



PULSE CROP DATABASE

Genomic, Genetic, and Breeding Resources
for Pulse Crop Improvement

Issue 2 | January 2022

What is PCD?

PCD is a centralized database containing genomics, genetics, and breeding data and analysis tools for pulse crops. 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.pulsedb.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 PCD

There are multiple genomes available on PCD. 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 [PathwayCyc](#); and view the synteny between the genome sequences with the [Synteny Viewer](#) tool.

Some of the available pulse genomes and associated metrics are listed on the table below. More genomes will be added to PCD over the next couple months, so stay tuned for more!

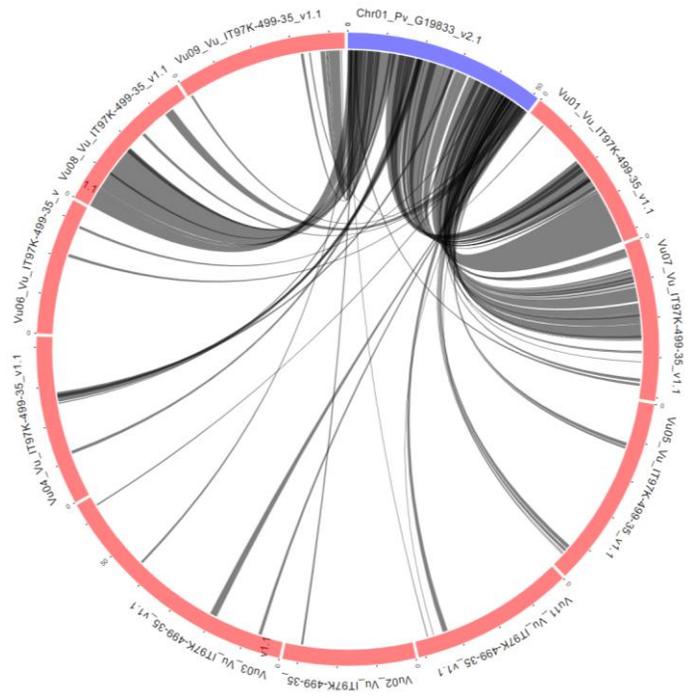
Genome	Number of scaffolds /contigs	Total size (Mb)	Number of transcripts	Number of genes
Vigna angularis cv. Jingnong6 v1.1	25,426	450	34,172	34,172
Cicer arietinum cv. CDC Frontier v1.0	7,163	532	28,269	28,269
Cicer arietinum cv. ICC 4958 v3.0	30,401	512	31,757	31,358
Cicer reticulatum L. cv. PI489777 v2.0	3,657	417	27,134	26,474
Phaseolus vulgaris cv. G19833 v2.1	478	537	36,995	27,433
Vigna unguiculata L. Walp cv. IT97K-499-35 v1.1	682	519	42,287	29,773
Pisum sativum L. cv. Cameor v1a	24,623	3920	57,835	44,756
Cajanus cajan acc. Asha v1.0	137,542	606	48,331	50,122
Lupinus albus cv. AMIGA v1.0	89	451	38,258	38,258
Lupinus angustifolius L. cv. Tanjil v1.0	14,378	609	33,083	33,074
Phaseolus acutifolius Frijol Bayo v1.0	155	513	50,635	27,538

Genome Synteny Viewer

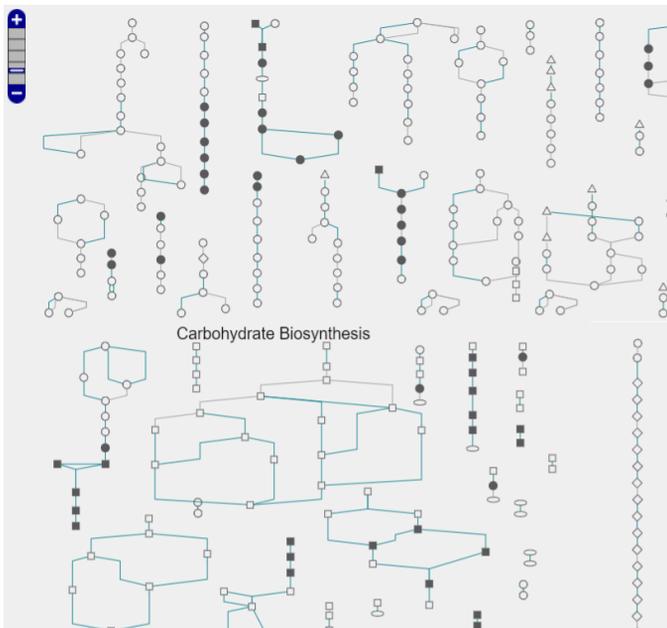
One of the additional analyses that PCD does with provided genome data is a synteny (gene order) comparison between the genomes. The analysis is conducted with [MCScanX](#) and displayed using the [Tripal Syntenic Viewer](#) module.

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 PCD. 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 *P. vulgaris* v2.1 compared to the *V. unguiculata* v1.1 genome.



Section of the Cellular Overview diagram from the *L. albus* genome.

PathwayCyc

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

PathwayCyc 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 PathwayCyc tool, watch the video on the MainLab Bioinformatics [YouTube channel](#).

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

Funded by: USDA National Research Project (NRSP10), USA Dry Pea & Lentil Council, Northern Pulse Growers Association