

Overview of NRSP10 and Planned Tripal Development

Tripal Database Network Initiatives Workshop

Plant and Animal Genome Conference 2016

San Diego

**Dorrie Main*, Sook Jung, Jim McFerson, Mike Kahn, Cameron Peace
Washington State University**

***dorrie@wsu.edu**

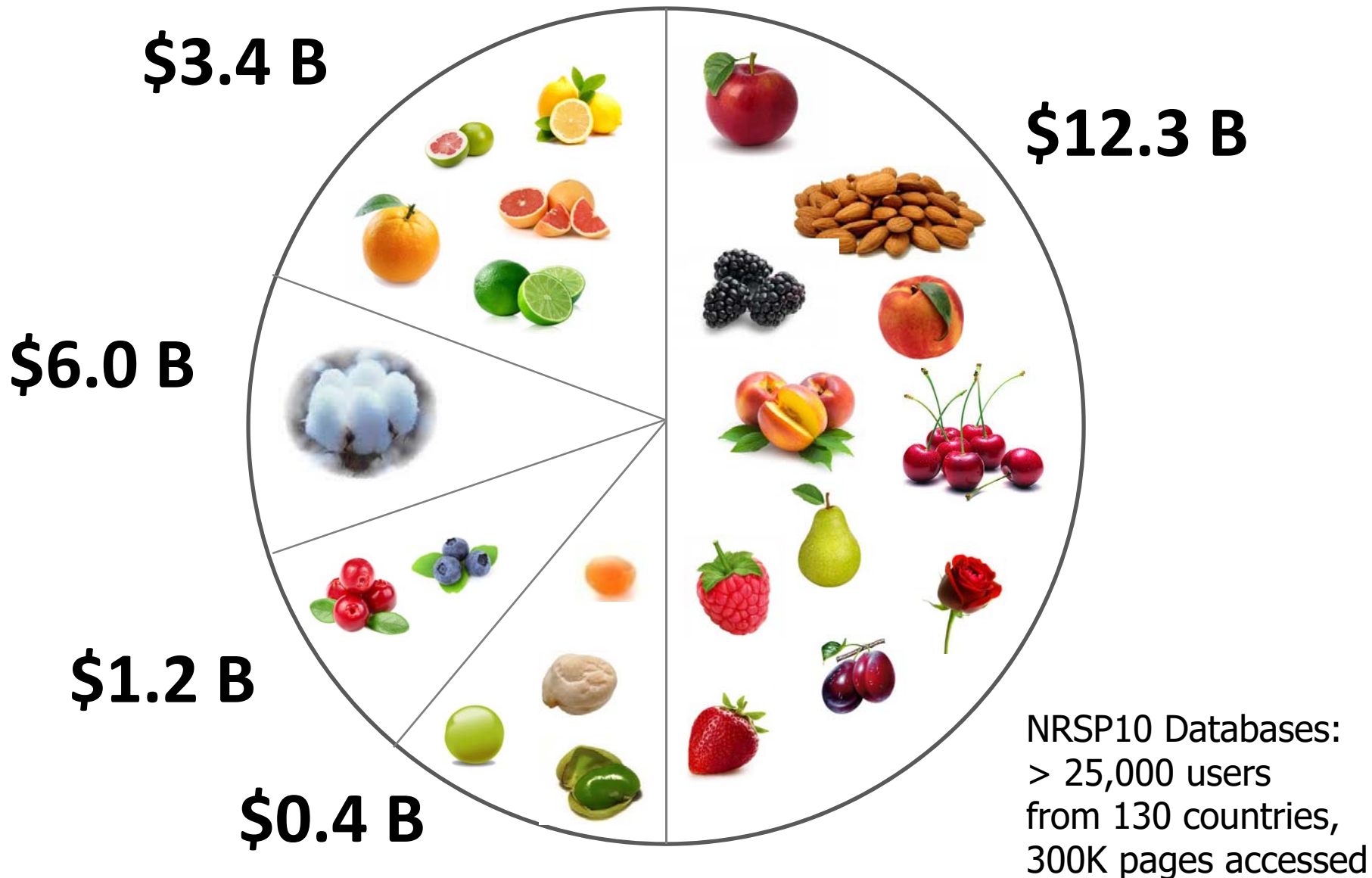
What is NRSP10?

