

Integrating Genome and Transcriptome Resources into TreeGenes

Jill Wegrzyn

David Neale

Doreen Main

Keithanne Mockaitis

University of California at Davis

Washington State University

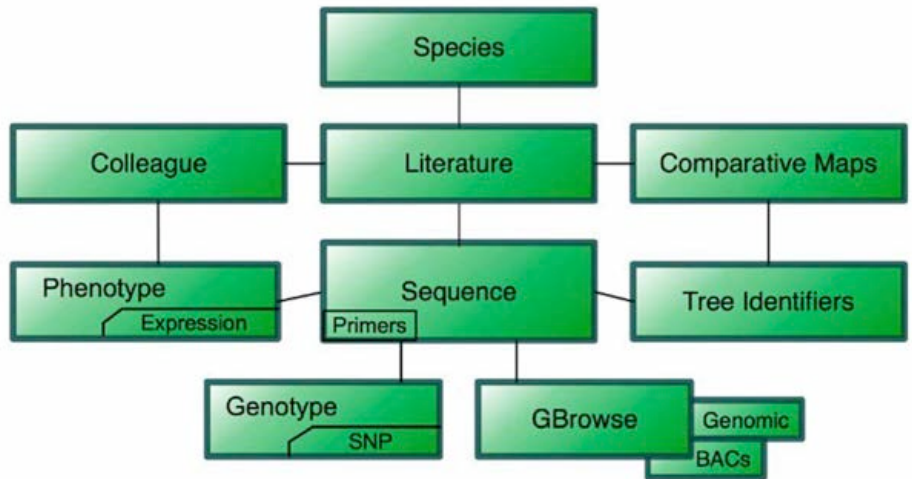
Indiana University/Center for Genomics and
Bioinformatics



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 |

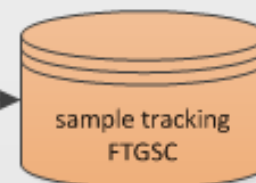
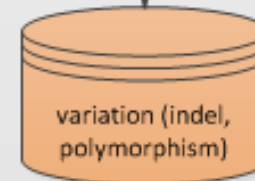
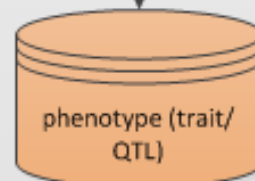
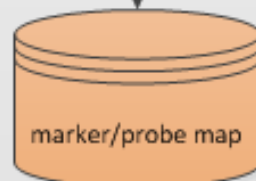
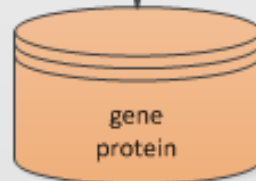
Online Analysis Tools

| BLAST | FASTA | SSR | CAP3 |
| Functional Annotation Pipeline |

Breeders Toolbox

| Sample Tracking Interface |
| FTGSC |

access



- data integration
- ontology associations
- computational analysis

TreeGenes Database in PostgreSQL

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 PINE BLISTER RUST FUNGUS CRONARTIUM RIBI COLA**
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 PINE 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 TEMPE**
Annals of Botany (London) 2000 86:1081-1086.

NAGEIA												74		
NEOCALLITROPIS												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](#)



Software

CMAF: Obtaining TreeGenes (TG) Accession Number

Obtain TreeGenes Accession Number

The following form is for the submission of genetic map files and will provide the submitter with a TreeGenes accession number that will reference the map in the TreeGenes database. This interface is intended for maps associated with a manuscript that will be submitted to a peer-reviewed journal.

If you would like to submit the mapping files for a manuscript that has already been published, please follow this [link here](#).

Please enter preliminary information about your manuscript and upload your genetic map as one file per linkage group. Detailed instructions on the nomenclature and the formatting of the files can be found [here](#).

Following successful submission, you will receive a TreeGenes accession number on the final screen and via e-mail. Your [colleague entry](#) must be up to date in our colleague database to begin this submission process.

Please reference the TreeGenes database and the reference number received here in the manuscript intended for submission.

For more information about how to complete this form. Please [click here](#).

Identification

✓ Primary Author: Eckert, Andrew
Not In Our Database?

✓ Author's Email: laeekert@ucdavis.edu

Other Authors: Joost van Heerwaarden, Jill L. Wegryn, C. Dana Nelson, Jeffrey Ross-Ibarra, Santiago C. González-Martínez, David B. Neale

Publication Information

✓ Title: Patterns of population structure and associations to broad scale environme

✓ Journal: Genetics

✓ Pinus taeda
 Keyword: Other EST

✓ Abstract: TBA

Genetic Map

✓ File: lb\genetic_map_eckert
 Organization: University of California at Davis

Submit Information

Add literature data and (first) map file

Add Supplemental Data

For instructions on the acceptable formats for these fields, [click here](#)

Organization Info

✓ Author: Eckert, Andrew

✓ Paper: Patterns of population structure and associations to broad scale enviro

✓ Supplement Type: Genetic Map

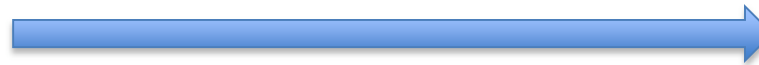
Genetic Map

✓ File: lb\genetic_map_eckert Browse...

