

The Sweetpotato Ontology

Vilma Rocio Hualla
International Potato Center,
Peru
v.hualla@cgiar.org

Reinhard Simon
International Potato Center,
Peru
r.simon@cgiar.org

Robert Mwangi
International Potato Center,
Nairobi
r.mawanga@cgiar.org

Henry Saul Juarez Soto
International Potato Center,
Peru
h.juarez@cgiar.org

Genoveva Rossel Montesinos
International Potato Center,
Peru
g.rossel@cgiar.org

Keywords: breeding; ontologies; sweetpotato; phenotypic; genotypic.

1. Introduction

The sweetpotato ontology is part of a community effort to establish a set standard nomenclature to describe crop development and agronomic traits to facilitate analyzing and sharing of phenotypic and genotypic information. The development and adoption of data standards is vital to the interoperability of sweetpotato data (Simon et. al. 2014). Phenotype ontologies are controlled, hierarchically-related phenotypic descriptions that enable large-scale computation among individuals, populations, and even multiple species (Hoehndorf et al., 2013). The International Potato Center (CIP) is currently pursuing the development of standards for plant phenotyping data in collaboration with other interested groups.

The advantage of ontology is that both humans and software applications can understand a data domain. This will allow the application of numerical or data mining techniques that may help to uncover previously unknown correlations. Building on previous draft versions, here we focus on traits important to breeding.

2. Materials and Methods

Through collaborations under the Generation Challenge Program, compatibility data is consolidated by ontologies. Descriptors used in morphology were taken from Huaman (2001). Descriptors used in evaluations were previously standardized (Grüneberg et al., 2009). Additionally, we used descriptors from the Catalogue of Orange-fleshed sweetpotato varieties for Sub-Saharan Africa (Kapinga et al., 2010).

We used the Crop Trait Dictionary Upload Template Version 4 to update the information in the web crop ontology. Terms in ontology were organized in the form of a tree. The nodes of the tree represent entities at greater or lesser levels of detail (Smith, 2004). The branches connecting the nodes represent the relation between two entities (ej. radicle emergence stage is a child of the parent term germination stage). Individual stages of a scale are then parts that can be related to the whole by their order of appearance during plant growth. Each term carries an unique identifier and strictly specified relationships between the terms allow systematic ordering of data within a database, this in turn improves input and retrieval of information (Bard and Rhee, 2004; Harris et al., 2004).

3. Results and Observations

The sweetpotato ontology currently describes 109 traits (Table 1). These include: morphological (28), agronomical performance (28), biochemical (23), reaction to biotic stress (7) and quality traits (23). These traits describe phenotypic variability for characteristics needed for crop improvement. We anticipate further refinements and cross-checks.

TABLE1. Frequency of sweetpotato ontology

Variable	Frequency absolute	Frequency relative
Agronomical traits	28	0.26
Morphological traits	28	0.26
Biochemical traits	23	0.21
Quality traits	23	0.21
Biotic stress	7	0.06
Total	109	1

The ontologies are managed in spreadsheets for ease of transfer from existing data dictionaries. The first version of sweetpotato ontology can be found online.¹

The screenshot shows the CO_331 ontology interface. On the left, a tree structure lists various traits under 'Morphological traits', including 'Abaxial Leaf Vein Pigmentation', 'Distribution of Secondary Flesh color', 'Flower color', 'General Outline of the Leaf', 'Ground Cover', 'Immature Leaf Color', 'Intensity of Predominant Skin color', 'Latex Production in Storage Roots', 'Leaf Lobe Number', 'Leaf Lobes Type', 'Mature Leaf Color', 'Mature Leaf Size', 'Oxidation in Storage Roots', 'Petiole Pigmentation', and 'Plant Type'. The 'Flower color' trait is expanded to show 'Observation of flower color' as a method. On the right, a detailed view of 'Observation of flower color' is shown, including its identifier (CO_331:0000047), description ('Visual categorization'), name of method ('Observation of flower color'), creation date (Mon Jun 16 08:40:50 UTC 2014), and a bibliographic reference: 'CIP, AVRDC, IBPGR. 1991. Descriptors for Sweet Potato. Huamán, Z., editor. International Board for Plant Genetic Resources, Rome, Italy'.

FIG 1: Sweetpotato trait ontology structure trait group relationships

4. Conclusions

The goal of ontologies is to construct a set of clearly defined vocabularies that can be used to construct queries between different crops linking the phenotypic and genetic data useful for integrated breeding through data annotation. In addition, the ontologies should enhance future

¹ http://www.croponology.org/ontology/CO_331/Sweet%20potato

efforts to explore the relationships among phenotypic similarity, gene function, and sequence similarity in plants, and to make genotype-to-phenotype predictions relevant to plant biology, crop improvement, and potentially even human health.

The use of ontological methods to organize biological knowledge is an active area of research and development. The definition of a set of common terms in sweetpotato is contributing in the development of the BioMart database. The datamart of sweetpotato is organized using the sweetpotato ontology.

Acknowledgements

Authors thanks financial support of Crop ontology community of practice of Bioversity International and Generation Challenge Program (Project Code CM-B7001).

References

- Bard J.B., Rhee S.Y. (2004). Ontologies in biology: design, applications and future challenges. *Nat Rev Genet* 5: 213–222
- Grüneberg, W. J., Eyzaguirre, R., Espinoza, J., Mwanga, R. O., Andrade, M., Dapaah, H. & Low, J. (2009). Procedures for the Evaluation and Analysis of Sweetpotato Trials.
- Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, et al. 2004. The gene ontology (GO) database and informatics resource. *Nucleic Acids Res* 32: D258–D261
- Hoehndorf R, Dumontier M, Gkoutos GV. (2013). Evaluation of research in biomedical ontologies. *Brief Bioinform.* 14:696–712.
- Huaman Z., 2001. Descriptores de la Batata. Retrieved March 20, 2015, from http://www.bioversityinternational.org/uploads/tx_news/Descriptors_for_sweet_potato_Descripteurs_pour_la_patate_douce_Descriptores_de_la_batata_263_ES.pdf.
- Kapinga, R., S.Tumwegamire, J. Ndunguru, M.I. Andrade, S. Agili, R.O.M. Mwanga, S. Laurie, and H. Dapaah. (2010). Catalogue of Orange-fleshed sweetpotato varieties for Sub-Saharan Africa. International Potato Center (CIP), Lima, Peru. 40p. Retrieved March 10, 2015, from: <http://sweetpotatoknowledge.org/germplasm/breeding/availableaterial/OFSP%20Catalogue%20final%20from%20website.pdf#>.
- Oellrich, A., Walls, R., Cannon, S., Cooper, L., Gardiner, J., Gkoutos, G., Harper, L., Hoehndorf, R., Jaiswal, P., Kalberer, Lloyd, J., Meinke, D., Menda, M.; Moore, L.; Nelson, R.; Pujar, A.; Lawrence, k.; Huala, E. (2015). An ontology approach to comparative phenomics in plants. *Plant Methods* 2015, 11:10 doi:10.1186/s13007-015-0053.
- Simon, R, Hualla, V.; Eyzaguirre, R.; Cordova, R.; Mwanga, R.; Rossel, G.; and Gruneberg, W. (2014). Progress in developing sweetpotato ontology for breeders. Poster present at Workshop on Crop Ontology and Phenotyping Data Interoperability. Conference, Montpellier, France, 31 March to 4 April, 2014.
- Smith, B. (2004). Beyond Concepts: Ontology as Reality Representation. In AVaL Vieu, ed, International Conference on Formal Ontology and Information Systems. Proceedings of FOIS 2004, Turin, Italy.