



United States
Department of
Agriculture

National Institute
of Food and
Agriculture



Solanaceae Coordinated
Agricultural Project



Introduction to the Tomato Genome Browser Part I

Heather L. Merk

The Ohio State University, OARDC

SolCAP workshop

Overview

Part 1

- 🍅 What is a genome browser?
- 🍅 Where can I find the tomato genome browser?
- 🍅 What kinds of information can I see and where does it come from?

Part 2

- 🍅 Demonstrate use of browser to identify markers linked to bacterial spot resistance

What is a Genome Browser?

- 🍅 We have a LOT of public tomato sequence data
 - > 2500 molecular markers on high density tomato map (*S. lycopersicum* LA925 x *S. pennellii* LA716)
 - > 300 000 ESTs (GenBank – an NIH public database)
 - 2 draft genome sequences (*S. lycopersicum*, *S. pimpinellifolium*)
- 🍅 We also have a lot of data from other species e. g. Arabidopsis, potato
- 🍅 Data rapidly generated
- 🍅 How can we integrate this data?

What is a Genome Browser?

- 🍅 Wikipedia Definition: “A graphical interface for display of information from a biological database for genomic data. Genome browsers enable researchers to visualize and browse entire genomes with annotated data including gene prediction and structure, proteins, expression, regulation, variation, comparative analysis, etc.” – en.wikipedia.org/wiki/Genome_browser (verified 26 Oct 2010)
- 🍅 There is a LOT of information available!
- 🍅 Easy to get overwhelmed!

The Tomato Genome Browser

- 🍅 Gbrowse – generic genome browser
- 🍅 Can easily be customized
- 🍅 Helpful introductory tutorial:
<http://www.openhelix.com/gbrowse>

GBrowse User Introductory Tutorial

Tutorial and training materials by OpenHelix



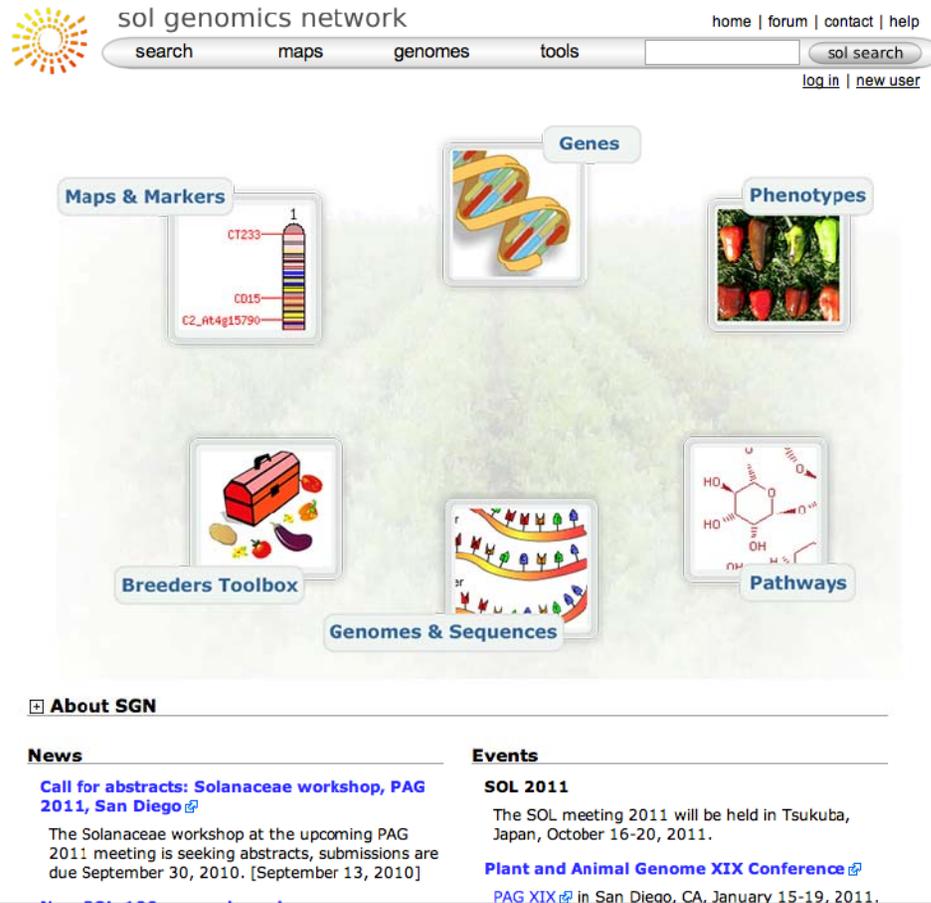
Learn to use [GBrowse](#), a web application that allows you to explore genomic sequences together with annotated data. GBrowse is rapidly becoming a genomic browser of choice among organism databases, because the browser is both universal and yet customizable. Once you learn to use GBrowse at one database, you'll be able to use it to view any genome. Results can be customized to show only the data you want to see. The tool is flexible to allow you to upload and incorporate your own unpublished data into the genomic viewer. You'll have fun as you explore a variety of genomes (from paramecia to personal genomics) with the new perspective and detailed annotations that GBrowse provides.

You'll learn:

- the basic layout and search methods at GBrowse
- how to access detailed annotation data tied to genomic sequences
- how to select and customize annotations using Tracks
- how to upload and incorporate your own data or other external data sources
- take a tour of different GBrowse installations at model organism databases

Where can I find the tomato genome browser?

<http://solgenomics.net>



The screenshot shows the homepage of the sol genomics network. At the top, there is a navigation bar with a sun icon, the text "sol genomics network", and links for "home | forum | contact | help". Below this is a search bar with "search" and "maps" tabs, a "genomes" tab, a "tools" tab, a search input field, and a "sol search" button. There are also links for "log in" and "new user".

The main content area features six interactive icons: "Maps & Markers" (showing a chromosome map with markers CT233, CO15, and C2_ft4g15790), "Genes" (showing a DNA double helix), "Phenotypes" (showing various tomatoes), "Breeder's Toolbox" (showing a toolbox with vegetables), "Genomes & Sequences" (showing a colorful genome map), and "Pathways" (showing a chemical structure of a sugar molecule).

At the bottom, there are sections for "About SGN", "News", and "Events".

About SGN

News

Call for abstracts: Solanaceae workshop, PAG 2011, San Diego

The Solanaceae workshop at the upcoming PAG 2011 meeting is seeking abstracts, submissions are due September 30, 2010. [September 13, 2010]

Events

SOL 2011

The SOL meeting 2011 will be held in Tsukuba, Japan, October 16-20, 2011.