✓ Organization: University of California at Davis

Submit Information

(optional) Add additional map files



Obtain TG Accession number!

Pita_UCD_genetic_2009_Eckert_225

Abn... genetic_2009

Accession ID: TG944

... (Loblolly Pine)

Map Type: Genetic
 Map Units: cM
 Linkage Groups: 12
 Published On: January, 2009
 Paper: High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (Pinus taeda L.)
 Mapping Pop: QTL and BASE pedigrees sensu Groover et al. (1994)

[Show Only This Set]
 [Download Map Set Data]
 [View Map Set In Matrix]
 [View Species Info]
 [View Map Type Info]
 [Display All Maps in Viewer]

Linkage Groups

Pita_UCD_2009_01	[Map Viewer]	[Map Details]	[Matrix Viewer]	[Download Map Data]
Pita_UCD_2009_02	[Map Viewer]	[Map Details]	[Matrix Viewer]	[Download Map Data]
Pita_UCD_2009_03	[Map Viewer]	[Map Details]	[Matrix Viewer]	[Download Map Data]
Pita_UCD_2009_04	[Map Viewer]	[Map Details]	[Matrix Viewer]	[Download Map Data]
Pita_UCD_2009_05	[Map Viewer]	[Map Details]	[Matrix Viewer]	[Download Map Data]

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:

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:
 Acacia mangium
 Cryptomeria japonica
 Eucalyptus globulus
 Eucalyptus grandis

Restrict feature types:
 AFLP
 Centromere
 Clone
 Contig

Map Search

Published:

Ref. Species:

Ref. Set:

Name (optional):

Min # of Related Maps:

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	per unit
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

Reference
Douglas Fir
Psme_UCD_genetic_2009

[i] [M]

Psme_UCD_2009_01
[26]

[?] [M] [F] [N]

- Mag +
0-140.70 cM

Feature Types:

— RFLP
— SNP

Menu Symbols:

- [i] Map Set Info
- [?] Map Details
- [M] Matrix View
- [L] Limit to One Map
- [X] Delete Map Set
- [X] Delete Map
- [F] Flip Map
- [UF] Unflip Map
- [N] New Map View

Individual features and their locations on map

List of features on map

GMOD Genome Browser



Showing 142.4 kbp from BAC3, positions 1 to 142,351

Instructions
[Bookmark this] [Upload your own data] [Hide banner] [Share these tracks] [Link to Image] [High-res Image] [Help] **Logout**

Search
Landmark or Region: BAC3:1..142351
Data Source: Pinus taeda (loblolly pine) BAC3

Reports & Analysis:
Annotate Restriction Sites
Scroll/Zoom: <<<

Overview

Region
Details

Masked Monocot/Dicot Repeat-Censor
LTR
Note:LTR_1 Note:LTR_11
Note:LTR_2 Note:LTR_4
Note:LTR_10 Note:LTR_6 Note:LTR_27
Note:LTR_12
Note:LTR_14 Note:LTR_18 Note:LTR_19 Note:LTR_21 Note:LTR_25
Note:LTR_15 Note:LTR_16 Note:LTR_3 Note:LTR_9 Note:LTR_23 Note:LTR_17
Note:LTR_26

Unmasked Monocot/Dicot BLASTX

Search and
Select data source

Tracks can be
reordered or
hidden as necessary

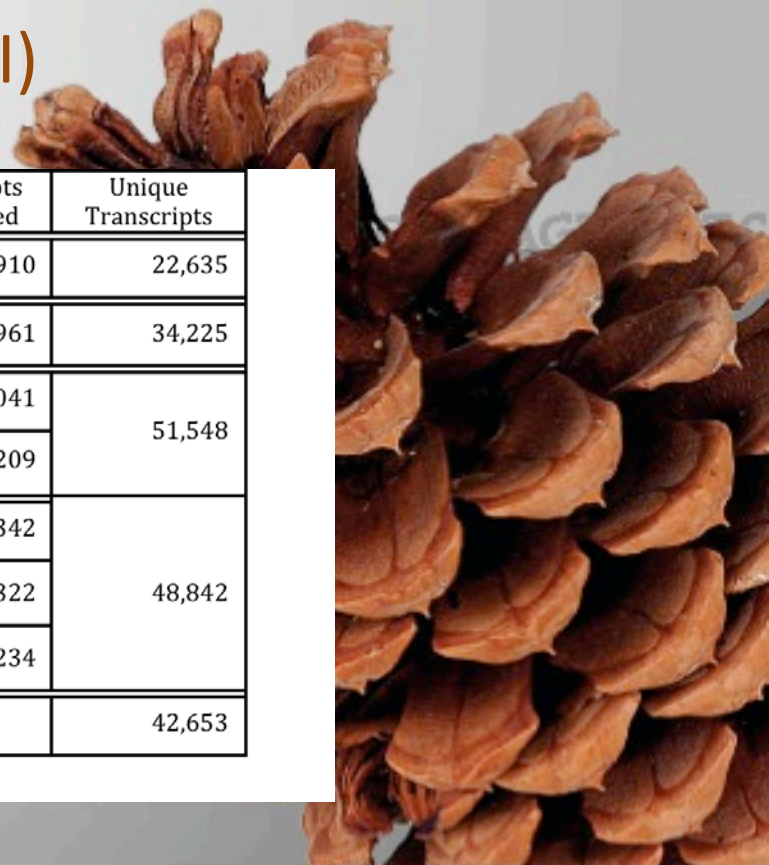
Transcriptome Assembly Summary

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



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 Transcriptome Assembly v1.0 [FS] (39,209)	More Info Download [Trinity]
------------------------------------------------------	---------------------------------

