**Tripal: a Construction Toolkit for Online Genome Databases**

Stephen P. Ficklin¹, Lacey-Anne Sanderson², Chun-Huai Cheng¹, Margaret Staton³, Sook Jung¹, Taein Lee¹, Il-Hyung Cho⁴, Kirstin E. Bett², Dorrie Main¹

¹ Washington State University, Pullman, WA, USA
² University of Saskatchewan, Saskatoon, SK, Canada
³ German University Genome Institute, IL, USA
⁴ Sagnaw Valley State University, University Center, MI, USA

**Brief Overview**

Tripal is a platform that simplifies construction of online genomic databases. With the increase in data from new sequencing technologies and downstream data analysis, the need for online visualization and data mining is ever increasing. The need for skilled web developers and IT professionals coupled with the complexities of project and data management creates an obstacle for many research communities or individual labs. Additionally, online genomic databases are expected to provide access to the raw data, analysis results, data mining tools, cross references to larger or companion databases, include outreach content and perhaps social networking capabilities. Tripal is intended to reduce these complexities by coupling the strengths of GMOD Chado, a relational database schema for biological data, and Drupal, a popular Content Management System (CMS).

Tripal provides a web interface that includes a Chado installer, data loaders for ontologies (controlled vocabularies), GFF files, and FASTA files. Web pages are automatically generated for organisms, genomic features, biological libraries, and stock collections. Web pages can be enriched with analysis results from BLAST, InterProScan, and Gene Ontology (GO). Tripal can be used "as is" but also allows for complete customization. PHP-based template files are provided for all data types to allow for precise customizations as required by the community. A well-developed Tripal API provides a uniform set of variables and functions for accessing any and all data within the Chado database.

Currently, Tripal only supports visualization of a subset of the current Chado schema, but further development is underway. Meanwhile, others can use the Tripal API to develop their own extensions. Those extensions can in turn be made available for anyone to use. These custom extensions, the Tripal package, and support resources such as an active mailing list can be found on the Tripal website (http://tripal.sourceforge.net/).

**Chado and Drupal**

Chado is a database schema to store and model biological data. It is designed to be generic enough to use for any organism. Chado has a modular design which is highly extensible. It is widely used and has good community support. A lot of popular tools, such as GBrowse and Apache Apollon, depend on Chado and Drupal as building blocks to deliver complex web applications.

Drupal is a popular content management system (CMS) that powers millions of websites. It is open source and is maintained and developed by a community of thousands of developers and users. Drupal also has a modular design so the functions of the website can be expanded easily by "plug-ins". Tripal is a collection of Drupal modules that interface with Chado tables. It can not only be used to insert, delete, update chado data, but also has visualization tools for the analysis results.

**References**


Tripal is funded indirectly through various agencies and groups. We are extremely thankful for this support. For a listing of these funding agencies please see the Tripal paper referenced above. The GMOD group provides logistical support.