The Plant Trait Ontology Links Wheat Traits for Crop Improvement and Genomics

Laurel COOPER[a,c] , Marie-Angélique LAPORTE[b], Justin ELSER[a], Victoria Carollo BLAKE [c], Taner Z. SEN [c], Chris MUNGALL [d], Elizabeth ARNAUD[b], Pankaj JAISWAL[a].

[a] *Planteome Database, Oregon State University, Corvallis, OR, USA*
[b] *Alliance Bioversity International-CIAT, Montpellier, France*
[c] *GrainGenes Database, USDA-ARS-WRRC, Albany, CA, USA*
[d] *Lawrence Berkeley National Laboratory, Berkeley, USA*

Global population growth and climate change result in the need to develop new and better-adapted wheat varieties. Three integrated resources, the Plant Trait Ontology (TO), the Crop Ontology (CO) and the GrainGenes[1] database provide researchers and plant breeders interconnected tools and resources to utilize the large amounts of genetic and genomic data that are available for plant genomics and crop improvement. The TO [1–3] is a reference-level ontology for plant traits, and is a key part of the Planteome Project[2] [4]. In the current Planteome Release Version 4.0, the TO consists of more than 1500 plant traits organized into nine upper-level categories: biochemical, biological process, plant growth and development, plant morphology, plant quality, plant vigor, plant stress and yield. As part of the Planteome, the TO is integrated with the Plant Ontology [5,6], and is used to annotate, or link to plant genomics and genetics data objects (e.g. germplasm, QTLs, genes, and proteins) from a wide variety of plant taxa, including important world crops and model plant species. In this Release, there are more than four million annotations linking TO terms to about 165,000 data objects in the Planteome database. Users are encouraged to submit requests for new TO terms or comments through the TO GitHub Issue Tracker [3].

The Planteome project has partnered with the CO [4] [7] to incorporate crop- and clade-specific traits of interest to plant breeders into the Planteome database. Terms from eleven CO trait dictionaries (banana, cassava, lentil, maize, pigeon pea, potato, rice, sorghum, soybean, wheat and yam) have been mapped to the equivalent reference traits in the TO in a semi-automated way, based on common design patterns [8], followed by manual curation and quality checks. This integration of traits across species based on the TO hierarchy allows users, for instance, to search for, or sort traits by plant structure or by trait category and facilitates cross-species queries and data discovery.

The CO Wheat Trait Dictionary, developed along with CIMMYT[5] and the Triticeae Toolbox Database[6], consists of traits of interest to the international wheat breeding

---

[1] https://wheat.pw.usda.gov/
[2] http://planteome.org/
[3] https://github.com/Planteome/plant-trait-ontology/issues
[4] https://www.cropontology.org/
[5] https://www.cimmyt.org/
[6] http://triticaetoolbox.org/

community, with methods and scales of measurement. It is incorporated into the Breeders Field Book[7] allowing scientists to record crop phenotype observations in the field. Of the 316 traits in the Wheat Trait Dictionary, approximately 50 are mapped to an exact matching class in the TO, while the others are mapped to a more general class.

The GrainGenes database [9] is an integrated relational database and internet resource for the international small grains community which provides curated genetic and genomic information about Triticeae species (mainly wheat, barley, rye and their wild relatives), and Avena species (mainly oat), with a diversity of data types including genome browsers, comparative linkage maps, sequence polymorphisms, and QTLs. To improve data standardization and interoperability, 246 traits in the GrainGenes database have been mapped to date to 155 unique TO terms, and 72 traits mapped to 42 unique wheat CO terms, all of which link back to the Planteome database.

As new wheat traits are characterized to adapt to increased demands for food and changes in the climate, these three integrated resources provide researchers and plant breeders tools to utilize the large amounts of genetic and genomic data that are available for crop improvement.

References:
1. Jaiswal P, Ware D, Ni J, Chang K, Zhao W, Schmidt S, Pan X, Clark K, Teytelman L, Cartinhour S, Stein L, McCouch S. Gramene: development and integration of trait and gene ontologies for rice. Comp Funct Genomics. 2002;3(2):132–6.
2. Arnaud E, Cooper L, Shrestha R, Menda N, Nelson RT, Matteis L, Skofic M, Bastow R, Jaiswal P, Mueller L, McLaren G. Towards a reference Plant Trait Ontology for modeling knowledge of plant traits and phenotypes. In: Proceedings of the International Conference on Knowledge Engineering and Ontology Development [Internet]. Barcelona, Spain: SciTePress; 2012. p. 220–5.
3. Yamazaki Y, Jaiswal P. Biological Ontologies in Rice Databases. An Introduction to the Activities in Gramene and Oryzabase. Plant Cell Physiol. 2005 Jan 15;46(1):63–8.
4. Cooper L, Meier A, Laporte M-A, Elser JL, Mungall C, Sinn BT, Cavaliere D, Carbon S, Dunn NA, Smith B, Qu B, Preece J, Zhang E, Todorovic S, Gkoutos G, Doonan JH, Stevenson DW, Arnaud E, Jaiswal P. The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics. Nucleic Acids Res. 2018 Jan 4;46(D1):D1168–80.
5. Cooper L, Walls RL, Elser J, Gandolfo MA, Stevenson DW, Smith B, Preece J, Athreya B, Mungall CJ, Rensing S, Hiss M, Lang D, Reski R, Berardini TZ, Li D, Huala E, Schaeffer M, Menda N, Arnaud E, Shrestha R, Yamazaki Y, Jaiswal P. The Plant Ontology as a tool for comparative plant anatomy and genomic analyses. Plant Cell Physiol. 2013 Feb 1;54(2):e1–e1.
6. Walls RL, Cooper L, Elser J, Gandolfo MA, Mungall CJ, Smith B, Stevenson DW, Jaiswal P. The Plant Ontology Facilitates Comparisons of Plant Development Stages Across Species. Front Plant Sci. 2019;10.
7. Shrestha R, Arnaud E, Mauleon R, Senger M, Davenport GF, Hancock D, Morrison N, Bruskiewich R, McLaren G. Multifunctional crop trait ontology for breeders' data: field book, annotation, data discovery and semantic enrichment of the literature. AoB Plants 2010 Jan 1.
8. Osumi-Sutherland D, Courtot M, Balhoff JP, Mungall C. Dead simple OWL design patterns. J Biomed Semant. 2017;8(1):18.
9. Blake VC, Woodhouse MR, Lazo GR, Odell SG, Wight CP, Tinker NA, Wang Y, Gu YQ, Birkett CL, Jannink J-L, Matthews DE, Hane DL, Michel SL, Yao E, Sen TZ. GrainGenes: centralized small grain resources and digital platform for geneticists and breeders. Database. 2019 Jan 1;2019.

---

[7] https://excellenceinbreeding.org/toolbox/tools/field-book