- National Research Support Project
- Research that is deemed national in scope that fulfills an unmet national research need
- Requires participation by scientists from a large majority of US Land Grant Universities Agricultural Experiment Stations
- NRSP10 is only one of 10 projects approved in the last 100 years

NRSP 10 Vision

- Enable basic, translational and applied crop research by expanding existing online community databases for underserved crops – **mostly specialty crops**
- Provide a comprehensive open-source, flexible, resource-efficient database solution – **Tripal**
- Develop a model for long term sustainability of community databases – **Stakeholder driven and supported**

NRSP10 Crops are Economically Important



NRSP 10 Specific Objectives

1. Expand online community databases currently housing high quality genomic, genetic and breeding data for Rosaceae, citrus, cotton, cool season food legumes and *Vaccinium* crops
2. Develop/Implement a tablet application to collect phenotypic data from field and laboratory studies – **FieldBook App (Jesse Poland Program)**
3. Develop a Tripal Application Programming Interface for building breeding databases – **NRSP10 Breeders Focus Group**
4. Convert GenSAS, a community genome annotation tool, to Tripal – **Completely rewritten, version 4.0 just released**
5. Develop Web Services to promote database interoperability



The USDA funded National Research Support Project 10 supports the development of database resources and tools to aid genomics, genetics and breeding research in under-represented crops.

News and Events

- NRSP10 to be presented at ASHS annual meeting in New Orleans (08/07/15)
- Breeding database needs workshop to be held at NAPB annual meeting (07/30/15)
- Archives

databases



[citrus](#)



[cool legumes](#)



[vaccinium](#)



[cotton](#)



[rosaceae](#)

tools and resources



computational annotation and curation of genome sequences



open-source, modular platform for genomics, genetics and breeding databases



NRSP10 scientists
without up to date,
comprehensive databases



Button-clicking energized NRSP10 scientists using up to date
databases to enable their research



Genomic, Genetic, and Breeding Resources for Rosaceae Crop Improvement



Welcome to the Genome Database for Rosaceae

Initiated in 2003, the Genome Database for Rosaceae (GDR) is a curated and integrated web-based relational database providing centralized access to Rosaceae genomics, genetics and breeding data and analysis tools to facilitate basic, translational and applied Rosaceae research. GDR is supported by grants from the NSF Plant Genome Program (2003-2008), USDA NIFA Specialty Crop Research Program (2009-2019), USDA NIFA National Research Support Project 10 (2014-2019), and the Washington Tree Fruit Research Commission (2008-2016), Clemson University, University of Florida and Washington State University.

What's new in GDR?

- Gene Naming Guideline is available (4/23/15)
- *F. vesca* genome V1.1.a2 and V2.0.a1 and *Malus x domestica* genome V3.0.a1 available (4/6/15)
- Peach genome V2.0.a1 available (3/18/15)
- New QTL, markers and genetic map data from apple and rose are available (3/15/15)
- New search interfaces for haplotype data available (2/9/15)

General Information

[Report a problem](#) | [Ask us a question](#) | [Post a job](#) | [Post a meeting or event](#) | [Used GDR data or tools in your research - how to reference us.](#)

News

- XI Annual International Rubus and Ribes Conference, Asheville, NC June 21-24, 2015
- Gene Naming Guideline available (4/23/15)
- Rose transcriptome data and 68K Axiom SNP array data (Koning-Boucoiran et al. 2015) available for download (4/23/15)
- GDR Newsletter April 2015 available (4/6/15)
- *F. vesca* genome V1.1.a2 and V2.0.a1 and *Malus x domestica* genome V3.0.a1 available (4/6/15)
- *P. persica* genome V2.0.a1 available (3/18/15)
- New QTL, markers and genetic map data for apple and rose (3/15/15)
- New search interfaces for haplotype data (2/9/15)
- Announcements archive



The Genomics, Genetics and Breeding Resource for Citrus

A collaboration of tree fruit breeders, genomicists, bioinformaticians, stakeholders, and extension educators



Welcome to the Citrus Genome Database

Funded through the USDA NIFA National Research Support Project, NRSP10 from 2014 to 2019 and the USDA NIFA Specialty Crop Research Initiative project, tree fruit Genome Database Resources (tfGDR, 2009-2014), the Citrus Genome Database houses the genomics, genetics and breeding data for sweet orange, trifoliolate orange, grapefruit, mandarin, tangerine, pummelo, lemon, lime and other related species from the Citrus family. In citrus, the fundamental genomic tools include linkage maps, BAC libraries, physical maps, EST libraries, microarray platforms, and whole genome sequencing of sweet orange (heterozygous diploid) and mandarin (haploid). To facilitate efficient application in molecular breeding programs we will integrate large scale sequences, genetic markers and trait loci with diversity data, assign controlled vocabularies to the datasets and develop breeder interface tools. In collaboration with the Citrus community, this database is housed at [Washington State University](#) (Dorrie Main), with support from the [University of Florida](#) (Fred Gmitter), [Clemson University](#) (Albert Abbott) and the [Boyce Thompson Institute for Plant Research](#) (Lukas Mueller).

