

Barley Coordinated Agricultural Project Work Plan FY06 (3/1/06 – 2/29/07) Julie A. Dickerson, Roger Wise, Iowa State University

1) Describe the research, education, and outreach activities you are planning for the next year (3/1/06 – 2/29/07)

Research The *Hordeum* Toolbox (THT), will serve as an integrated public resource for discovery and dissemination of large-scale SNP data sets as it applies to genetic, phenotypic, and trait data. Once the database has been established in THT; SNP data, pedigree data, and phenotypic data from breeding germplasm will be uploaded to the database.

Establish infrastructure for housing BarleyCAP data. THT will store all the SNP and trait data, view SNPs, haplotypes, and sort genotypes. This data will be cross-referenced to the Barley1 GeneChip contigs with integration to the physical and genetic maps, and expression data. Germplasm, phenotype, and trait data will be stored using the Passport Module of the GERMINATE database developed at the Scottish Crop Research Institute. THT will be Genomic Diversity and Phenotype Connection (GDPC) compatible, providing access to genomic diversity data such as SNPs, SSRs, sequences, etc. and phenotypic data that will be collected in field, genetic, or physiological experiments.

Integrate SNP and gene expression data. Every SNP marker will be mapped back to the sequences on the Barley1 GeneChip (using Vmatch) as a permanent sequence based link to the rapidly accumulating expression profile data sets. SNP expression data sets from Illumina bead array technology will also be made available and linked to Barley1 exemplars if possible.

Education and Outreach: THT toolbox will make SNP, breeding, and marker data from this project publicly available to researchers and students around the world.

2) List specific outcomes and deliverables that will be accomplished in the first 6 months (3/1 – 8/31). These will be used as benchmarks for your bi-annual progress report.

- Set up preliminary web site for The Hordeum Toolbox.
- Finalize plans for submission of breeder data from Jennifer Kling using a spreadsheet interface. Develop an initial list of measured traits and their units and establishing identifiers for all participants using a uniform numbering scheme with Kevin Smith.
- Detailed THT database schema published for Illumina SNP data and plant breeding data that uses the Germinate database design (ported to MySQL) for its initial development.
- Preliminary web site design published giving detailed website structure.
- Research off-site data storage options for geographically diverse data storage at the web portal and develop a data life profile.
- Port the Genome Viewer Tool (GVT) from SCRI and the GBROWSE tool to THT.

3) List specific outcomes and deliverables that will be accomplished in the second 6 months (9/1 – 2/29). These will be used as benchmarks for the bi-annual progress report.

- Make first year breeding and pedigree data yield trial data available on THT web portal.
- Provide tool for formatting an input file to the QTLMiner tool using data in THT.
- Plant and Animal Genome Conference presentation about THT.
- Select an off-site data storage option for the web database backup for the duration of the project.
- THT web portal database links on-line with searches for publicly available:
 - Breeding data by trait, breeder and other fields.
 - Illumina SNP data in both raw and normalized form with links to Barley1 GeneChip expression data in PLEXdb.

Barley Coordinated Agricultural Project Biannual Progress Report FY06 (4/1/06 – 9/30/06)

Julie A. Dickerson, Roger Wise, Iowa State University

1) Describe the research, education, and outreach activities you completed in the first half of the FY06 (4/1/06 – 9/30/06)

Research. Began the process of developing The *Hordeum* Toolbox (THT), an integrated public resource for discovery and dissemination of large-scale SNP data sets as it applies to genetic, phenotypic, and trait data. Use cases for two different users, plant breeders and academic researchers were developed and circulated to the project team. Comments from the potential user base were incorporated into these documents. User scenarios have also been developed to guide website development.

File format development. File formats for pedigree information, QTLMiner, and SNP data were agreed upon with the respective research groups.

Integrate SNP and gene expression data. Every SNP marker will be mapped back to the sequences on the Barley1 GeneChip using the assemblies in the HarvEST database and the gene locus names. All SNPs will also be labeled according to the proposal from David Marshall at the SCRI.

Education and Outreach: Nothing to report.

2) List specific outcomes and deliverables accomplished in the first half of FY06 (4/1 – 9/30).

- Web site design based on use case analysis is underway. Registered the domain names *hordeumtoolbox.org* and *thehordeumtoolbox.org* for The Hordeum Toolbox web site.
- Finalized plans for submission of breeder data from Jennifer Kling using a spreadsheet interface.
- Developed an initial list of measured traits and their units and establishing identifiers for all participants using a uniform numbering scheme with Kevin Smith.
- File formats for data exchange with QTL Miner established.
- Reviewed Germinate schema 1.9, waiting for revised schema based on a MySQL Platform.
- Developed table design for SNP data results. Reported allelic data for each line will be searchable by users and raw data files will be available in zip files if users wish to apply their own environment to the data.
- Researched off-site data storage options for geographically diverse data storage. All data tables will be backed up monthly and stored at a commercial vendor; estimated cost is around \$13/month for this service.