



Welcome to the 3rd issue of the CottonGen newsletter in 2022. This newsletter is issued to inform users about **new or updated data and tools in CottonGen**. In addition to new and updated data, each issue will provide more information on data or tools on the **featured tool/data** section.

## New Genome Data

Data from **three whole genome assemblies**:

[G. tomentosum \(AD3\) genome HAU v1 \(Shen 2021\)](#)

[G. anomalum \(B1\) genome JAAS v1.2 a1.1 \(Xu 2022\)](#)

[G. herbaceum \(A1\) 'Wagad' genome USDA v1 \(in press\)](#)

## New Functional Analysis

**CottonGen Functional Analysis** (InterProScan and KEGG), Protein Homologies, Cottongen Marker and RefTran Alignments, and so on) **added to the below genomes**:

[G. thurberi \(D1-5\) genome CRI v1 \(Yang 2021\)](#)

[G. hirsutum \(AD1\) 'NDM8' genome HEAU v1 \(Ma 2021\)](#)

[G. barbadense \(AD2\) 'P90' genome HEAU v1 \(Ma 2021\)](#)

**Access or download genome data** from [Synteny Viewer](#), [Data Download](#), or from the individual genome pages listed at above.

**COTTONGEN** Species - Data - Search - Tools - ICGI - General - He

### Data Download

Please note, other data can be downloaded through the [link] is available through this downloads

- Whole Genomes Sequence Data (FASTA or G[...])
  - From original genome sequencing project:
  - From CottonGen: functional annotations, p[...]
  - Click to view and download data of:
    - Diploid genomes
      - Diploid genome A & D5 last update 04/21/22
      - All D genomes except D5 last update 04/21/22
      - Other diploid genomes last update 07/02/22
    - Tetraploid genomes
      - G. hirsutum (AD1): TM-1 | ZM24 | NDM8 | B713 | Bar32 last update 04/21/22
      - G. barbadense (AD2): 3-79 | Hai-7124 | Pima90 & Other Tetraploids last update 07/02/22

## Other New Data

- CottonSNP63K array used to identify 17,441 polymorphic markers in a panel of 114 diverse Pee Dee genotypes. ([Billings 2022](#))
- Fiber quality, yield, and verticillium wilt disease severity data screened on 100 early varieties and 100 modern varieties phenotype data ([Ma 2021](#))
- Yield component and fiber quality QTLs and genetic map ([Kumar 2019](#))
- Germplasm, pedigrees from USDA Cotton Collection and publications
- Structured vocabularies for some new trait and trait descriptors

**Access and download CottonSNP63K-PD Genotype data** via [Genotype Search](#) (search results on page 2)

Species - Data - Search - Tools - ICGI - General - Help - Data Curation

### Search Genotype

SNP Genotype | SSR Genotype

Search SNP Genotype is a page where users can search for the search for SSR Genotype | [Text tutorial](#) | [Email us with probl](#)

Dataset:

Species:

Germplasm Name:

SNP:

Genome:

Chr/Scaffold:  between  and  bp

Gene Model:  +/-  bp



## Results of Genotype Search

Link to download  
All results as file

Link to download  
Polymorphic results only

55733 records were returned

Download Table | Table (Polymorphic)

| # | Array ID | Marker                 | Location           | Allele | Line F | Line FJA | Line FTA | PD 2165-65 | PD 2165-77 |
|---|----------|------------------------|--------------------|--------|--------|----------|----------|------------|------------|
| 1 | i21830Gh | USDA_CFB4521           | A01:16101..16188   | G/T    | -      | -        | -        | -          | -          |
| 2 | i17382Gh | CSIRO_D5chr10_61514109 | A01:16874..17074   | A/G    | G      | G        | G        | G          | G          |
| 3 | i35736Gh | TAMU_GH_TBh007G10F34   | A01:39179..39262   | A/G    | R      | R        | R        | R          | R          |
| 4 | i66691Ga | TAMU_Arm_004996        | A01:58953..59053   | A/G    | G      | G        | G        | G          | G          |
| 5 | i66854Ga | TAMU_Arm_014454        | A01:64293..64393   | A/G    | G      | G        | G        | G          | G          |
| 6 | i20730Gh | USDA_CFB0732           | A01:125962..126051 | T/C    | C      | C        | C        | C          | C          |
| 7 | i48914Gh | TAMU_GH_TBh086D20r342  | A01:126504..126604 | T/C    | C      | C        | C        | C          | C          |
| 8 | i66587Ga | TAMU_Arm_077711        | A01:131345..131445 | A/G    | G      | G        | G        | G          | G          |
| 9 | i67050Ca | TAMU_Arm_025343        | A01:150880..150980 | T/C    | Y      | Y        | Y        | Y          | Y          |

## MegaSearch verticillium wilt QTLs

### Tripal MegaSearch

Tripal MegaSearch is a tool for downloading biological data. (Current limit per download: 1,000,000 records. 200,000 FASTA sequences.)

Video tutorial: 2020 presentation | 2021 tutorial (Gene/Transcript) | 2022 tutorial (QTL)

Select a data type to start building your own query and download data in bulk:

Data Type

Refer to the short video tutorial for more about how to use MegaSearch QTL

7,009 QTL. Note: actual rows in downloaded file depend on the selected fields.

Query

Type

Organism

Trait

Trait Category

Downloadable Fields

All Fields

QTL Name

Trait Name

Organism

Type

Published Symbol

LOD

R2

Category

Map

Join the [CottonGen Mailing List](#)  
and follow us on [Twitter](#)

Funded by:  
Cotton Inc; USDA-ARS; Bayer CS; CORTEVA;  
USDA National Research Project (NRSP10)