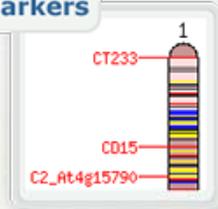
Plant and Animal Genome XIX Conference

PAG XIX in San Diego, CA, January 15-19, 2011.

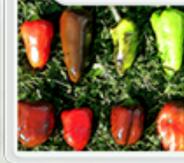


- Browse**
- Tomato genome data
- Projects**
- Solanaceae project (SOL)
- International Tomato Sequencing Project
- Solanum pimpinellifolium* Project (CSHL)
- U.S. tomato sequencing project

Maps & Markers



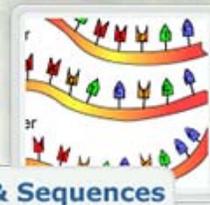
Phenotypes



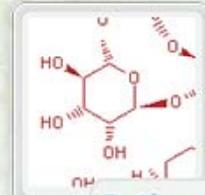
Breeders Toolbox



Genomes & Sequences



Pathways



About SGN

News

Call for abstracts: Solanaceae workshop, PAG 2011, San Diego

The Solanaceae workshop at the upcoming PAG 2011 meeting is seeking abstracts, submissions are due September 30, 2010. [September 13, 2010]

New SOL-100 page released

Events

SOL 2011

The SOL meeting 2011 will be held in Tsukuba, Japan, October 16-20, 2011.

Plant and Animal Genome XIX Conference

PAG XIX in San Diego, CA, January 15-19, 2011. Solanaceae session call for abstracts



Tomato Genome Data

Tomato genome sequence builds

Release	Date	Description	Annotation	Download
1.00	Dec 2009	initial build, based on the Newbler assembler and containing only 454 sequencing data	ITAG1	scaffolds proteins cds
1.03	Jan 2010	like 1.00, but with additional 454 runs and improved contamination screen	Not annotated	scaffolds
cabog1.00	Mar 2010	All 454 data, bac end and fosmid end data, assembled using the CABOG assembler	Not annotated	scaffolds
1.50	Apr 2010	Includes all 454 data, bac ends, fosmid ends, polishing with Solexa and SOLID data	Not annotated	scaffolds
2.00	Jun 2010	Release withdrawn.	Not annotated	-
2.10	Jun 2010	Additional scaffold merging using clone end sequences. Scaffolds placed and oriented using multiple physical maps, first release to include chromosome pseudomolecule sequences.	Not annotated	scaffolds, chromosomes
2.30	Aug 2010	Integration and polishing of tomato BAC sequences	in progress	scaffolds, chromosomes

Official annotation

[browse genome contigs and official annotations](#)

The official annotation for the tomato genome is provided by the [International Tomato Annotation Group \(ITAG\)](#), a multinational consortium, funded in part by the [EU-SOL project](#).



ITAG1 (official release)

last modified Oct 22, 2010

Bulk download

[\[FTP\]](#)

ITAG1 Release: genomic annotations

Browse or

ITAG1 Release: protein annotations

Browse or



At Last... The Tomato Genome Browser

ITAG1 Release: genomic annotations: 969 bp from SL1.00sc06004:457,840..458,808

http://solgenomics.net/gbrowse/bin/gbrowse/ITAG1_genomic/ sol genomics network

sol genomics network home | forum | contact | help

search maps genomes tools sol search log in | new user

File Help

ITAG1 Release: genomic annotations: 969 bp from SL1.00sc06004:457,840..458,808

Browser Select Tracks Upload and Share Tracks Preferences

Search

Landmark or Region: SL1.00sc06004:457840.. Search Annotate Restriction Sites Configure... Go

Examples: SL1.00sc00002_16.1.1, Sulfite oxidase, TG154, C2_At3g02060.

Data Source ITAG1 Release: genomic annotations Scroll/Zoom: Show 969 bp Flip

Overview SL1.00sc06004

Region 360k 370k 380k 390k 400k 410k 420k 430k 440k 450k 460k 470k 480k 490k 500k 510k 520k 530k 540k 550k

Genes SL1.00sc06004_46.1 SL1.00sc06004_51.1 SL1.00sc06004_56.1 SL1.00sc06004_58.1 SL1.00sc06004_61.1 SL1.00sc06004_66.1 SL1.00sc06004_47.1 SL1.00sc06004_53.1 SL1.00sc06004_48.1 SL1.00sc06004_54.1 SL1.00sc06004_59.1 SL1.00sc06004_63.1 SL1.00sc06004_67.1 SL1.00sc06004_49.1 SL1.00sc06004_55.1 SL1.00sc06004_60.1 SL1.00sc06004_62.1 SL1.00sc06004_50.1 SL1.00sc06004_57.1 SL1.00sc06004_64.1 SL1.00sc06004_65.1 SL1.00sc06004_52.1

Details SL1.00sc06004: 969 bp 200 bp

Gene models SL1.00sc06004_58.1.1 Helix-loop-helix DNA-binding KAHFD V1 *-*NG Q2HMX8_MECTR); ccontains In...

CDS - click to browse protein SL1.00sc06004_58.1.1

ESTs and cDNAs - Tomato

There's a lot of information!

What am I looking at?

ITAG1 Release: genomic annotations: 969 bp from SL1.00sc06004:457,840..458,808 ← Overview

Browser [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

Search

Landmark or Region: ← Search

Examples: [SL1.00sc00002_16.1.1](#), [Sulfite oxidase](#), [TG154](#), [C2_At3g02060](#).

Data Source: ← Source

Annotate Restriction Sites

Scroll/Zoom:

I'm scared! Get me outta here!



Scales

ITAG1 Release: genomic annotations: 969 bp from SL1.00sc06004:457,840..458,808

Browser [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

Search

Landmark or Region:

Examples: [SL1.00sc00002_16.1.1](#), [Sulfite oxidase](#), [TG154](#), [C2_A13g02060](#).

