

GENOME DATABASE FOR ROSACEAE



Resources for Rosaceae Research Discovery and Crop Improvement

October 2021

Welcome to the first issue of the GDR newsletter in 2021. This newsletter is issued to inform users about **new or updated data and tools in GDR**. 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 Assemblies

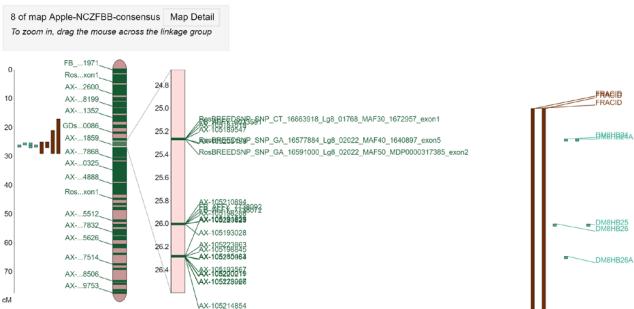
Data from **20 whole genome assemblies** have been added to GDR. It includes data from **diploid strawberry** (*Fragaria vesca* CFRA 2339), **apple** (*Malus x domestica* Gala haploid and **diploid**, *Malus sieversii* haploid and **diploid**, and *Malus sylvestris* haploid and **diploid**), **apricot** (*Prunus armeniaca* Marouch n14, *Prunus armeniaca* Stella, *Prunus mandshurica* CH264 4, *Prunus sibirica* CH320 5, *Prunus armeniaca* Sungold, *Prunus armeniaca* Longwangmao, and *Prunus sibirica* F106), **peach** (*Prunus persica* Chinese Cling), **plum** (*Prunus salicina* Sanyueli FAAS), **pear** (*Pyrus pyrifolia*), **rose** (*Rosa rugosa*), and **other Rosaceae species** used for medicinal purposes (*Rubus chingii* Hu, and *Gillenia trifoliata*). Access data from **Gene Search**, **MegaSearch**, **Synteny Viewer**, **JBrowse**, **BLAST** as well as the individual genome pages linked above.

GDR Gene Database Updated

This update included additional 43,692 genes and 35,284 mRNAs in NCBI Rosaceae gene and mRNA sequences in GDR and 42,231 genes in GDR Gene Database. **NCBI Rosaceae gene and mRNA** is gene-coding sequence data from NCBI nr database excluding predicted genes from the whole genome assembly. Distinct gene symbols for each species are stored as **GDR Gene Database** and linked to multiple gene/mRNA sequences from NCBI nr. Predicted genes from whole genome assemblies are also linked when users submit those data. See an **example page of Malus x domestica ABP gene page** with associated mRNA from nr and predicted genes from whole genome assemblies. **Please submit your gene data!**

New QTL, SNP array, Genotype, and Haplotype data

New **apple 50K SNP array data** as well as **haplotype** and **QTL** data from **Rymenants et al. (2021)** are available. View **SNP**, **QTL**, and **haplotype block** in **Mapviewer**, **search haplotype block**, and **SNP** data.



MapViewer with SNPs, haplotype blocks and QTLs.

New data also includes **QTL and haplotype data for peach** from Rawandoozi et al. (2020 and 2021), **apple SNP genotype data** (Howard et al. 2021), and **apple SSR genotype data** (National Laboratory for Genetic Resources Preservation, USDA). **IRSC 1K SNP array for pear**, part of 9K apple/pear SNP array, is searchable separately from apple SNPs.

Haplotype Block									
Project: Rosaceae-phylogenetic, Iris-Rawandoozi-Haplotype-2021									
Genome: Prunus persica Whole Genome Assembly v2.0 & Annotation v2.1 (v2.0.a1)									
Map: RosBREED-4K-2019									
Marker	SSID	Genome location	Map position	h4	h4	Galaxy	TX2B136	TXW1490_1	VICTOR
snr_1_46757382	Pp014305300		1:172.23	A A	B B	B B	A B	A B	A B
SNP_GA_135737	Pp0143109690		1:172.44	B B	B B	B B	B A	B A	B B
SNP_GA_135737	Pp0143109690		1:172.35	A A	A A	A A	B A	A B	A A
SNP_GA_134730	Pp0144576596		1:174.31	B B	B B	B B	B B	B B	B B
SNP_GA_133408	Pp0144068867		1:176.28	B B	B B	B B	B B	B B	B B
SNP_GA_132001	Pp0144350295		1:177.4	B B	B B	B B	B B	B B	B B
SNP_GA_132047	Pp0144909698		1:179.62	B B	B B	B B	A B	A A	A B
SNP_GA_131988	Pp0144913729		1:179.65	B B	B B	B B	B B	A A	A B
SNP_GA_129512	Pp0145448596		1:181.79	A A	A A	A A	B A	A A	A A
SNP_GA_128625	Pp0145581205		1:182.32	A A	A A	A A	B A	A B	B A
SNP_GA_128603	Pp0145586061		1:182.34	B B	A A	A A	A A	B B	B A

