

**Minutes of the First Plant Ontology Consortium Meeting
XII PAG Conference, San Diego CA, Jan 11, 2004**

Esquire Room, Town & Country Convention Center, San Diego, CA

8:30-11:00 am, and 3:30-6:00 pm on Sunday, January 11, 2004

PARTICIPANTS:

Cornell University
Pankaj Jaiswal

CSHL
Doreen Ware, Shuly Avraham

Carnegie Institution of Washington
Sue Rhee, Leonore Reiser, Katica Ilic

University of Missouri, St. Louis
Toby Kellogg

University of Missouri, Columbia & MaizeGDB
Mary Polacco, Leszek Vincent

Maize Genetics Cooperative (USDA)
Marty Sachs

Pioneer Hi-bred International
Mark Whitsitt

Other participants

Plant Research International/PlaNet Consortium, NL
Paulien Adamse,

NASC/PlaNet Consortium, UK
Beatrice Schildknecht

AGENDA

Morning Session, 8:30-11 AM

Group Introductions (15 minutes)

Brief report from St. Louis meeting (overview 5-10 minutes)

Discussion of Big Issues from St. Louis Meeting

New Relationship types (is sometimes a part of, contained in)

How to address 'combinatorial annotations (e.g. YFG1 expressed in tapetum during floral development stage 7).

Afternoon Session, 3:30-6:00 PM

Discussion continued (1 hour)

Formation of ontology working groups

Allocating additional (non-ontology related) responsibilities

- Website

- Documents

- Database

- Outreach

Logistics/Future planning

- Deadlines

- Release policy (Website + Ontology)

- Future plan for meetings- what meetings to attend, when/where workshops

MINUTES

1. Group Introductions
2. Plant ontologies contributed by Pioneer Hi-bred and Monsanto (report by Doreen). Each company has an FTE to dedicate to the PO project.
3. Brief report from St. Louis meeting (overview by Pankaj)

Major issues addressed: high nodes of the ontology, relationship types.

4. Discussion on relationship types. What kind of relationships types will be used?

Suggestions from the meeting in St. Louis:

In addition to PART_OF relationship, there is a need for alternative relationship types, such as SOMETIMES_A_PART_OF. Example: embryonic leaf in maize is sometimes_a_part_of embryo. Not all embryos have embryonic leaves, and therefore, (always) a_part_of would violate true path rule.

Another alternative type of relationship suggested is CONTAINED_IN. Example: An embryo is contained in embryo sac.

Sue raised an issue of the concept of mereology - where for example the relationship PART_OF could have multiple representations, e.g., the blade is a physical component of the leaf, embryo is contained within the embryo sac, and the petal is a member of the perianth. In this case, a_part_of relationship type would have a broader meaning and would be sufficient to describe the relationships such as sometimes_a_part_of and contained_in.

5. Three ontologies (Arabidopsis, rice and maize) are being merged into one. Goal is to begin building single shared plant ontology.

Q: What is the general framework of the ontology?

How to built plant anatomy ontology initially and to what extent do we want to cover all Angiosperms?

Toby: Start with a structure that can be easily expanded to a broader range of Angiosperm species. We need to look down the road at other plants in the works.

Two issues to consider: variation between organisms and how the ontologies are going to be used.

Q: What are the criteria for including a term in the ontology?

We agreed on the 4 criteria to be considered when defining the terms: morphological features, anatomical/histological features, location and derivation (always keeping in

mind the true path rule). There would be no restriction on the current terms in use by the 3 contributing ontologies. In part because these terms are in fact in use by member groups. Some decisions can be made during the merge about which terms constitute a synonym and which is the primary term. The definition will determine what terms are synonyms (same definition) and what are not (different meanings of the term).

For each term, we need to include what criteria were used, and it should be a part of definition, therefore keeping a strict format for definition.

6. How to address 'combinatorial annotation'?

Example: A search for genes that are expressed in cotyledons (anatomy term) at the late heart-shape embryo stage (developmental term).

Currently, there is no facility to do either composite annotation or combinatorial terms.

Composite annotation - using terms from several ontologies in conjunction to make an annotation.

Combinatorial terms - taking terms from two orthogonal (independent) ontologies and creating a third term. For example, taking embryo term and a developmental stage term and making a third term.

Leonore: Right now we don't have tools to do either composite annotation or combinatorial terms programmatically. It may be that we will have to include combinatorial terms and then dissect them later. From recent GO meeting, it sounds like the data structure may be close for accommodating composite annotations. It is unlikely that the combinatorial terms will ever come to pass.

We need to know how the ontology will be applied to come up with the idea of how it would be built. How the query is done is critical for defining the structure. Two ways: How do you conduct a query, and how do you structure the ontology.

Mike: (on how to structure the ontologies to work with the query tools that we have) It is better to develop a query tool so that users can redefine a query, than opposite.

This would require adaptations to the software.

