Integrating Genome and Transcriptome Resources into TreeGenes

Jill Wegrzyn

David Neale

Doreen Main

Keithanne Mockaitis

University of California at Davis

Washington State University

Indiana University/Center for Genomics and Bioinformatics

USDA

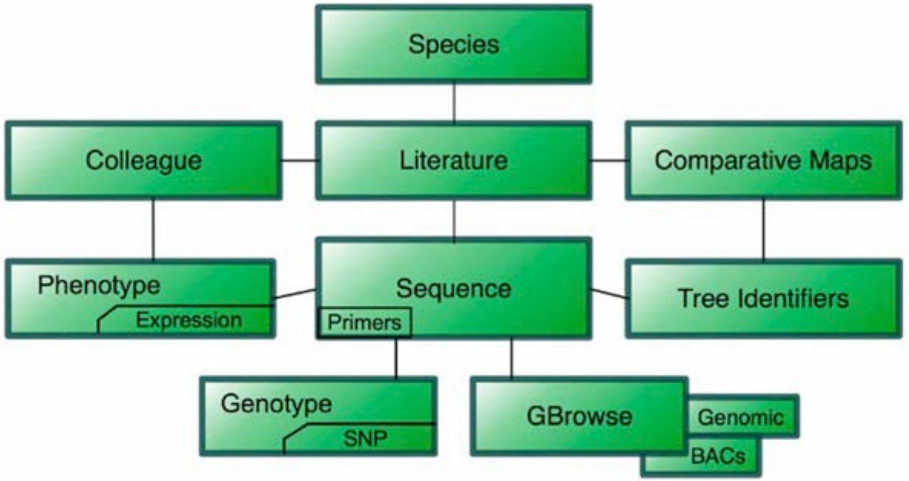
PineRefSeq

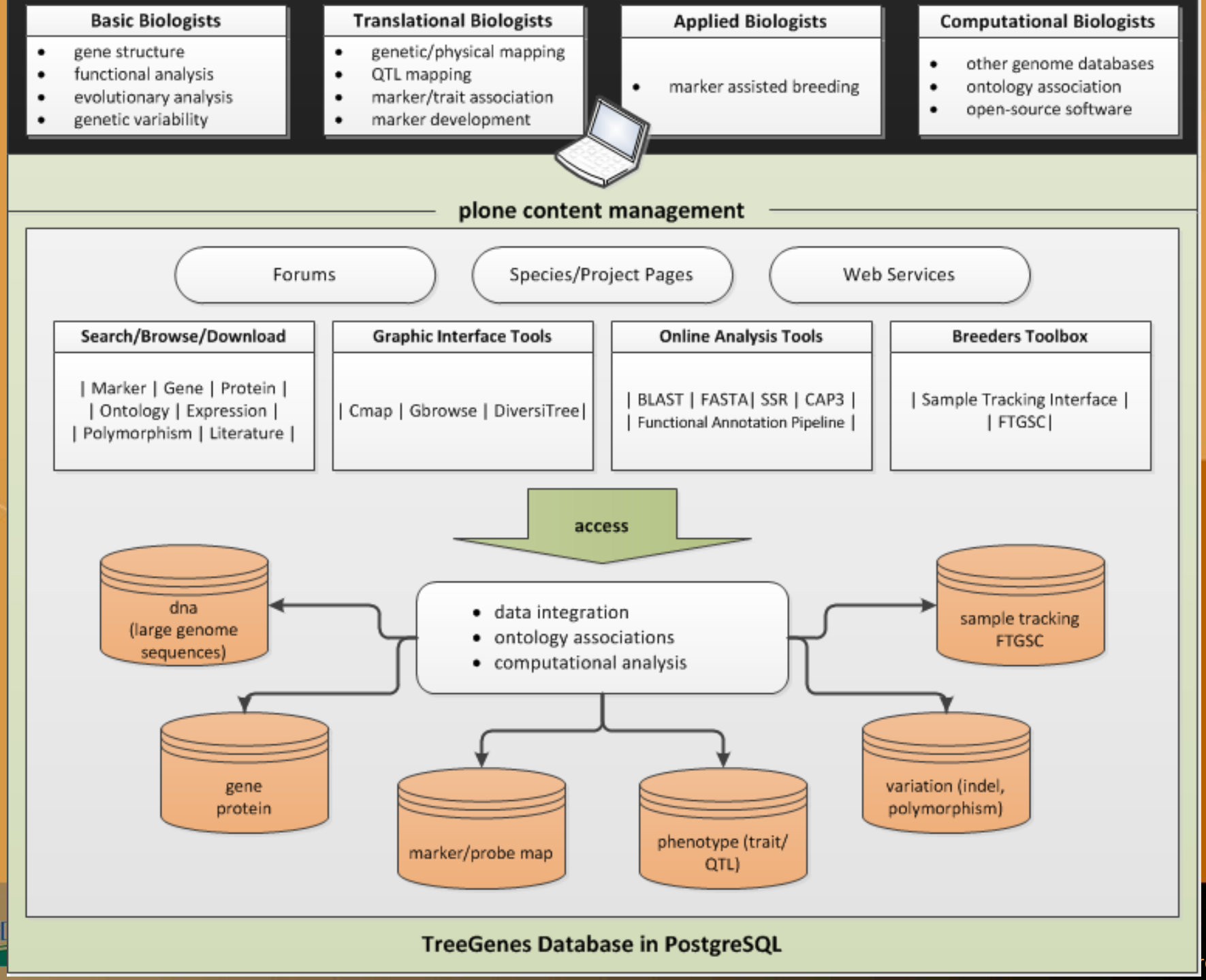
http://pinegenome.org/pinerefseq/

TreeGenes Database

Encompasses Dendrome Resources, DendromePlone, TreeGenes Database &DiversiTree

- Ten modules to store and interrelate data for query and analysis in PostgreSQL
  - Direct resource for nearly 2,500 forest geneticists representing 800 organizations worldwide. Over 6,000 unique visitors in December 2011.
    - Forest Geneticists Colleague module
    - Literature module
    - EST annotation pipeline and module
    - Comparative map module
    - Species module
    - Sequencing module
    - Primers module
    - Genotype/EST module
    - Phenotype/Expression module
    - Sample tracking module




Basic Biologists
gene structure
functional analysis
evolutionary analysis
genetic variability
Translational Biologists
genetic/physical mapping
QTL mapping
marker/trait association
marker development
Applied Biologists
marker assisted breeding
Computational Biologists
other genome databases
ontology association
open-source software
plone content management
Forums
Species/Project Pages
Web Services
Search/Browse/Download
Marker	Gene	Protein
Ontology	Expression
Polymorphism	Literature
Graphic Interface Tools
Cmap	Gbrowse	DiversiTree
Online Analysis Tools
BLAST	FASTA| SSR	CAP3
Functional Annotation Pipeline
Breeders Toolbox
Sample Tracking Interface
FTGSC
access
dna
(large genome
sequences)
data integration
ontology associations
computational analysis
sample tracking
FTGSC
gene
protein
marker/probe map
phenotype (trait/
QTL)
variation (indel,
polymorphism)
TreeGenes Database in PostgreSQL

refseq/

Genomic Resources

678 Species Representing 77 Genus

Pinus lambertiana | Sugar Pine

Information | Resources | Map | Publications

Baker, S. M. | White, Eleanor
A CHALCONE SYNTHASE/STILBENE SYNTHASE DNA PROBE FOR CONIFERS
Theoretical and Applied Genetics 1996 92:827-831.

Contreras-medina, Raul | Vega, Isolda Luna | Morrone, Juan J.
APPLICATION OF PARSIMONY ANALYSIS OF ENDEMICITY TO MEXICAN GYM BIOGEOGRAPHICAL PROVINCES AND TRACK ANALYSIS
Biological Journal of the Linnean Society 2007 92:405-417.

Ekramoddoullah, Abul Kalam Mohammed | Hunt, Richard S.
CHANGES IN PROTEIN PROFILE OF SUSCEPTIBLE AND RESISTANT SUGAR PI BLISTER RUST FUNGUS CRONARTIUM RIBICOLA
Canadian Journal of Plant Pathology 1993 15:259-264.

Yu, Xueshu | Ekramoddoullah, Abul Kalam Mohammed | Misra, Santosh
CHARACTERIZATION OF PIN M III CDNA IN WESTERN WHITE PINE
Tree Physiology 2000 20:663-671.

Kinloch, Bohun B., Jr.
DISTRIBUTION AND FREQUENCY OF A GENE FOR RESISTANCE TO WHITE PI SUGAR PINE
Canadian Journal of Botany 1992 70:1319-1323.

Kinloch, Bohun B., Jr. | Dupper, Gayle E.
EVIDENCE OF CYTOPLASMIC INHERITANCE OF VIRULENCE IN CRONARTIUM PINE
Phytopathology 1999 89:192-196.

Lipow, Sara R. | Bradley St. Clair, J. | Johnson, G. R.
EX SITU GENE CONSERVATION FOR CONIFERS IN THE PACIFIC NORTHWEST 2002.

Hall, Sarah E. | Dvorak, William S. | Johnston, J. Spencer | Price, H. James | Williams C
FLOW CYTOMETRIC ANALYSIS OF DNA CONTENT FOR TROPICAL AND TEMPI
Annals of Botany (London) 2000 86:1081-1086.

