

CottonGen: Updates to the Cotton Community Database for Genomics, Genetics and Breeding Research

Jing Yu¹, Sook Jung¹, Chun-Huai Cheng¹, Taeh Lee¹, Ping Zheng¹, Katheryn Buble¹, James Crabb¹, Jodi Humann¹, Heidi Hough¹, Don Jones², Todd Campbell³, Josh Udall⁴, Dorrie Main¹

1. Washington State University, 2. Cotton Incorporated, 3. USDA-ARS, SC, 4. USDA-ARS, TX



Visit CottonGen



Join in the Mailing List

ABSTRACT

Established in 2012 through a merger of CottonDB and the Cotton Marker Database, CottonGen is the established genomics, genetics and breeding database for the cotton community. It provides a comprehensive collection of data, various analysis tools, Breeding Information Management System (BIMS), and links to external resources of interest to cotton researchers. CottonGen currently contains 53 (21 tetraploids and 33 diploids) annotated genome sequences; 5,725,571 genes/transcripts, 117 genetic maps; 665,269 markers; 7,199 QTLs; 20,187 germplasm; metabolic pathways for 13 species (AD1-AD5, A, B, D, E, F, G, K, and kirkii); 31,148,121 SNP and 14,284 SSR genotype measurements; 540,171 phenotype measurements (mainly from RBTN and NCGC projects), 45,155 images (mainly of NCGC); and synteny data for 53 genomes with links to genes, mRNA, orthologs and function. Analysis and visualization tools in CottonGen include the genome browser JBrowse, Synteny Viewer, MapViewer, CottonCyc, BLAST+, BIMS, and MegaSearch, a powerful search engine, both with recently added new features and functions. All the data are integrated within the database and can easily be queried through its various search engines. Serving as a portal for communication, CottonGen also houses all ICGI information for the community.

GENETIC TOOLS

MegaSearch

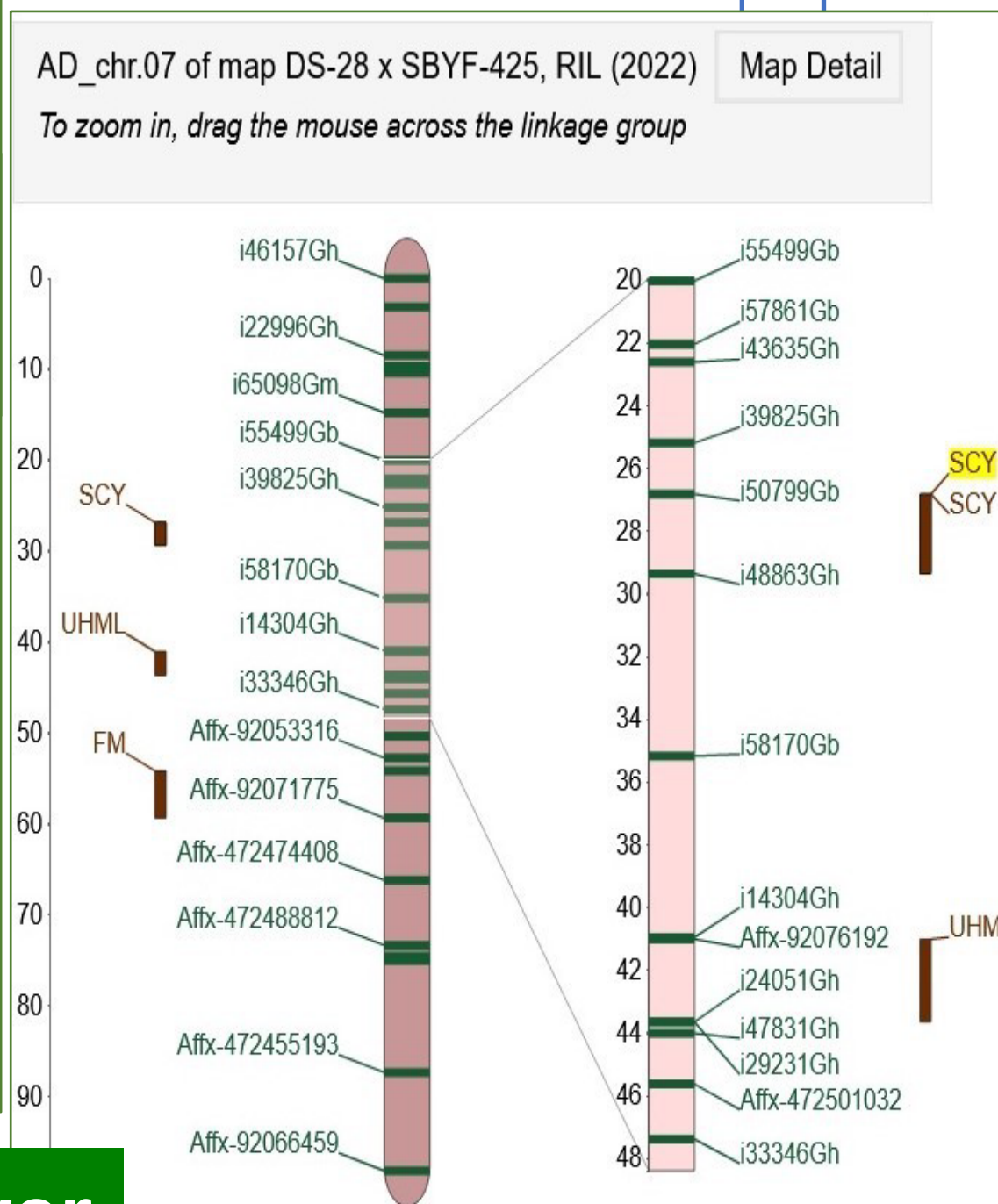
7,436 QTL. Note: actual rows in downloaded file depend on the selected fields.

