

GENOME DATABASE FOR ROSACEAE



Resources for Rosaceae Research Discovery and Crop Improvement

January 2022

Welcome to the first issue of the GDR newsletter in 2022. 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 detailed information in the specific **featured tool/data** section.

New Genome Assemblies

Data from **12 whole genome assemblies** have been added to GDR this past quarter. Access data from [Gene Search](#), [MegaSearch](#), [Synteny Viewer](#), [JBrowse](#), [BLAST](#) as well as the individual genome pages linked below.

Four wild relatives of peach

- [Prunus mira Genome v2.0](#)
- [Prunus kansuensis Genome v2.0](#)
- [Prunus ferganensis Genome v2.0](#)
- [Prunus davidiana Genome v2.0](#)

Fragaria x ananassa Royal Royce

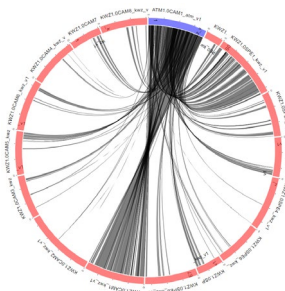
- [Fragaria x ananassa Royal Royce Genome v1.0](#)
Assembly & Annotation available from genome page, JBrowse, and BLAST under Ft. Lauderdale agreement.

Five diploid strawberry species (Qiao et al. 2021)

- [Fragaria daltoniana YNU Genome v1.0](#)
- [Fragaria pentaphylla YNU Genome v1.0](#)
- [Fragaria mandschurica YNU Genome v1.0](#)
- [Fragaria nilgerrensis YNU Genome v1.0](#)
- [Fragaria viridis YNU Genome v1.0](#)

Two flowering cherry species

- [Cerasus x kanzakura cv. 'Kawazu-zakura' v1.0](#)
- [Cerasus x kanzakura cv. 'Atami-zakura' v1.0](#)



Synteny between the chromosome 1 of the two flowering cherry genomes

BIMS manual updated

[BIMS \(Breeding Information Management System\) manual](#) has been updated. It now has extensive information on [Entering Data into Templates](#).

Submitted by user on Thu, 12/30/2021 - 12:12

There are two templates that can be used to enter phenotype data. Two templates are the same except that descriptors are entered as column headings in phenotype_bims and as data in the trait column in phenotype_long_form_bims. As a result, there will be one row per each sample for phenotyping in phenotype_bims and there will be multiple rows per each sample for phenotyping in phenotype_long_form_bims as shown in the diagram below. As shown below, the descriptors in phenotype_bims template should have # prefix.

Phenotype_name	Accession	Species	Phenotype_ID	Primary_order	Secondary_order	Phenotype_name	Accession	Species	Phenotype_ID	Primary_order	Secondary_order
1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	10	10
11	11	11	11	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	14	14
15	15	15	15	15	15	15	15	15	15	15	15
16	16	16	16	16	16	16	16	16	16	16	16
17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19	19
20	20	20	20	20	20	20	20	20	20	20	20
21	21	21	21	21	21	21	21	21	21	21	21
22	22	22	22	22	22	22	22	22	22	22	22
23	23	23	23	23	23	23	23	23	23	23	23
24	24	24	24	24	24	24	24	24	24	24	24
25	25	25	25	25	25	25	25	25	25	25	25
26	26	26	26	26	26	26	26	26	26	26	26
27	27	27	27	27	27	27	27	27	27	27	27
28	28	28	28	28	28	28	28	28	28	28	28
29	29	29	29	29	29	29	29	29	29	29	29
30	30	30	30	30	30	30	30	30	30	30	30
31	31	31	31	31	31	31	31	31	31	31	31
32	32	32	32	32	32	32	32	32	32	32	32
33	33	33	33	33	33	33	33	33	33	33	33
34	34	34	34	34	34	34	34	34	34	34	34
35	35	35	35	35	35	35	35	35	35	35	35
36	36	36	36	36	36	36	36	36	36	36	36
37	37	37	37	37	37	37	37	37	37	37	37
38	38	38	38	38	38	38	38	38	38	38	38
39	39	39	39	39	39	39	39	39	39	39	39
40	40	40	40	40	40	40	40	40	40	40	40
41	41	41	41	41	41	41	41	41	41	41	41
42	42	42	42	42	42	42	42	42	42	42	42
43	43	43	43	43	43	43	43	43	43	43	43
44	44	44	44	44	44	44	44	44	44	44	44
45	45	45	45	45	45	45	45	45	45	45	45
46	46	46	46	46	46	46	46	46	46	46	46
47	47	47	47	47	47	47	47	47	47	47	47
48	48	48	48	48	48	48	48	48	48	48	48
49	49	49	49	49	49	49	49	49	49	49	49
50	50	50	50	50	50	50	50	50	50	50	50