NAGEIA									74		
NEOCALLITROPSIS									4		
NOTHOTSUGA		1							19		
PAPUACEDRUS									6		
PARASITAXUS									1		
PHYLLOCLADUS									44		
PICEA	514139	1021			4	753	8777	19	30467		877
PILGERODENDRON		1							9		
PINUS	453932	453	75513	283479	32	20583	103919	37	22087		19003
SPECIES	EST	cDNA	TSA	TG Transcr Assembly	cMAP	ReSeqAmp	SNP	ExpStudies	Protein	Gbrowse	StockCenter
Pinus albicaulis						657			87		25
Pinus aristata						623			89		10
Pinus armandii									140		
Pinus attenuata									70		
Pinus ayacahuite									86		
Pinus balfouriana		15				628			207		27
Pinus banksiana	36379	1							164		

dendrome.ucdavis.edu/treegenes/species/species_detail.php?id=49

dendrome.ucdavis.edu/treegenes/species/species_detail.php?id=49

Pinus lambertiana | Sugar Pine

Information | Resources | Map | Publications

Genus
Pinus

Chromosomes
12

TG Species Code
Pila

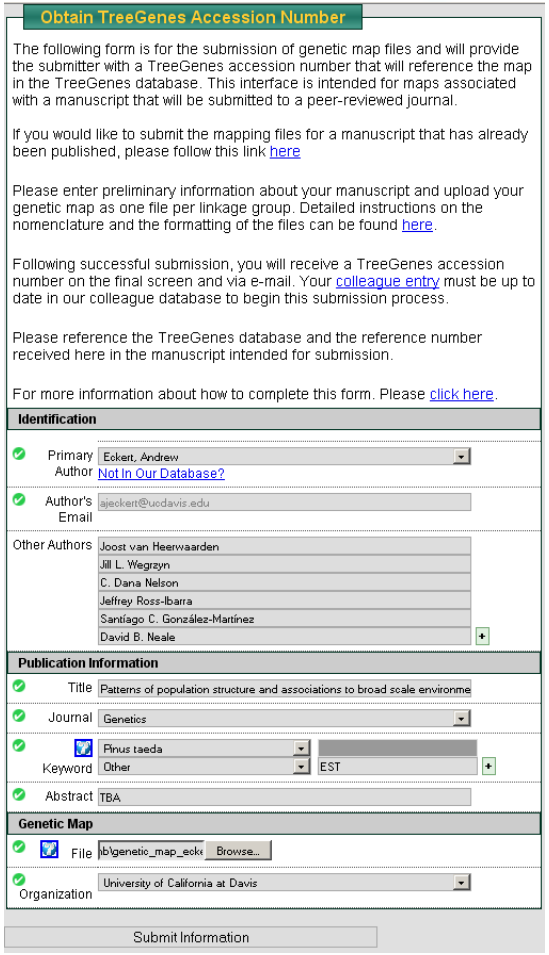
More Info
NCBI Taxonomy

Softwar

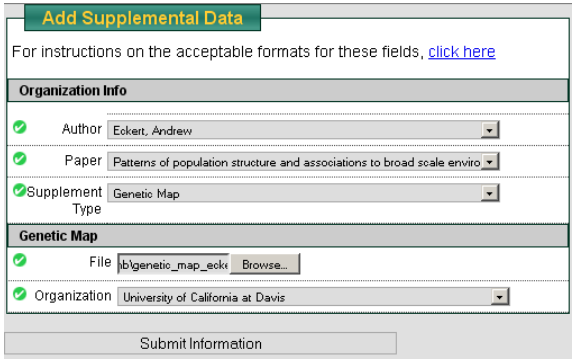
USDA

PineRefSeq
http://pinegenome.org/pinerefseq/

CMAP: Obtaining TreeGenes (TG) Accession Number

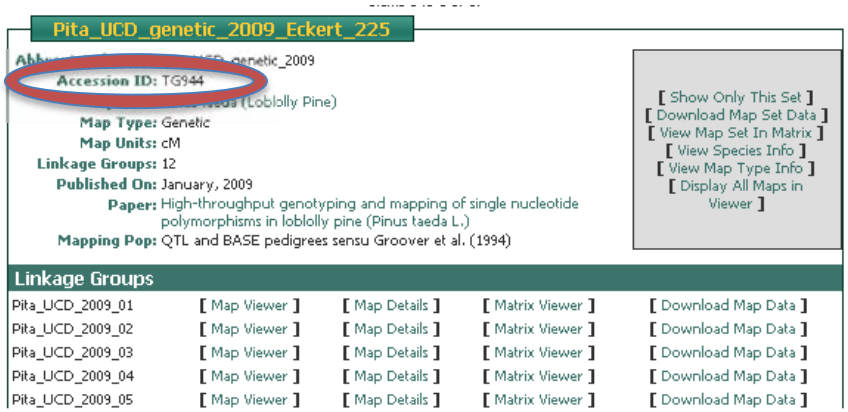


Add literature data and (first) map file



(optional) Add additional map files

Obtain TG Accession number!



Feature Search

A **feature** in CMap is any element that can be placed on a map, either as a point or an interval.

Feature Search

Search On: Name

Search For:

Separate multiple names with commas or whitespace. Use "" or "%" for wildcards. To find features with spaces in the name, surround the name in double quotes, e.g., "abc 123."

Restrict species:
–All Species–
Acacia mangium
Cryptomeria japonica
Eucalyptus globulus
Eucalyptus grandis

Restrict feature types:
–All Feature Types–
AFLP
Centromere
Clone
Contig

Submit Reset

Map Search

Published: Both

Ref. Species: Black Cottonwood (Populus trichocarpa) Change Species

Ref. Set: Genetic : Potr_Mansfield_genetic_2000_Chen_837

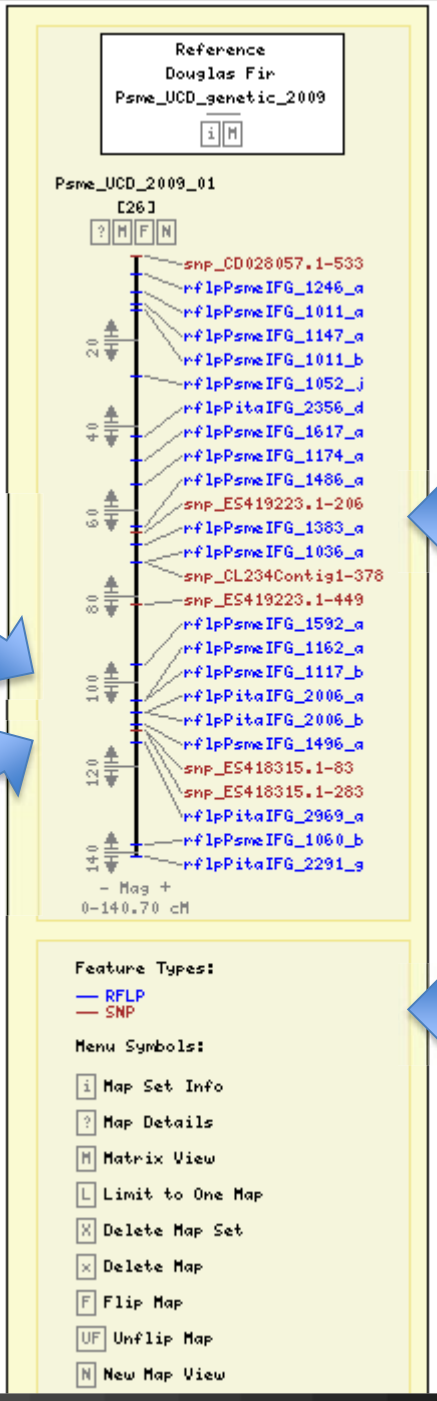
Name *(optional)*:

Min # of Related Maps: 0

Submit

Items 1 to 9 of 9.

Map Name	Related Maps	Related Maps per unit	Correspondences	Correspondences per unit	Start	Stop	SSR		GENE	
							total	per unit	total	p u
Potr_OSU_2000_01 (spider)	0	0/unit	0	0/unit	0.00	59.00	0	0	1	16.
Potr_OSU_2000_02 (spider)	0	0/unit	0	0/unit	0.00	83.00	3	36.14/K	1	12.
Potr_OSU_2000_03 (spider)	0	0/unit	0	0/unit	0.00	103.00	2	19.42/K	1	9.7
Potr_OSU_2000_04 (spider)	0	0/unit	0	0/unit	0.00	44.00	1	22.73/K	0	0
Potr_OSU_2000_05 (spider)	0	0/unit	0	0/unit	0.00	17.00	0	0	0	0
Potr_OSU_2000_06 (spider)	0	0/unit	0	0/unit	0.00	175.00	1	5.71/K	0	0
Potr_OSU_2000_07 (spider)	0	0/unit	0	0/unit	0.00	101.00	0	0	1	9.9
Potr_OSU_2000_08 (spider)	0	0/unit	0	0/unit	0.00	170.00	0	0	1	5.8
Potr_OSU_2000_09 (spider)	0	0/unit	0	0/unit	0.00	68.00	0	0	0	0