Haplotype block search result page

GENOME DATABASE FOR ROSACEAE

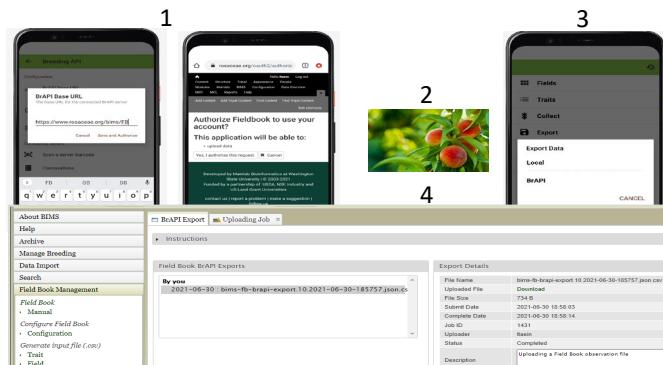


Resources for Rosaceae Research Discovery and Crop Improvement

September 2021

BIMS is made BrAPI compliant

The Breeding Information Management System (BIMS) in GDR is made BrAPI compliant. BrAPI is a web service API specification for communicating plant breeding data. With this new functionality, BIMS users can import data in Field Book app to BIMS just by clicking buttons!



1. Connect to BIMS from FieldBook App and receive an input file through BrAPI 2. Collect phenotype data. 3. Export phenotype data to BIMS. 4. View data in BIMS!

A manuscript on BIMS has been published ([Jung et al. 2021](#)). Visit BIMS website [breedwithbims.org](#), join the [BIMS Mailing List](#) and follow us on [Twitter](#) for more information!

More..

- Updated [DNA Testing Handbook for Strawberry available](#)
- [Dynamic Rosaceae Trait Ontology page available](#)
- Genes in *Prunus persica* v1.0 linked to gene family tree data in [PhyloGenes](#). Check out an [example gene page](#).
- Video tutorials from GDR workshop at RGC10 available. Go to [the manual page](#) or our [YouTube channel](#).

Retrieve Sequences around SNPs and Genes

Sequence retrieval with user-defined length of upstream and downstream available in Marker Search and Gene Search! Users can filter the list of genes and markers or load a file with gene names or marker names before retrieving sequences around them.

SNP Marker Search

Marker Search | SNP Marker Search | Search Nearby Markers

Search for SNP markers in GDR. Some of the SNP positions in the genome are >1 when the alignment was done using flanking sequences. Sequences including flanking sequences can be downloaded by opening Sequence retrieval and entering the number of nucleotides in the Upstream and Downstream boxes. | Email us with problems and suggestions

SNP Name: contains Choose File | No file chosen Name *contains uploaded words (limit: 1000 lines)

Array Name: IRSC 16K SNP array for peach

Genome: Prunus persica Whole Genome Assembly v2.0 & Annotation v2.1 (v2.0.a1)

Chr/Scaffold: Pp04 between 10582000 and 10582100 bp

Sequence retrieval

Warning: This may take hours if too many sequences are being downloaded! Please do not start two sequence retrieval jobs at the same time. To create a job, specify upstream and downstream bases to include with the sequences. The job won't be started until the Download 'Sequence Retrieval' link is clicked. (Limit for job creation: 50000 records)

Upstream bases: 10

Downstream bases: 10

Search | Reset | Download | Table | Sequence Retrieval

#	Name	SNP Array Name	SNP Array ID	dbSNP ID	Allele	Location	Flanking Sequence
1	SNP_JGA_409901	IRSC 16K SNP array for peach	SNP_JGA_409901		A/G	Pp04:10582092..10582092	AGAATGAAATTGTTGAGCTTGACACTGAAGACTCGGAAACAAAAACCGA

SNP search page with sequence retrieval option

MainLab Bioinformatics
39 subscribers [SUBSCRIBE](#)

