



From Plant Ontology to Gene Ontology and back

Dennis Stevenson^{a,*}, Cecilia Zumajo-Cardona^{a,b}

^a The New York Botanical Garden, Bronx, NY, 10458, USA

^b The Graduate Center, City University of New York, New York, NY, 10016, USA

ARTICLE INFO

Keywords:

Angiosperm
Gene ontology
Gymnosperm
Integument
Ontology
Ovule
Plant ontology

ABSTRACT

Our recent goal for the Plant Ontology (PO) is to have it integrated with the Gene Ontology (GO). The simplest aspect of this is to link morphological and anatomical images of structures for PO terms with genes involved in the development of those structures (GO terms). This, also, would include images of expression analyses by *in situ* hybridization. By using a process of reciprocal illumination, we will be able to clarify and/or redefine PO terms. In particular, an example of this is the integument development during ovule maturation in seed plants. Gymnosperms have one integument; whereas, Angiosperms (flowering plants) have two integuments, i.e., an inner integument surrounded by an outer integument. The question that arises is which of the two integuments in the Angiosperms is the equivalent of the single integument in the gymnosperms? In Angiosperms, the gene *INNER NO OUTER (INO)* is involved in the proper development of the outer integument but not in the inner integument and in *ino* mutants there is no outer integument. *INO* genes are angiosperm specific as no orthologs have been found in gymnosperms. Thus, it appears that the inner integument of the Angiosperms is equivalent –homologous– to the single integument of the gymnosperms and the PO terms can be revised accordingly.

1. Introduction

Due to the vast terminology available to describe the different plant characteristics, the Plant Ontology (PO; <http://planteome.org>) [1] arose as an initiative to develop a controlled database of plant terminology used to describe the morphology, anatomy and development of plants. The words are additionally arranged in ontologies that is, one term is associated with another term [2–6]. In order to better describe the development, the PO database is also linked with databases that contain genomic information [5] such as The Arabidopsis Information Resource (TAIR; <https://www.arabidopsis.org>) for *Arabidopsis*, Gramene (<http://www.gramene.org>) for rice, MaizeGDB (<https://www.maizegdb.org>) for maize and the Gene Ontology (<http://www.geneontology.org>). The Gene Ontology (GO) database is a valuable source for molecular biology as it provides information about the gene's molecular activities including function, role in a biological process and location in a cellular component [7].

In spite of the enormous diversity of extant plant lineages, most molecular and developmental studies are focused on flowering plants (angiosperms, with 53 genomes annotated and publicly available; <https://phytozome.jgi.doe.gov/pz/portal.html>) and more specifically on *Arabidopsis thaliana*. Outside angiosperms, few genomes have been

sequenced making the GO annotations mostly *Arabidopsis* centered and, furthermore, developmental studies are generally based on model species [8]. Considering the morphological changes through plant evolution and in consequence also in the genetic mechanisms, the available information for *Arabidopsis*, maize and rice, constitute only three distant points across plant evolution and it does not allow us to make a full association between the PO and GO. To better illustrate the idea, we are going to use one of many examples: ovule development in seed plants.

2. A Case study: ovule development in seed plants

After a search in the Planteome database (<http://planteome.org/>) for PO term '*plant ovule development*' (ID: GO:0,048,481; <http://browser.planteome.org/amigo/term/GO:0,048,481>; Fig. 1), the definition of this term is linked with genes involved in its development only for the model species *Arabidopsis* (http://browser.planteome.org/amigo/gene_product/TAIR:locus:2,028,020; Fig. 1d,e). The ovule is a salient feature of seed plants (Fig. 2a) and it has been defined as a megasporangium (nucellus) surrounded by the integument(s) [9,10]. Gymnosperm ovules only have one integument that surrounds the nucellus leaving a small opening called the micropyle distal to the stalk

Abbreviations: GO, Gene Ontology; *INO*, *INNER NO OUTER*; PO, Plant Ontology

* Corresponding author.

E-mail addresses: dws@nybg.org (D. Stevenson), czumajo@nybg.org (C. Zumajo-Cardona).

