

CITRUS GENOME DATABASE



Resources for citrus genomics, genetics, breeding and disease research

Issue 2 | January 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.

Genome Resources on CGD

There are multiple genomes available on CGD. For all the genomes, you can view the genome sequence and annotations in [JBrowse](#); search the scaffold, transcripts, and protein sequences using the [BLAST tool](#); view the different biochemical pathways using [CitrusCyc](#); and view the synteny between the genome sequences with the [Synteny Viewer](#) tool.

Some of the available citrus genomes and associated metrics are listed on the table below. CGD also has some *Ca. Liberibacter* sp. genomes available within the database, JBrowse, and BLAST.

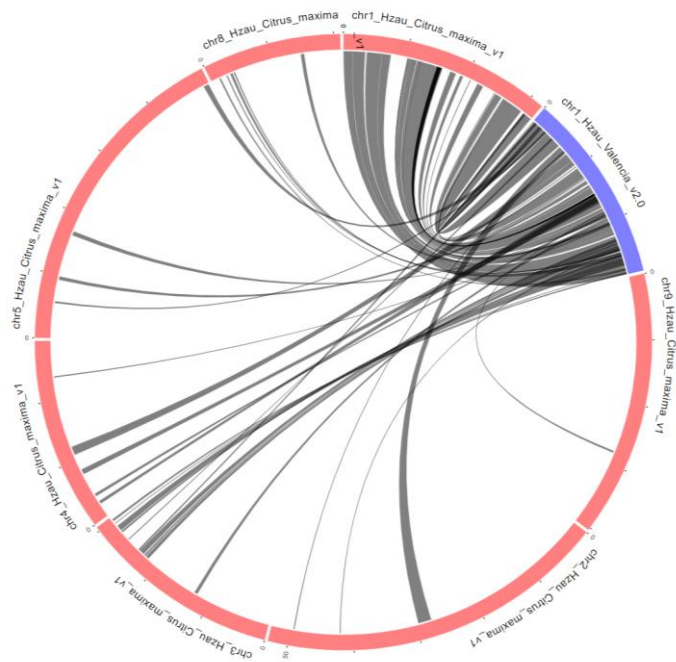
Genome	Number of scaffolds /contigs	Total size (Mb)	Number of transcripts	Number of genes
C. clementina v1.0	1,398	301	33,929	24,533
C. sinensis v1.1	12,574	319	46,147	25,379
C. sinensis v2.0	10	328	44,275	29,655
C. maxima v1.0	9	302	42,886	30,123
C. ichangensis v1.0	14,915	357	43,103	32,067
C. medica v1.0	32,731	405	47,506	32,579
C. reticulata v1.0	42,714	334	42,676	28,833
C. limon v1.0 – Primary	811	313	35,020	32,894
Poncirus trifoliata v1.3.1	152	265	33,229	25,537
Fortunella hindsii v1.0	1,226	374	52,686	32,257
Atalantia buxifolia v1.0	25,600	316	65,499	28,412

Genome Synteny Viewer

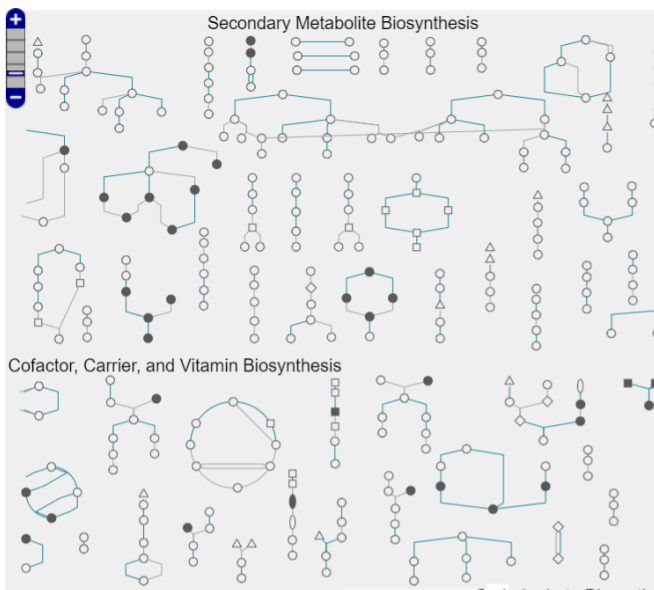
One of the additional analyses that CGD does with provided genome data is a synteny (gene order) comparison between the genomes. The analysis is conducted with [MCScanX](#) and displayed using the [Tripal Syntenic Viewer](#) module.

Using the [simple web interface](#), select the first genome and a chromosome or scaffold, and then select one or more genomes to compare against. The pairwise comparison data is displayed in a circos plot and syntenic blocks can be selected and viewed in a more detailed, linear side-by-side view.

In the detailed view, there is also an accompanying table that lists the corresponding mRNAs and has hyperlinks to the corresponding mRNA feature page on CGD. Each mRNA feature page has the associated sequences and details about homology to the Swiss-Prot protein and InterPro protein family databases.



Syntenic plot of Chromosome 1 from *C. sinensis* v2.0 compared to the *C. maxima* v1 genome.



Section of the Cellular Overview diagram from the primary haplotype of *C. limon* genome.

CitrusCyc

Genomes added to CGD are also analyzed with [Pathway Tools](#) to identify metabolic pathways. The data is then displayed with an embedded instance of the software on CGD under the [CitrusCyc tool](#).

Users can also overlay data from transcriptomics or metabolomics experiments onto the cellular overview graphic. The pathways that have corresponding data are color coded to show over- or under-expression.

CitrusCyc accounts can also be requested to allow users to upload, sort, and save data using the SmartTables feature. For more details on how to use the CitrusCyc tool, watch the video on the MainLab Bioinformatics [YouTube channel](#).

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

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