As explained in the Creating a New Breeding Program section, the names of the four columns, accession, unique_id, primary_order, and secondary_order, can be changed and the downloaded template will reflect the change. The unique_id represents a unique phenotyping sample and it should be unique within the dataset name. The columns primary_order and secondary_order are for the plot design such as plot and row. If those are not relevant for the program, breeders can just add the same number for each row as shown in the figure above. If there are any data for specific samples, such as specific sample treatment, users can specify it in the phenotype_bims unit as it is a column heading with ## as a prefix.

Below are descriptions for each of the columns in phenotype_bims. Columns with * are required.

- * dataset_name: name of the phenotyping dataset. It should match a 'dataset_name' column entry in the dataset_bims.
- * accession_ID: ID of the accession that has been phenotyped. It should match an 'accession_bims' or 'progeny_name' column entry in progeny_bims.

More tutorial videos are available

[How to use MegaSearch](#) and [How to use Synteny Viewer](#) short tutorials were made available this past quarter. Go to [the manual page](#) or our [YouTube channel](#).

YouTube channel with video tutorials

GENOME DATABASE FOR ROSACEAE



Resources for Rosaceae Research Discovery and Crop Improvement

January 2022

Featured Data and Tools

SNP Genotype data

Did you know? The volume of SNP genotype data housed in GDR **increased more than ten times during the last two years** (from 11 million to 131 million data points). We also introduced haplotype data a little over two years ago. The majority of the data have been produced through the efforts of the RosBREED project ([lezzoni et al. 2020](#)). Below are currently available datasets. [Search SNP genotype](#), [view the list of the project](#), or click the individual project pages.

Apple

- [Apple SNP genotype Howard 2021](#)
- [RosBREED Apple CRS SNP Genotype 2020](#)
- [Dessert and cider core collections 8k SNP](#)

Apricot

- [Apricot-SNP genotype-Dondini-2018](#)

Peach

- [RosBREED Peach CRS SNP Genotype 2020](#)
- [Peach FB SNP genotyping 2015](#)

Pear

- [Pyrus NCRG collection SNP genotyping data](#)

Strawberry

- [RB CRS Strawberry SNP genotype 2019](#)

Sweet Cherry

- [RosBREED Sweet Cherry CRS SNP Genotype 2020](#)
- [Sweet Cherry Wiersma 2018](#)
- [Sweet cherry RosCOS SNP 2011](#)

Tart Cherry

- [RB CRS Tart Cherry SNP genotype 2019](#)

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

Search Genotype

SNP Genotype | SSR Genotype

Search SNP Genotype is a page where users can search for the SNP genotype dataset based on the germplasm and SNP markers used in the dataset. Click the next tab to search for SSR Genotype. To search for SNP genotype data only for cultivars and breeding selections please visit the 'Search Genotyping Data' page in the Breeders Toolbox. Click the question mark next to 'Dataset' to view the details of the dataset. | [Text tutorial](#) | [Email us with problems and suggestions](#)

Dataset: RosBREED_Apple_CRS_SNP_Genotype_2020

Species: Any, Malus floribunda, Malus hybrid, Malus x domestica

Germplasm Name: Winter Banana, Worcester Pearmain, Yellow Newtown, Yellow Transparent

SNP: contains

Genome: Malus x domestica GDH13 v1.1 Whole Genome Assembly & Annotation

Chr/Scaffold: Chr02 between 1 and 1000000 bp

Gene Model: +/- bp

Search | Reset

You can search SNP data using **gene model of interest** as well as SNP name, location, germplasm, and dataset name

