

CottonGen: A Central Data Repository and Analysis Resource for the Cotton Community www.cottongen.org



Jing Yu¹, Sook Jung¹, Chun-Huai Cheng¹, Taein Lee¹, Ping Zheng¹, Katheryn Buble¹, James Crabb¹, Jodi Humann¹, Heidi Hough¹,

Don Jones², J. Todd Campbell³, Josh Udall⁴, Dorrie Main¹

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

Abstract

CottonGen is a genomics, genetics and breeding database for the cotton community. It provides a comprehensive collection of data, various analysis tools, Breeding Information Management System, and links to external resources of interest to cotton researchers. CottonGen currently contains 55 (21 tetraploids) annotated genome sequences; 5,725,571 genes/transcripts, 117 genetic maps; 690,386 markers; 7,436 QTLs; 20,177 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; 540,457 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 (the Breeding Information Management System), and MegaSearch, a powerful search engine, both with recently added new features and functions. All the data are integrated within CottonGen's search engines. In this presentation, examples of using CottonGen search engines to find marker/QTL/trait/Genome location and integration among them.

MegaSearch to find all markers that associate with seed cotton yield, have genetic map position on **Chromosome 23, and identify their genomic locations.**

Refresh Count

View CSV TS

Sequence re va

Mapped Organism

Marker Type

Linkage Group

Standardized Linkage Group

Standardized Chromosome

🔽 Map

Location

Step1. Go to CottonGen's MegaSearch, select 'Data Type' = Marker

BIMS to store, manage, archive and analyze public or private breeding data (genotype/phenotype/germplasm)





Step2. In sections of 'Genetic Position' and 'Trait', select 'Standardized' Linkage Group' and 'Trait' of interest (red arrows), then from 'Downloadable Field', select contents that you want to 'View' or 'Download'

Note: Some of the SNP positions in t	the genome are >1 wh	hen the alignment was done using flanking sequences	5.		
Marker Type	Any	\checkmark	Clear	Refresh Count	
Marker Name	contains	✓			All Fields
Browse No file selected.					Unique Name
					🗹 Marker Name
SNP Array Name	Any	~			Organism



View Search Result. Genome locations are hyperlinked to a genome browser. In example below, click on green text at red arrow, then its genome location in JBrowse view appears above

502 records were returned

	Marker Name	Marker Type	Мар	Linkage Group	Standardized Linkage Group	Genome	Standardized Chromosome	Landmark	Location	Trait
1	NAU3052	SSR	CCRI-35 x Yumian-1, RIL (2017)	chr23	AD_ch23_Dt.09	Gossypium hirsutum (AD1) 'TM-1' genome NAU-NBI_v1.1	AD_ch09_At.09	A09	A09:6605722766057399	seed cotton yield
2	NAU3052	SSR	CCRI-35 x Yumian-1, RIL (2017)	chr23	AD_ch23_Dt.09	Gossypium hirsutum (AD1) 'TM-1' genome NAU-NBI_v1.1		scaffold4321_D09	scaffold4321_D09:7202272206	seed cotton yield
v	NAU3052	SSR	CCRI-35 x Yumian-1, RIL (2017)	chr23	AD_ch23_Dt.09	Gossypium hirsutum (AD1) 'TM-1' genome UTX-JGI-Interim- release_v1.1	AD_ch09_At.09	A09	A09:7112185771122029	seed cotton yield
4	NAU3052	SSR	CCRI-35 x Yumian-1, RIL (2017)	chr23	AD_ch23_Dt.09	Gossypium hirsutum (AD1) 'TM-1' genome UTX-JGI-Interim- release_v1.1	AD_ch23_Dt.09	D09	D09:4297090442971088	seed cotton yield
5	NAU3052	SSR	CCRI-35 x Yumian-1, RIL (2017)	chr23	AD_ch23_Dt.09	Gossypium hirsutum (AD1) 'TM-1' genome UTX_v2.1	AD_ch19_Dt.05	D09	D09:4425132344251507	seed cotton yield
6	NAU3052	SSR	CCRI-35 x Yumian-1, RIL (2017)	chr23	AD_ch23_Dt.09	Gossypium hirsutum (AD1) 'TM-1' genome UTX_v2.1	AD_ch22_Dt.04	A09	A09:7403589374036065	seed cotton yield

Choose a program or a trial to view stats for each trait



Select trials and compare their trait measurements





Download Search Result. Click on 'CSV' to download data as comma-separated values table (below).