Data Source: Flip

Overview

SI 1.00sc06004

Markers

Region

Genes

Details

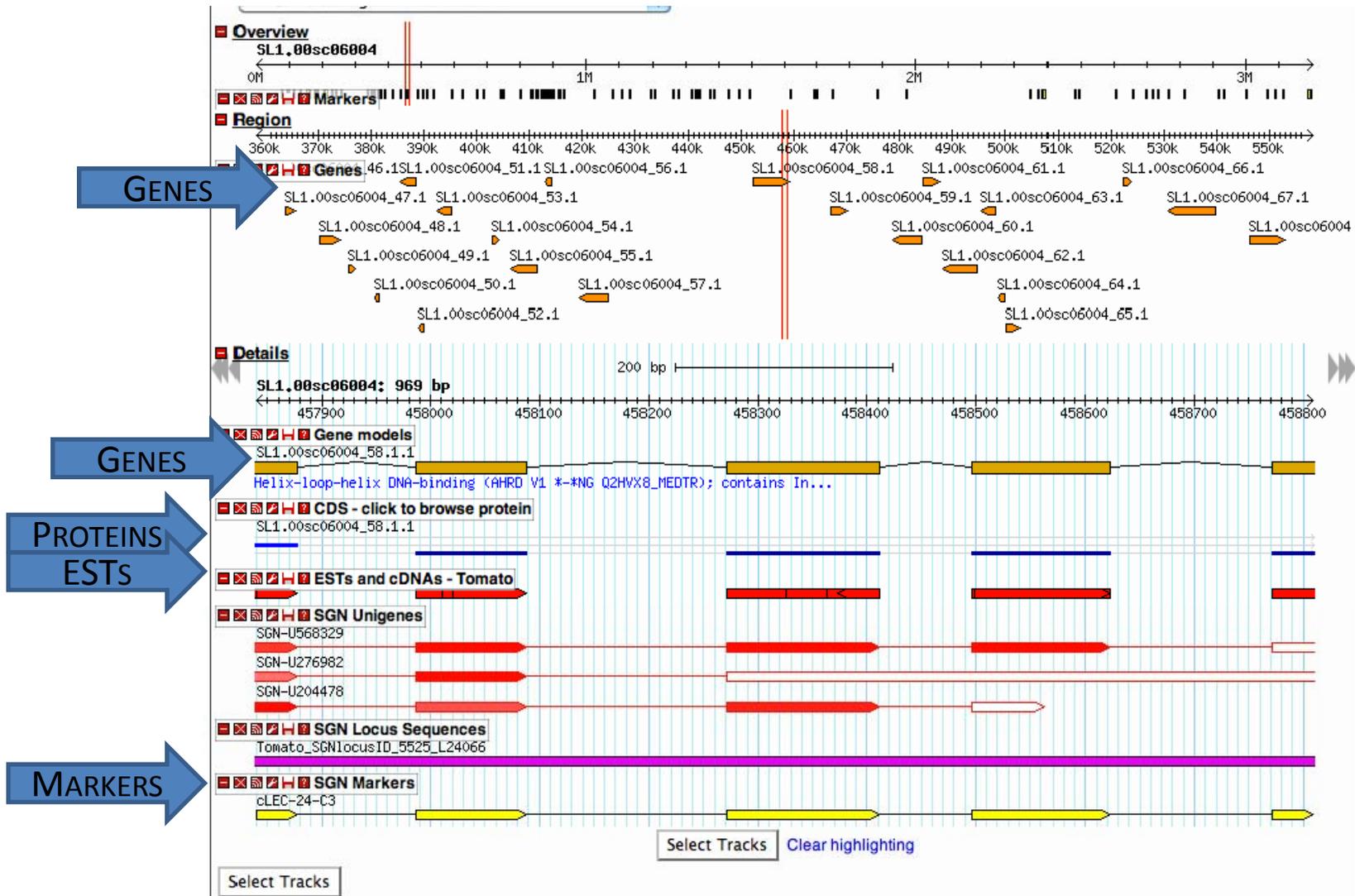
SL1.00sc06004: 969 bp

Gene models

CDS - click to browse protein

ESTs and cDNAs - Tomato

Tracks – Types of Data to View



Wait! – Where Does the Data Come From?

ITAG – International Tomato Annotation Group

- Predicted genes and proteins – using prediction software
- Based on tomato, other Solanaceae, Arabidopsis