454 Transcriptome Assembly [JGI] (25,041)	Download
-------------------------------------------	----------

Download Annotations Annotations by Species

Click the picture below for GBrowse interface

TreeGenes UniGenes (EST,cDNA,RNASeq,454) (51,548)



Project Links Pine Reference Sequence Project Douglas-fir Transcriptome Project

Proteome Resources

Protein (3,902)	Search Download
-----------------	--------------------

Transcriptome Resources in TreeGenes

Forest Tree Genetic Stock Center



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 UC Davis 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 Population Sets page.

Sample Submission

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



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?

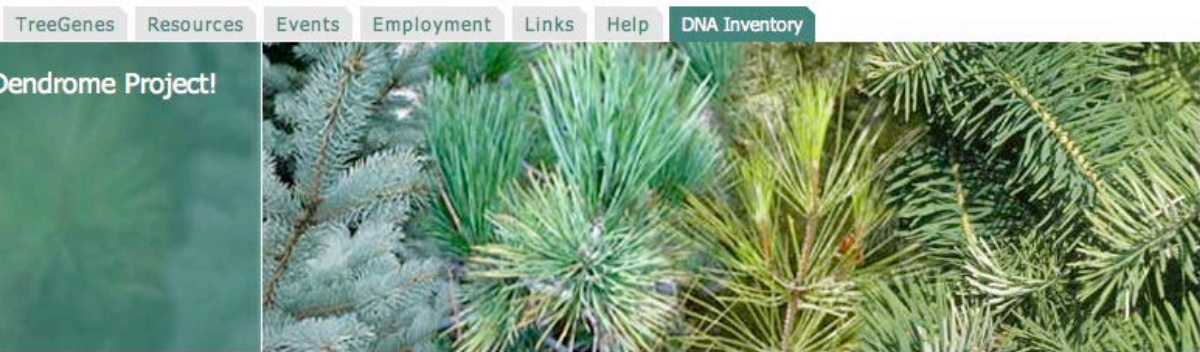
Please provide the total number of individual (genetically distinct) trees, not the total number of clones.

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)

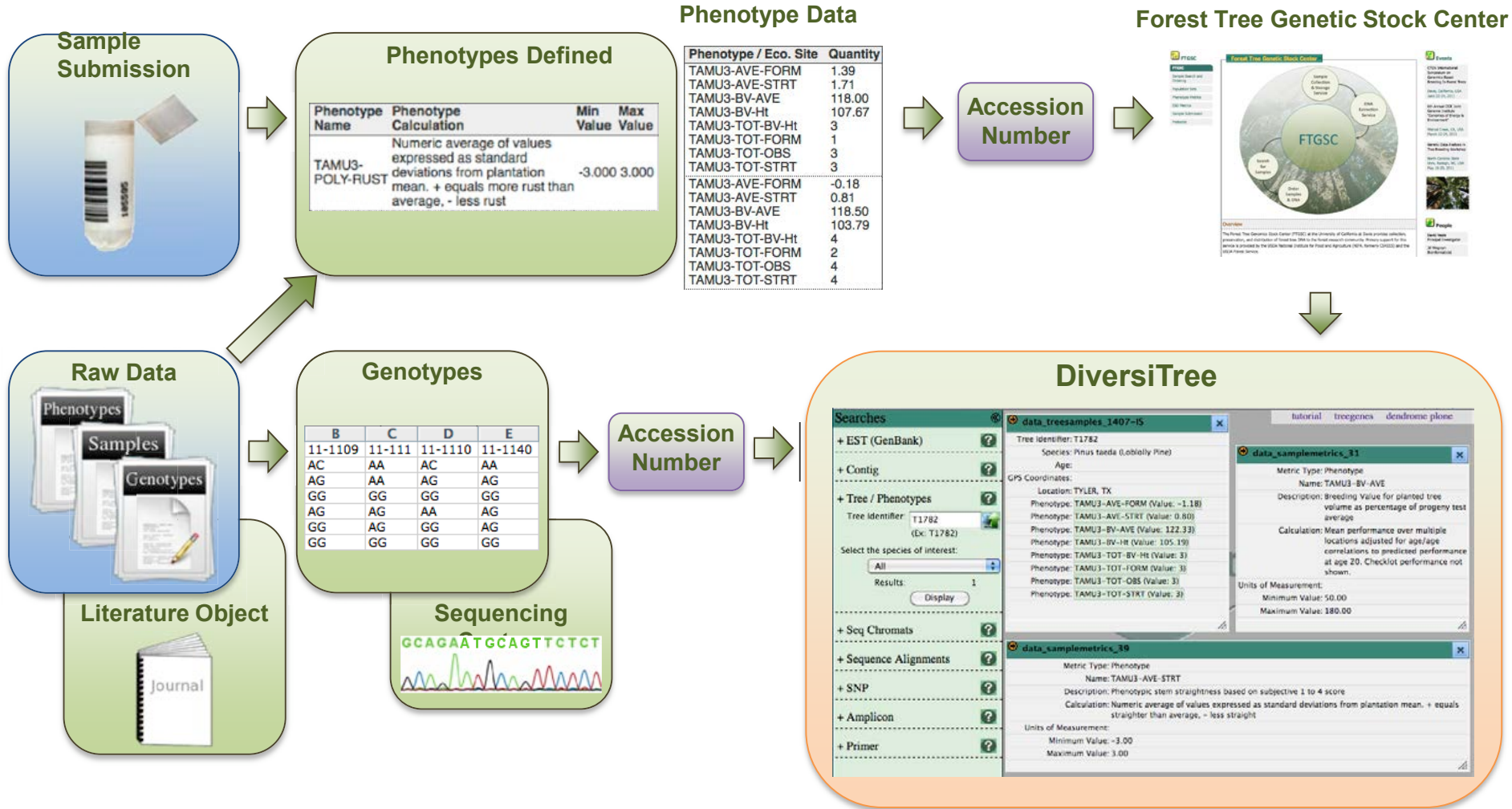
Provide interfaces and database backend to support a DNA distribution center (UCD)

