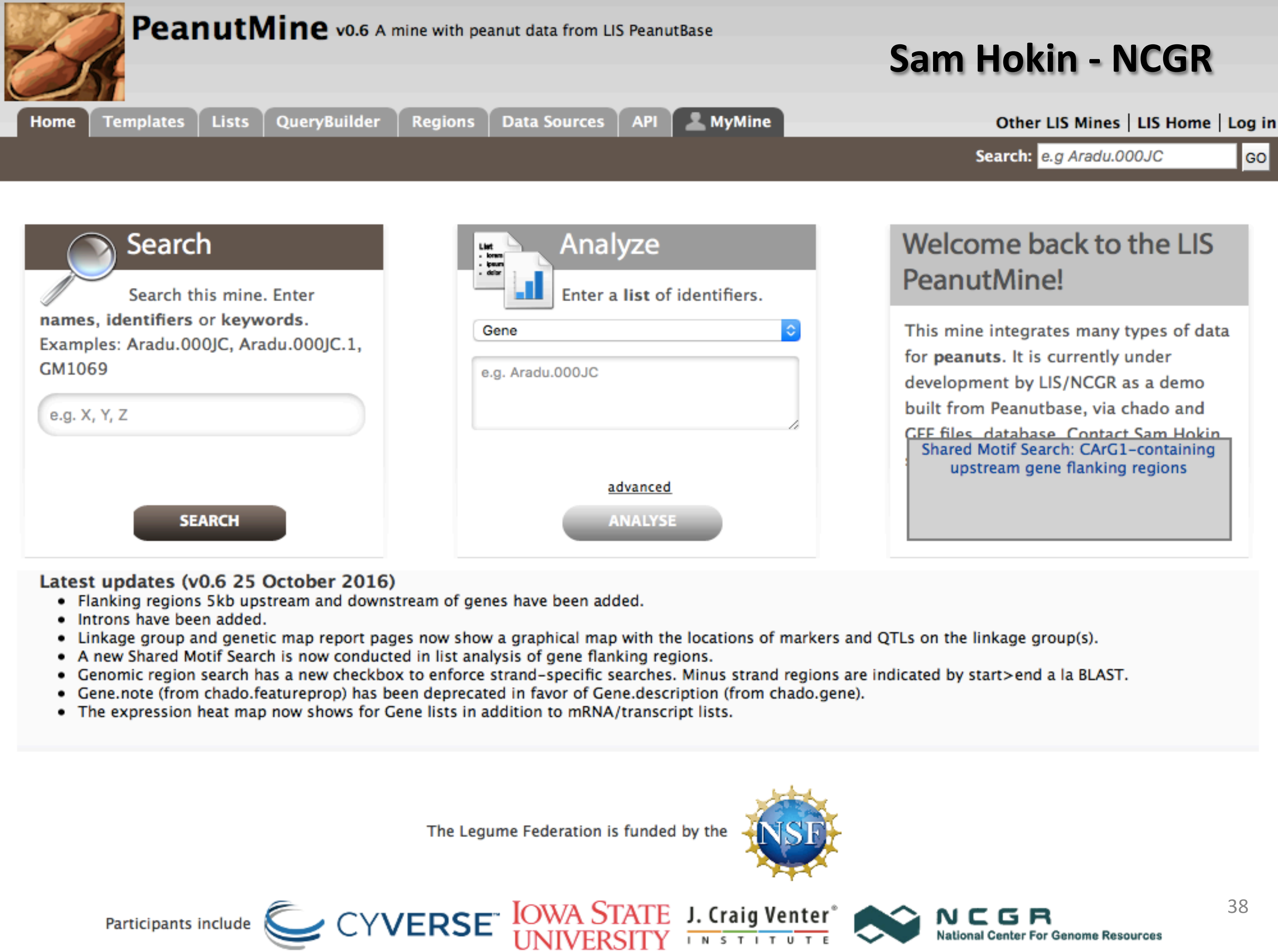
PeanutBase

Genetic and genomic data to enable more rapid crop improvement in peanut.