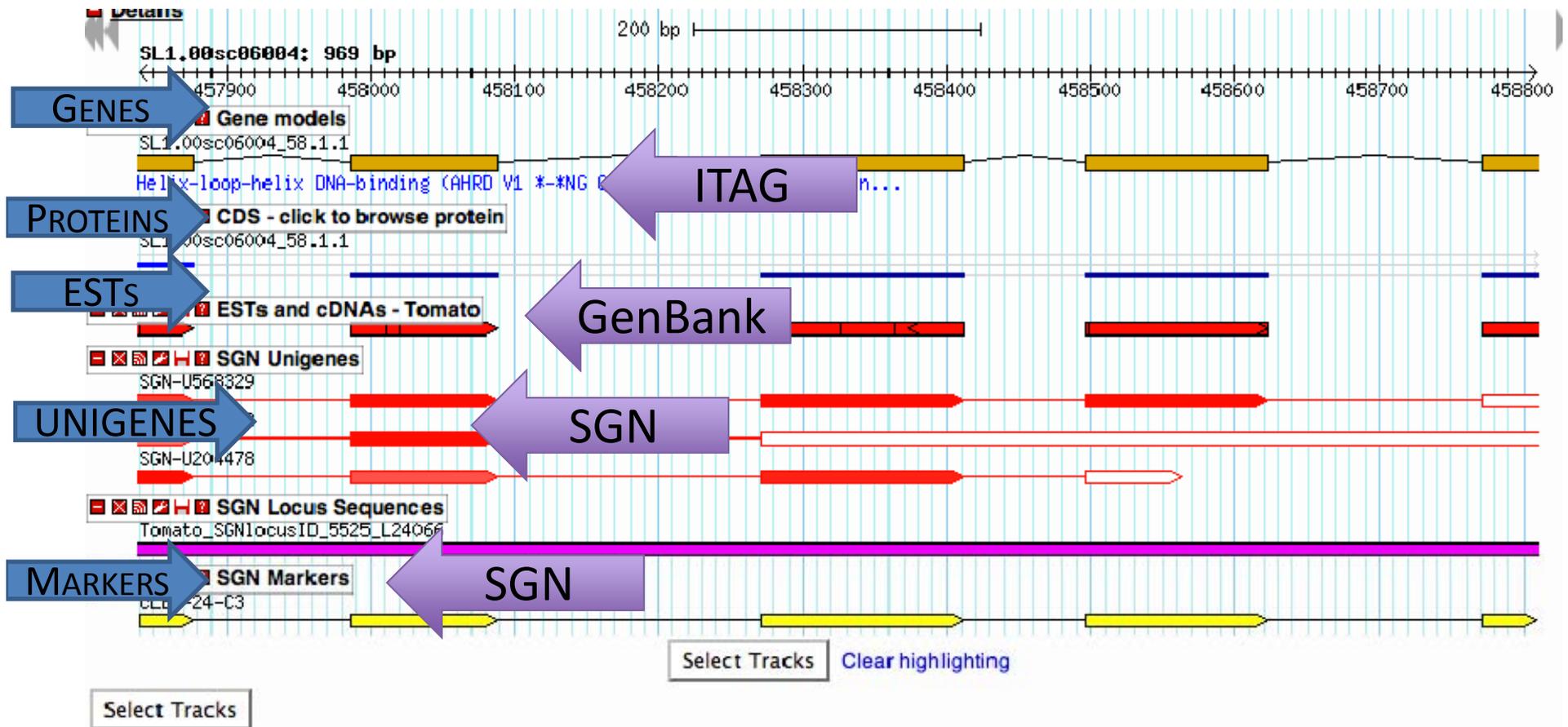
GenBank

- ESTs
- May or may not support predicted genes/proteins

SGN – Sol Genomics Network

- Markers

Tracks – Data Sources



Select Tracks

Overview
SL1.00sc06004

Region
360k 370k 380k 390k 400k 410k 420k 430k 440k 450k 460k 470k 480k 490k 500k 510k 520k 530k 540k 550k

Genes
46.1SL1.00sc06004_51.1 SL1.00sc06004_56.1 SL1.00sc06004_58.1 SL1.00sc06004_61.1 SL1.00sc06004_66.1
SL1.00sc06004_47.1 SL1.00sc06004_53.1 SL1.00sc06004_59.1 SL1.00sc06004_63.1 SL1.00sc06004_67.1
SL1.00sc06004_48.1 SL1.00sc06004_54.1 SL1.00sc06004_60.1 SL1.00sc06004_62.1
SL1.00sc06004_49.1 SL1.00sc06004_55.1 SL1.00sc06004_64.1
SL1.00sc06004_50.1 SL1.00sc06004_57.1 SL1.00sc06004_65.1
SL1.00sc06004_52.1

Details
SL1.00sc06004: 969 bp
200 bp

Gene models
SL1.00sc06004_58.1.1
Helix-loop-helix DNA-binding (AHRD V1 *-*NG Q2HVX8_MEDTR); contains In...

CDS - click to browse protein
SL1.00sc06004_58.1.1

ESTs and cDNAs - Tomato

SGN Unigenes
SGN-U568329
SGN-U276982
SGN-U204478

SGN Locus Sequences
Tomato_SGNlocusID_5525_L24066

SGN Markers
cLEC-24-C3

Select Tracks Clear highlighting

Select Tracks

Select Tracks

ITAG1 Release: genomic annotations: 20 kbp from SL1.00sc06004:490,001..510,000

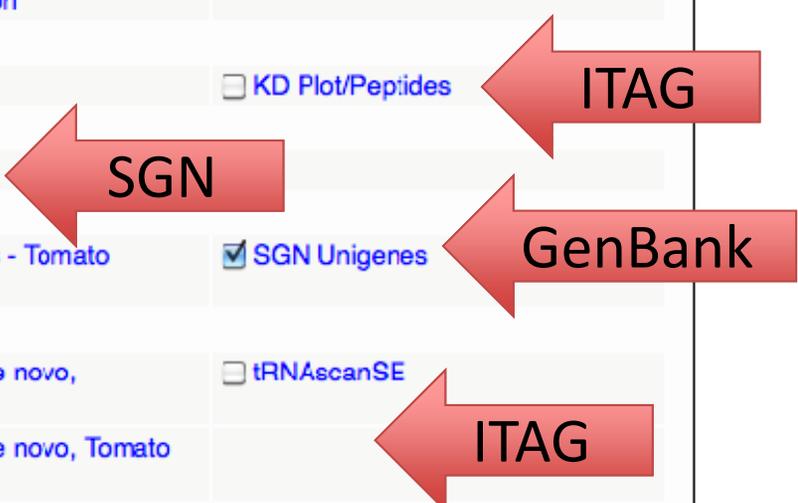
Browser **Select Tracks** Upload and Share Tracks Preferences

