

NRSP10 Progress in the last 12 months (April 1, 2017- March 31, 2017)

(1) Tripal Progress

- Core Tripal releases (3 releases in the past 12 months):
 - Tripal v2.1 (2.1) released on April 1, 2017
This is the version most Tripal databases currently use. They will upgrade to Tripal 3 as it becomes a stable release.
 - Tripal v3.0 beta 2 (3.0-rc1) released on June 30, 2017
Tripal v3 represents a major change over Tripal v2. It changes the way data types are referenced, using drupal entities instead of being “chado centric”. This provides more flexibility in how to access data for development and easier customization. It also allows non-Tripal databases to implement Tripal while continuing to use their current database structure if so desired, thus significantly expanding the adoptability of Tripal for older databases.
 - Tripal v3.0 beta 2 (3.0-rc2) released on September 28, 2017
Fixes some bugs with release candidate 1
- Tripal Extension modules:
 - Tripal MapViewer v1.0 for Tripal v2.1 developed and implemented on all 5 NRSP10 databases – code not yet released to the CSFL and CGD Databases
 - Tripal MapViewer v1.0 for Tripal v3.0 developed and being tested by the TreeGenes database before code release to the community in summer 2018
 - Elastic search module released enabling within database and cross-database querying capability (Staton Program, UTenn)
 - Many new modules being developed by 8 research groups in 3 countries. Examples include the SyntenyViewer and Expression Analysis (Fei Program, BTI), Tripal BrAPI (Valentin, INRA France), Tripal Galaxy (Ficklin Program, WSU), Tripal Breeding Information Management System (Main Program, WSU), etc.
- Usage, training, support and usage for Tripal:
 - Tripal downloaded/used for more than 130 species/clade databases
 - Monthly online conference calls
 - Quarterly 2 hr online training sessions held
 - 281 correspondences answered through the Tripal support mailing list
 - Full copy of CottonGen (less data) provided for SCRI CarrotRep project with support from the Mainlab team for implementation
 - 2 day hackathon at PAG, Jan 2018 (attended by 28)
 - Tripal workshop at PAG, Jan 2018 (attended by 65) – Presentation titles (Table 1) highlight the growth of the Tripal platform
 - Tripal.info website upgraded to a new design and kept up-to-date (Figure 1)

Table 1: Presentations at the Tripal Networks and Initiatives Workshop, PAG Jan 2018

Cucurbit Genomics Database: Integration Genetic and Genomics Resources for Cucurbit Breeding - <i>Zheng et al., Cornell University</i>
Updates to the Tripal Mapviewer and the Tripal Breeding Information Management System - <i>Main et al., Washington State University</i>
The Tripal Gateway Project: Supporting Exchange, Transfer and Analysis of Large-Scale Data for Online Biological Databases - <i>Ficklin et al., Washington State University</i>
Tripal Extensions Facilitate Association Mapping Studies on Forest Trees with CartograTree - <i>Wegrzyn et al., University of Connecticut</i>
Tripal Elasticsearch: Bringing Simple and Powerful Sitewide Search to Tripal Websites – <i>Staton et al., University of Tennessee</i>
National Center for Genome Analysis Support (NCGAS) use and development of Tripal Genome Browsers on XSEDE’s Jetstream – <i>Sanders et al., Indiana University</i>
An Open and Community Oriented Web Portal for Subtropical Fruit Trees Genomic, Phenotype and Breeding Data Analysis - <i>Talevera et al., University of Malaga</i>

Figure 1: Tripal Informational Site


v2.x User's Guide
v3.x User's Guide
v3.x Developer's Handbook
Tripal @ [gitHub.com](https://github.com)
Tripal @ drupal.org

Tripal is a toolkit for construction of online biological (genetics, genomics, breeding, etc), community database, and is a member of the [GMOD](#) family of tools. Tripal provides by default integration with the [GMOD Chado database schema](#) and [Drupal](#), a popular Content Management Systems (CMS). Tripal is licensed under the [GNU General Public License version 2](#).