HOME VIDEOS PLAYLISTS CHANNELS DISCUSSION ABOUT

Uploads PLAY ALL

Using the GDR BIMS for Efficient Management and...	Using the Genome Database for Genetics...	BIMS ASHS 2021	GDV Webinar: Breeding Data and Tools	GDV Webinar: Genomics Data and Tools	GDV Webinar: Genetics Data and Tools
2 views • 2 weeks ago	2 views • 3 weeks ago	4 views • 4 weeks ago	8 views • 1 month ago	6 views • 1 month ago	6 views • 1 month ago
View video	View video	View video	View video	View video	View video

GDV Webinar: Introduction 4 views • 1 month ago
How to Use BIMS (Breeding Information Management) 38 views • 4 months ago
How to Use PathwayCyc in GDR 13 views • 8 months ago
How to Use MapViewer in GDR 42 views • 8 months ago
How to Use Synteny Viewer in GDR 104 views • 8 months ago
How to Use MegaSearch in GDR 104 views • 8 months ago

BIMS webinar - May 21, 2018 25 views • 3 years ago
CSFL_MapViewer 105 views • 4 years ago
CSFL_SequenceSearch 70 views • 4 years ago
CSFL_QTLSearch 27 views • 4 years ago
CSFL_PublicationSearch 16 views • 4 years ago
CSFL_MarkerSearch 51 views • 4 years ago

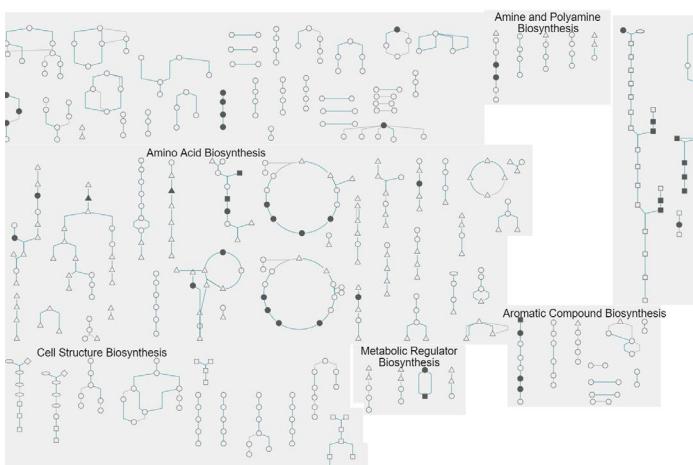
YouTube channel with video tutorials

Genome Synteny Viewer

One of the additional analyses that GDR 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.

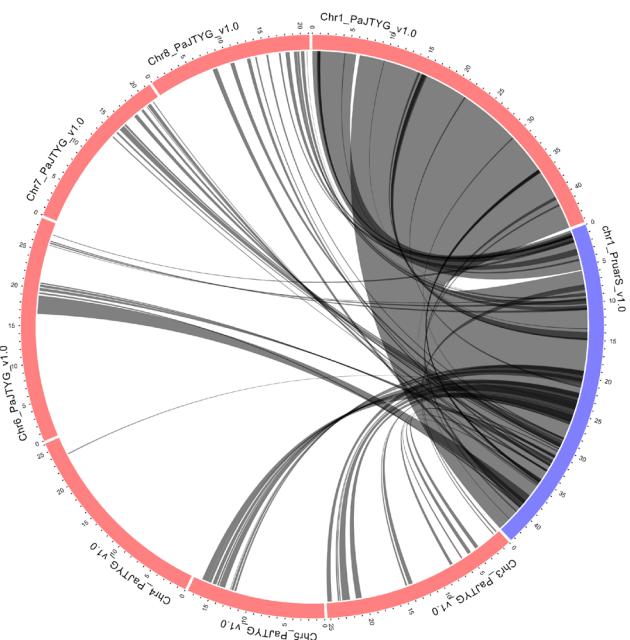
Using the [simple web interface](#), select the first genome and a chromosome or scaffold, and then select one or more genomes to compare against. The pairwise comparison data is displayed in a circus plot and syntenic blocks can be selected and viewed in a more detailed, linear side-by-side view.

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 GDR. Each mRNA feature page has the associated sequences and details about homology to the Swiss-Prot protein and InterPro protein family databases.



Section of the Cellular Overview diagram from the *Fragaria x ananassa* genome.

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



Synteny plot of Chromosome 1 from *Prunus armeniaca* cv.*Stella* compared to the *Prunus armeniaca* *Sungold* genome.

PathwayCyc

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

Funded by:
USDA National Research Project
(NRSP10)