<https://doi.org/10.1016/j.cpb.2018.09.009>

Received 19 September 2018; Received in revised form 30 September 2018; Accepted 30 September 2018
2214-6628/© 2018 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

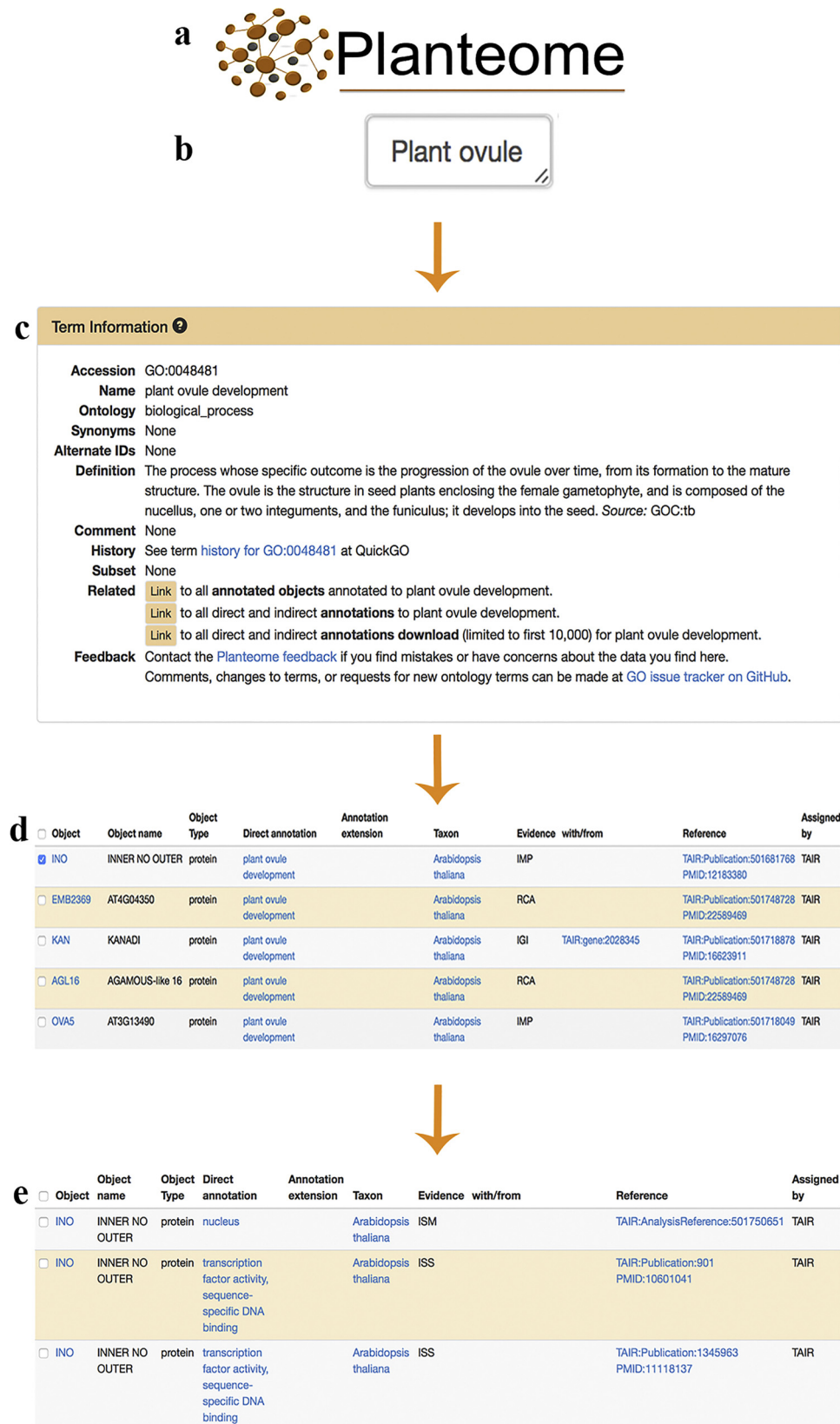


Fig. 1. Example of Plant Ontology (PO) **a.** <http://planteome.org> for **b.** *plant ovule* (PO:0,020,003). **c.** The information includes the Gene Ontology (GO) ID number, definition, related terms (<http://browser.planteome.org/amigo/term/GO:0,048,481>) and **d.** a list of genes involved in the process, five genes are shown here, all of them associated with a taxon: *A. thaliana* found with TAIR database. **e.** Specific information for each of the genes is provided indicating its function and location, information found throughout the GO (http://browser.planteome.org/amigo/gene_product/TAIR:locus:2,028,020).