Genomics, genetics, breeding and other biological data are increasingly complicated and time consuming to publish online for other researchers to search, browse and make discoveries. Tripal provides a framework to reduce the complexity of creating such a site, and provides access to a community of similar groups that share community-standards, and interact to address questions and learn best practices for sharing, storing and visualizing complex biological data.

Compatible With

Tripal directly supports or can be installed side-by-side with these popular tools:



News

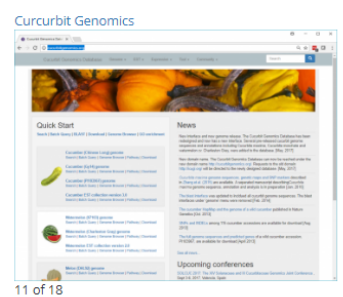
[Tripal Help Desk Schedule](#)
02/22/2018 - 18:35
Times and meeting links for the Tripal Help Desk meetings for Friday, February 23, 2018.

[Online Help Desk Session](#)
02/20/2018 - 22:17
Tripal Help Desk, a series of virtual collaboration meetings, kicks off this Friday, February 23, 2018.

[October Training](#)
10/25/2017 - 16:25
The Tripal Fall Training series will continue this Friday with an overview of the efforts to integrate Tripal and Galaxy.

[Online Training Event: Tripal v3](#)
09/28/2017 - 19:58
Join us Friday Sept 29th for an online training event to learn how to setup and manage a Tripal v3 site. [More](#)

Sites Using Tripal



11 of 18
[\[View more\]](#)

Contributors

The following organizations contribute to development of the core Tripal package or to extension modules that are freely available. Funding agencies that contribute funds for core Tripal development are also shown.

Tripal Tweets

Tweets by [@TripalProject](#)

(2) NRSP10 Databases

- Usage over the last 12 months of the five NRSP10 databases has significantly increased (number of visits and pages served almost doubled)

Table 2: Usage by Database, April 1, 2017 – March 31, 2018 (by Google Analytics)

<u>Database</u>	<u># Visitors</u>	<u># Visits</u>	<u># Pages Served</u>	<u># Countries</u>
Rosaceae	21,498	60,168	474,157	154
Cotton	11,253	31,263	241,588	136
Citrus	5,289	9,840	72,425	128
Legumes	3,262	5,972	41,912	106
Vaccinium	2,339	3,889	29,198	94
Total	43,641	111,132	859,280	154

- All of the 5 NRSP10 databases are now implemented in Tripal v2.1 with the new easier-to-use design. Cross database querying capability has been implemented in the Rosaceae and Citrus databases and these databases can now query each other and the other Tripal-based tree databases, TreeGenes and Hardwood Genomics. A significant volume of new data has been curated and added to all of the 5 databases (as documented in the individual databases).
- Training and outreach activities :

Table 3: List of outreach activities of NRSP10 from April 1, 2017 to March 31, 2018

March 2018: Plant Breeding Capacity Survey developed and hosted on NRSP10 website for the Plant Breeding Coordinating Committee. Enables searching by crop, state, organization and viewing on a map
March 2018: RosBREED Participants Meeting - Full Day Breeders Information Management System Training Workshop
Jan 2018: Cotton Beltwide Conference - CottonGen Presentation and Poster, CottonGen users meeting
Jan 2018: Plant and Animal Genome Conference - GenSAS, GDR, CottonGen, NRSP10, Tripal, TripalMap, BIMS presentations, dedicated Tripal Workshop (65 participants), dedicated AgBioData Workshop (~ 80 participants), GDR Advisory Board and Users Meeting, RosBREED Advisory Committee Meeting , ICGI annual in-person meeting, database booth, brochures
Nov 2017: North American Pulse Improvement Biannual Conference – CSFL presentation, discussions with researchers
Sept 2017: Cotton Breeders Tour – CottonGen presentation and Overview of BIMS by a Cotton Breeder presentation, discussion of breeder software needs and NRSP10

Sept 2017: American Society for Horticultural Science Annual Conference – NRSP10 workshop (60 participants), presentations by users of NRSP10 resources and open discussion of NRSP10 and researchers needs.

