Legumes: From Beans To Trees

- Legumes -- third largest plant family -- beans, peas, ground nuts, forage plants and trees.

- SoyBase -- USDA ARS Soybean Genetics and Genomics Database (soybase.org)

- Primarily a species DB, but strive to establish a controlled vocabulary (ontology) of traits that could be extended for the benefit of other legume communities specifically the beans, peas and peanut.

- Two types of users -- breeders and molecular biologists -- one goal (crop improvement) -- two methodologies and more or less separate literature and traditions.

Challenges

- Accurate and precise description of phenotypes -- "dwarf", "witches broom"

- Accommodation of the instability of taxonomy and nomenclature -- taxonomic revisions, teleomorphic vs. anamorphic binomials in fungi

- Semantic reasoning -- can we construct a "high level" ontology that others could assign subclass associations?

- Accommodation of common names and 'field' terms -- "pod maturity", "brown spot"

SoyBase Soybean Ontologies

- A first suggestion of a controlled vocabulary for soybean and other legume communities
- Four divisions
  - Whole plant growth
    - Keyed to the BBCH scale and the PO whole plant growth ontology where applicable
  - Developmental ontology for soybean growth stages
  - Structure ontology to tag each structural metamer
  - Trait ontology for soybean growth and developmental traits
    - Trait ontology titles and definitions are as organism neutral as possible (node number, inter-node length not "dwarf")
    - Associated with the Plant Ontology as closely as possible by use of xrefs and synonyms
    - Titles do not use common names but rather binomials or anatomic designations
    - Linked to the soybase trait class by xref assignments -- trait class contains the common names expected by breeders
    - Single inheritance -- to allow reasoning