Samples

A listing of each individual samples and their Phenotypes / Ecological Site Descriptors.

Date Range	Project	Institution	Tree ID	Barcode	Sample Type	Phenotype / Eco. Site	Quantity
Start: 01/01/09 End: 06/15/10	All Projects	All Institutions			All	All	
2009.9.23	CTGN - Pita - Loblolly Pine	North Carolina State University (Whetten)	SOURCE_NC000009	NC000009	Needle	<ul style="list-style-type: none"> All NCSU1-BV-Cellulose NCSU1-BV-Density NCSU1-BV-Form NCSU1-BV-Ht NCSU1-BV-Lignin NCSU1-BV-Rust NCSU1-BV-Strt NCSU1-BV-Taper-B NCSU1-BV-Vol NCSU1-NC-region NCSU1-r-BV-Cellulose NCSU1-r-BV-Density NCSU1-r-BV-Form NCSU1-r-BV-Ht NCSU1-r-BV-Lignin NCSU1-r-BV-Rust NCSU1-r-BV-Strt NCSU1-r-BV-Taper NCSU1-r-BV-Vol 	
2009.9.23	CTGN - Pita - Loblolly Pine	North Carolina State University (Whetten)	SOURCE_NC000023	NC000023	Needle	<ul style="list-style-type: none"> NCSU1-NC-region NCSU1-r-BV-Form NCSU1-r-BV-Ht NCSU1-r-BV-Rust NCSU1-r-BV-Strt NCSU1-r-BV-Vol 	<ul style="list-style-type: none"> P 0.790 0.690 0.840 0.910 0.660

Data Submission



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

Forest Tree Genetic Stock Center



DiversiTree

Searches

- + EST (GenBank) ?
- + Contig ?
- + Tree / Phenotypes ?
- + Seq Chromats ?
- + Sequence Alignments ?
- + SNP ?
- + Amplicon ?
- + Primer ?

data_treesamples_1407-IS

Tree Identifier: T1782
 Species: Pinus taeda (Loblolly Pine)
 Age:
 GPS Coordinates:
 Location: TYLER, TX
 Phenotype: TAMU3-AVE-FORM (Value: -1.18)
 Phenotype: TAMU3-AVE-STRT (Value: 0.80)
 Phenotype: TAMU3-BV-AVE (Value: 122.33)
 Phenotype: TAMU3-BV-Ht (Value: 105.19)
 Phenotype: TAMU3-TOT-BV-Ht (Value: 3)
 Phenotype: TAMU3-TOT-FORM (Value: 3)
 Phenotype: TAMU3-TOT-OBS (Value: 3)
 Phenotype: TAMU3-TOT-STRT (Value: 3)

Select the species of interest:

 Results: 1

data_samplemetrics_31

Metric Type: Phenotype
 Name: TAMU3-BV-AVE
 Description: Breeding Value for planted tree volume as percentage of progeny test average
 Calculation: Mean performance over multiple locations adjusted for age/age correlations to predicted performance at age 20. Checklist performance not shown.
 Units of Measurement:
 Minimum Value: 50.00
 Maximum Value: 180.00