NEWS

- (June, 2017) PeanutBase now hosts Peanut Bioscience content
- (May, 2017) Peanut downloads in the data store, more marker data
- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis

Sam Hokin - NCGR



PeanutMine v0.6 A mine with peanut data from LIS PeanutBase

Search

Search this mine. Enter **names, identifiers** or **keywords**. Examples: Aradu.000JC, Aradu.000JC.1, GM1069

Analyze

Enter a **list** of identifiers.

Welcome back to the LIS PeanutMine!

This mine integrates many types of data for **peanuts**. It is currently under development by LIS/NCGR as a demo built from Peanutbase, via chado and GFF files, database. Contact Sam Hokin

Shared Motif Search: CArG1-containing upstream gene flanking regions

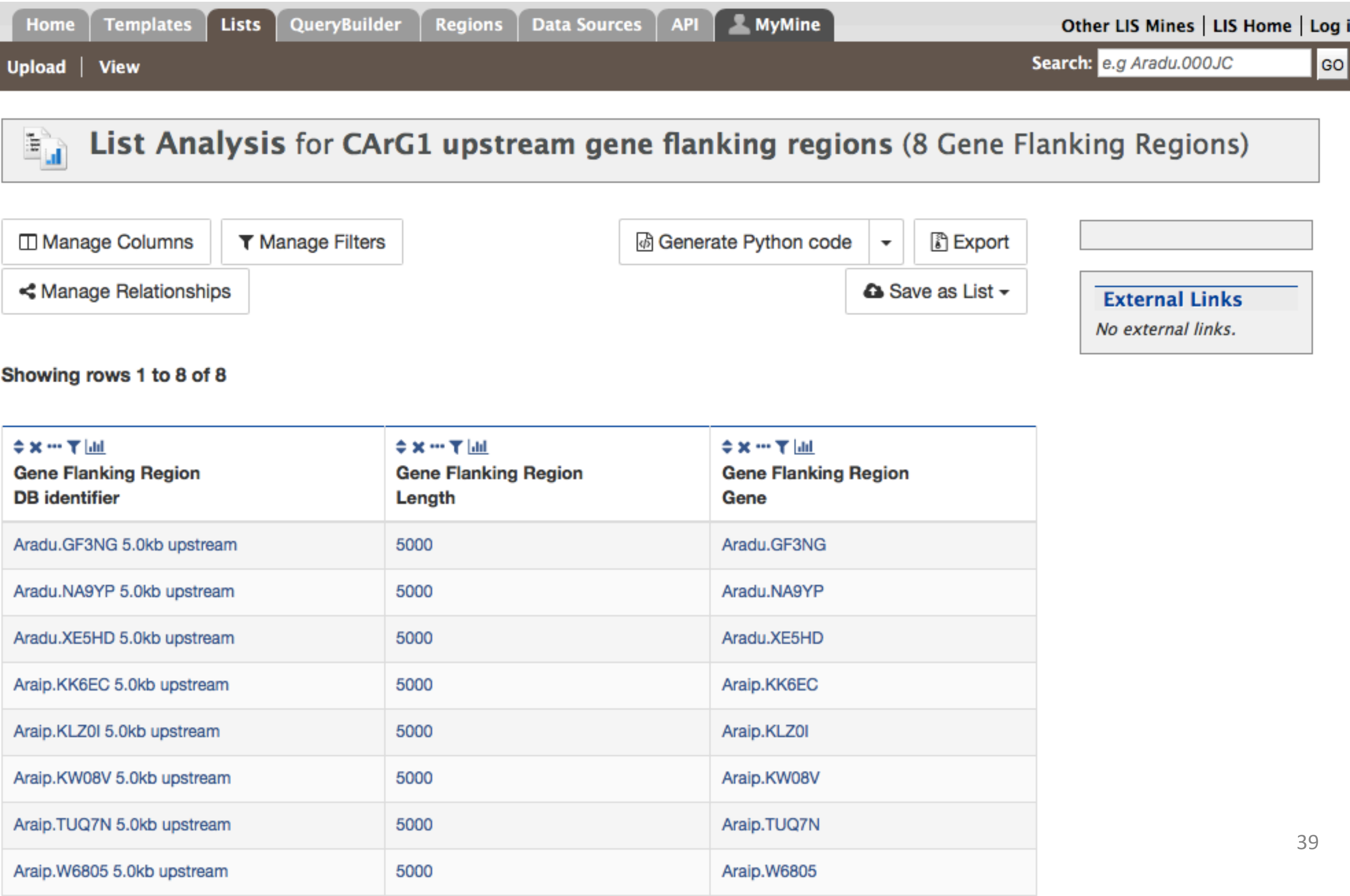
Latest updates (v0.6 25 October 2016)