<< Back to Browser

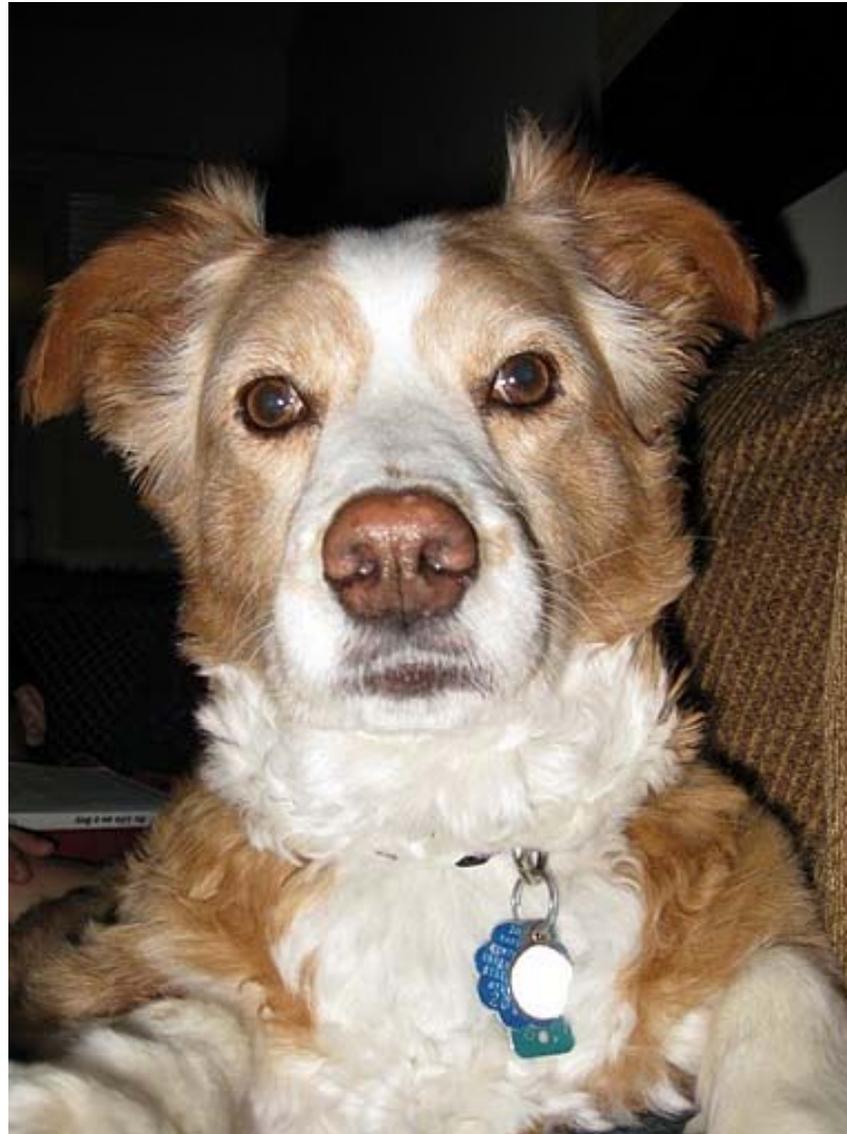
Tracks

- Overview** All on All off
 - Markers tRNAscanSE
- Region** All on All off
 - GC Content Genes
- DNA** All on All off
 - DNA/GC Content 6-frame translation
- Gene models** All on All off
 - CDS - click to browse protein Gene models KD Plot/Peptides
- Genetic Loci** All on All off
 - SGN Locus Sequences SGN Markers
- Genome data and reagents** All on All off
 - ESTs and cDNAs - Other Solanaceae ESTs and cDNAs - Tomato SGN Unigenes
- Prediction Features** All on All off
 - AUGUSTUS (de novo, Tomato trained) GlimmerHMM (de novo, Arabidopsis trained) tRNAscanSE
 - GeneID (de novo, Tomato trained) GlimmerHMM (de novo, Tomato trained)
 - GeneMark (de novo, Arabidopsis trained) Infernal
- Analysis** All on All off
 - Restriction Sites

Back to Browser



Much Better! I can do this!



Conclusion – What We Learned

- 🍅 What a genome browser is
- 🍅 Where to find the tomato genome browser
- 🍅 The kinds of information we can look at using the tomato genome browser
- 🍅 Where information in the browser comes from



United States
Department of
Agriculture

National Institute
of Food and
Agriculture



Solanaceae Coordinated
Agricultural Project



Using the Tomato Genome Browser: An Example Part 2

Heather L. Merk

The Ohio State University, OARDC

SolCAP workshop



Overview

- 🍅 How can the genome browser help ME?
- 🍅 Example – finding markers associated with bacterial spot resistance

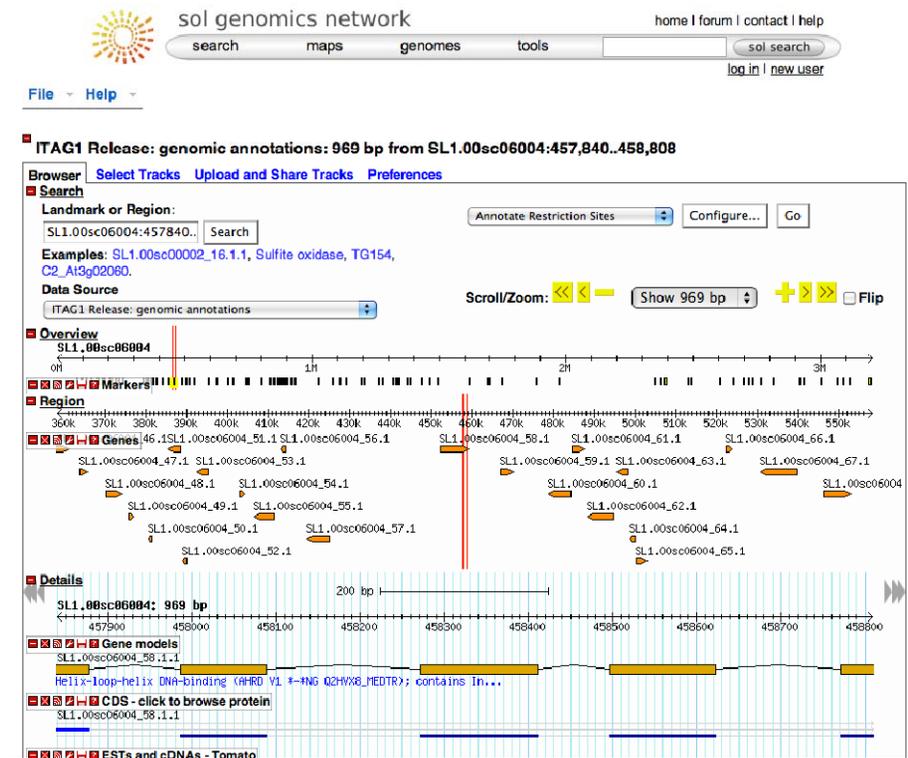


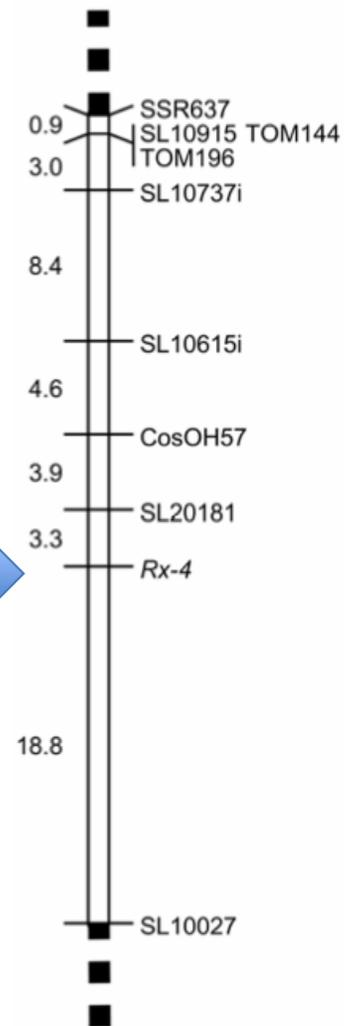
Photo courtesy of D. Francis, Ohio State University

How can the genome browser help ME?

- 🍅 Information from multiple databases/data sources is combined in browser
- 🍅 What does the genome look like around a marker, EST, BAC, other sequence, etc?
- 🍅 Tool to help guide further research – e.g. identify candidate genes, identify potential molecular markers

Tomato Bacterial Spot Resistance

- 🍅 Robbins et al. (2009)
- 🍅 PI 128216 confers resistance to race T3
- 🍅 Rx-4 mapped to chr. 11
- 🍅 SL20181 – SNP in tomato unigene SGN-U567870
- 🍅 We can use genome browser to find more markers – for fine mapping, marker-assisted selection



Step 1: Search

- 🍅 A. Type SGN-U567870 in search box
- 🍅 B. Data source - ITAG1 Release: Genomic Annotation
- 🍅 C. Select tracks to view
- 🍅 D. Search

[File](#) ▾ [Help](#) ▾

ITAG1 Release: genomic annotations

Browser [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

Search

Landmark or Region:

Examples: [SL1.00sc00002_16.1.1](#), [Sulfite oxidase](#), [TG154](#), [C2_At3g02060](#).