Individual features and their locations on map

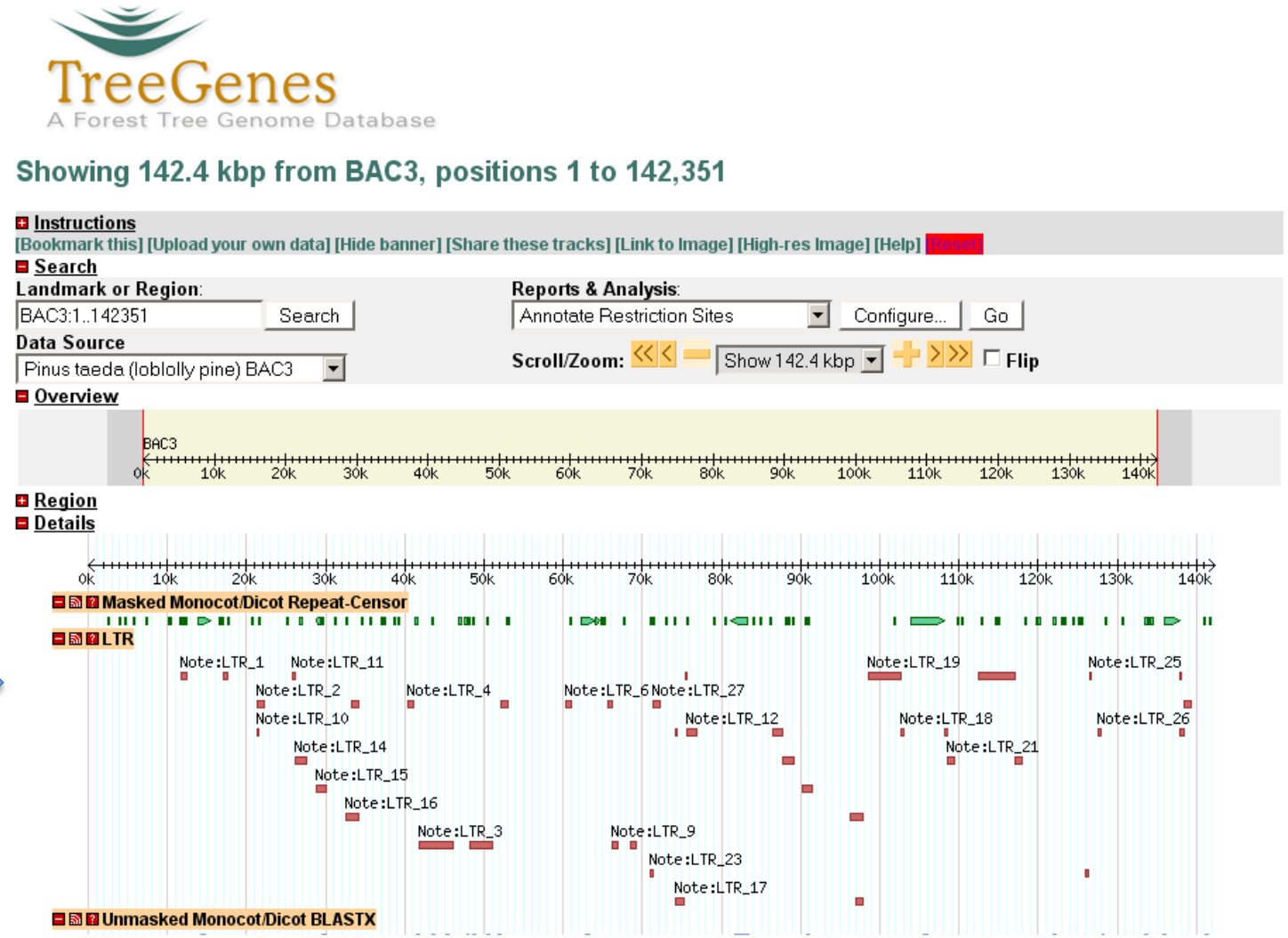
List of features on map

PineRefSeq
http://pinegenome.org/pinerefseq/

GMOD Genome Browser

Search and
Select data source

Tracks can be
reordered or
hidden as necessary



Transcriptome Assembly Summary

Loblolly pine 454 (JGI and CGB)

Douglas-fir RNASeq (FS) and 454 (JGI)

Sugar pine RNASeq (FS) and 454 (JGI)

Library	N, quality filtered	length mean	Nucleotides	Transcripts Assembled	Unique Transcripts
Pila Needles & Candles 454 (Nblr)	1,096,017	353	387,174,063	30,910	22,635
Pila Needle RNASeq (Trinity)				38,961	34,225
Psme Needles and Candles 454 (Nblr)	1,216,156	345	419,643,998	25,041	51,548
Psme Needle RNASeq (Trinity)				39,209	
Pita Shoot 454 (Nblr)	874,971	235	205,284,775	62,342	48,842
Pita Callus 454 (Nblr)	882,199	391	344,842,307	37,322	
Pita Stem 454 (Nblr)	934,760	332	310,498,816	43,234	
Co-assembly		501			42,653

Transcriptome Resources in TreeGenes

Pseudotsuga menziesii | Douglas Fir

Information | Resources | Map | Publications

» Genome Resources

Sanger Resequenced Amplicons (720)	DiversiTree
SNP (5,988)	DiversiTree
Genetic Maps (5)	Search

External Sequence Data

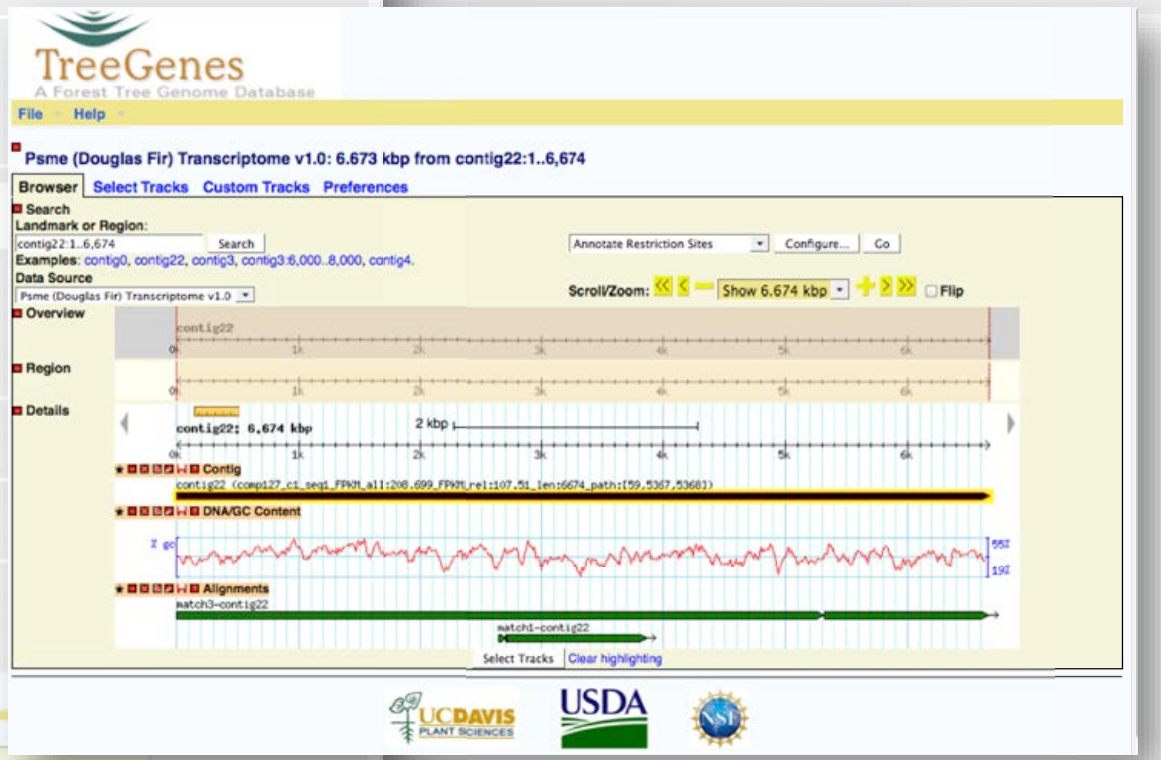
NCBI	Nucleotide (2,821) SRA (1) Popset (164) UniSTS (13)

» Transcriptome Resources

EST (18,142)	Search Download
Full Length cDNA (209)	Search Download
Expression (1)	Search
Short Read Trancriptome Assembly v1.0 [FS] (39,209)	More Info Download [Trinity]
454 Trancriptome Assembly [JGI] (25,041)	Download
TreeGenes UniGenes [EST,cDNA,RNASeq,454] (51,545)	Download Annotations Annotations by Species Click the picture below for GBrowse interface
Project Links	Pine Reference Sequence Project Douglas-fir Transcriptome Project

» Proteome Resources

Protein (3,902)	Search Download



PineRefSeq
http://pinegenome.org/pinerefseq/

Forest Tree Genetic Stock Center

FTGSC

- FTGSC
- Sample Search
- Order Form
- Population Sets
- Phenotypes
- Ecological Site Descriptors
- Sample Submission
- Protocols

Forest Tree Genetic Stock Center

Overview

The Forest Tree Genomics Stock Center (FTGSC) at the University of California at Davis provides collection, preservation, and distribution of forest tree DNA to the forest research community. Primary support for this service is provided by the USDA National Institute for Food and Agriculture (NIFA, formerly CSREES) and the USDA Forest Service.

The Forest Tree Genomics Stock Center performs four primary services:

1. Conifer Sample Collection & Storage
2. DNA Extraction
3. Web-based Sample Search & Information Retrieval
4. Web-based Sample Ordering

DNA Extraction and Storage

The Neale Lab at UCDavis is responsible for coordinating the collection of samples from a variety of tissues within important association and diversity study populations from several different forest tree species. Collection of critical populations, barcoding of the samples, and DNA extractions are currently underway.

Information Retrieval & Sample Ordering

Sample information retrieval and sample ordering services are available through the Sample Search pages. Information about the population sets provided by the stock center are available through the Populations Sets page.

