CottonGen, a Central Platform to Facilitate Knowledge Discovery in Cotton Research

www.cottongen.org

WASHINGTON STATE UNIVERSITY
World Class. Face to Face.

Jing Yu[1], Sook Jung[1], Chun-Huai Cheng[1], Taein Lee[1], Ping Zheng[1], Katheryn Buble[1], Jodi L. Humann[1],

James Crabb[1], Heidi Hough[1], B. Todd Campbell[2], Don C. Jones[3] Josh Udall[4], and Dorrie Main[1]

1. Washington State University, Pullman, WA, 2. USDA-ARS, Florence, SC, 3.Cotton Incorporated, Cary, NC, 4. USDA-ARS, College Station, TX

Abstract

CottonGen (www.cottongen.org) integrated significantly more data in 2019: New whole genome assembly data include those from 5 diploids (*G. arboreum, G. raimondii, G. turneri, G. austral, Gossypioides kirkii*) and 5 tetraploids (*G. hirsutum, G. barbadense, G. mustelinum*). Cotton sequences, curated from NCBI, has also been updated and now contains over 53K new genes and over 205K new proteins. Seven new (*G. arboretum, G. barbadense, G. turneri, G. kirkii, G. austral, G. mustelinum*) or updated (*G. hirsutum*) pathways are available from CottonCyc database. 282 cotton trait descriptors, developed from QTL publications and evaluation data, have been submitted to Crop Ontology (www.cropontology.org). Other new data added in 2019 include 8000 non-fiber traits and 5,000 new fiber trait data from RBTN, 321 QTLs, 594 QTL trait data, 2,230 SSR markers collected from peer-viewed publications. New and updated tools include the interactive and customizable search tool, MegaSearch and new functionalities including SNP genotype data management in BIMS.
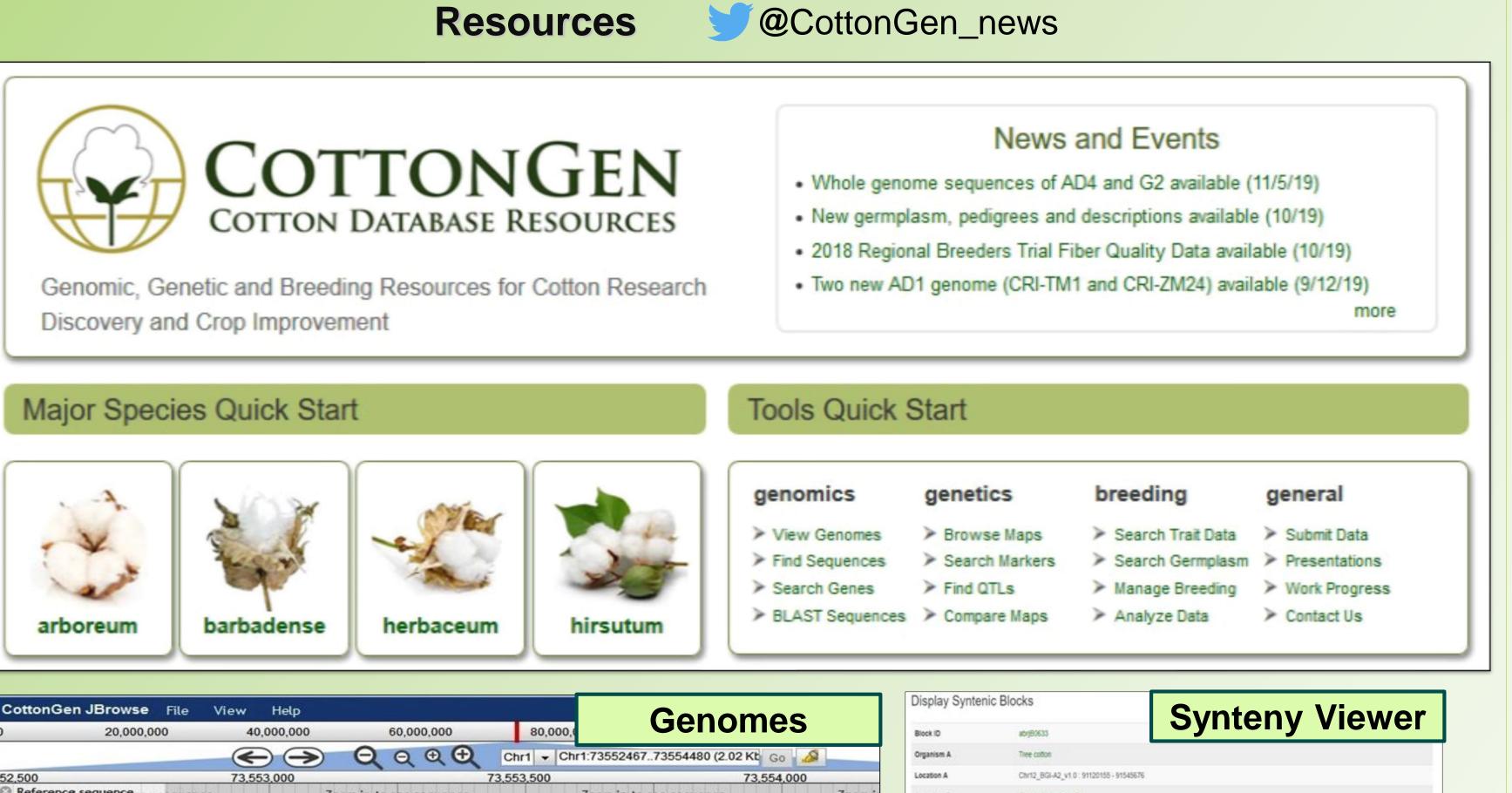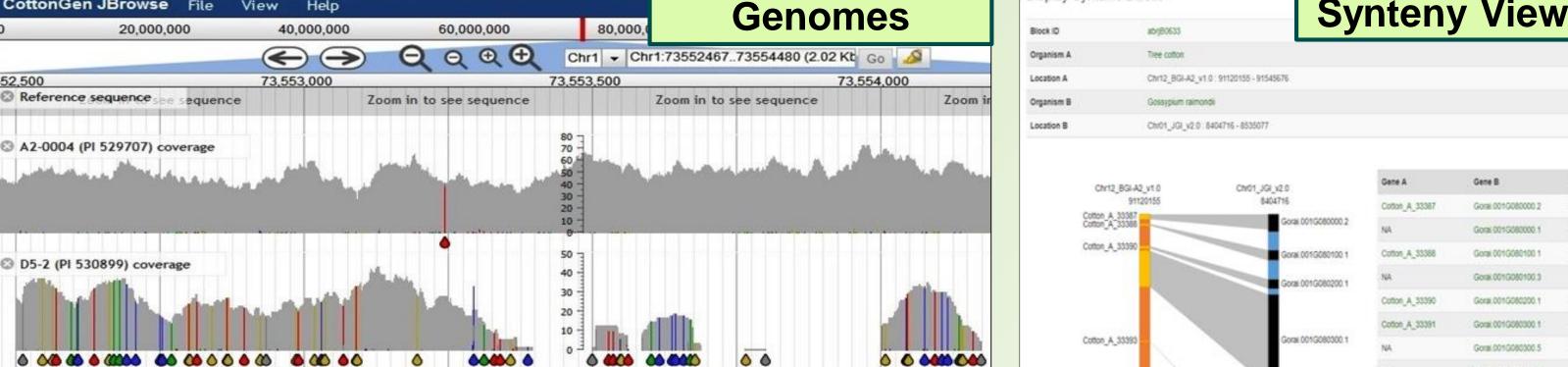
Key Features