Data Source

Step 1c: Select Tracks



sol genomics network

[home](#) | [forum](#) | [contact](#) | [help](#)

[search](#)

[maps](#)

[genomes](#)

[tools](#)

[log in](#) | [new user](#)

[File](#) ▾ [Help](#) ▾

ITAG1 Release: genomic annotations: 6.94 kbp from SL1.00sc06004:224,489..231,428

[Browser](#) [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

<< [Back to Browser](#)

Tracks

Overview All on All off

Markers

tRNAscanSE

Region All on All off

GC Content

Genes

DNA All on All off

DNA/GC Content

6-frame translation

Gene models All on All off

CDS - click to browse protein

Gene models

KD Plot/Peptides

Genetic Loci All on All off

SGN Locus Sequences

SGN Markers

Genome data and reagents All on All off

ESTs and cDNAs - Other
Solanaceae

ESTs and cDNAs - Tomato

SGN Unigenes

Prediction Features All on All off

AUGUSTUS (de novo, Tomato trained)

GlimmerHMM (de novo, Arabidopsis trained)

tRNAscanSE

GeneID (de novo, Tomato trained)

GlimmerHMM (de novo, Tomato trained)

GeneMark (de novo, Arabidopsis trained)

Infernal

Analysis All on All off

Restriction Sites

MARKERS - SGN

ESTs - GENBANK

Step 2: Orientation

🍅 Look at scales for orientation

ITAG1 Release: genomic annotations: 6.94 kbp from SL1.00sc06004:224,489..231,428

Browser [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

Search

Landmark or Region:

Examples: [SL1.00sc00002_16.1.1](#), [Sulfite oxidase](#), [TG154](#), [C2_At3g02060](#).

Data Source: Flip

Overview
SL1.00sc06004 **LARGE**

Region
130k 140k 150k 160k 170k 180k 190k 200k 210k 220k 230k **MEDIUM**

Details
SL1.00sc06004: 6.94 kbp 2 kbp **SMALL**

SGN Unigenes
SGN-U567870 **UNIGENE**
SGN-U282751
SGN-U2

SGN Locus Sequences

SGN Markers
TG286

ESTs and cDNAs - Tomato

Step 3: Zoom Out

ITAG1 Release: genomic annotations: 6.94 kbp from SL1.00sc06004:224,489..231,428

Browser [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

Search

Landmark or Region:

Examples: [SL1.00sc00002_16.1.1](#), [Sulfite oxidase](#), [TG154](#), [C2_At3g02060](#).

Data Source Flip

Overview

SL1.00sc06004

Region

130k 140k 150k 160k 170k 180k 190k 200k 210k 220k 230k 240k 250k 260k 270k 280k 290k 300k 310k 320k

Details

SL1.00sc06004: 6.94 kbp

2 kbp

225k 226k 227k 228k 229k 230k 231k

SGN Unigenes

SGN-U567870

SGN-U282751

SGN-U2

SGN Locus Sequences

SGN Markers

TG286

TG286

Select Tracks

Select Tracks

Step 4: Locate Markers

ITAG1 Release: genomic annotations: 289.9 kbp from SL1.00sc06004:105,056..394,945

Browser [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

Search

Landmark or Region:

Examples: [SL1.00sc000C2_16.1.1](#), [Sulfite oxidase](#), [TG154](#), [C2_At3g0206C](#).

Data Source:

Scroll/Zoom:

Overview

Region

Details

SL1.00sc06004: 289.9 kbp

SGN Unigenes

SGN Locus Sequences

SGN Markers

← UNIGENE →

22 MARKERS →

Select Tracks Clear highlighting

Select Tracks

Step 5: Click on Marker

ITAG1 Release: genomic annotations: 289.9 kbp from SL1.00sc06004:105,056..394,945

Browser [Select Tracks](#) [Upload and Share Tracks](#) [Preferences](#)

Search

Landmark or Region:

Examples: [SL1.00sc00002_16.1.1](#), [Sulfite oxidase](#), [TG154](#), [C2_At3g02060](#).

Data Source:

Scroll/Zoom: Flip

Overview

SL1.00sc06004

Region

Details

SL1.00sc06004: 289.9 kbp

SGN Unigenes

SGN Locus Sequences

SGN Markers

SSR406

Potato_SGNlocusID_2191_SGN-U273550

Tobacco_SGNlocusID_8183

C2_At2g27470

C2_At2g27450

TG121

TG46

TG286

T1297

C2_At2g27310

C2_At2g27290

cLEF-51-69

C2_At2g27730

C2_At3g44380

TG46

TG286

C2_At5g22340

cLET-10-011

T1094

cLEX-14-

SSR406

T0979

C2_At1g30825

C2_At5g22350

C2_At5g22210

C2_At5g40950

C2_At3g54470

Select Tracks

Clear highlighting

Step 7: Locate primers



sol genomics network

[home](#) | [forum](#) | [contact](#) | [help](#)

[search](#)

[maps](#)

[genomes](#)

[tools](#)

[sol search](#)

[log in](#) | [new user](#)

SGN Marker SSR406

SGN-M1079

SSR info

Annealing temperatures: Low: 50 High: 55
Repeat motif: TC Repeat number: 15
Forward primer: ACCTGTGGGATCGACCTAGT
Reverse primer: GCTTGTGGGTGCATAACCTT
Predicted size: 198



Derivations

This marker was derived from [EST read SGN-E8107](#)

Mapped locations

None

Other PCR data

PCR data	Exp. ID 32092	
Forward primer (5'-3') ACCTGTGGGATCGACCTAGT	Accessions and product sizes	Approximate temperature 55°C
Reverse primer (5'-3') GCTTGTGGGTGCATAACCTT		Mg²⁺ concentration Unknown

Genomic locations

Annotation Dataset

ITAG1 Release: genomic annotations

Location(s)

SL1.00sc06741:390261..390359

BUT...