Sample Submission

To submit samples to the FTGSC, please fill out the form found on the Sample Submissions page.

FTGSC

- FTGSC
- Sample Search
- Order Form
- Population Sets
- Phenotypes
- Ecological Site Descriptors
- Sample Submission
- Protocols

Sample Submission Form

Please fill out and submit the form below. After the form has been submitted, you will receive shipping instructions along with a Shipping Receipt which must be included with your shipment. Depending on how you fill out the form below, you may or may not receive barcodes for your samples.

Select your institution or company name:

If not listed, register your lab here.

Contact information:

Principal Investigator's Name:
Principal Investigator's E-mail:
Project Coordinator's Name:
Project Coordinator's E-mail:
Project Coordinator's Phone:
Shipping Address:
(for parcel packages)

What is the name of this population set?
Please enter a name for the set of samples being shipped to the FTGSC. For example, a population set may be name "Loblolly QTL Study - Set 1" or "Smith Lab - 2011 Lignin Study Samples". The name entered here will be used to describe your population set from within the FTGSC tracking system.

Please describe the population set and the research it has been used for.
Please enter a description for your population set, and provide information about the research studies that have involved these samples. Please feel free to mention prior publications in this description. This description will be used to help other researchers determine whether or not this sample set may be useful for their studies.

Select the tree species?

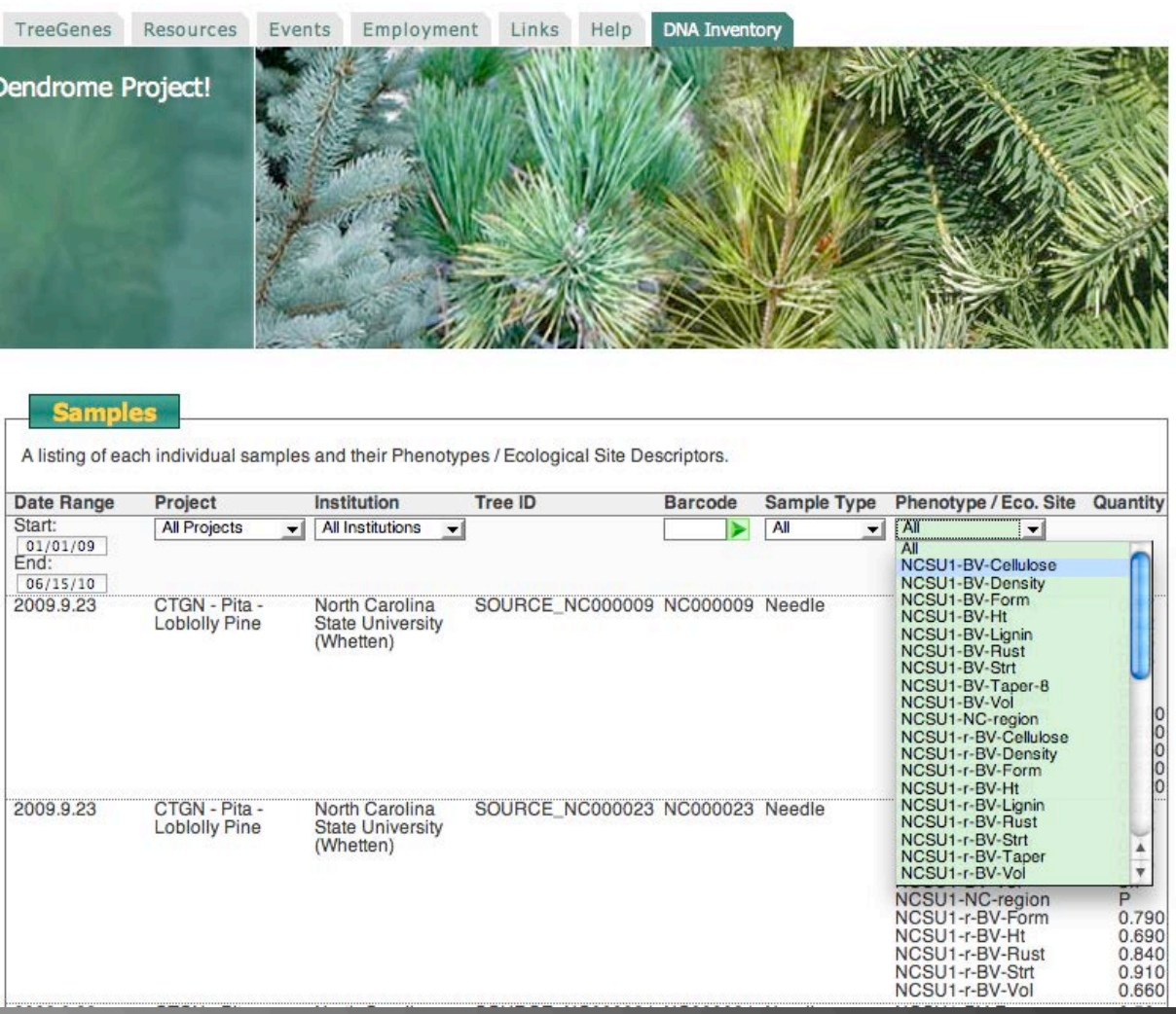
Select the preservation type?

Select your sample type?

How many individual trees will you be sending samples for?
Please provide the total number of individual (genetically distinct) trees, not the total number of clones.
NOTE: All genetically identical samples MUST share the same FTGSC Barcode.

Have any of these samples been submitted to the FTGSC in the past?

TreeGenes Sample Tracking System



- Accurately track samples through collection, DNA extraction, and genotyping
- Provide a standard and efficient method to collect and store phenotypic data
- Provide a public interface to readily query raw genotype, phenotype, and association results (DiversiTree)
- Provide interfaces and database backend to support a DNA distribution center (UCD)

Data Submission

Sample Submission

Phenotypes Defined

Phenotype Name	Phenotype Calculation	Min Value	Max Value
TAMU3-POLY-RUST	Numeric average of values expressed as standard deviations from plantation mean. + equals more rust than average, - less rust	-3.000	3.000

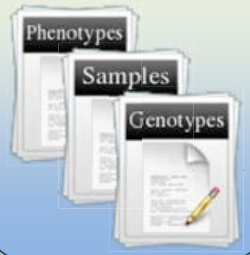
Phenotype Data

Phenotype / Eco. Site	Quantity
TAMU3-AVE-FORM	1.39
TAMU3-AVE-STRT	1.71
TAMU3-BV-AVE	118.00
TAMU3-BV-Ht	107.67
TAMU3-TOT-BV-Ht	3
TAMU3-TOT-FORM	1
TAMU3-TOT-OBS	3
TAMU3-TOT-STRT	3
TAMU3-AVE-FORM	-0.18
TAMU3-AVE-STRT	0.81
TAMU3-BV-AVE	118.50
TAMU3-BV-Ht	103.79
TAMU3-TOT-BV-Ht	4
TAMU3-TOT-FORM	2
TAMU3-TOT-OBS	4
TAMU3-TOT-STRT	4