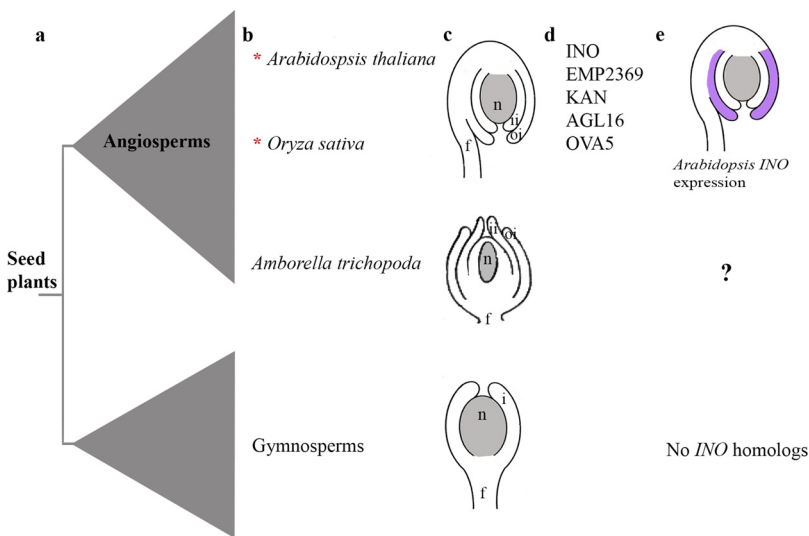


Fig. 2. a Simplified phylogeny of the seed plants. b. selected species or groups with distinctive ovule morphologies. c. schematic representation of the different ovule morphologies here presented. d. genes found with the PO search. e. available information for the gene *INO*, purple indicating expression of *INO* in the Arabidopsis ovule, restricted to the outer integument. No *INO* homologs have been identified in gymnosperms [16]. Red asterisks, species for which molecular information is included in the PO and GO databases; f, funiculus; i, integument; ii, inner integument; n, nucellus; oi, outer integument.

(i.e., the funiculus) bearing the ovule (Fig. 2a-c). In contrast, angiosperm ovules have two integuments and in derived angiosperms such as *Arabidopsis* the ovule is reflexed so that the micropyle is next to the base of the funiculus (Fig. 2c) [9,10]. One question that is still unresolved is which of the two integuments from angiosperms is homologous to the single integument in gymnosperms? The genetic network involved in ovule development has been relatively well studied in *A. thaliana* where expression and functional analyses have been done for multiple ovule developmental genes [11]. *INNER NO OUTER (INO; Arabidopsis thaliana* sequence: AT1G23420) belongs to the *YABBY* transcription factors and it is a gene that is specifically involved in the development of the outer integument of *Arabidopsis* ovules [11–15]. In angiosperm species with only one integument *INO* shows ambiguous expression patterns but it is generally found in the outer most layer of the integument [16,17]. The *YABBY* gene family has suffered multiple duplication events before the diversification of angiosperms, in fact *INO* seems to be angiosperm specific as no *INO* orthologs have been found in gymnosperms [18]. The same pattern is found in other genes of the ovule developmental network. For example, *ABERRANT TESTA SHAPE (ATS = KANADI 4; Arabidopsis thaliana* sequence: AT5G42630) is a gene involved in the separation layer between the two integuments of *Arabidopsis*, *ats* mutants develop both integuments fused [19] and it is also angiosperm specific (Zumajo-Cardona and Ambrose *in prep.*). With the data gathered so far, it is likely that the genes involved in the proper development of the outer integument are the result of a neo-functionalization.

The genes associated with ovule development in plants found with the PO search are from *Arabidopsis* (Fig. 2d) thus, it is still unclear if the function of these genes is conserved across seed plant evolution, or which genes are involved in ovule development in species with different ovule morphologies such as *A. trichopoda* or gymnosperms. With the expression and functional results from *Arabidopsis* and available data about the evolution of these genes (Fig. 2e), one can hypothesize that the integument of gymnosperms is homologous to the inner integument of angiosperms and that the outer integument is the result of a neo-functionalization event in these gene families. However, gene expression data constitute a useful tool to hypothesize homologous relationships but if used in isolation it can be misleading [20], which corroborates the need for standardizing genomic tools, such as *in situ* hybridization expression analyses and gene knockdowns, in non-model organisms [8].

3. Conclusion

Here we present one of many instances for which having genetic information across different plant lineages permit us to have a better

understanding of the different processes that lead to the morpho-anatomical evolution and development of certain traits. The PO constitute a very important tool for all plant biologists, in consequence, the next goal for the database is to be able to provide information about homology as well. It is crucial extend different genomic tools to other non-model species, including genome and transcriptome annotations, expression analyses and gene knockdowns. In doing so we will be able to better integrate the PO with the GO making both more useful to a broader user group of plant biologists.

Conflict of interest

None.

Acknowledgment

The authors thank the two anonymous reviewers for their helpful suggestions and comments on the initial version of this manuscript. This work was supported in part by the National Science Foundation - NSF Grants: DBI-0822201, IOS-1340112 EF-0629817, IOS-0421604, and IOS-0922738 to DWS.