July 2017: International Peach Symposium – Resources in GDR for peach research (user presentation)

June 2017: American Society of Plant Biology Annual Conference – NSRP10 and BIMS presentations, AgBioData Booth, brochures

April 2017: AgBioData Consortium Workshop – attended by 43 scientists representing over 20 databases and allied resources. White paper finalized for submission to peer-reviewed journal and NSF Research Network Coordination proposal being planned by AgBioData executive committee

(3) Development of Breeding Tools (TriPalBIMS) progress

- **Data in BIMS:**
1 cotton, 2 Rosaceae, and 1 legume demo breeding program data now available in BIMS as well as public data from RosBreed in public GDRBIMS. Demo program breeders now loading their own data directly from FieldBook or Excel files. Following a RosBREED BIMS training workshop several other breeders are testing the system with their own data.
- **BIMS Functionality**

Figure 2: Functionality of BIMS v1.0, implemented in GDR, CottonGen and CSFL

Current Functionality of BIMS

In BIMS v1.0 Breeders can

- Create programs, add members and control permission levels
- Add their own data (accessions, cross, phenotype data from trials) in template
- Add phenotype data exported from the Field Book App
- Create input files for the Field Book App
- Sync data among multiple devices with the Field Book App through BIMS
- Perform/view statistical analysis on any list of accessions they created
- Perform search with various categories and trait cutoff values and save the list and phenotype data
- Archive the entire data

- **Training and Feedback:**
 - Full day BIMS Training Workshop (Figure 3) held at RosBREED Participants Meeting in March 2018. Over 40 participants, users surveys collected. Goals

of workshop in Figure 3. Response from breeders was generally very positive and next steps are to work with those who attended to provide individualized training on getting started using BIMS, plus providing monthly online help desk webinars.

- BIMS presented at 6 conference and meetings in the last 12 months to highlight progress and solicit further feedback.
- Worked with 4 demo breeders (cotton, peach, apple, pea, lentil, chickpea) extensively on usability and functionality of BIMS. Now being extended to other breeders.

Figure 3: Goals and Expected Outcomes of BIMS Workshop for Rosaceae Breeders

Goals and Expected Outcomes

- Demonstrate current and future utility of BIMS
 - Rosaceae breeding program personnel have a understanding of what BIMS can provide as a resource for Rosaceae breeding
- Provide hands-on training of BIMS v1.0
 - Rosaceae breeding program personnel know how to use BIMS v1.0
- Solicit feedback on future direction of BIMS
 - BIMS continues to develop functionality that is useful for Rosaceae breeding

