

## **GDPC: The Genomic Diversity and Phenotype Connection: open-source platform for accessing public data via XML web services**

*Terry M. Casstevens<sup>1</sup> and Edward S. Buckler<sup>2</sup>*

<sup>1</sup> *Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853-2703,* <sup>2</sup> *USDA-ARS, Institute for Genomic Diversity, Cornell University, Ithaca, NY 14853-2703*

If all existing genomic diversity and phenotypic data were publicly accessible, this resource would revolutionize the analysis of quantitative traits and facilitate the conversion of genomics to improvements in plant breeding. The Genomic Diversity and Phenotype Connection (GDPC) already provides access to some of this data, and this platform will continue to make data sources publicly available. GDPC is software written in Java that integrates data related to taxa, loci, experiments, localities, genotypes, and phenotypes from multiple databases.

GDPC provides the infrastructure to create database connections that mask the complexities of any database's underlying schema. These "GDPC connections" communicate with databases and are designed as web services that transfer XML-formatted data to users.

GDPC also provides a Java API that allows software developers to create "GDPC aware" analysis applications. The GDPC Browser and TASSEL ([www.maizegenetics.net/bioinformatics/tasselindex.htm](http://www.maizegenetics.net/bioinformatics/tasselindex.htm)) are applications that currently use this API. The GDPC Browser is a front-end application that allows users to retrieve, view, and group data based on search criteria.

"GDPC connections" already exist to an earlier generation maize diversity database, Panzea ([www.panzea.org](http://www.panzea.org)) and to the new diversity datasets in Gramene ([www.gramene.org](http://www.gramene.org)). Work is in progress to create access to databases that implement the Germinate schema ([bioinf.scri.sari.ac.uk/cgi-bin/germinate/germinate.cgi](http://bioinf.scri.sari.ac.uk/cgi-bin/germinate/germinate.cgi)). Other collaborative efforts include connections to MaizeGDB ([www.maizegdb.org](http://www.maizegdb.org)) and GRIN ([www.ars-grin.gov](http://www.ars-grin.gov)). In addition, any organization wanting to create public access to their data can develop their own "GDPC connection" using the GDPC Java library.

The source, binaries, documentation, etc. are freely available at:  
[www.maizegenetics.net/gdpc/index.html](http://www.maizegenetics.net/gdpc/index.html).