- Flanking regions 5kb upstream and downstream of genes have been added.
- Introns have been added.
- Linkage group and genetic map report pages now show a graphical map with the locations of markers and QTLs on the linkage group(s).
- A new Shared Motif Search is now conducted in list analysis of gene flanking regions.
- Genomic region search has a new checkbox to enforce strand-specific searches. Minus strand regions are indicated by start>end a la BLAST.
- Gene.note (from chado.featureprop) has been deprecated in favor of Gene.description (from chado.gene).
- The expression heat map now shows for Gene lists in addition to mRNA/transcript lists.

The Legume Federation is funded by the NSF

Participants include CYVERSE, IOWA STATE UNIVERSITY, J. Craig Venter INSTITUTE, NCGR National Center For Genome Resources

9. PeanutMine (example: motifs common among a list of genes)



List Analysis for **CArG1 upstream gene flanking regions** (8 Gene Flanking Regions)

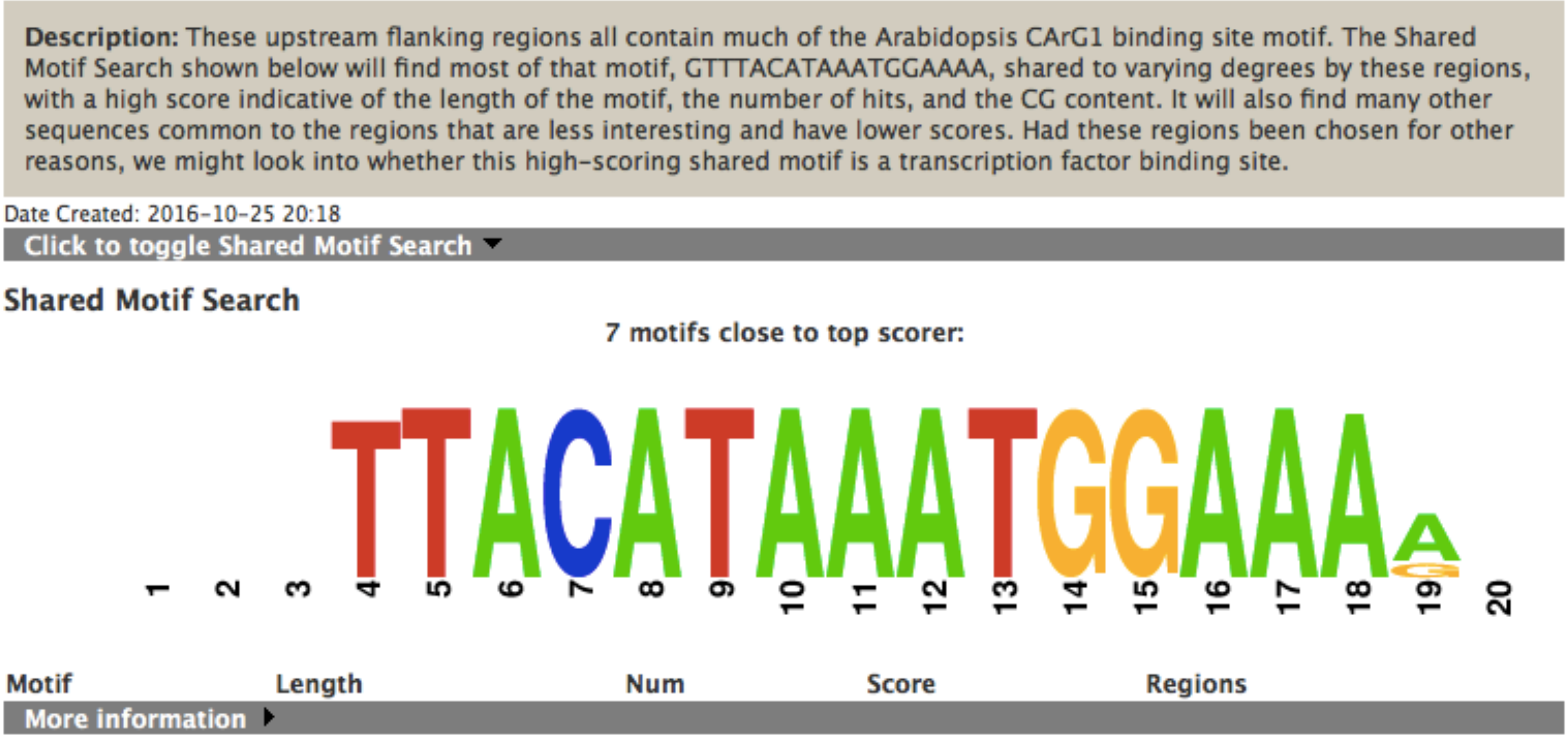
External Links

No external links.

Showing rows 1 to 8 of 8

Gene Flanking Region DB identifier	Gene Flanking Region Length	Gene Flanking Region Gene
Aradu.GF3NG 5.0kb upstream	5000	Aradu.GF3NG
Aradu.NA9YP 5.0kb upstream	5000	Aradu.NA9YP
Aradu.XE5HD 5.0kb upstream	5000	Aradu.XE5HD
Araip.KK6EC 5.0kb upstream	5000	Araip.KK6EC
Araip.KLZ0I 5.0kb upstream	5000	Araip.KLZ0I
Araip.KW08V 5.0kb upstream	5000	Araip.KW08V
Araip.TUQ7N 5.0kb upstream	5000	Araip.TUQ7N
Araip.W6805 5.0kb upstream	5000	Araip.W6805

9. PeanutMine (example: motifs common among a list of genes)



Description: These upstream flanking regions all contain much of the Arabidopsis CArG1 binding site motif. The Shared Motif Search shown below will find most of that motif, GTTTACATAAATGGAAAA, shared to varying degrees by these regions, with a high score indicative of the length of the motif, the number of hits, and the CG content. It will also find many other sequences common to the regions that are less interesting and have lower scores. Had these regions been chosen for other reasons, we might look into whether this high-scoring shared motif is a transcription factor binding site.

Date Created: 2016-10-25 20:18

Shared Motif Search

7 motifs close to top scorer:

Motif	Length	Num	Score	Regions

In progress

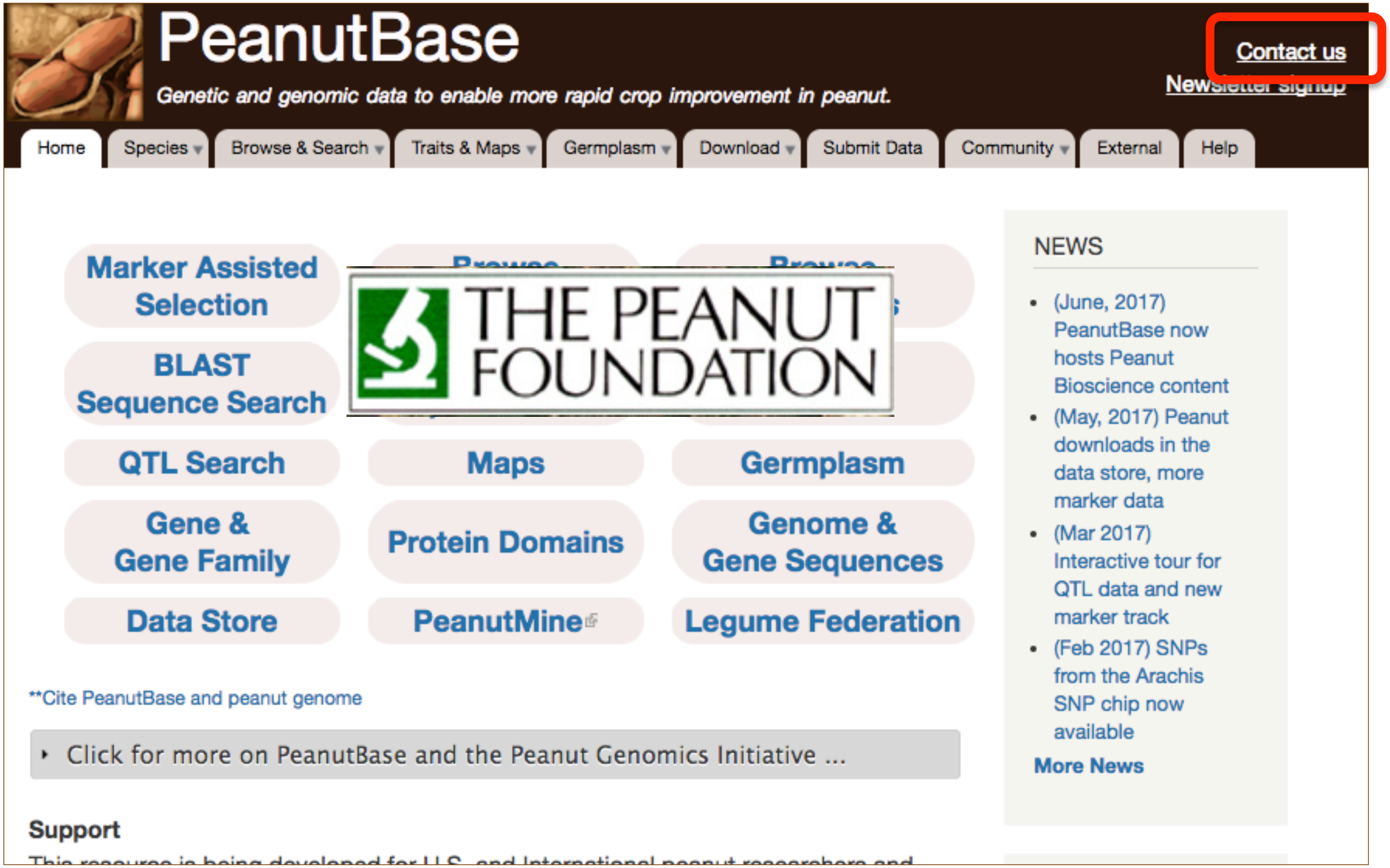
- Continued curation of marker data
- Tetraploid genome assembly
- Continued curation of marker-trait information from literature

Future plans

- Bring in all publicly-available trait data for germplasm and link to stock records
- Align publicly-available genotype-phenotype data and provide results of population structure, LD, and GWAS analyses
- Provide more integrated access to data from other legumes

PeanutBase is a community resource

Let us know what you need



Genetic and genomic data to enable more rapid crop improvement in peanut.

NEWS

- (June, 2017) PeanutBase now hosts Peanut Bioscience content
- (May, 2017) Peanut downloads in the data store, more marker data
- (Mar 2017) Interactive tour for QTL data and new marker track
- (Feb 2017) SNPs from the Arachis SNP chip now available

More News

**Cite PeanutBase and peanut genome

‣ Click for more on PeanutBase and the Peanut Genomics Initiative ...

Support

This resource is being developed for U.S. and International peanut researchers and

Contributors to PeanutBase

Iowa State University/PeanutBase

- **Ethy Cannon** – lead
- Wei Huang – computational biologist
- Paul Otyama – PhD student

Iowa State University/Legume Federation

- Jacqueline Campbell – data curator and outreach

USDA-ARS at Ames, IA

- **Steven Cannon** – lead scientist, ARS legume database project
- Nathan Weeks – IT specialist and computational biologist
- Scott Kalberer – data curator
- Andrew Wilkey – programmer

National Center for Genomic Resources

- **Andrew Farmer** – lead, Legume Information System
- Sudhansu Dash - geneticist and computational biologist
- Alex Rice - programmer