

TUTORIAL 1: About GDR/ Getting Started

1. The GDR homepage can be reached by going to <http://www.rosaceae.org>



The screenshot shows the homepage of the Genome Database for Rosaceae (GDR). At the top left is the GDR logo with the text "Genome Database for Rosaceae". To the right of the logo is a search bar with a "Search" button and a "Login" link. Below the logo is a dark green navigation bar with white links: Home, General Info, Species, Projects, Maps, Search, Tools, Community, Calendar, and Contact. Below the navigation bar is a light green banner with the text "Genomic, Genetic, and Breeding Resources for Rosaceae Crop Improvement" on the left and "Peach Genome Sequence Released" on the right, accompanied by an image of a peach. Below the banner is a main content area with a dark green header "Welcome to the Genome Database for Rosaceae". The main content area is divided into two columns. The left column contains a paragraph about the GDR database, a section "Whats new in GDR?" with a bulleted list of updates, and a partially visible section "Batch BLAST". The right column contains a dark green header "Community News (8)" and a bulleted list of news items, including a WSU Press Release, the release of the peach genome sequence, a Fruit and Nut Workshop, a highlighted publication, a RosEXEC meeting, and the election of new members to the RosEXEC committee.

The homepage has a number of important features.

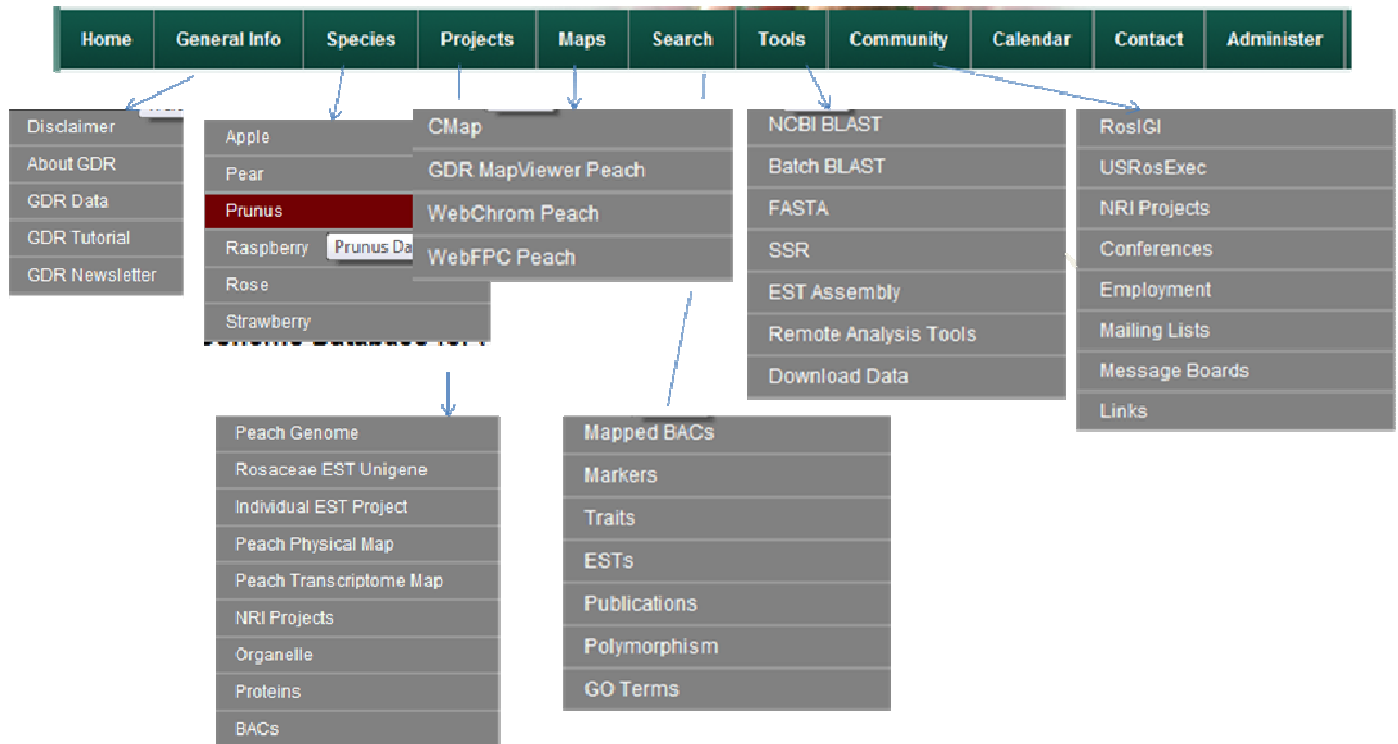
What's New in GDR lists new features of the database and is regularly updated.