References

- [1] L. Cooper, A. Meier, M.A. Laporte, J.L. Elser, et al., The Planteome database: an integrated resource for reference ontologies, plant genomics and phenomics, *Nucleic Acids Res.* 46 (2017), <https://doi.org/10.1093/nar/gkx1152>.
- [2] The Gene Ontology Consortium, Gene Ontology: tool for the unification of biology, *Nat. Genet.* 25 (2000) 25–29.
- [3] The Gene Ontology Consortium, Creating the gene ontology resource: design and implementation, *Genome Res.* 11 (2001) 1425–1433.
- [4] The Gene Ontology Consortium, The plant ontology consortium and plant ontologies, *Comp. Funct. Genom.* 3 (2002) 137–142.
- [5] P. Jaiswal, S. Avraham, K. Ilic, E.A. Kellogg, S. McCouch, A. Pujar, L. Reiser, S.Y. Rhee, M.M. Sachs, M. Schaeffer, L. Stein, P. Stevens, L. Vincent, D. Ware, F. Zapata, Plant Ontology (PO): a controlled vocabulary of plant structures and growth stages, *Comp. Funct. Genom.* 6 (2005) 388–397.
- [6] R.L. Walls, B. Athreya, L. Cooper, J. Elser, et al., Ontologies as integrative tools for plant science, *Am. J. Bot.* 99 (8) (2012) 1263–1275.
- [7] Gene Ontology Consortium, The Gene Ontology (GO) database and informatics resource, *Nucleic Acids Res.* 32 (2004), <https://doi.org/10.1093/nar/gkh036>.
- [8] S.A. Rensing, Why we need more non-seed plant models, *New Phytologist* 216 (2017) 355–360.
- [9] A. Foster, E.M. Gifford, *Comparative Morphology of Vascular Plants*, 2nd edition, W. H. Freeman & Co., 1974.
- [10] D. Soltis, P. Soltis, P. Endress, M. Chase, S. Manchester, W. Judd, L. Majure, E. Mavrodiev, *Phylogeny and Evolution of the Angiosperms*, The University of Chicago Press, 2018.
- [11] S.C. Baker, K. Robinson-Beers, J.M. Villanueva, J.C. Gaiser, C.S. Gasser, Interactions among genes regulating ovule development in *Arabidopsis thaliana*, *Genetics* 145 (1997) 1109–1124.

- [12] J.M. Villanueva, J. Broadhvest, B.A. Hauser, R.J. Meister, K. Schneitz, C.S. Gasser, *INNER NO OUTER* regulates abaxial-adaxial patterning in *Arabidopsis* ovules, *Genes Dev.* 13 (1999) 3160–3169.
- [13] R.J. Meister, L.M. Kotow, C.S. Gasser, *SUPERMAN* attenuates positive *INNER NO OUTER* autoregulation to maintain polar development of *Arabidopsis* ovule outer integuments, *Development* 129 (2002) 4281–4289.
- [14] R.J. Meister, L.A. Williams, M.M. Monfared, T.L. Gallagher, E.A. Kraft, C.G. Nelson, C.S. Gasser, Definition and interactions of a positive regulatory element of the *Arabidopsis* *INNER NO OUTER* promoter, *Plant J.* 37 (2004) 426–438.
- [15] J.M. McAbee, T.A. Hill, D.J. Skinner, A. Izhaki, B.A. Hauser, R.J. Meister, G. Venugopala Reddy, E.M. Meyerowitz, J.L. Bowman, C.S. Gasser, *ABERRANT TESTA SHAPE* encodes a KANADI family member, linking polarity determination to separation and growth of *Arabidopsis* ovule integuments, *Plant J.* 46 (2006) 522–531.
- [16] J. Lora, J.I. Hormaza, M. Herrero, Transition from two to one integument in *Prunus* species: expression pattern of *INNER NO OUTER* (*INO*), *ABERRANT TESTA SHAPE* (*ATS*) and *ETTIN* (*ETT*), *New Phytologist* 208 (2015) 584–595, <https://doi.org/10.1111/nph.13460>.
- [17] D.J. Skinner, R.H. Brown, R.K. Kuzoff, C.S. Gasser, Conservation of the role of *INNER NO OUTER* in development of unitegmic ovules of the Solanaceae despite a divergence in protein function, *BMC Plant Biol.* 16 (2016).
- [18] C. Bartholmes, O. Hidalgo, S. Gleissberg, Evolution of the YABBY gene family with emphasis on the basal eudicot *Eschscholzia californica* (Papaveraceae), *Plant Biol.* 14 (2012) 11–23.
- [19] J.M. McAbee, T.A. Hill, D.J. Skinner, A. Izhaki, B.A. Hauser, R.J. Meister, et al., *ABERRANT TESTA SHAPE* encodes a KANADI family member, linking polarity determination to separation and growth of *Arabidopsis* ovule integuments, *Plant J.* 46 (2006) 522–531.
- [20] M.A. Jaramillo, E.M. Kramer, The role of developmental genetics in understanding homology and morphological evolution in plants, *Int. J. Plant Sci.* 168 (1) (2007) 61–72.