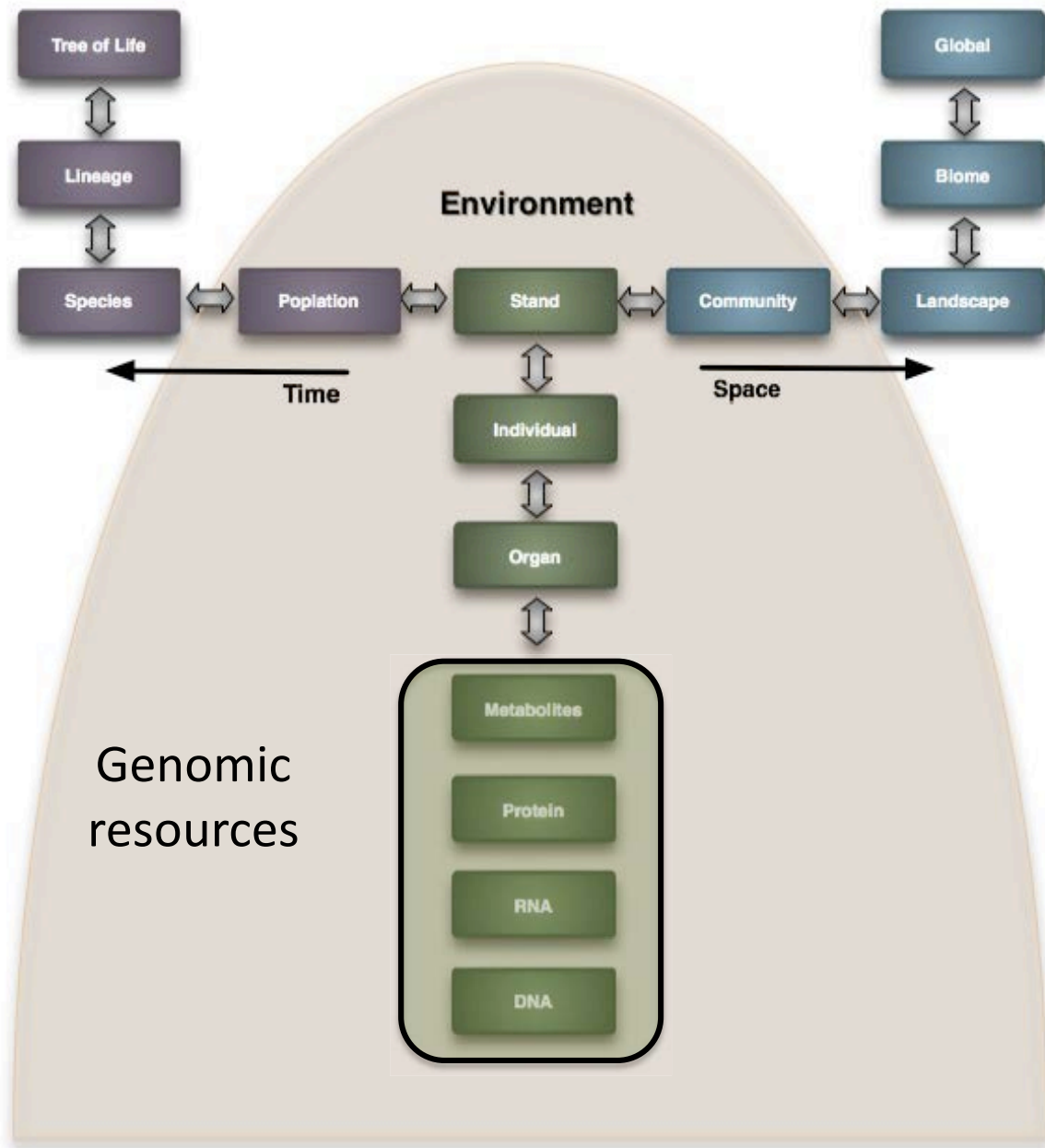
data_samplemetrics_39

Metric Type: Phenotype
 Name: TAMU3-AVE-STRT
 Description: Phenotypic stem straightness based on subjective 1 to 4 score
 Calculation: Numeric average of values expressed as standard deviations from plantation mean. + equals straighter than average, - less straight
 Units of Measurement:
 Minimum Value: -3.00
 Maximum Value: 3.00

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



Genomic, Genetic, and Breeding Resources for Rosaceae Crop Improvement



Welcome to the Genome Database for Rosaceae

The Genome Database for Rosaceae (GDR) is a curated and integrated web-based relational database providing centralized access to Rosaceae genomics and genetics data and analysis tools to facilitate cross-species utilization of data. Supported by the NSF Plant Genome Program from 2003 to 2008, GDR is now funded until 2013 by the USDA Specialty Crop Research Initiative. As part of tree fruit Genome Database Resources (tfGDR), GDR will be expanded to include whole genome sequences and annotations for apple, peach, and strawberry, transcript data, metacyc pathways, large scale phenotype and genotype data, breeding data, controlled vocabularies, and new analysis tools, all of which will be implemented within the open source infrastructure of drupal and chado (tripal).

What's new in GDR?

- The **strawberry genome v1.0 (build 8)**. View and download the genome sequence, predicted genes, homologs and more. Additionally, blast your sequences against this assembly.
- SNP markers included in the **apple 9K**, **peach 9K** and **cherry 6K** illumina SNP arrays available to download in excel and view in GBrowse. Additionally **2.1 million candidate SNPs for apple** are also available.
- The **apple genome v1.0**. View and download the genome sequence, predicted genes, homologs, markers and blast your sequences.
- The **peach genome v1.0**. View and download the predicted genes, genome sequences, putative protein and EST homologs, Rosceae genera mapped transcripts, mapped markers, and blast your sequences against the datasets.
- [Report a problem](#) | [Ask us a question](#) | [Post a job](#) | [Post a meeting or event](#) | [What's been fixed in GDR?](#) | [What are we working on?](#) | [Used GDR data or tools in your research? how to reference us](#)