Query: QTL

Downloadable Fields: QTL Name, Trait Name, Organism, Genome Group, Type, Published Symbol, LOD, Category, Map, Standardized Linkage Group

#	QTL Name	Trait Name	Genome Group	LOD	Map	Standardized Linkage Group	Neighboring Marker
1	qSCY.DS-RIL.ch07.Dha18	seed cotton yield	AD	3.577103	DS-28 x SBYF-425, RIL (2022)	AD_ch07_A1.07	TAMU_Gb379_022646
2	qSCY.DS-RIL.ch07.Dha18	seed cotton yield	AD	3.577103	DS-28 x SBYF-425, RIL (2022)	AD_ch07_A1.07	TAMU_GH_TBH061P12r297
3	qSCY.DS-RIL.ch07.Dha18-20	seed cotton yield	AD	6.78	DS-28 x SBYF-425, RIL (2022)	AD_ch07_A1.07	TAMU_Gb379_022646
4	qSCY.DS-RIL.ch07.Dha18-20	seed cotton yield	AD	6.78	DS-28 x SBYF-425, RIL (2022)	AD_ch07_A1.07	TAMU_GH_TBH061P12r297
5	qSCY.DS-RIL.ch10.Dha18	seed cotton yield	AD	3.801727	DS-28 x SBYF-425, RIL (2022)	AD_ch10_A1.10	BYU_ca_026223
6	qSCY.DS-RIL.ch10.Dha18	seed cotton yield	AD	3.801727	DS-28 x SBYF-425, RIL (2022)	AD_ch10_A1.10	USDA_CFB3365
7	qSCY.DS-RIL.ch13.Dha19	seed cotton yield	AD	3.243706	DS-28 x SBYF-425, RIL (2022)	AD_ch13_A1.13	UCD_ch_M_013103
8	qSCY.DS-RIL.ch13.Dha19	seed cotton yield	AD	3.243706	DS-28 x SBYF-425, RIL (2022)	AD_ch13_A1.13	USDA_CFB1405
9	qSCY.DS-RIL.ch19.Dha20	seed cotton yield	AD	3.62019	DS-28 x SBYF-425, RIL (2022)	AD_ch19_D1.05	Aflx-290808224
10	qSCY.DS-RIL.ch19.Dha20	seed cotton yield	AD	3.62019	DS-28 x SBYF-425, RIL (2022)	AD_ch19_D1.05	Aflx-472468315