- 🍅 We still need to determine whether these markers are polymorphic in our population
- 🍅 We can use BLAST (Basic Local Alignment Search Tool) to look for sequence differences between our sequence and others
 - May give us an indication of likelihood of polymorphism in our population

Locating sequence

Marker was derived from an EST
Click on EST read to obtain sequence

 sol genomics network [home](#) | [forum](#) | [contact](#) | [help](#)

[search](#) [maps](#) [genomes](#) [tools](#) [sol search](#)

[log in](#) | [new user](#)

SGN Marker SSR406

SGN-M1079

SSR info

Annealing temperatures: Low: 50 High: 55
Repeat motif: TC **Repeat number:** 15
Forward primer: ACCTGTGGGATCGACCTAGT
Reverse primer: GCTTGTGGGTGCATAACCTT
Predicted size: 198

Derivations

This marker was derived from [EST read SGN-E8107](#) 

Mapped locations

None

Other PCR data

PCR data	Exp. ID 32092	
Forward primer (5'-3') ACCTGTGGGATCGACCTAGT	Accessions and product sizes	Approximate temperature 55°C
Reverse primer (5'-3') GCTTGTGGGTGCATAACCTT		Mg²⁺ concentration Unknown

Genomic locations

Annotation Dataset	Location(s)
ITAG1 Release: genomic annotations	SL1.00sc06741:390261..390359



EST details - SGN-E8107

Search information

Request: SGN-E8107
Request From: web user

Match: SGN-E8107
Match Type: EST sequence internal identifier

Clone information

SGN ID: SGN-C9975
Library Name: cLEC

Clone name: cLEC-67-I15
Organism: *Solanum lycopersicum* (formerly *Lycopersicon esculentum*)

Order Clone

Tissue: combined undifferentiated and shooting callus
Development Stage: 7-10 days post germination

Microarray: This clone is not found on any microarray
There is no map position defined on SGN for this EST or others in the same unigene.

Additional sequencing

Clone: SGN-C9975 [cLEC-67-I15] **Trace:** SGN-T33613 **EST:** SGN-E211387 **Direction:** 5' **Facility:** TIGR

[\[Show information hierarchy\]](#)

Sequence

SEQUENCE

Sequence ID: SGN-E8107
Status: Legacy Chimera not assessed

Length: 703 bp (called/trimmed by facility)
Direction: 5'

```
>SGN-E8107 [] (called/trimmed by facility)
CATCACCACCACCGCGCCGCGCTGCGCCACCACGAGTTCAGAGAAGCGCCTGAATTCTACAACCTCAGCTGAATGTCCTTCAATTGAAAATA
CCCATGAATTAATTTGCTCCGATGAAGCTGTTTCATGCAGCAATGACACTCGATTCCGCTTATATCCGTGGATCAATGGCGGCTATTCTCTCAATT
CTTCAACATTTCTTGGCCACAAAACATAATTTTCACTTCGTTGCTTCCGCATCCGCAGACGGTCACATCTACGGCCACCATTTTAGCTTC
ATTCCTTATTTGAAATTCGAAGTTTATCGATTCGACGATTCCTCTGTTTCCGGTTTAAATTTCAACTTCTATTCGTTTCAGCTCTCGATTGCCCTT
TGAACATGCGGAGAAGTTATTTAGCTAATATTTTACCTTTATGTTGTTCCGAAAGTTGTTGTTATTTGGATTCCGATTTGGTTTATTTAGTTGACGATATT
CGGAAGCTATCAGAACTCCACTCGGGGAAGATCATATACTTGCTGCTCCTGAATACTGTAATGCGAATTTCACTTCGTATTTTACTCCGACTTT
CTGGTCCAACCCATCTCTATCTTACATTTGCGGATCGGAAAGCTTGTATTTTAAATACAGGGGTAATGGTAATCGATCTCGATCGATGGGAGA
ATGGTGATTATACCAGGAAGATTGAAGAATGGATGGAA
```

[\[BLAST\]](#)

BLAST

Consider what to BLAST against...



sol genomics network

[home](#) | [forum](#) | [contact](#) | [help](#)

[search](#)

[maps](#)

[genomes](#)

[tools](#)

[sol search](#)

[log in](#) | [new user](#)

NCBI BLAST

Simple

Advanced

Sequence Set

SGN Tomato Combined - WGS, BAC, and unigene sequences

[db details](#)

Program

BLASTN (nucleotide to nucleotide)

Query sequence

single sequence only, use Advanced for multiple

```
>SGN-E8107
CATCACCACCACCGCGGCCGCCCTGGCCACCAGTTTCAGAGAAGCGCCTGAATTCTACAAC
TCAGCTGAATG
TCCTTCAATTGAAAATACCCATGAATTAATTTGCTCCGATGAACCTGTTTCATGCAGCAATGACTCTCG
ATTCCGCTTA
TATCCGTGGATCAATGGCGGCTATTCTCTCAATCTTCAACATTCTTCTTGCCACAAAACATAATTTT
TCACTTCGT
TGCTCCGCATCCGCAGACCGGCTCACATCTACGGCCACCATTTAGCTTCATTTCTTATTGAAATT
```

Expect (e-value) Threshold

1e-10

Clear

Search

Substitution Matrix

BLOSUM62 (default)

Show Graphics

all

Max. hits to show

100

There are a LOT of options I chose Lycopersicon mRNA

sol genomic [home](#) | [forum](#) | [contact](#) | [help](#)