News

- RosEXEC 2012 membership announced.
- Breeders Toolbox now available on GDR.
- QTL data available for apple, cherry, peach and strawberry.
- 2.1 million apple SNPs now available.
- Peach-Strawberry Synteny now available on GBrowse-Syn.
- Strawberry genome sequence and annotation available on GDR.
- IRSC release data for illumina Infinium apple, peach and cherry SNP arrays.
- ASHS Conference Sept 25-28, 2012
- EU FruitBreedomics project is funded.
- Apple genome sequence and annotation available in GDR.
- Peach genome sequence and annotation available on GDR.
- [Announcements archive](#)



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 & domains 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 incompatibilities
 - 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

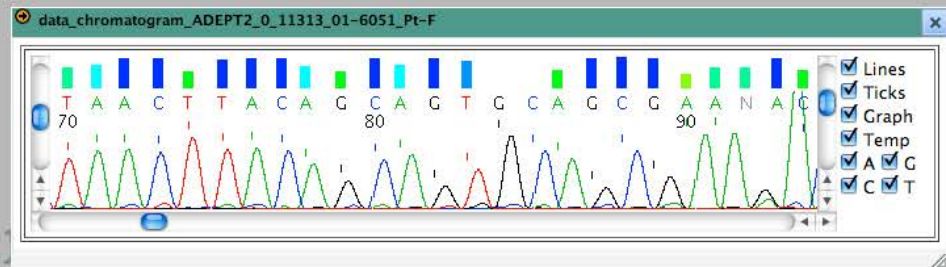
- + EST (GenBank) ?
- + Contig ?
 - Contig Name: (Ex: 0_8156)
 - BLAST GenBank Accession: (Ex: ABC59093)
 - BLAST GenBank GI: (Ex: 224284550)
 - BLAST Hit: (Ex: Kinase)
 - GO Term: (Ex: Gene)
 - InterPro Term: (Ex: Kinase)
 - Enzyme: (Ex: Protein)
- Mark all species of interest:
 - Picea abies
 - Pinus elliotii
 - Pinus lambertiana
 - Pinus pinaster
 - Pinus radiata
 - Pinus sylvestris
 - Pinus taeda
 - Pseudotsuga menziesii
- Results: 8
- + 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

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
 agagatcatttcagttacatgtggttgaagtXXagtggggtaagg
 ggcgaccaattgggtaagcaggctgctggatcccgaggtcatcaaaXXX
 ggatgtgcacatgcttcaagctgtatattcagtgaaatcctttagaat
 gtccc.aaaaacagagggaaacatctgcaactatatttcaaaagaactc
 aatgtgttcctagttgcaaacctgtacaaatgtatcaactaacagcagt
 gcaagcaagactcgtgtagtctgagatccttgagaagaatcattctg
 cctctgXXXtctcgcctgtgXXXXgcttctgccaagggtggcctt
 attctggaacctaaacctgataatgattaccatXtatctc
 ctctcgagttaggtcccaaggacct



data_seqassembly_4630

ACE File: [Download](#)
 FASTA File: [View](#)
 FASTA File: [Download](#)

ADEPT2 Chromatogram Archive: [0_11313_01-ADEPT2.tar.gz](#)

Sequencing Chromat	Tree Sample	NCBI Trace	View Chromatogram
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

http://dendrome.ucdavis.edu - List Uploader

EST Contig List Uploader

Upload File:
Select an Excel CSV file for upload. This file should contain a single column list of Contig Names.

Browse...

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

Species Filter:
Select the species to include in the final output file.

All

Submit

Done

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

login/signup panel

User Information

Welcome *leetaei*

email

registered user [logout](#)

non-registered user [sign up](#)

data retrieval panel

Retrieve Saved Data

task queue panel

Task Queue

g	   
g1	   
a	   
aaa	   
afsad	   

query sequence panel


















Sequence Information

sequence

```
>gi|13907843|ref|NG_000007.1| Homo sapiens genomic beta
globin region (HBB) on chromosome 11
GAATTCTAATCTCCCTCTCAACCCTACAGTCACCCATTGGTATATTAAGATGTGTTGT
CTACTGCTAGTATCCCTCA
AGTAGTGCAGGAATTAGTCATTTAAATAGTCTGCAAGCCAGGAGTGGTGGCTCATGTCT
GTAATTCAGCACTGGAGAG
GTAGAAGTGGGAGGACTGCTTGAGCTCAAGAGTTTGATATTATCCTGGACAACATAGCAA
```

tool selection panel

Tool Information

<i>gene prediction programs</i>		
<input checked="" type="checkbox"/>	Genscan	  
<input type="checkbox"/>	Fgenesh	 
<input type="checkbox"/>	Glimmer M	 
<input type="checkbox"/>	GeneMark	 
<i>genetic features</i>		
<input checked="" type="checkbox"/>	SSR	
<input checked="" type="checkbox"/>	Open Reading Frame	 
<input checked="" type="checkbox"/>	RepeatMasker	  
<i>homology search</i>		
<input type="checkbox"/>	EST	
<input type="checkbox"/>	Protein	

