

CITRUS GENOME DATABASE



Resources for citrus genomics, genetics, breeding and disease research

Issue 5 | October 2022

What is CGD?

CGD is a centralized database containing genomics, genetics, and breeding data and analysis tools for citrus research. Genome sequences are available to view and search and there is also information about genetic maps, molecular markers, and QTL. If you are a breeder who needs to manage private breeding program data, access to the Breeding Information Management System (BIMS) can be requested. Visit us at www.citrusgenomedb.org to see everything that is available. Each issue of the newsletter will focus on a different type of data and what features are available.

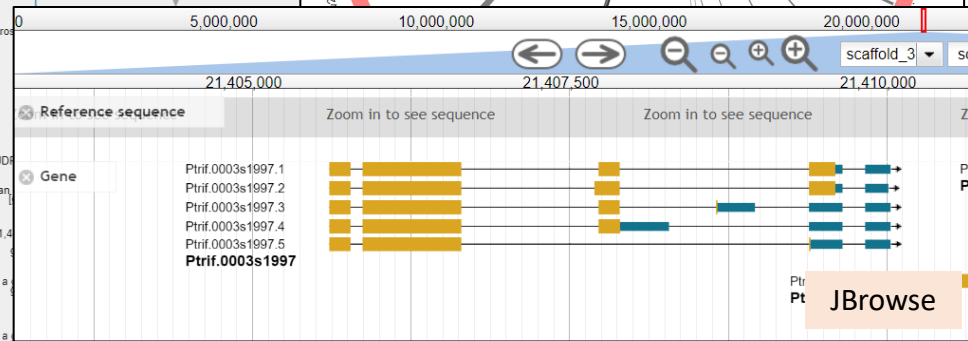
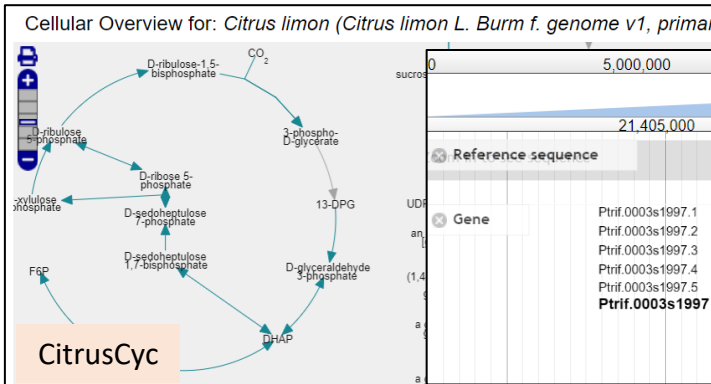
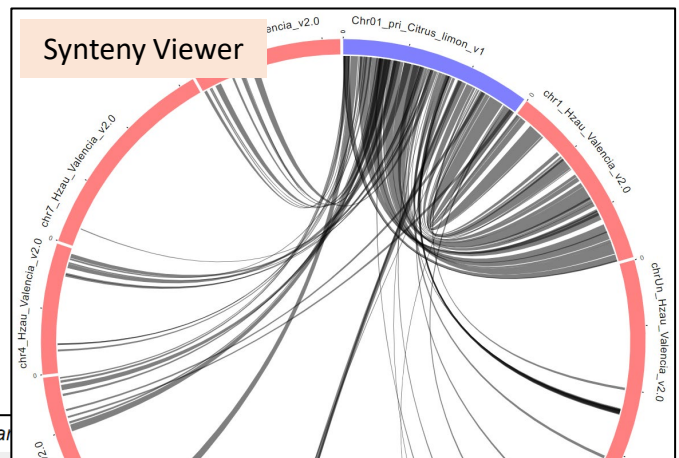
More CGD training and tools

Thanks to a new award from the USDA-SCRI, CGD and the other MainLab specialty crop databases will continue to grow to provide the tools, data, and training that breeders and researchers need. You can read more about the award here: [Enabling Genomics-Assisted Specialty Crop Breeding and Research Through Advanced Database Resources](#).

The continued funding will provide new tools and resources for big data, data management, and staff for in-person training with CGD users. We will be announcing all the new developments via the CGD mailing list, so please subscribe if you have not done so already.

Genomic resources on CGD

There are genomes from 11 *Citrus sp.* and related species on CGD. These genomes are available to view in [JBrowse](#), sequences can be searched in [BLAST](#), and the mRNA and gene annotations are available to search in the database via [MegaSearch](#). These genomes are also available in [CitrusCyc](#) and [Synteny Viewer](#). More genomes will be added soon, so keep watch for those.



Ortholog/Paralog MegaSearch

We have made a new MegaSearch for the data that is stored during the [Synteny Viewer](#) analysis. Now you can search that data without going through the Synteny Viewer.

The new [Ortholog/Paralog Search](#) interface is simple to use. First select the primary genome and chromosome/scaffold of interest. Or provide a single gene/transcript ID or a file of them.

Once parameters are set, you can preview the number of results by clicking “Refresh Count”. And then you can configure what data will be returned and available for download.

When viewing the results on the website, there are blue hyperlinks that open details about the associated transcripts including the annotated terms, homology to SwissProt, InterPro results, and mRNA and protein sequences.

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

1,729 Ortholog/Paralog. *Note: actual rows in downloaded file depend on the selected fields.*

Options to view and download data

Query

Genome

Chromosome/Scaffold

Start

Stop

Gene/Transcript Name

No file chosen

Compare to
Citrus clementina genome v1.0 (JGI)
Citrus limon L. Burm f. genome v1.0 - Alternative
Citrus limon L. Burm f. genome v1.0 - Primary

Chromosome/Scaffold

Downloadable Fields

All Fields

Genome1

Chromosome1

Location1

Ortholog1

Genome2

Chromosome2

Location2

Ortholog2

Associated Gene

1802 records were returned

#	Genome1	Chromosome1	Location1	Ortholog1	Genome2	Chromosome2	Location2	Ortholog2
1	Citrus limon L. Burm f. genome v1.0 - Primary	Chr01_pri	Chr01_pri:10004596..10008589	CL1G001587011.t1_pri	Citrus sinensis genome v2.0 (HZAU)	chr1	chr1:15922673..15929983	Cs1g12880.1
2	Citrus limon L. Burm f. genome v1.0 - Primary	Chr01_pri	Chr01_pri:1000710..1002829	CL1G000192011.t1_pri	Citrus sinensis genome v2.0 (HZAU)	chr1	chr1:27829956..27832843	Cs1g25500.1

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

Funded by: USDA National Research Project (NRSP10), USDA-NIFA Specialty Crop Research Initiative, NSF Plant Genome Research Program