CottonGen User Tutorial: Browse Genome Sequences

This tutorial describes how to:
- *Browse genome sequence projects that are available on CottonGen*
- *View detailed information of a genome sequencing project*
- *Download genome sequence and sequence alignment data*
- *View genome data and alignments via GBrowse*

Browse Genome Sequencing Projects and Data

1. Open the web page http:www.cottongen.org
2. Select **Data** from **Navigation Bar** then click on **Genome**
   ➢ A page of all CottonGen Collected genome sequence projects will be displayed
3. Click on a sequence project name to view the project details page.

Download genome sequence and sequence alignment data

This area shows summary information about the selected genome .sequencing

Click here to download the fasta file of assembled genome sequences.

Click HERE to BLAST your dataset against the assembled genome sequences or predicted genes.

Click HERE will redirect you to the GBrowse web interface that displays the assembled and annotated data tracks of the genome sequences aligned to it.

Gossypium raimondii (D5) genome JGI assembly v2.0 (annot v2.1)

Details

Analysis Name	Gossypium raimondii (D5) genome JGI assembly v2.0 (annot v2.1)
Software	Arachne2 (modified)
Source	Sanger based sequence, Roche 454, and Illumina based short reads
Date performed	2013-02-18
Materials & Methods	**Please Note:** This genome assembly is made available through a "Reserved Analysis" restriction Please see the usage policy below for further details.

The following text comes from phytozome.org:

Overview
This v2.1 annotation release is on genome assembly v2.0, a high quality version of the Cotton D (*Gossypium raimondii*) genome sequenced from DNA provided by Andrew Paterson at Univ. GA. It was sequenced with a combination of Sanger, Roche 454 pyrosequencing and Illumina read pairs. This release includes additional screening of small repetitive contigs and a new map integration that corrects several orientation issues within scaffolds.

Statistics

Assembly Summary	
Scaffold total	1,033
Contig total	19,735

Resources

- View G. raimondii in GBrowse
- NCBI BLAST
- Batch BLAST
- Details
- Assembly
- Genes
- Repeats
- Markers
- SNPs
- Transcript Alignments
- Protein Alignments
- Protein Homology
- Functional Annotation
- Downloads

To view and download fasta file of predicted genes and/or GFF3 file of predicted gene positions on the genome, click HERE

To view and download text format of InterPro domains, Gene Ontology (GO) terms, KEGG pathways and KEGG orthologs associated with proteins of this genome, click HERE

To view and download all kinds of data related to this genome sequencing project,

To download GFF3 file of aligned protein sequences and positions on the genome, click HERE

View genome sequencing data and alignments via GBrowse (refer to picture on next page):

1. **Landmark or Region** window shows the chromosome name and the coordinates for the region that is currently being displayed below.

2. **Data Source** window shows the name of the genome dataset currently displayed in GBrowse. The dropdown list on the right side allows users to switch to other genome datasets.

3. The 'blue vertical bar' in the **Overview** indicates the current region that is being viewed in relation to the whole chromosome.

4. **Track(s)** from analyses are also displayed in the **Overview** session. An overview of all the alignments between the analysis results and the displayed chromosome will be displayed under the Track name.

5. **Details** session displays a detailed view of the selected coordinates.

6. The **Track** name in the **Details** session indicates the alignments under it are Arabidopsis TAIR10 proteins aligned to the genome.

7. Move the cursor over the small icons in the track name and a pop-up window will explain what function is associated with the icon

8. Move cursor to a feature on the track, a pop-up window will appear and display the protein name and chromosome name with position the protein aligned to the chromosome.

9. Click on a feature in the track and a new window will show up to display more detailed information about this alignment

10. To annotate restriction sites, download sequences, or alignment data, use the drop down menu on the right side of the Search section. Select the function and click on Configure to select options.

11. Click **on Select Tracks** to view all available tracks and select which of them to show under the **Details** region.