		A B	С	D	E	F	G	Н	T	J	К	L
	#	rker Name	Туре	Map	Linkage G	Standardi	Genome	Standardized Ch	Landmark	Location	Trait	
29	5 2	94 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch19_Dt.05	Dt_chr5	Dt_chr5:82961738296357	seed cotton	yield
29	6 2	95 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch09_At.09	A09	A09:7265780872657980	seed cotton	yield
29	7 2	96 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch23_Dt.09	D09	D09:4254802242548206	seed cotton	yield
29	8 2	97 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch09_At.09	Ghir_A09	Ghir_A09:7089312070893292	seed cotton	yield
29	9 2	98 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch23_Dt.09	Ghir_D09	Ghir_D09:4281687542817059	seed cotton	yield
30	0 2	99 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch09_At.09	A09	A09:6605722766057399	seed cotton	yield
30	1 3	00 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	NAU-NBI_v1.1	scaffold43	scaffold4321_D09:7202272206	seed cotton	yield
30	2 3	01 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch09_At.09	A09	A09:7112185771122029	seed cotton	yield
30	3 3	02 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch23_Dt.09	D09	D09:4297090442971088	seed cotton	yield
30	4 3	03 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch19_Dt.05	D09	D09:4425132344251507	seed cotton	yield
30	5 3	04 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch22_Dt.04	A09	A09:7403589374036065	seed cotton	yield
30	6 3	05 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	ZJU-improved_v2	A09	A09:7255168472551856	seed cotton	yield
30	7 3	06 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	ZJU-improved_v2	D09	D09:4237517542375359	seed cotton	yield
30	8 3	07 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'ZM24' genome	AD_ch09_At.09	A09	A09:6927104869271220	seed cotton	yield
30	9 3	08 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium hirsutum (AD1) 'ZM24' genome	AD_ch23_Dt.09	D09	D09:4064825440648438	seed cotton	yield
31	0 3	09 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium mustelinum (AD4) genome JGI	AD_ch09_At.09	CM017644	CM017644.1:7138017671380350	seed cotton	yield
31	1 3	10 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium mustelinum (AD4) genome JGI	AD_ch23_Dt.09	CM017657	CM017657.1:4307558743075771	seed cotton	yield
31	2 3	11 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium tomentosum (AD3) genome H0	AD_ch09_At.09	CM017618	CM017618.1:6848041268480584	seed cotton	yield
31	3 3	12 NAU3052	SSR	0-153 x sGK9708,	chr23-1	AD_ch23_	Gossypium tomentosum (AD3) genome H0	AD_ch23_Dt.09	CM017631	CM017631.1:4166466141664839	seed cotton	yield
31	4 3	13 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium barbadense (AD2) '3-79' genor	AD_ch16_Dt.07	Dt07	Dt07:1742585317426037	seed cotton	yield
31	5 3	14 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium barbadense (AD2) '3-79' genor	ne HAU_v1	scaffold23	scaffold2301:195168195352	seed cotton	yield
31	6 3	15 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium barbadense (AD2) '3-79' genor	ne HAU_v1	scaffold60	scaffold6024:333922334097	seed cotton	yield
31	73	16 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium barbadense (AD2) '3-79' genor	AD_ch23_Dt.09	Gbar_D09	Gbar_D09:4145323241453416	seed cotton	yield
31	8 3	17 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium barbadense (AD2) '3-79' genor	AD_ch23_Dt.09	CM018223	CM018223.1:4076886940769053	seed cotton	yield
31	9 3	18 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium barbadense (AD2) 'H7124' gen	ome ZJU_v1.1_a1	D09	D09:4392214043922324	seed cotton	yield
32	.0 3	19 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium darwinii (AD5) genome HGS_v2	AD_ch23_Dt.09	CM017709	CM017709.1:4225605042256234	seed cotton	yield
32	1 3	20 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD_ch23_	Gossypium hirsutum (AD1) 'TM-1' genome	AD_ch19_Dt.05	Dt_chr5	Dt_chr5:82961738296357	seed cotton	yield
32	2 3	21 NAU3052	SSR	CCRI-35 x Yumiar	chr23	AD ch23	Gossypium hirsutum (AD1) 'TM-1' genome	AD ch09 At.09	A09	A09:7265780872657980	seed cotton	vield

Search specific region of micronaire and save the result 🗏 🎮 Hello jing | Site Home | Log out 🗖 Search Accession 🛋 Search Results Save Results / Download Data Name Save Description Click 'Save' to keep the results Download Click 'Download' to download data Results : matched 101 distinct accession Descriptions Filtered By Values MIC (unit) numeric : MIN 4 - MAX 5 BWT (g) LY (kg/ha) E1 (%) MIC (unit 4.448 Acala Maxxa 1063.69343777 27.65 3.848 Acala Maxxa 7.25 802.7502730892 25.9 4.392 Acala Maxxa 1160.433973092 4.748 Acala Maxxa 1242.363246378 22.65 6.5 5.792 22.55 NM24016 994.6134004578 4.952 NM24016 6.75 1020.385069659 23 3.766 24.53 NM24016 729.4137751534 6.095 NM24016 5.316 7.75 1325.3077467662 22.95 4.2

BIMS ONLINE WORKSHOP WILL BE SCHEDULED, WE WILL SEND OUT A NOTICE SOON

Acknowledged with Thanks



seed number per boll









Field Book Managem

Data Analysis

The Cotton Research Community