20 records were returned Download Table | Table (Polymorphic)

#	Marker	Location	Allele	Winter Banana	Worcester Pearmain	Yellow Newtown
1	RosBREEDSNP_SNP_TC_1219423_Lg2_00348_MAF20_1649423_exon1	Chr02:246562..246562	T/C	A/A	A/B	-/-
2	RosBREEDSNP_SNP_TC_1229074_Lg2_01134_MAF10_1625597_exon3	Chr02:257110..257110	T/C	A/A	A/B	A/B
3	RosBREEDSNP_SNP_CT_1237816_Lg2_01134_MAF20_1625352_exon1	Chr02:265853..265853	T/C	B/B	B/B	A/B
4	RosBREEDSNP_SNP_AG_1241012_Lg2_01134_MAF30_215646_exon1	Chr02:269036..269036	T/C	A/A	A/A	A/B
5	RosBREEDSNP_SNP_GA_1260072_Lg2_01134_MAF40_1677893_exon1	Chr02:285752..285752	R	A/B	A/B	A/A
6	RosBREEDSNP_SNP_TG_1270705_Lg2_01134_MAF10_484885_exon2	Chr02:295747..295747	T/G	A/A	A/A	A/A
7	RosBREEDSNP_SNP_CT_1272762_Lg2_01134_MAF50_617054_exon1	Chr02:297798..297798	T/C	A/A	A/A	A/B
8	GD9rg01134	Chr02:298306..298406	T/G	B/B	B/B	A/B
9	RosBREEDSNP_SNP_GA_1278998_Lg2_01134_MAF30_MDP0000233214_exon4	Chr02:304024..304024	A/G	B/B	B/B	A/B
10	RosBREEDSNP_SNP_GA_1285091_Lg2_01134_MAF50_864622_exon1	Chr02:310112..310112	A/G	-/-	A/A	A/B
11	RosBREEDSNP_SNP_CT_8352817_Lg15_226490_226490_exon1	Chr02:873327..873327	T/C	A/A	A/A	A/B
12	snpCO903605	Chr02:715208..715308	A/G	B/B	B/B	B/B
13	RosBREEDSNP_SNP_AG_1969486_Lg2_01735_MAF20_477197_exon1	Chr02:821789..821789	A/G	A/A	A/A	A/A
14	RosBREEDSNP_SNP_CT_1879153_Lg2_01735_MAF40_554134_exon1	Chr02:829353..829353	T/G	A/B	A/B	B/B
15	RosBREEDSNP_SNP_TC_1879153_Lg2_01735_MAF30_MDP0000516949_exon1	Chr02:831453..831453	Y	A/A	A/B	B/B
16	RosBREEDSNP_SNP_TC_1910740_Lg2_01735_MAF10_458777_exon1	Chr02:850465..850465	T/C	A/B	A/B	A/A
17	RosBREEDSNP_SNP_TG_1926287_Lg2_01735_MAF50_1659427_exon1	Chr02:866925..866925	T/A	A/B	A/B	A/B
18	RosBREEDSNP_SNP_GA_1956821_Lg2_AAT1_MAF20_1677156_exon1	Chr02:910864..910864	T/C	A/B	A/B	A/B
19	RosBREEDSNP_SNP_GA_1979331_Lg2_AAT1_MAF50_MDP0000307573_exon4	Chr02:939240..939240	A/G	A/B	A/B	B/B
20	RosBREEDSNP_SNP_AC_1873474_Lg2_01735_MAF20_MDP0000289720_exon10	MDC010540:327:13137..13137	A/C	A/A	A/B	A/B

The result page and downloaded file shows marker sorted by genome location, allele, and SNP variation for each germplasm

SSR Genotype data

GDR also has SSR genotype data from 20 datasets from apple, apricot, ground cherry, peach, pear, sweet cherry, and tart cherry. [Search SSR genotype](#), [view the list of the project](#) to access the individual project pages.

Funded by:
USDA National Research Project
(NRSP10)