- New MegaSearch interfaces for map, marker and QTL searches and data downloads
- MapViewer tool to view and interact with genetic maps



MapViewer

Viewing genetic map DS-28 x SBYF-425, RIL (2022)

AD_ch01 AD_ch02 AD_ch03 AD_ch04 AD_ch05 AD_ch06 AD_ch07 AD_ch08 AD_ch09 AD_ch10 AD_ch11 AD_ch12

BREEDING TOOLS

- Breeding Information Management System (BIMS) available to breeders through CottonGen
- Allows to store, manage, archive and analyze public or private breeding data
- Multiple users can manage each program and upload data
- Compatible with Android app Field Book for electronic data collection and with BrAPI
- comparable with public data

BIMS

The US Regional Breeders Testing Net

Manage Trait: MIC (unit)

Definition: MIC (micronaire) is an HVI measurement, the fineness of the sample taken from the ginned lint, measured by a Fibronaire and expressed in standard (curvilinear scale) micronaire units.

Properties: data_unit = unit

Histogram: MIC (unit)

# Data	Max	Min	Mean	STD
5465	6.17	2.83	4.84	0.40

GENOMIC TOOLS

- Genome sequences available to search using BLAST tool and mRNA and genes can be searched using MegaSearch
- JBrowse displays the genome assemblies with links to more information about gene, mRNA, marker, and transcript features
- Explore the metabolic pathways of genomes and overlay transcriptomics and metabolomics data using PathwayCyc tool
- Synteny Viewer to view genome syntenic blocks and Ortholog/Paralog MegaSearch as another way to explore the data

690,405 Marker. Note: actual rows in downloaded file depend on the selected fields.

Query: Marker Type: SNP

Marker Name: contains

SNP Array Name: Any

Genome Group: AD

Standardized Linkage Group: AD_ch07_A1.07

Chromosome/Scaffold: Any

80 records were returned

#	Marker Name	Marker Type	Map	Linkage Group	Genome	Standardized Chromosome	Landmark	Location	Trait
1	TAMU_Gb379_022646	SNP	3-79 x TAMU, F2 (2015 SNP)	ch07	Goosypium barbadense (AD2) 3-79 genome HAU_v2_a1	AD_ch07_A1.07	Gbar_A07	Gbar_A07:8930434..8930534	seed cotton yield
2	TAMU_Gb379_022646	SNP	3-79 x TAMU, F2 (2015 SNP)	ch07	Goosypium barbadense (AD2) 3-79 genome HGS_v1.1	AD_ch07_A1.07	CM018208	CM018208:1.8972503..8972603	seed cotton yield
67	TAMU_GH_TBH061P12r297	SNP	Phytogen 72 x Stoneville 474, constin (2017)	ch07	Goosypium tomentosum (AD3) genome HGS_v1.1	AD_ch07_A1.07	CM017616.1	CM017616.1:11088591..11088691	seed cotton yield
68	TAMU_GH_TBH061P12r297	SNP	Phytogen 72 x Stoneville 474, F2 (2015 SNP)	c07	Goosypium barbadense (AD2) 3-79 genome HAU_v2_a1	AD_ch07_A1.07	Gbar_A07	Gbar_A07:11276466..11276566	seed cotton yield