- Easy-to-use search interfaces for retrieving data on markers, genes, transcripts, sequences, QTLs, and traits
- JBrowse and Syntenic Viewer with genome comparisons with orthologs data
- MapViewer for interactive map viewing of genetic maps
- BIMS, Breeder Data Management and Analysis System
- Assembled and annotated reference transcriptome (RefTrans) which is searchable using BLAST
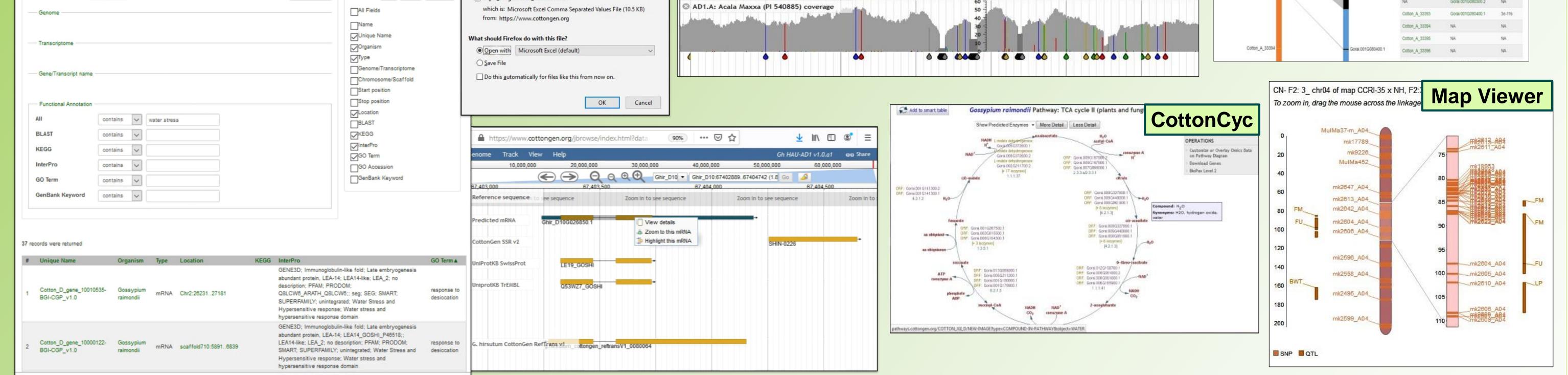- PathwayCyc with eight species (5 diploids and 3 tetraploids)

Available Data

- 20 whole genome sequences for 8 diploid and 12 tetraploid cotton
- 575,844 genetic markers and 111 genetic maps
- 5,522 QTLs, 12,337 images, and 19,563 germplasm
- 25,227,605 genotype scores and 383,431 phenotype scores

Resources
@CottonGen_news

CottonGen Cotton Database Resources
Genomic, Genetic and Breeding Resources for Cotton Research Discovery and Crop Improvement

News and Events
- Whole genome sequences of AD4 and G2 available (11/5/19)
- New germplasm, pedigrees and descriptions available (10/19)
- 2018 Regional Breeders Trial Fiber Quality Data available (10/19)
- Two new AD1 genome (CRI-TM1 and CRI-ZM24) available (9/12/19)
- more

Major Species Quick Start: arboreum, barbadense, herbaceum, hirsutum

Tools Quick Start
- genomics: View Genomes, Find Sequences, Search Genes, BLAST Sequences
- genetics: Browse Maps, Search Markers, Find QTLs, Compare Maps
- breeding: Search Trait Data, Search Germplasm, Manage Breeding, Analyze Data
- general: Submit Data, Presentations, Work Progress, Contact Us

Gene Search

Genomes

Synteny Viewer

CottonCyc

Map Viewer

Trait Search

BIMS

BRACT COLOR
GREEN (1) SUN RED (3) RED (2)

Acknowledged with Thanks