Sequence Set

Program

Expect (e-value) Threshold

Substitution Matrix

Max. hits to show

- Potato BAC Sequences
- S. penellii Genome
 - S. penellii BAC Ends - LpenBAC
 - S. penellii Cosmid Ends - LpenCOS
- Tobacco Genome
 - Assembly of tobacco methylation filtered genome sequences (from TGI)
 - Tobacco genome sequences (methylation filtered from TGI)
- Organelle Genomes
 - Arabidopsis thaliana Mitochondria
 - Nicotiana tabacum Chloroplast
 - Tomato chloroplast genome, cultivar IPA-6
- SGN ESTs
 - All SGN mRNA sequences
 - Genus: Nicotiana SGN mRNA sequences
 - Species: Antirrhinum majus SGN mRNA sequences
 - Species: Capsicum annuum (pepper) SGN mRNA sequences
 - Species: Coffea arabica SGN mRNA sequences
 - Species: Coffea canephora SGN mRNA sequences
 - Species: Hedyotis terminalis SGN mRNA sequences
 - Species: Ipomoea batatas (sweet potato) SGN mRNA sequences
 - Species: Kadua centranthoides SGN mRNA sequences
 - Species: Nicotiana benthamiana SGN mRNA sequences
 - Species: Nicotiana sylvestris (wood tobacco) SGN mRNA sequences
 - Species: Nicotiana tabacum (tobacco) SGN mRNA sequences
 - Species: Petunia hybrida SGN mRNA sequences
 - Species: Solanum cheesmaniae SGN mRNA sequences
 - Species: Solanum habrochaites SGN mRNA sequences
 - Species: Solanum lycopersicoides SGN mRNA sequences
 - Species: Solanum lycopersicum SGN mRNA sequences
 - Species: Solanum melongena (eggplant) SGN mRNA sequences
 - Species: Solanum pennellii SGN mRNA sequences
 - Species: Solanum peruvianum SGN mRNA sequences
 - Species: Solanum pimpinellifolium SGN mRNA sequences
 - Species: Solanum tuberosum (potato) SGN mRNA sequences
 - Subgenus: Lycopersicon (all tomato species) SGN mRNA sequences**
 - SGN Unigenes
 - All SGN Unigene sequences
 - Antirrhinum majus Unigenes
 - Capsicum annuum (pepper) Unigenes
 - Coffea arabica (arabica coffee) Unigenes
 - Coffea canephora (robusta coffee) Unigenes
 - Coffee species Unigenes
 - Hedyotis species Unigenes
 - Ipomoea batatas (sweet potato) Unigenes
 - Lycopersicon Combined (Tomato) Unigenes
 - Nicotiana benthamiana Unigenes

```
>SGN-E8107
CATCACCACCACCGCG
TCAGCTGAATG
TCCTTCAATTGAAAATA
ATTCCGCTTA
TATCCGTGGATCAATGC
TCACTTCGT
TGCTTCCGCATCCGCA
```

BLAST Results

BLAST Search Report - Sol Genomics Network

http://solgenomics.net/tools/blast/view_result.pl?output_graphs=bioperl_histogram&filterq: Google

log in | new user

BLAST Results

Note: Please **do not bookmark** this page. BLAST results are automatically deleted after 7 days.

Graphics

Alignment Summary

All hits shown.

Mouse-overs require JavaScript

Fwd: >> Rev: << Neg P Exponent: -> 0-10 10-50 50-200 200-> 4/Nov/2010

Conservedness Histogram

Conservedness Histogram

The histogram shows a count of hits for each base in the query sequence, but counts only the domains BLAST finds, meaning this is really more a function of region than of individual base. Within the graph, green shows exact base matches within conserved regions; blue shows non-matching bases within conserved regions. Gaps introduced into the query by BLAST are ignored; gaps introduced into target sequences are not.

Query region [388 -> 676]: conservedness in (3, 9)

Query 351 703 Conservedness

BLAST Report

[View / download raw report] (9.6K)

CONSERVEDNESS HISTOGRAM

Our EST is highly conserved

BLAST Results

```
http://solgenomics.net/static/documents/tempfiles/blast/Xejwrt
http://solgenomics.net/static/documents/tempfiles/blast/Xejwrt&flv
Apple Yahoo! Google Maps YouTube Wikipedia News (376) Popular v

Sequences producing significant alignments:

Score      E
(bits)     Value

SGN-E211387 #Solanum lycopersicum (formerly Lycopersicon esculen... 1318 0.0
SGN-E1255387 #Solanum lycopersicum (formerly Lycopersicon esculen... 821 0.0
SGN-E205127 #Solanum lycopersicum (formerly Lycopersicon esculen... 593 e-168

>SGN-E211387 #Solanum lycopersicum (formerly Lycopersicon esculentum)
   [cLEC-67-I15](5')
   Length = 730

Score = 1318 bits (665), Expect = 0.0
Identities = 674/677 (99%)
Strand = Plus / Plus

Query: 1   catcaccaccaccgcccggccgcccgtgcccaccgcaacagttcagagaagcgcoctgaatt 60
          |||
Sbjct: 54  catcaccaccaccgcccggccgcccgtgcccaccgcaacagttcagagaagcgcoctgaatt 113

Query: 61  ctacaactcagctgaatgtccttcaattgaaaatacccatgaattaatttgcctgatga 120
          |||
Sbjct: 114 ctacaactcagctgaatgtccttcaattgaaaatacccatgaattaatttgcctgatga 173

Query: 121  agctgttcatgcagcaatgacactogattccgcttatatccgtggatcaatggcgctat 180
          |||
Sbjct: 174  agctgttcatgcagcaatgacactogattccgcttatatccgtggatcaatggcgctat 233

Query: 181  tctctcaattcttcaacattcttcttgcccacaaaacataaattttcacttctgttcttc 240
          |||
Sbjct: 234  tctctcaattcttcaacattcttcttgcccacaaaacataaattttcacttctgttcttc 293

Query: 241  cgcataccgcagacgcgtcacatctacggccaccattttagttcatttcottatattgaa 300
          |||
Sbjct: 294  cgcataccgcagacgcgtcacatctacggccaccattttagttcatttcottatattgaa 353

Query: 301  attcgaagtttatcgatttcgacgattcctctgtttccggtttaatttcaacttctattcg 360
          |||
Sbjct: 354  attcgaagtttatcgatttcgacgattcctctgtttccggtttaatttcaacttctattcg 413
```

 Look for mismatches between query and subject sequences

 No mismatches here

 Doesn't provide evidence for marker polymorphism

Conclusion – What We Learned

- 🍅 How to use the tomato genome browser to identify molecular markers that may be useful in fine-mapping and marker-assisted selection
- 🍅 Important to remember that the browser is a work in progress – e.g. don't have all marker data currently

sol genomics network home | forum | contact | help

search maps genomes tools sol search

log in | new user

File - Help -

ITAG1 Release: genomic annotations: 969 bp from SL1.00sc06004:457,840..458,808

Browser Select Tracks Upload and Share Tracks Preferences

Search

Landmark or Region: SL1.00sc06004:457840.. Search Annotate Restriction Sites Configure... Go

Examples: SL1.00sc00002_16.1.1, Sulfite oxidase, TG154, C2_A13g02060.

Data Source ITAG1 Release: genomic annotations Scroll/Zoom: Show 969 bp Flip

Overview SL1.00sc06004

Region

Genes

Details

SL1.00sc06004: 969 bp

Gene models

Helix-loop-helix DNA-binding (APRD V1 *-NG Q2HWI8_MEDTR); contains In...

CDS - click to browse protein

ESTs and cDNAs - Tomato



Photo courtesy of D. Francis, Ohio State University