Dated: January 22, 2010

Ref: Letter of collaboration with the NSF funded Gramene Database and the Plant Ontology Projects

Dear Andrew and Alan,

We are writing this letter as PIs of the Gramene database project (Doreen Ware) and Plant Ontology Projects (Pankaj Jaiswal) for support and collaboration on the continuation project on sorghum genome function and evolution that you and your colleagues are planning to submit to the NSF Plant Genome Research Program. This is so because it involves the curation of datasets using the Plant (PO) and Trait (TO) Ontologies, integration of the datasets in ontology and germplasm databases at Gramene, building associations to the Plant Ontology database and genotype data to the Gramene's Diversity section.

It is our understanding that you will be generating the following data sets:

(1) Photos and descriptive information for about 5000 M3 lines that will be used in TILLING. We will set up a system in which the photos are physically stored at your project's site and you send the URLs to the pictures, along with the ontology based trait descriptors, to Gramene. This way people will be able to search for traits at Gramene, get a report that includes the image (which is coming to them from your site), and then link to your site for more details. Your use of the 'Trait Ontology' descriptors for these, to the degree possible, will facilitate integration of this information with other trait-related data in Gramene. Where no trait descriptor is available, you will work with staff of the Gramene and the Plant Ontology Consortium to develop the appropriate trait description and the associated plant structure and growth stages at which they are evaluated for incorporation into the existing ontologies. It is anticipated that this resource will be brought on line in late 2011 and that updates will be warranted annually.

(2) Genome-wide large-scale SNP diversity data for a small sampling of sorghum genotypes (approximately early 2011), and locus-specific data for about 500 genes across a diversity panel of about 240 genotypes (approximately 2012). Your use of the GDPDM schema will simplify the incorporation of these data into the Gramene Genetic Diversity database. We will work with your group to prepare Ensembl variations databases to allow the display of variation in reference to the *Sorghum bicolor* sequence resulting from your prior work. We appreciate that you also intend to deposit your variation data at dbSNP. Because we have been involved with the curation of diversity data and sequence data on a global scale we look forward to your results and placing them in the context of genomic data and the comparative framework Gramene provides. It is anticipated that the data will be made available by your group in the desired format at a mutually agreeable date.

We are pleased that co-PI Alan Gingle, the curator of the Saccharinae-specific database that you have created and propose to expand on, will facilitate the data transfer, curation and integration with the Gramene and with the Plant Ontology (only ontology based annotations) . Alan will be responsible for assuring the quality of the data and generation

of copies, maintaining continuity with updates to the reference genome assembly, generating the file format requested by the Gramene and Plant Ontology databases, and coordinating upload with a our respective staff members; and assuring that the integrity of the data is reproduced in our databases, working with appropriate members of your group to make any corrections or modifications.

We will also be pleased for you to include a link back to the integrated annotations in the Gramene and Plant Ontology databases from the "Sorghum Portal" that you propose. Please be aware that we can only commit to serving the data for the length of the current Gramene grant that will be ending in October of 2011 and Plant Ontology that will end in 2013. If our grant is refunded we will be happy to continue to serve the data as long as funding for our projects exists.

Sincerely,

Doreen Ware
USDA-ARS and Cold Spring Harbor Laboratories

Gramene Database PI

Pankaj Jaiswal
Assistant Professor
Dept. of Botany and Plant Pathology
Oregon State University

Plant Ontology PI
Co-PI Gramene database