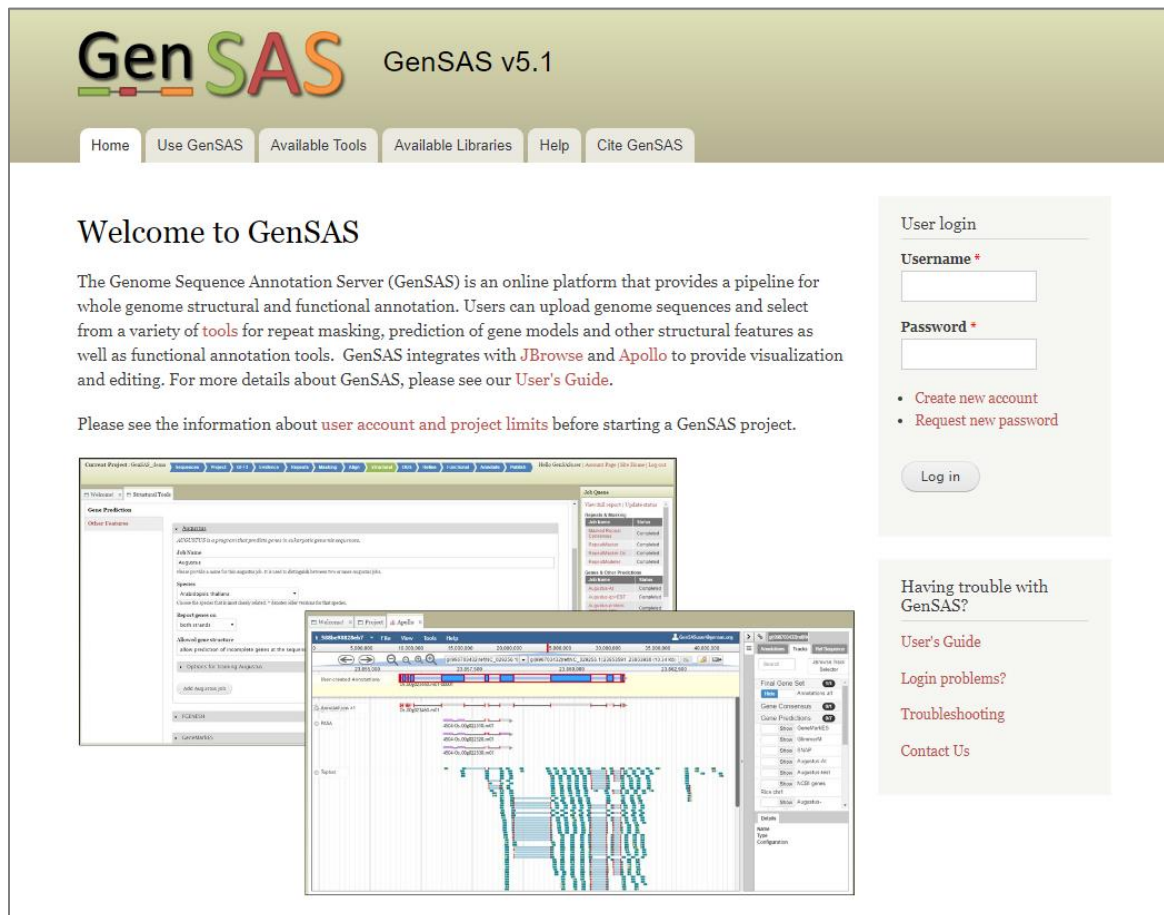
- Further functionality due to be implemented in BIMS by July 2018:
 - Ability to edit data within BIMS
 - Generate individual accession report with selected data
 - Additional data types: Ability to load, search, display and download pictures, raw SNP genotype, haplotype, trait associated allele or haplotype and breeding values.
 - Ability to download selected data in format for input to pedimap software for Rosaceae breeders.

(4) Genome Sequence Annotation Server (GenSAS)

- GenSAS v5.1 released in January 2018 (Figure 5). Improvements in the last year have included completed include (1) upgraded the integrated Apollo to the latest version (2) manually curated data with Apollo now merged into a final official gene set (3) functional analysis of the user curated data and the annotations now merged into the final gene set (4) enhanced GFF3 importer so that GenSAS can handle a GFF3 file that has more features. It also scans a GFF3 file for various types of errors before

- upload into GenSAS (5) RNAmmer tool added (6) optimized code for running on a cluster to speed up processing (6) introduced a pre-step where sequences uploaded are ordered by size and only those above a certain size are able to be analyzed.
- Over the year GenSAS was presented at three conferences and accessed by 2,813 visitors from 60 countries, with 8,999 sessions (as reported by Google Analytics)
- Currently have 250 users actively using GenSAS for annotation and community curation. Of these genomes currently being annotated in GenSAS, 54 are plants, 54 are fungi, 47 are bacteria, 12 are invertebrate, 4 are vertebrates and 12 are viral species.
- Finalizing some further performance tuning before submitting for publication by May 31st, 2018