Community News lists upcoming conferences, job opportunity, and various documents for the community.

The dark green bar and white links found at the top of the page are navigational tools. Similar links are expanded and listed for easy access along the light green bar at the left of the page.

The GDR provides the same navigation bar at the top of all of its pages to allow users to easily move between sections of the database.

The seven main sections, general info, species, projects, maps, search, tools, and community, each have a drop down menu with further options. To see this menu, move your pointer over the title. Additional section in the header bar includes calendar and contact. When you login, you will see the last section called 'Administer' where it allows you to change the content of GDR depending on your login privilege.



The tutorials provided systematically explore each of the sections available through this main menu, but are able to be done individually as well.

2. Please continue by clicking on the “About GDR” link under “general info” in the dark green navigation bar.

About GDR

The Genome Database for Rosaceae (GDR) is a curated and integrated web-based relational database. GDR contains comprehensive data of the genetically anchored peach physical map, annotated EST databases of apple, peach, almond, cherry, rose, raspberry and strawberry, Rosaceae maps and markers and all publicly available Rosaceae sequences. Annotations of ESTs include contig assembly, putative function, simple sequence repeats, ORFs, Gene Ontology and anchored position to the peach physical map where applicable. Our integrated map viewer provides graphical interface to the genetic, transcriptome and physical mapping information. We continue to add Rosaceae map data to CMap, a web-based tool that allows users to view comparisons of genetic and physical maps. ESTs, BACs and markers can be queried by various categories and the search result sites are linked to the integrated map viewer or to the WebFPC physical map sites. In addition to browsing and querying the database, users can compare their sequences with the annotated GDR sequences via a dedicated sequence similarity server running either the BLAST or FASTA algorithm, search their sequences for microsatellites using the SSR server or assemble their ESTs using the CAP3 Server.

Funding source: GDR has been financed by the NSF PGRP - Award #0320544.

Contact Us: If you have any problem, please contact us either by email or through our feedback form, we welcome your comments.

GDR Usage: In 2007, GDR was accessed by researchers from 44 countries with over 3.9 million pages accessed.

How to reference GDR:

- Sook Jung , Margaret Staton , Taein Lee, Anna Blenda, Randall Svancara, Albert Abbott , and Dorrie Main. GDR (Genome Database for Rosaceae): integrated web-database for Rosaceae genomics and genetics data. *Nucleic Acids Research*, 2008 January; 36(Database issue): D1034–D1040.
- Sook Jung , Christopher Jesudurai , Margaret Staton , Zhidian Du , Stephen Ficklin , Ilhyung Cho , Albert Abbott , Jeffrey Tomkins and Dorrie Main. 2004. GDR (Genome Database for Rosaceae): integrated web resources for Rosaceae genomics and genetics research. *BMC Bioinformatics* 2004, 5:130.

Other GDR publications:

- S. Jung, D. Main, M. Staton, I. Cho, T. Zhebentyayeva, P. Arus, A. Abbott. Synteny conservation between the prunus genome and both the present and ancestral Arabidopsis genomes. *BMC Genomics* 2006 7(1):8.
- Kevin M Folta , Margaret Staton, Philip J Stewart, Sook Jung , Dawn H Bies, Christopher Jesudurai and Dorrie Main. 2005. Expressed sequence tags (ESTs) and simple sequence repeat (SSR) markers from octoploid strawberry (*Fragaria × ananassa*). *BMC Plant Biology* 2005, 5:12
- Horn R, Lecouls AC, Callahan A, Dandekar A, Garay L, McCord P, Howad W, Chan H, Verde I, Main D, Jung S, Georgi L, Forrest S, Mook J, Zhebentyayeva T, Yu Y, Kim HR, Jesudurai C, Sosinski B, Arus P, Baird V, Parfitt D, Reighard G, Soorza R, Tomkins J, Wing R, Abbott AG. 2005. Candidate gene database and transcript map for peach, a model species for fruit trees. *Theor Appl Genet.* 2005 May;110(8):1419-28.
- Jung S, Abbott A, Jesudurai C, Tomkins J, Main D. 2005. Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs. *Funct Integr Genomics.* 2005 Jul;5(3):138-43.