GenSAS development

Multiple Gene Prediction Tracks

The screenshot displays the GENome Sequence Annotation Server interface. At the top, a navigation bar includes the title "GENome Sequence Annotation Server" and a search bar. Below this is a horizontal scale from 1 to 120. The main area contains several tracks: a "control track" with a zoom level of 6; a "sequence track" showing DNA sequence; "evidence tracks" including "Fgenesh" (with an "Open Reading Frame" sub-track), "Microsatellite", and "EST"; a "custom track" with a blue bar; and a "function track" with a red bar. A control bar at the bottom features buttons for "reset track", "save track", "show info", "config tracks", and "upload file". A "message box" at the very bottom shows a system message: "Message: working folder> tmp/2010-02-26101037525928".

overview track

control track

sequence track

evidence tracks

custom track

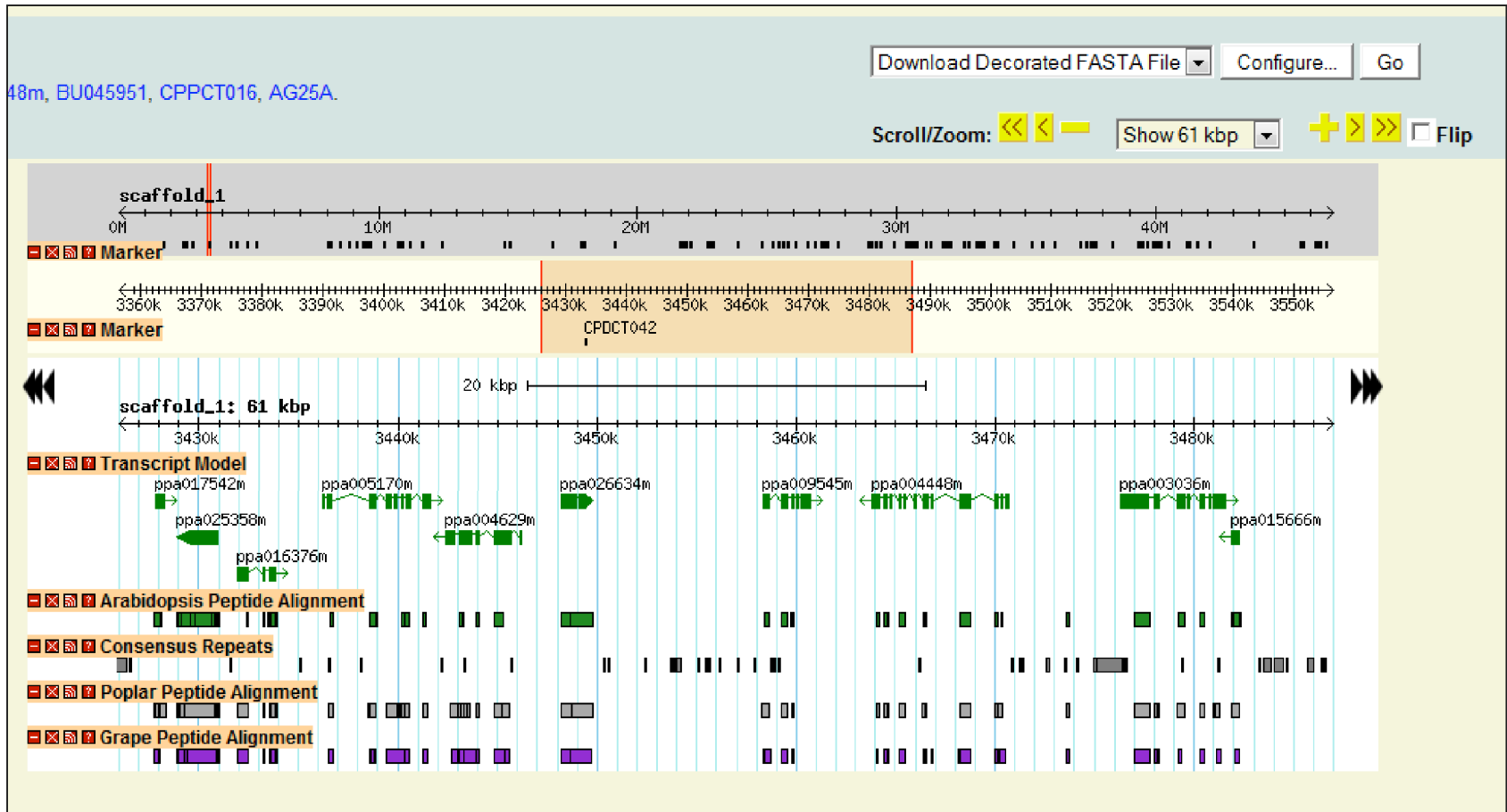
function track

message box

Message:
working folder> tmp/2010-02-26101037525928

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, discontinuous meg
- protein blast - Search protein database using a protein qu
 - Algorithms: blastp, psi-blast, phi-blast
- blastx - Search protein database using a translated nucle
- 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 seq 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

Galaxy / Dendrome Bioinformatics Analyze Data Workflow Shared Data Help User

Tools Options

search tools

Get Data
Send Data
ENCODE Tools
Lift-Over
Text Manipulation
Filter and Sort
Join, Subtract and Group
Convert Formats
Extract Features
Fetch Sequences
Fetch Alignments
Get Genomic Scores
Operate on Genomic Intervals
Statistics
Wavelet Analysis
Graph/Display Data
Regional Variation
Multiple regression
Multivariate Analysis
Evolution
Motif Tools
Multiple Alignments
Metagenomic analyses
FASTA manipulation
NCBI BLAST+
NGS: QC and manipulation
NGS: Picard (beta)
NGS: Mapping


Dendrome
A Forest Tree Genome Database

Welcome to our BETA page of Galaxy on Dendrome.