Accession Number

Forest Tree Genetic Stock Center

Raw Data

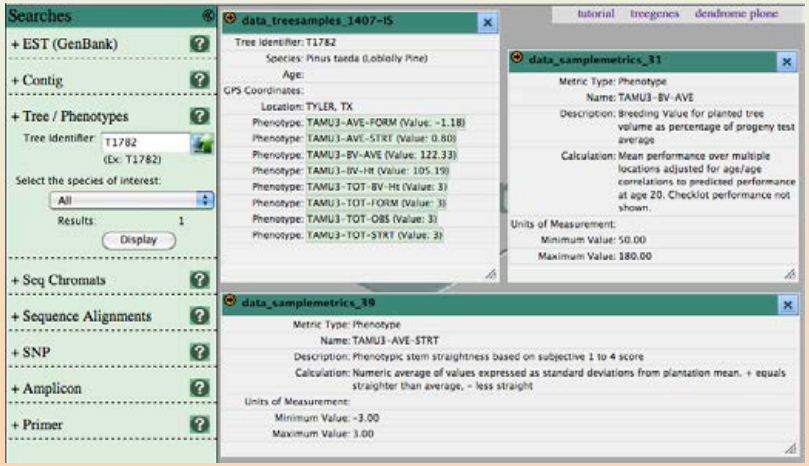


Genotypes

B	C	D	E
11-1109	11-111	11-1110	11-1140
AC	AA	AC	AA
AG	AA	AG	AG
GG	GG	GG	GG
AG	AG	AA	AG
GG	AG	GG	AG
GG	GG	GG	GG

Accession Number

DiversiTree



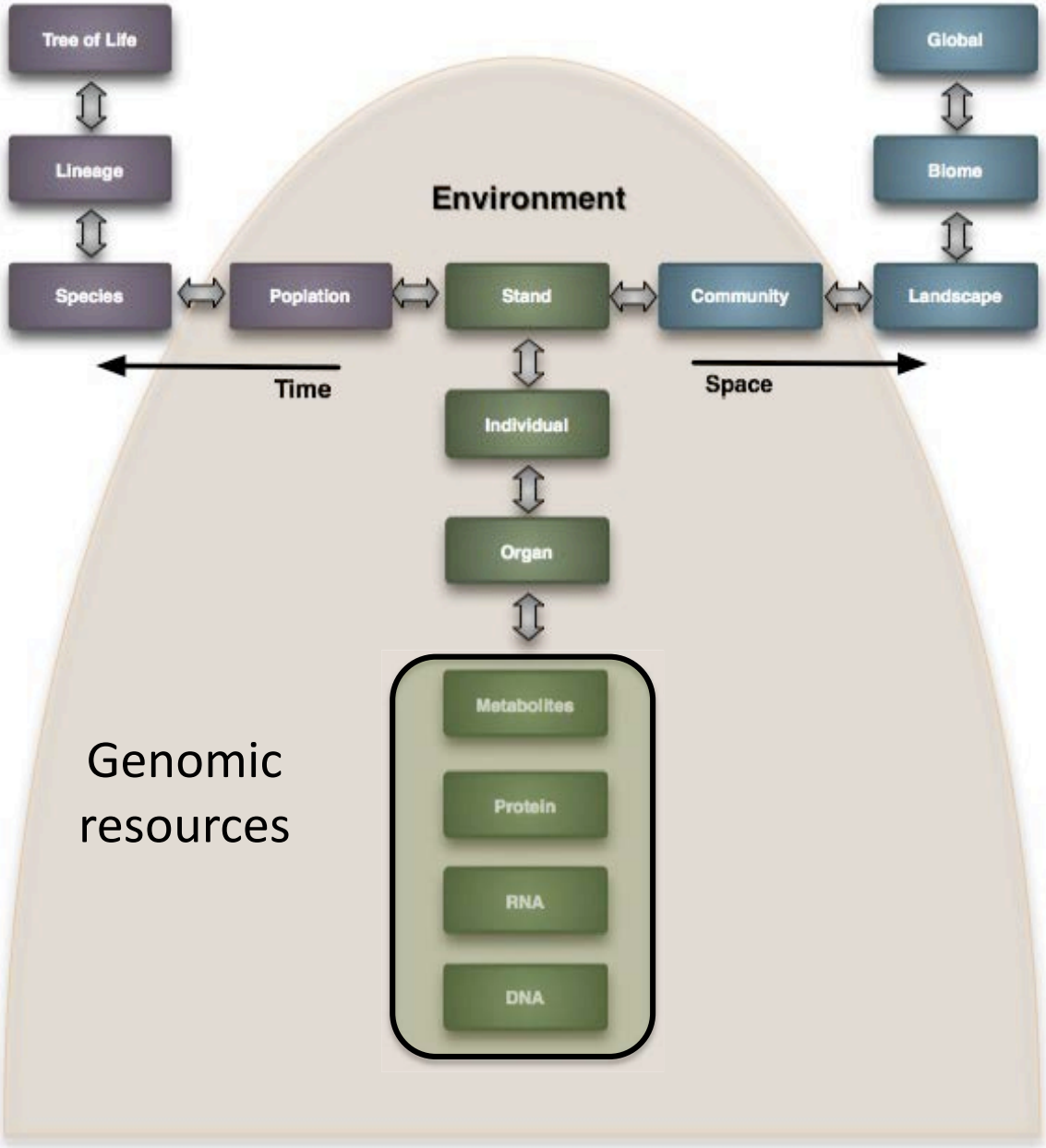
Literature Object

Sequencing

Ontology Development

Plant Ontology and Trait Ontology

- Plant Ontology
  - Structure
    - Needle, Cambium
  - Growth stages
- Trait Ontology
  - Forest Tree Specific Phenotypes
    - Wood Density
- PATO
  - Phenotypic Qualities
- Forest Tree Ontology Meeting (PO and TO) – Feb. 2012
  - NSF Funded Plant Genome Research Resource


Tree of Life
Lineage
Species
Environment
Popiation
Stand
Community
Global
Biome
Landscape
Time
Space
Individual
Organ
Metabolites
Protein
RNA
DNA
Genomic resources

USDA

PineRefSeq
http://pinegenome.org/pinerefseq/

Fagaceae Genomics Web

genomic tools for chestnut, oak, beech, and other trees.

Home | Search | BLAST | DNA Libraries | Sequences | Markers | Genetic Maps | Physical Map

Organisms

- American Beech
- American Chestnut
- Chinese Chestnut
- Oak
- Red Oak
- White Oak

Project Information

- Project Background
- Project Phases and Objectives
- Bioinformatic Methods
- Outreach
- Progress Reports
- Project Team
- Publications
- Photo Gallery
- Links

User login

Username: *

Welcome to the Fagaceae Project

The Fagaceae Genome Web is the web front-end for a multi-institutional NSF-funded project dedicated to the development of genomic resources for the *Fagaceae*. This site disseminates data and analyses by providing genetic and physical maps, transcriptomic data, functional analyses, a listing of available resources and data mining tools. Additionally, background and progress information is available.

Announcements

New Tools & Analyses

New analyses and tools are available for contigs of the most recent unigenes. The analyses include annotation of KEGG pathways, Interpro protein motifs & domians and Gene Ontology (GO) terms. GO and KEGG browsers are now available.

Conference: Genomics of Forest and Ecosystem Health in the Fagaceae

A successful international conference on the genomics of forest and ecosystem health in the Fagaceae was held Nov 10th-13th in Raleigh, NC, USA.

Web Services Development

Communication within TreeGenes

- Development of Web Services in cooperation with NSF's iPlant Cyberinfrastructure Project
  - Software system to support interoperable machine to machine interaction over a network regardless of platform incompatabilities
  - Web service descriptive language (WSDL) is implemented to relate operations

Service Oriented Architecture (SOA)	Remote Procedure Call (RPC)	Representational State Transfer (REST)
With SOAP, the basic unit of communication is a message	RPC Web services define a call interface which the basic unit is the WSDL operation.	REST use HTTP by constraining the interface to standard operations (like GET, POST, PUT, DELETE for HTTP). The focus is on interacting with stateful resources, rather than messages or operations.

SSWAP Ontology

Creating and Contributing to Existing Servlets for Common Genomic Types

Biology Core Set

Ontologies containing core classes and properties useful for establishing a minimum vocabulary with relations to the Open Biomedical Ontologies and others.

OWL Ontology	Ontology Browser	Source File	Ontology Provider	Description
genome	Browse genome	genome.owl	SSWAP	Genome-related terms
map	Browse map	map.owl	SSWAP	Map-related terms
marker	Browse marker	marker.owl	SSWAP	Marker-related terms
qtl	Browse qtl	qtl.owl	SSWAP	QTL-related terms
sequence	Browse sequence	sequence.owl	SSWAP	Sequence-related terms
sequenceServices	Browse sequenceServices	sequenceServices.owl	SSWAP	Sequence Services' terms for BLAST, ClustalW, etc.
taxa	Browse taxa	taxa.owl	SSWAP	Taxa-related terms
trait	Browse trait	trait.owl	SSWAP	Plant trait-related terms

Searches

+ EST (GenBank)

+ Contig

Contig Name: (Ex: 0_8156)

BLAST GenBank Accession: (Ex: ABC59093)

BLAST GenBank GI: (Ex: 224284550)

BLAST Hit: Glycogen (Ex: Kinase)

GO Term: (Ex: Gene)

InterPro Term: (Ex: Kinase)

Enzyme: (Ex: Protein)

Mark all species of interest:

- Picea abies
- Pinus elliottii
- Pinus lambertiana
- Pinus pinaster
- Pinus radiata
- Pinus sylvestris
- Pinus taeda
- Pseudotsuga menziesii

Results: 8

Display

+ Tree Samples
+ DNA Extractions
+ Seq Chromats
+ Sequence Assembly
+ SNP
+ Amplicon
+ Primer
+ Sequencing Rxn
+ Phenotype

rslt_contig_|||Glycogen||||Paab,Piel,Pila,Pipn,Pira,Pisy,Pita,Psme

ID	Contig_Name	Species	BLAST Hit	Has Assemblies	Has SNPs
1410	0_11313	Pinus taeda (Loblolly Pine)	glycogen synthase kinase 3	Yes	Yes
11699	0_4981	Pinus taeda (Loblolly Pine)	glycogenin	No	No
12334	0_5883	Pinus taeda (Loblolly Pine)	glycogen synthase	No	No
14315	0_8526	Pinus taeda (Loblolly Pine)	glycogen phosphorylase	No	No
21987	2_10067	Pinus taeda (Loblolly Pine)	glycogen phosphorylase	No	No
26227	2_4736	Pinus taeda (Loblolly Pine)	glycogenin-like protein	No	No
28155	2_6578	Pinus taeda (Loblolly Pine)	glycogen synthase	No	No