Figure 5: GenSAS Home Page

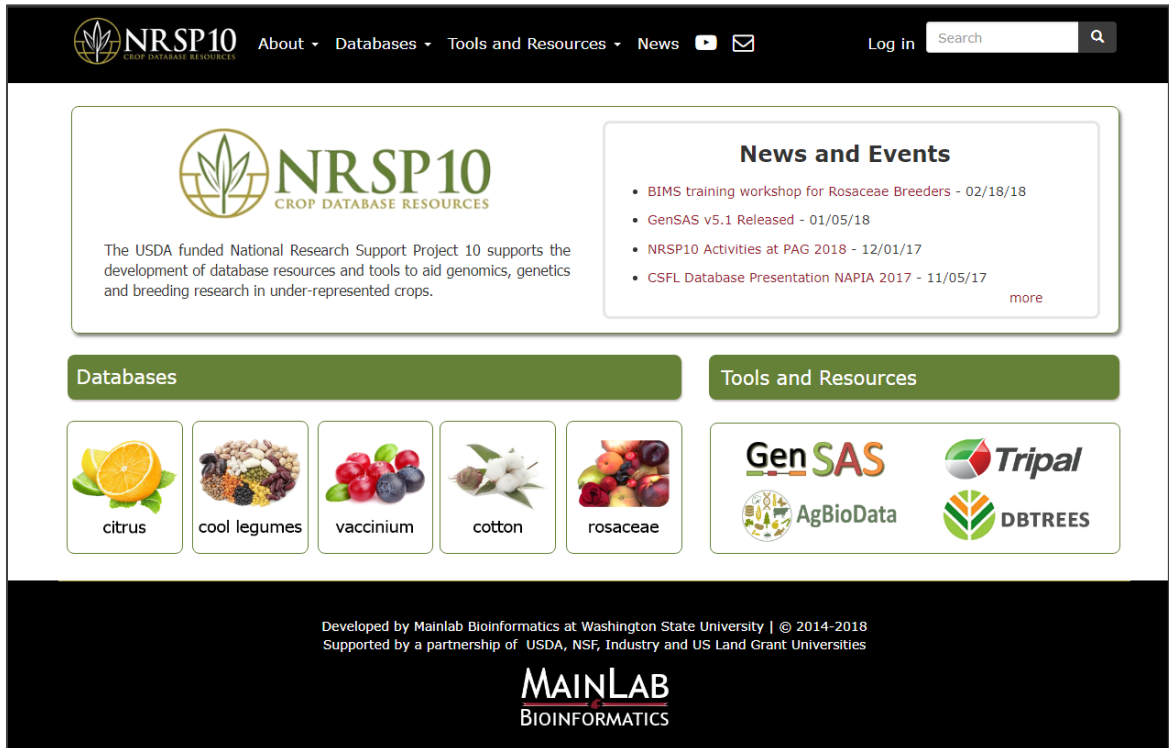


(6) Other activities

- Four manuscripts associated with NRSP10 activities from the Mainlab published and eight manuscripts (GenSAS, GDR, CottonGen, CSFL, CGD, TripalMap, CGD and Tripal v3.0) being prepared for submission to peer-reviewed journals by July 15th.

- AgBioData Consortium whitepaper “The AgBioData Consortium Recommendations for Sustainable Genomics and Genetics Databases for Agriculture” ready for submission to the oxford journal ‘Database’. This involved 43 representatives from over 20 databases and allied resources coming together for a workshop in April 2017 (funded by NSF PGRP to Dorrie Main) and then writing this 50 page whitepaper document.
- NRSP10 updated with participants, publications, presentations, funding, tools etc. (Figure 6)

Figure 6: NRSP10 Website Home Page



- Requested another user-driven NRSP10 workshop at ASHS annual conference (DC, July 2018). The workshop, if approved, will provide updates on the NRSP10 project and will include a discussion session followed by a participants meeting. Abstracts for the following talks have been submitted with the workshop request (Table 4). We will be notified by April 15 if our application is successful.

Table 4: Submitted Talks for Proposed NRSP10 Workshop at ASHS 2018 in DC

Welcome and goals of workshop - <i>Mike Kahn, Washington State University</i>
Tripal v3, the collaborative online database platform supporting an international community of plant genome databases – <i>Meg Staton University of Tennessee</i>
New data and functionality in NRSP10 Databases – <i>Sook Jung, Washington State University</i>

Integrating free mobile apps into specialty crop breeding and horticultural programs –
Trevor Rife – Kansas State University

Using the free tripal breeding information management system to improve Breeding
Efficiency – *Ksenija Gasic, Clemson University*

Extending the tripal breeding information system to combine data for global
performance predictions – *Cameron Peace, Washington State University*

NRSP10 and the next 5 year plan: community driven specialty crop database
functionality and data – *Dorrie Main, Washington State University*