Galaxy is an open, web-based platform for data intensive biomedical research. This platform is developed by the Galaxy Team, part of the BX at Penn State, and the Biology and Mathematics and Computer Science departments at Emory University.

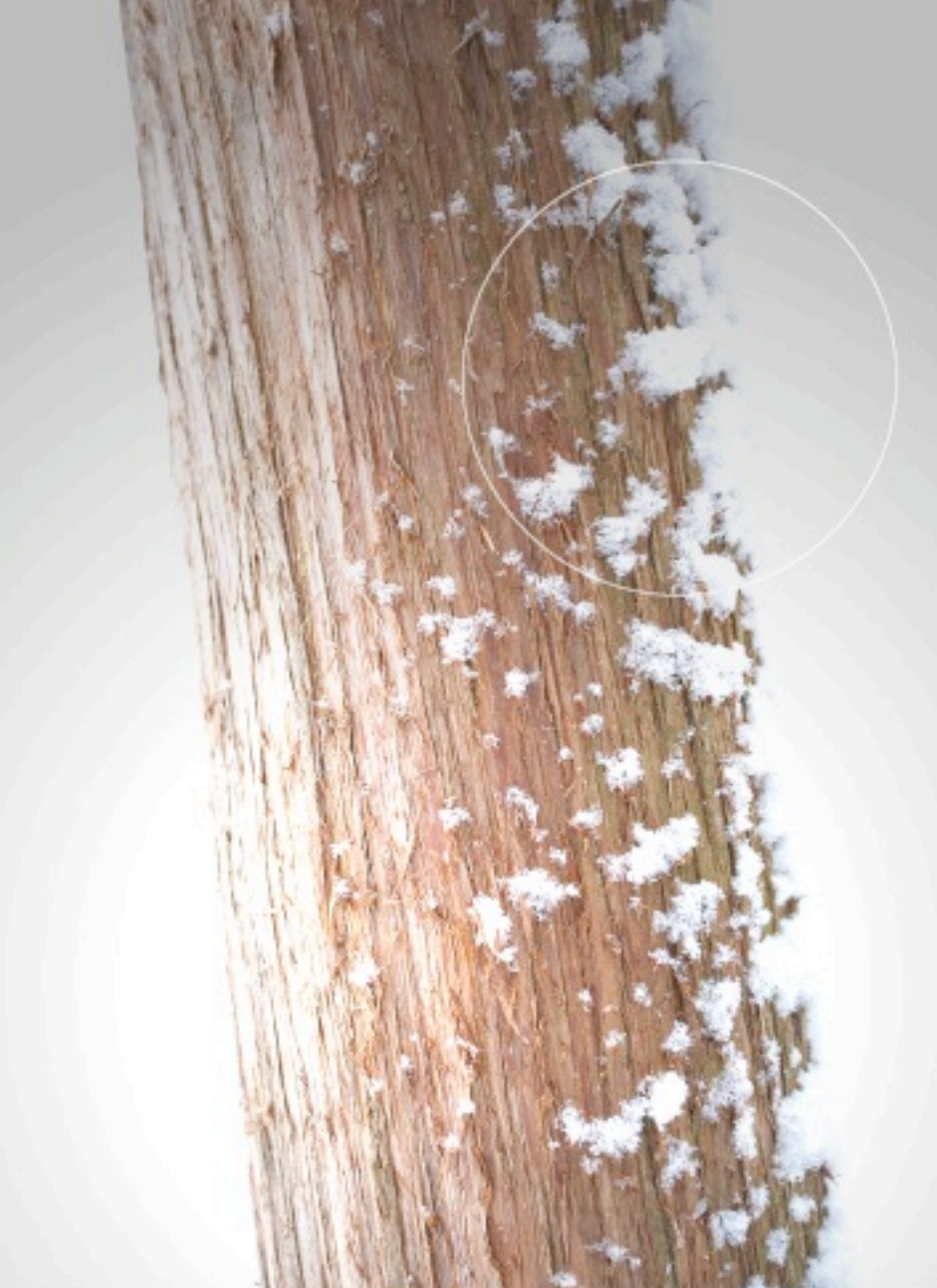
All tools are found in the left navigation bar. You may search for a specific tool in the 'search tools' box. The tools range from manipulating the data to performing blast to mapping to different analysis on your data.

Any data that are uploaded and any output from the tools will be shown in the right navigation bar. You may preview the data/output (if allowed) or you can click on the eye to display the data/output on the center browser window. You can also rerun the tool that was performed on data that produced the output. You can download the output onto your local machine.

Dendrome Features

At the bottom of the left navigation bar, we have **BETA Dendrome Tools** that are featured here. Currently, we have DnaSAM and more information about this tool can be found here. More tools will be added in the near future.

This project is supported in part by NSF, NHGRI, and the Huck Institutes of the Life Sciences.



* 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