Updates will be made available via this website, to keep informed check back regularly or join [Citrus Genome Database mailing list](#). If you have any questions, feedback or problems, please contact us through the [contact link](#) on the navigation bar.

News

- [Citrus Genome Database activities at PAG 2015](#)
- [PAG 2015 - Citrus Genome Workshop](#)
- [8 New Citrus Genomes Sequenced from Mandarin, Pummelo and Orange](#)
- [USDA SCRI Citrus Greening Program](#)
- [tfGDR Website Launched to Aid Fruit Growers and Scientists](#)
- [News Archive](#)

Genome Database for *Vaccinium*

Search this site:

Search

Login

General

Species

Search

Tools

Education

Developing Genomic, Genetic and Breeding Resources
for Blueberry, Cranberry and Other *Vaccinium* sp.



Welcome to the Genome Database for *Vaccinium*

Vaccinium is an ancient genus of the family *Ericaceae*. Member species include blueberry (*V. corymbosum*, *angustifolium* and *virgatum*), cranberry (*V. macrocarpon*) and lignonberry (*V. vitis-idaea*). The Genome Database for *Vaccinium* (GDV) is being developed to house and integrate genomic, genetic and breeding data for blueberry, cranberry and other *Vaccinium* species. The GDV will include the blueberry genome being sequenced by North Carolina State University, and annotated transcripts, traits, maps and markers being generated by *Vaccinium* researchers. The GDV is implemented using Chado and Drupal (Tripal) and will include public and private sites to meet individual research group needs.

News

- National Research Support Project (NRSP 10), funded from 2014-2019.
- Generating Genomic Tools for Blueberry Improvement

Supported by the Plants for Human Health Institute, North Carolina State University and Washington State University.

Leveraging database resource development funded by the USDA NIFA Specialty Crop Research Initiative.

Copyright© 2010-2014. This site is designed to work with IE8, Mozilla, Safari and Opera.

Building a Genomics, Genetics and Breeding Resource for Cool Season Food Legume Improvement



Welcome to the Cool Season Food Legume Genome Database

Pea (*Pisum sativum* L.), **lentil** (*Lens culinaris* Medik.), **chickpea** (*Cicer arietinum* L.) and **faba bean** (*Vicia faba* L.) are cool season food legume crops. Members of the Fabaceae family, these legumes are distinguished by providing delicious and nutritious grains to the human diet. These crops serve as an important part of sustainable agricultural systems by providing biologically fixed nitrogen, made possible through symbiosis with rhizobia bacteria, and serving as a key rotation crop. Genomic information from the sequenced legume crops (Chickpea, *Medicago*, Lotus and Soybean) along with genomic and genetic data from the cool season food legumes provide a framework and road map for crop improvement. This database is being developed to serve as a resource to facilitate cool season food legume plant breeding and research by providing centralized, integrated access to relevant genetic and genomic information.

Quick Links

[Chickpea GBrowse](#) | [QTL search](#) | [Marker search](#) | [Sequence search](#) | [Franssen et al., 2012 Pea Transcriptome](#)

[Data submission](#) | [Report a problem](#) | [Ask us a question](#) | [What's been added/fixd?](#) | [What are we working on?](#) | [How to reference us](#)

News

- ChickpeaCyc now available
- NAPIA meeting to be held at Niagara Falls, Ontario, November 5-6, 2015
- EST libraries for BLAST tool updated April 2015
- Cool Season Legume Workshop at PAG 2015
- Mendel's Legacy Conference in the Czech Republic, Sept. 7-10, 2015
- Mendel's legacy: International Pea Genome Sequencing Consortium formed!





The Genomics, Genetics and Breeding Resource for Cotton

A collaboration of cotton breeders, genomicists, bioinformaticians, stakeholders, and extension educators



Welcome to CottonGen

CottonGen is a new cotton community genomics, genetics and breeding database being developed to enable basic, translational and applied research in cotton. It is being built using the open-source [TriPal database infrastructure](#). CottonGen consolidates and expands the data from [CottonDB](#) and the [Cotton Marker Database](#), providing enhanced tools for easy querying, visualizing and downloading research data. This project is funded by Cotton Incorporated, the USDA-ARS Crop Germplasm Research Unit at College Station, TX, the Southern Association of Agricultural Experiment Station Directors, Bayer CropScience, Dow/Phytogen, Monsanto and Washington State University.

What's New in CottonGen?