Disclaimer:

Technical information contained herein is provided by Washington State University and Clemson University. While the authors make every effort to present accurate and reliable information on this Internet site, Washington State University and Clemson University do not certify such information, nor do they guarantee the accuracy, completeness, efficacy, timeliness, or correct sequencing of such information. Use of such information is voluntary, and reliance on it should only be undertaken after an independent review of its accuracy, completeness, efficacy, and timeliness. Washington State University and Clemson University assume no responsibility for results from the use of the information herein, or from use of the information obtained at linked Internet addresses, or in any respect for the content of such information, including (but not limited to) errors or omissions, the accuracy or reasonableness of factual or scientific assumptions, studies or conclusions, ownership of copyright or other intellectual property rights, and the violation of property, privacy, or personal rights of others. Washington State University and Clemson University are not responsible for, and expressly disclaim all liability arising out of use, reference to, or reliance on such information.

This page provides a basic description of the GDR project and what types of resources it makes available to Rosaceae family researchers.

The funding sources are listed and links to the corresponding site.

GDR publication is also listed. This is another good resource for learning more about the GDR.

This page has hyperlinks to the contact page. The GDR team welcomes any and all questions relating to the website or Rosaceae research.

3. Please continue by clicking on the “GDR Data” link under “general info” in the dark green navigation bar.

Data Overview

Rosaceae Mapping Data:

- Genetically Anchored Peach Physical Map:**
 The latest peach physical map (Apr 2008) contains 18387 BACs and 1387 contigs. A total of 149 contigs are anchored to the TxE map.
- Peach Transcriptome Map:**
 A total of 387 peach unigenes are anchored to genetic maps by hybridization to genetically anchored BACs. Of these, 216 unigenes are anchored to the TxE map. An additional 382 peach unigenes are anchored to genetic maps by hybridization to BACs that belong to genetically anchored BAC contigs.
- Rosaceae Genetic Maps:** Rosaceae genetic maps are available through CMap, the comparative map viewer. A total of over 1000 molecular markers and traits from various Rosaceae genetic maps and diversity studies are available from the marker or trait search page.
- Data available for the three peach BACs that are completely sequenced.**
 Anchored ESTs/markers and the sequences for the three BACs (082118, 028F08 and pPn31C7) can be accessed from BAC search site. An online annotation pipeline is currently under development by the GDR team.

Rosaceae EST Data:

- SNP Analysis:** The publicly available autoSNP software was used to analyze the most recent versions of the Rosaceae unigene and the individual genera unigenes. A summary of the results and downloadable data can be found here.
- EST Assemblies:** Assemblies of publicly available ESTs are available for the family Rosaceae and the genera Malus, Prunus, Fragaria, Rosa, Pyrus, and Rubus. Homology data for each assembly with SWISS-PROT is available to browse and download. Homology with the NCBI nr protein database and other rosaceae ESTs will be available soon.
- EST Analysis:** Full bioinformatics analysis, including quality trimming, assembly, homology searching, microsatellite discovery and putative unigene development, is currently available for Clemson Peach and Almond, UF Octoploid Strawberry, and CRA ISF Peach ESTs. Rose and blackberry projects will be available soon.

Organism	Project	EST Number
Rosaceae	Rosaceae Assembly	359001
Fragaria	Fragaria Assembly	49132
x ananassa (octoploid strawberry)		
whole plant:	Kevin Folta - University of Florida	1505
vesca (diploid strawberry)		
unopened flower buds:	Tom Davis - University of New Hampshire	2717
Malus	Malus Assembly	249195
Prunus	Prunus Assembly	89168
amygdalus (almond)		
developing seed:	Albert Abbott - Clemson University	2794
persica (peach)		
fruit mesocarp:	Albert Abbott - Clemson University	9984
root:	Albert Abbott - Clemson University	Available Soon
shoot:	Albert Abbott - Clemson University	7085
fruit mesocarp:	Elisa Vendramin - CRA ISF	1887

This page provides an overview of some of the projects being done by the GDR bioinformatics team and other collaborating researchers.

The grey chart displays the different EST projects that are available through the GDR. Clicking on an EST project will take you to the main page for that project.

The Rosaceae Mapping Data in available in GDR are listed and the appropriate links are provided. The mapping section of the database is also linked through the main navigation bar and has its own tutorial, “Using GDR Mapping Resources.”

4. Please continue by clicking on the “Apple” link under “species” in the dark green navigation bar.

Each species page lists of and provides links to all the data for the species available in GDR, such as mapping, BAC, marker, and sequence data. Species page also provides links to remote sites for funded projects and other related links for the species.

5. Please continue by clicking on the “Links” link under “Community” in the dark green navigation bar.

This page is good resource for finding information not available on the GDR website. The links are broken into three main categories, plant genomics and/or germplasm databases, genetics and genomics institutions, and journals.