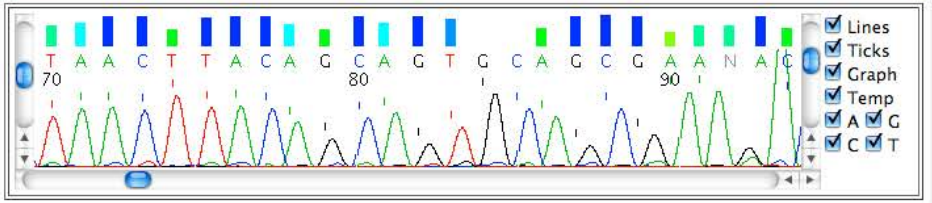
View Selected

data_contig_1410

Contig Name: 0_11313
ID: 1410
Sequence File: 0_11313.fasta
Genus/Species: Pinus / Pinus taeda
ESTs: RTCU1_32_H11.b1_A029
Forward Primer: 0_11313_01_F
Reverse Primer: 0_11313_01_R
Amplicon: 0_11313_01
Seq Assembly: 0_11313_01 20061004 Pita
BLAST Hit: glycogen synthase kinase 3
BLAST GenBank GI: 226519682
BLAST GenBank Accession: ACO65672
InterPro Term: Information not available at this time.
Function GO Term: kinase activity
Function GO Term: protein kinase activity
Function GO Term: ATP binding
Function GO Term: transferase activity
Function GO Term: nucleotide binding
Process GO Term: protein amino acid phosphorylation
Function GO Term: protein serine/threonine kinase activity
ExPASy Enzyme: Information not available at this time.
Sequence: >0_11313(1356234) SuperScript=21352

```
agagatcatttcagttaccatgtggtttgaagttXXXagtgggggtaagg
ggcgaccattgggtaagcaggtgctgggatcccggagttcatcaaaXXXX
ggatgtgcacatgcttcaagctgtatatccagtgaatatccttttagaat
gtcccaaaaacagagggaacatctgcaatctatttattcaaagaaactac
aatgttgtccatgttgcaaacatgtacaaatgtatctaacttacagcagt
gcagcgaagactcggtgagtactggagaatccttgagagaagatctatcg
cctctgcXXXtgctcgcctgtgXXXXXgcttctgccaagggtgggccttt
atctgtggaaacttaaactctggataattaggattcatgcattXtatctc
ctcgcgagtaggagtcccaaggacct
```

data_chromatogram_ADEPT2_0_11313_01-6051_Pt-F



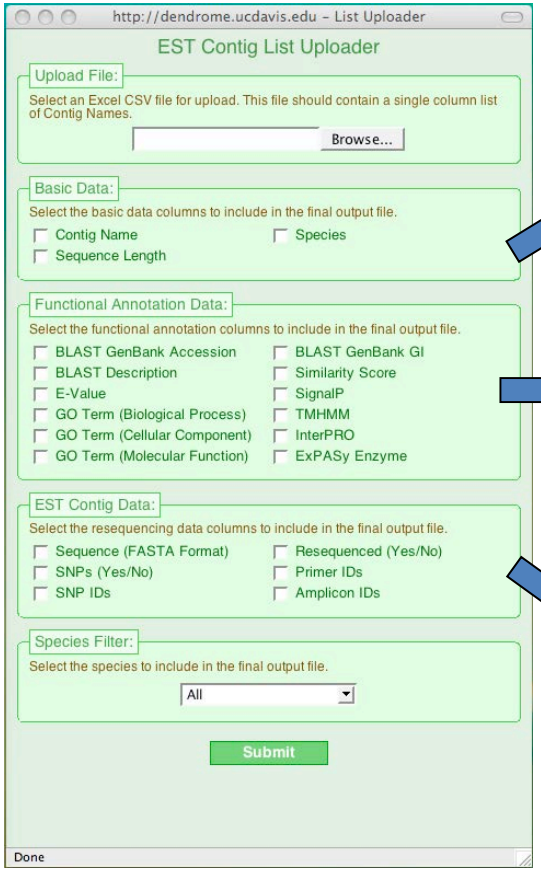
data_seqassembly_4630

ACE File: Download
FASTA File: View
FASTA File: Download
ADEPT2 Chromatogram Archive: 0_11313_01-ADEPT2.tar.gz

Sequencing Chromat: 0_11313_01-6051_Pt-F	Tree Sample: 17016	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6051_Pt-R	Tree Sample: 17016	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6052_Pt-F	Tree Sample: 22005	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6052_Pt-R	Tree Sample: 22005	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6053_Pt-F	Tree Sample: 25051	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6053_Pt-R	Tree Sample: 25051	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6054_Pt-F	Tree Sample: 01011	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6055_Pt-F	Tree Sample: 23009	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6055_Pt-R	Tree Sample: 23009	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6056_Pt-F	Tree Sample: 19002	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6056_Pt-R	Tree Sample: 19002	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6057_Pt-F	Tree Sample: 08103	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6057_Pt-R	Tree Sample: 08103	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6058_Pt-F	Tree Sample: 09010	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6058_Pt-R	Tree Sample: 09010	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6059_Pt-F	Tree Sample: 10005	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6059_Pt-R	Tree Sample: 10005	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6060_Pt-F	Tree Sample: 07056	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6060_Pt-R	Tree Sample: 07056	NCBI Trace	View Chromatogram
Sequencing Chromat: 0_11313_01-6061_Pt-F	Tree Sample: 18001	NCBI Trace	View Chromatogram

Bulk Retrieval Window Components

Bulk Retrieval Window



Data & Annotation Selection Fields

Basic Data:

Select the basic data columns to include in the final output file.

- Contig Name
- Species
- Sequence Length

Functional Annotation Data:

Select the functional annotation columns to include in the final output file.

- BLAST GenBank Accession
- BLAST GenBank GI
- BLAST Description
- Similarity Score
- E-Value
- SignalP
- GO Term (Biological Process)
- TMHMM
- GO Term (Cellular Component)
- InterPRO
- GO Term (Molecular Function)
- ExPASy Enzyme

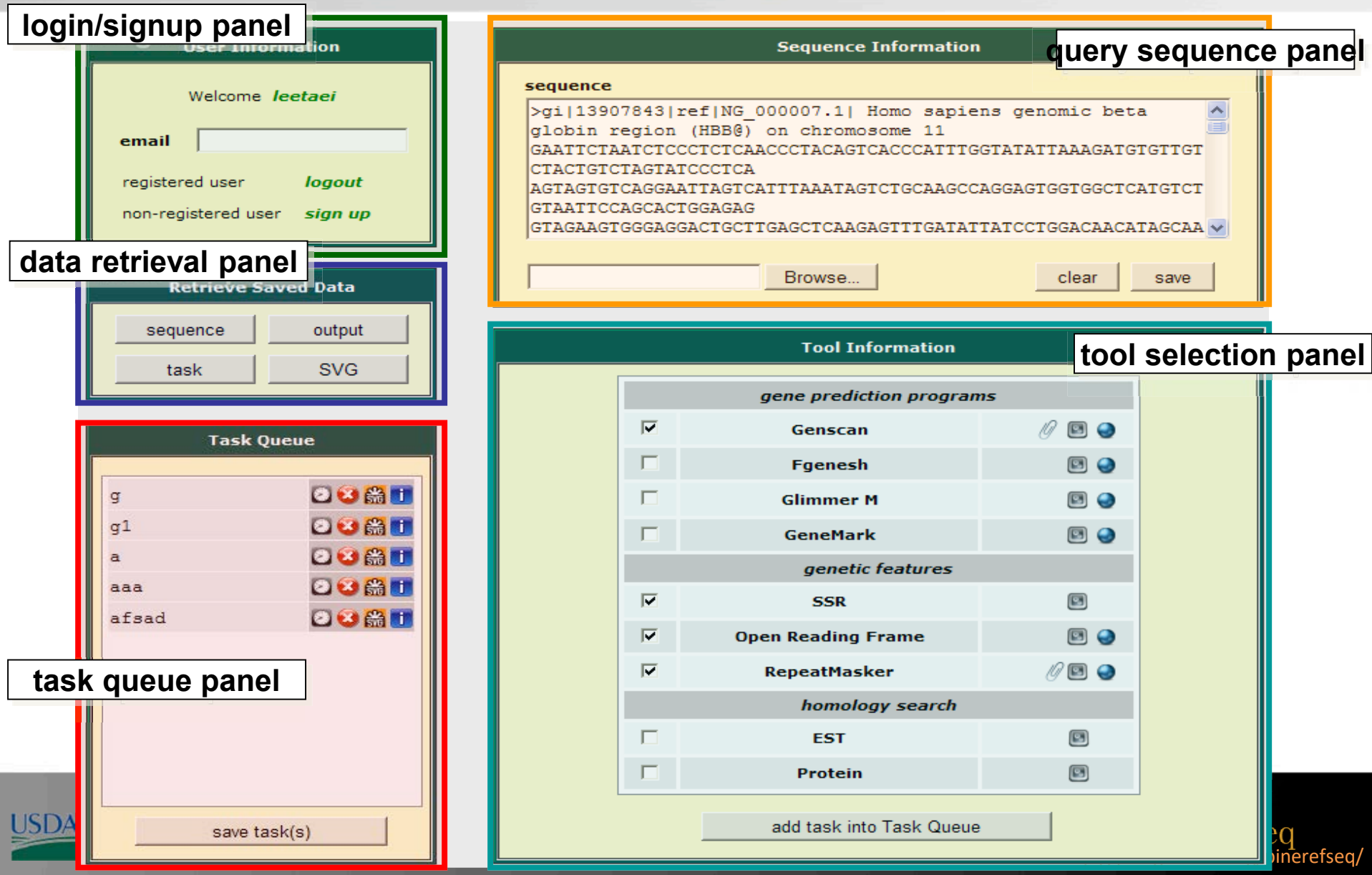
EST Contig Data:

Select the resequencing data columns to include in the final output file.