- [TM-1 genome sequence and annotation data available on CottonGen BLAST servers](#) (5/1/15)
- [TM-1 genome sequence and annotation data available on CottonGen Data Download](#) (4/24/15)
- [CGP-BGI *G. hirsutum* TM-1 genome sequence published on Nat Biotechnol.](#) (4/20/15)
- [NAU-NBI *G. hirsutum* TM-1 genome sequence published on Nat Biotechnol.](#) (4/20/15)
- [ICGI 2015-2017 Officers Elected.](#) (3/29/15)

General Information

[| Report a problem](#) | [| Ask us a question](#) | [| Post a job, meeting, or event](#) | [| Used CottonGen data or tools in your research - Reference us](#) |

News

- [May 2 - 6, 2016: World Cotton Research Conference - 6, Goiânia Convention Center, Goiânia - Goiás, Brazil](#)
- [TM-1 genome sequence data available on CottonGen BLAST Servers](#)
- [TM-1 genome sequence data available on CottonGen Data Download](#)
- [Sequencing of allotetraploid cotton \(*Gossypium hirsutum* L. acc. TM-1\) provides a resource for fibre improvement.](#) Nat Biotechnol
- [Genome sequence of cultivated Upland cotton \(*Gossypium hirsutum* TM-1\) provides insights into genome evolution.](#) Nat Biotechnol.
- [News archive](#)

Sustainability

- **Federal – Infrastructure Development (> \$10M)**
 - NRSP10 - **\$2M** Jan 1, 2015 for 5 years(PI Main)
 - USDA SCRI - **\$2.7M** Sept 1, 2014 for 5 years (PI Main)
 - NSF DIBBS - **\$1.5M** Jan 1, 2015 for 3 years (PI Ficklin)
 - NSF PGRP - **\$3.3M** Nov 15, 2015 for 3 years (PI Main)
- **Industry (>\$1M) – Curation**
 - Cotton
 - Washington Tree Fruit Research Commission
 - US Dry Pea and Lentil Council, Northern Pulse Growers
 - Citrus
- **Universities – Scientist salaries**
 - WSU, Clemson, UF, UT, UK, UConn and many others

Breeders Tools

- Held a NRSP10 Breeders Focus Group Workshop in Pullman at the National Association of Plant Breeding Annual Meeting to discuss breeder database needs
- Presented current breeding tools in NRSP10 databases and reviewed other breeding database resources
- Discussed current methods for data collection
 - Presented the FieldBook App for data collection (Trevor Rife, Ksenjia Gasic)
 - Provide focus group breeders with a tablet with FieldApp installed to test in their program
- Develop a plan for development for review by all NRSP10 Breeder

Functionality being considered

- Data upload capability from Field Book App, Allegro, AgroBase, Excel template (dataset, germplasm, trait, marker, phenotype and genotype)
- Add/edit individual data
- Manage germplasm data (create a list of germplasm, etc)
- Manage breeding experiments (create/store field maps, crosses, etc)
- Compare traits of two germplasm
- Search/download for germplasm with certain phenotype(s) and genotype(s)
- Download of data in different formats
- Calculate mean and other statistical values for traits
- calculate potential QTL from genotypic and phenotypic data
- Design markers around QTL of interest to use in certain population
- Predicting best combination of parents to achieve breeding goals

Presentations

- ASHS 2015 Annual Conference Computer Applications in Horticulture/Teaching Methods Workshop, New Orleans, August 8 2015
 - [Introducing NRSP10: Database Infrastructure for Specialty Crops](#)
 - [GenSAS: An Easy-to-Use, Web-Based Solution for Specialty Crop Genome Annotation](#)
- NRSP10 Breeders Database Needs Assessment Focus Group Workshop, National Association of Plant Breeders 2015 Annual Meeting, Pullman, July 30 2015
 - [NRSP10 Breeders Workshop Introduction](#)
 - [NRSP10 Breeding Databases and other Breeding Database Resources](#)
 - [NRSP10 FieldBook App](#)
- Plan and Animal Genome 2015 Annual Conference, January 2015
 - [Tripal Overview](#)
 - [GDR: What's New and What's Next](#)
 - [GDR Computer Demonstration](#)
 - [GDR in Tripal](#)
 - [GenSAS](#)
 - [GenSAS Computer Demonstration](#)

Acknowledgements

- Mainlab Bioinformatics Team
- Project coPIs/Pis
 - tfGDR (GDR and Citrus); Cacao Genome Database; Pine Genome Sequencing Project; Genome Database for Vaccinium; Cool Season Food Legume Database; CottonGen
- Rosaceae, Citrus, Cacao, Blueberry, Legume, Cotton and Bioinformatics Communities
- USDA NIFA SCRI, USDA DOE, NSF Plant Genome Program, USDA-ARS, SAAEDS, Mars Inc, Washington Tree Fruit Research Commission, Cotton Incorporated, USA Dry Pea and Lentil Commission, Northern Pulse Growers,
- US Land Grant University researchers and extension agents

Thanks for your attention
and support



NRSP 10
CROP DATABASE RESOURCES