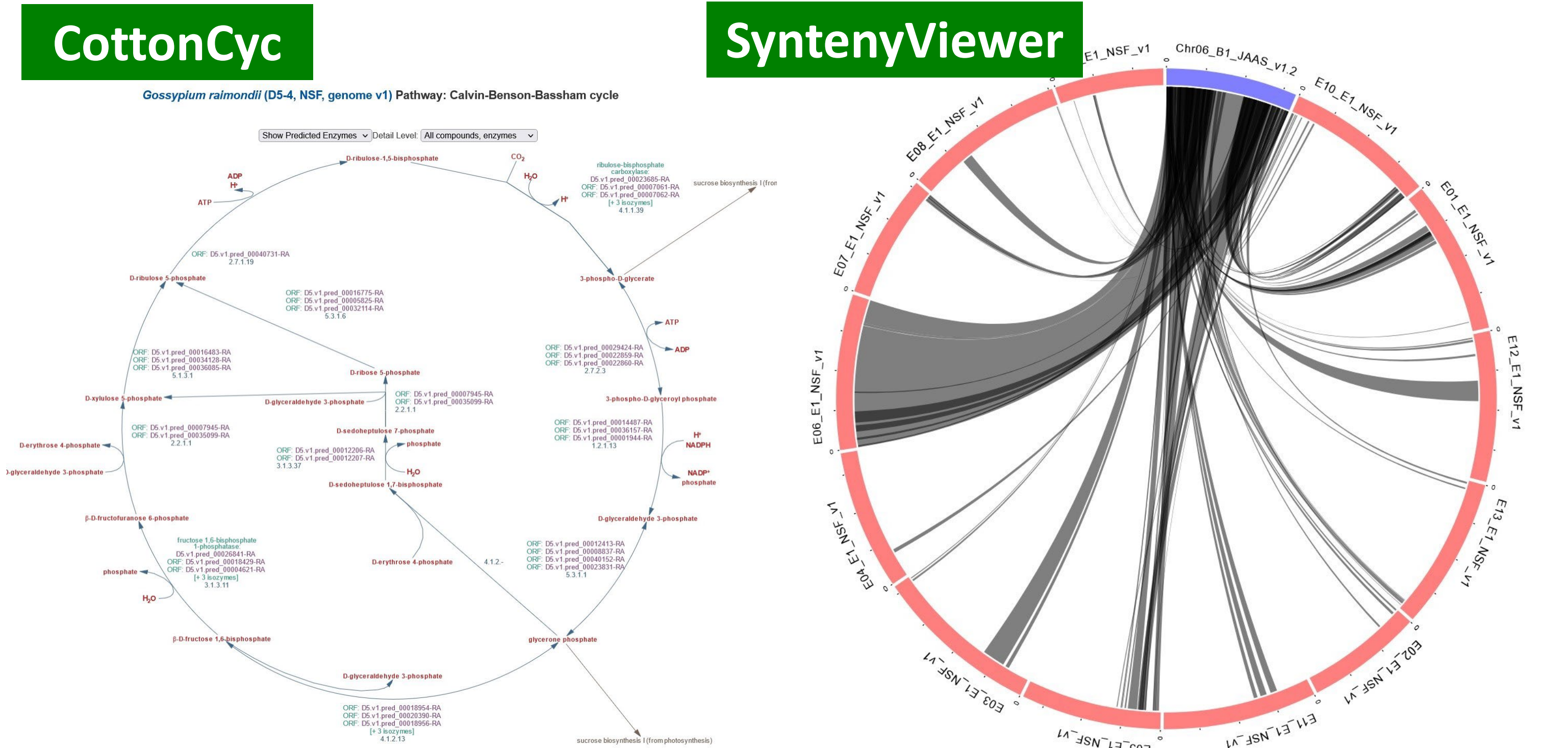
JBrowse

CottonCyc

Goosypium raimondii (DS-4, NSF, genome v1) Pathway: Calvin-Benson-Bassham cycle

Available Tracks: CottonGen RFLP/RAD, CottonGen SNP/Indel, CottonGen SSR v2, CottonGen Protein Alignments, CottonGen Transcript Alignments

SyntenyViewer



FUNDING