- Sequence (FASTA Format)
- Resequenced (Yes/No)
- SNPs (Yes/No)
- Primer IDs
- SNP IDs
- Amplicon IDs

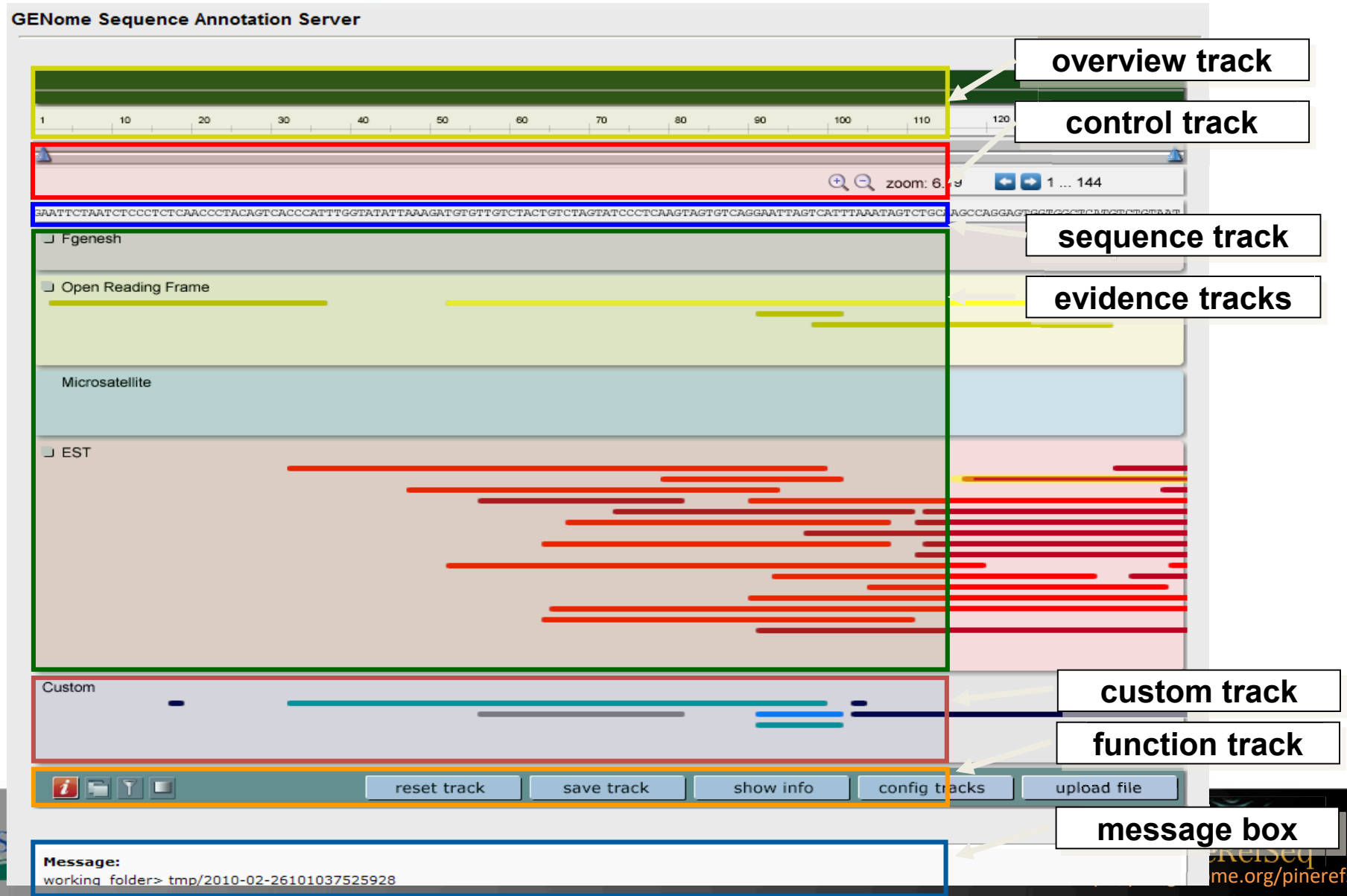
GenSAS development with Content Management

Plone and Drupal



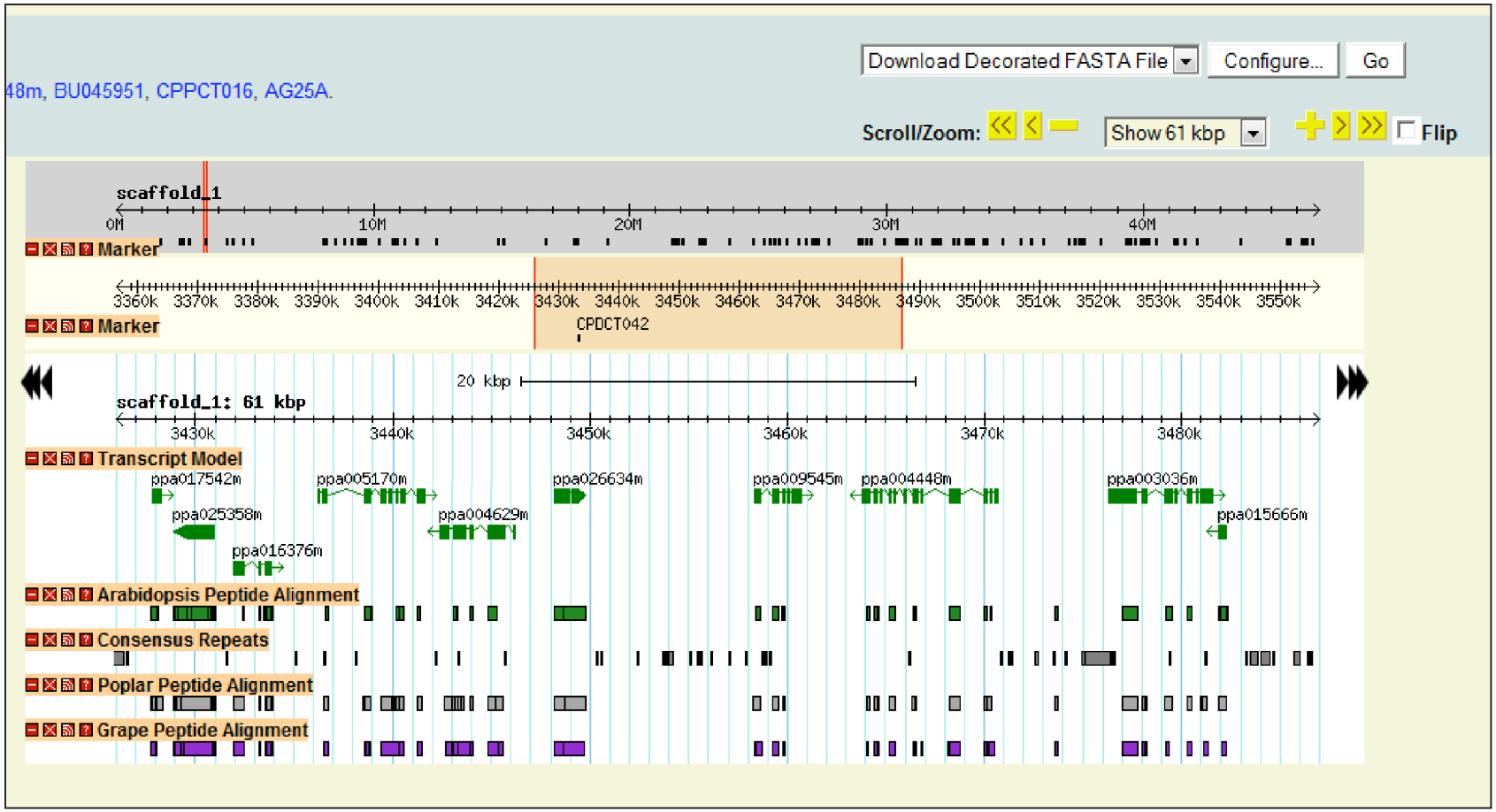
GenSAS development

Multiple Gene Prediction Tracks



GenSAS integration with Gbrowse

Prototyped with Peach Genome in GDR



Analysis Resources

Custom Databases

BLAST

Basic BLAST

Use BLAST by pasting in a nucleotide or amino acid sequence in simply using an accession or GI number.

- nucleotide blast - Search a nucleotide database using a n
  - Algorithms: blastn, megablast, discontiguous meg
- protein blast - Search protein database using a protein qu
  - Algorithms: blastp, psi-blast, phi-blast
- blastx - Search protein database using a translated nuclec
- tblastn - Search translated nucleotide database using a pr
- tblastx - Search translated nucleotide database using a tra query

PSI-BLAST

(PSI) Position Specific Iterated BLAST - Search databases of sequ members of the same protein family. First round runs BLAST and specific score matrix (PSSM) from significant hits while iterations query on profile databases. Uncovers more distant protein relatio