Doreen: How much development will we need to do in order to be able to satisfy the queries we want to have? How much development on the structures and curation tools? Are we asking for things - will this require adaptations to the software or can we think about how we will eventually define this in the future.

7. Species-specificity issue (an issue of granularity)

Q: How far do we go with granular terms?

Toby (on generic ontology and species-specific ontology): There is a problem between maize and Arabidopsis that needs to be dealt with immediately (one example is maize spikelet).

We do need to go down to granularity to allow species-specific annotation.

If the structure exists in one species but not in the other, that structure should be included in ontology. There is a need for using some sort of association to taxonomic terms that would bring out the species specificity. This would enable people to derive a species-specific ontology from the generic PO ontology. The level of granularity that is required could be decided on a species basis, in contrast to GO where the cut off is a few taxonomic levels up at the level of family.

Idea of introducing SLIM terms, as higher-level terms, defined on GO project (to check out what are the GO criteria for SLIM terms).

8. Serial homology

This issue was also brought up at the meeting in St. Louis.

Example: Petal is an instance of modified leaves. The relationship is included in the definition of the term.

9. Issue of synonyms

To define what is instance and what is synonym, and also to define what we mean by synonym (see GO web site <http://www.geneontology.org/GO.synonyms.html#type>).

Toby: An example with lodicules vs. petals. Is it instance or synonym? Lodicule can be instance of petals and it can also be a synonym of petal. The question is, which one is better? This can be decided later.

Examples: (1) Primary term is inflorescence, and tassel is then a synonym of this general term. (2) Leaf for instance can be a general term and it can be a specific term.

SUMMARY:

- Building the plant anatomy ontology - idea is a broad-spectrum angiosperm anatomical ontology.

- To use explicit wording and four criteria for defining terms (morphological features, anatomical/histological features, location and derivation (the criteria will be included as a part of the definition in a format that will be easily parsable later. Something like CRITERIA: XXX, XXX, XXX, XXX).

- There is no need for using SOMETIMES_A_PART_OF, instead, use A_PART_OF with a broader meaning (sometimes can be a part of it, and thus we will not be able to hold the true-path rule here). Also, CONTAINED-IN type of relationship will not be

used at this point. Suggestion to look at Stanford Encyclopedia of Philosophy (mereology) to find out to see if A_PART_OF is sufficient or whether we want to add additional types of part-of relationship.

- The level of species-specificity/granularity that is required could be decided on a species basis. Granularity should be sufficient to allow species-specific annotation.

- We will create terms within the anatomy ontology that may later need to be separated into separate ontologies rather than deciding on combinatorial terms or composite annotations at this point.

Action items

1. Formation of ontology working groups

In addition to the working groups that were formed at St Louis meeting in December, we need to actively include Monsanto/Pioneer people into project and get them in working groups (Katica to contact both Monsanto - Alice Augustine, and Pioneer Hi-bred International - Mark Whitsitt).

2. Allocating additional (non-ontology related) responsibilities

A. Website

- Needs documentation (see below).

- Katica has access to html documents from CVS repository.

- SOP - first make changes, update the CVS on dev and then make updates on the production server.

B. Documents to be included/modified on the web site:

- Project overview - to be expanded
- Developer's guide needs to be edited
- DB cross references to be updated
- Help documents for AmiGO (Shuly/Katica)

C. Database:

- Shuly will be working on developing scripts for quality assurance, data checks, loading and visualization. This includes updating the AMIGO browser to show what is currently available for the ontologies and to make it PO rather than GO specific. She will work with Katica to ensure the formats.

3. Logistics/Future planning/Deadlines

3.1. After first three months, get out the new PO and test the annotation.

Version 1.0 is to be released before first users workshop meeting (ASPB meeting, July 24-28). This version should be browsable and would only address Anatomy not Development/Growth stages.

Specific tasks:

- Needs to be three-collaborative effort
- Species level terms added
- Sort synonyms
- Incorporate Pioneer/Monsanto ontologies (some 4000 terms available from Monsanto)
- Software and ontology testing period (to be completed by mid-June)
- What is the definition for the first version release? It will also include Monsanto/Pioneer ontologies.
- Release policy (Website + Ontology)

3.2. Within next 6 months all three Dbs will gradually convert to POC. TAIR will start working on this right away to minimize error propagation and to avoid duplication of efforts.

4. Non-ontology related issues

- Archive: To keep track of all the contacts and announcements
- Active bug tracking tool
- Next in-person meeting to be held in St. Louis by the end of May 2004.
- User's workshops
- First PO Workshop for users (1 to 1 1/2 hours), at the next ASPB meeting, July 24-28, 2004.

Who will be the audience and what will the participants get at this workshop?

Lenore: Developers and annotators and not explicitly consumers (e.g., end-users) of annotations. This defines the content of the workshop.

Next NSF PGRP meeting (presentation).