PHI-BLAST

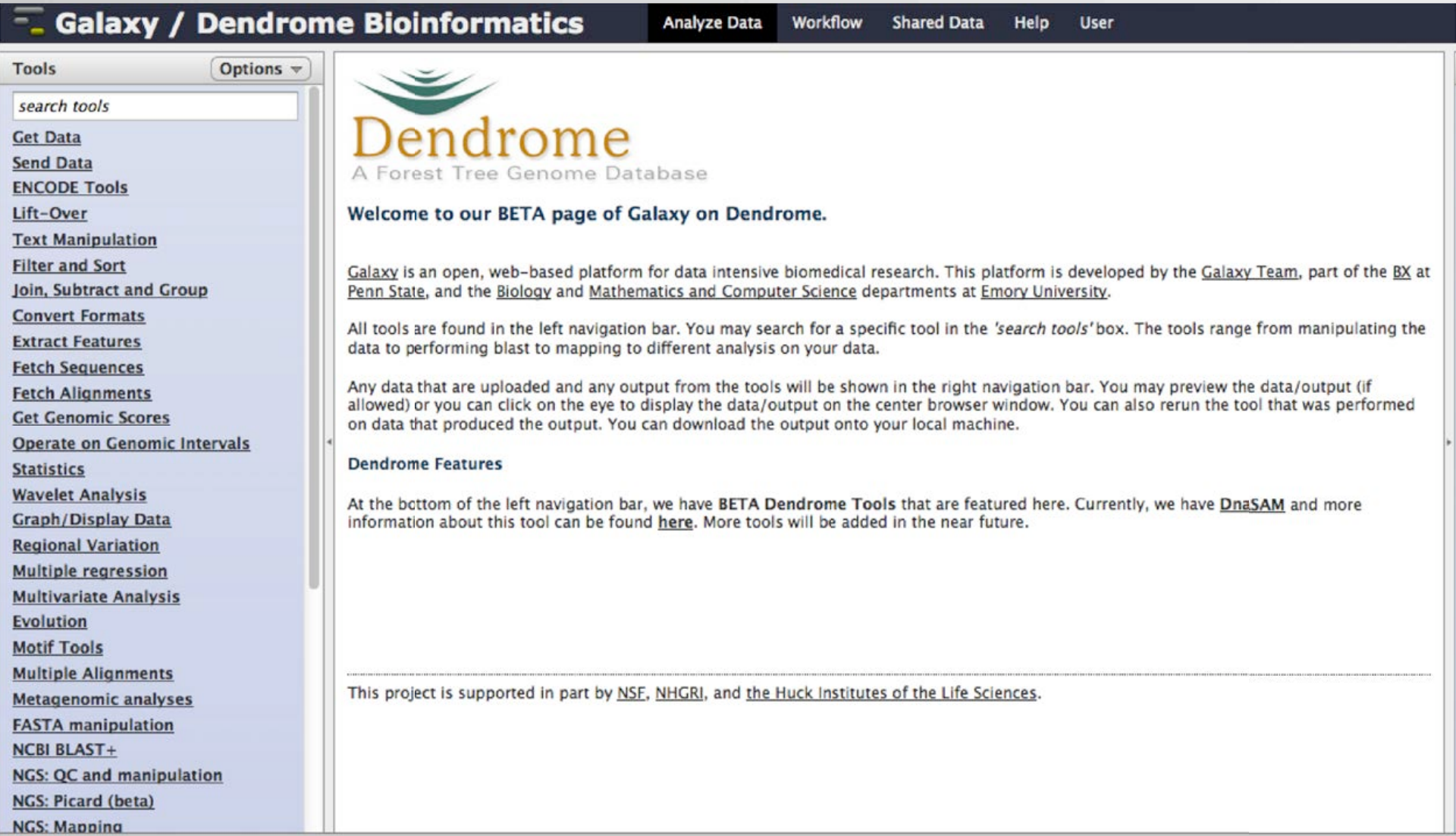
(PHI) Pattern Hit Initiated BLAST - Identify proteins that preserve contained within your query. Query includes a protein sequence expression defining the pattern of interest

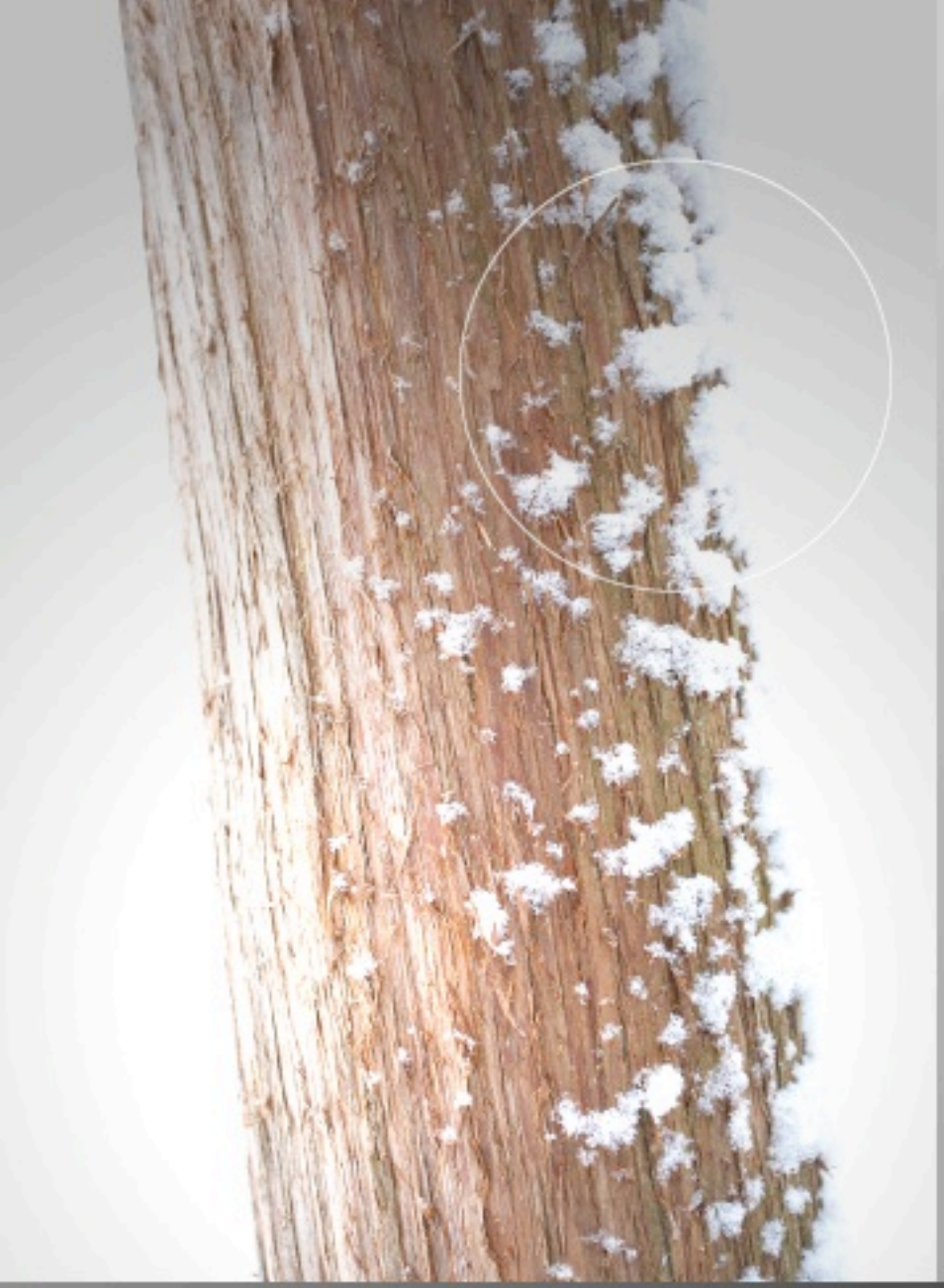
Species-Specific Nucleotide Blast Databases

Database	Description	Links	Last Updated
AllSpeciesAmplicons.nt	Sanger resequenced amplicons in # conifers from EST contigs	[More_Info]	01/13/2012
At.seq.uniq	Unigene set for Arabidopsis thaliana from NCBI	[More_Info]	10/29/2011
Mtr.seq.uniq	Unigene set for Medicago truncatula from NCBI	[More_Info]	09/02/2011
Nta.seq.uniq	Unigene set for Nicotiana tabacum from NCBI	[More_Info]	09/02/2011
Os.seq.uniq	Unigene set for Oryza sativa from NCBI	[More_Info]	12/15/2011
Pgl.seq.uniq	Unigene set for Picea glauca from NCBI	[More_Info]	11/08/2011
pila_transcriptomev1.nt	TreeGenes Transcriptome Assembly v1.0 for Pinus lambertiana		01/13/2012
PitaSangerAmplicons.nt	Sanger resequenced amplicons from EST contigs in Pinus taeda	[More_Info]	01/13/2012
pita_transcriptomev1.nt	TreeGenes Transcriptome Assembly v1.0 for Pinus taeda	[More_Info]	01/13/2012
Ppa.seq.uniq	Unigene set for Physcomitrella patens from NCBI	[More_Info]	09/02/2011
Ppe.seq.uniq	Unigene set for Prunus persica from NCBI	[More_Info]	09/02/2011
Psi.seq.uniq	Unigene set for Picea sitchensis from NCBI	[More_Info]	09/02/2011
psme_transcriptomev1.nt	TreeGenes Transcriptome Assembly v1.0 for Pseudotsuga menziesii	[More_Info]	01/13/2012
Pta.seq.uniq	Unigene set for Pinus taeda from NCBI	[More_Info]	09/01/2011
Pth.seq.uniq	Unigene set for Populus trichocarpa from NCBI	[More_Info]	09/02/2011
Ptp.seq.uniq	Unigene set for Populus tremula X Populus tremuloides from NCBI	[More_Info]	12/15/2011
Qro.seq.uniq	Unigene set for Quercus robur from NCBI	[More_Info]	09/02/2011
Sbi.seq.uniq	Unigene set for Sorghum bicolor from NCBI	[More_Info]	10/02/2011
tair.cds	TAIR10 CDS from Arabidopsis	[More_Info]	08/12/2011
tair.seq	TAIR10 Nucleotide from Arabidopsis	[More_Info]	08/12/2011
Vvi.seq.uniq	Unigene set for Vitis vinifera from NCBI	[More_Info]	09/02/2011
Zm.seq.uniq	Unigene set for Zea mays from NCBI	[More_Info]	12/15/2011

Integrating Tools into TreeGenes

Galaxy





Acknowledgements

University of California at Davis

Ben Figueroa

John Liechty

John Yu

Hans Vasquez-Gross

USFS

Rich Cronn

Jessica Wright

Brian Knaus

Hardeep Rai

UGA

Jeffrey Dean

Walt Lorenz

More Information

C19 Monday, 1:50 pm

P064

USDA

PineRefSeq

http://pinegenome.org/pinerefseq/