Ontology - Crop Breeding Perspective

Peter Jack

Ontology Workshop
Hinxton, 8-9th Dec 2011

My Role & Limitations of Research in Breeding Companies…

- Genotyping lead for RAGT cereal crops
  - Development of markers
  - Application in breeding programmes
- Main crop is hexaploid wheat
  - But comments will apply to most field crops
- Royalty return of inbred crops limits in house research
  - Public sector research collaborations are critical
  - Need to be able to integrate results across labs and across traits
  - Consistent *trait* description critical

Breeding & Trait Specification

- Crop Breeding is a Commercial Activity
  - Failure to sell varieties to farmers > insolvency

- Farmer informed by HGCA Recommended List (RL) of Varieties
  - This defines variety specification for breeder to select for
  - Essentially performance traits
    - Agronomic: Yield, Disease Resistance, Phenology, Lodging (often relative scale eg 1-10)
    - End user requirements: Functionality of grain
    - Very few relate to variation in Ontologic descriptors (grain size may be exception)
  - Downloadable from HGCA site

- Varieties also need to pass DUS (Distinctness, Uniformity, Stability)
  - Morphologic traits captured in detail (every minute part of ear named!)
  - Although closer to Ontologic characters, important to remember that they do not relate to productivity – just descriptions of a variety.

Breeding, Genomics & Need for Ontology?

- Marker-assisted breeding becoming increasingly important
- Key traits usually quantitative. Steps in marker development:
  - Develop segregating populations & phenotype
    - Sub-traits sometimes identified: eg yield > grain/ear or grain/m2, grain size, biomass...
  - Genotype & QTL identification
  - QTL validation (BC generation of NILs, use of advanced breeding lines…)
  - Fine mapping (ideally gene identification!)
  - > Diagnostic Marker
  - Long process!
- High marker density SNP platforms > AG approaches feasible
- And NGS/BSA
  - May be possible to get close to genes in 1 step
  - Useful to be able to identify candidate genes in a region
    - from gene/trait relationship in other species eg rice>wheat grain size.
  - Consistent ontology may be important to make correct queries?

Future….?

- If
  - productivity traits can be reduced to ontologic entities
  - and genes for these entities identified
  - and mechanism of action/biochemistry of genes determined
- Then
  - can improved alleles be predicted (sequence change)
  - and improved alleles be mined or generated (mutagenesis)??

QTL database

- Not specifically an ontology issue but illustrates common problem
- Pressing need to collate UK public QTL data into single database.
  - Whole is greater than sum of parts
  - Query 1 – show me all QTLs in specific region
    - To look for trait interactions (eg yield/quality/disease)
  - Query 2 – show me all regions for specific trait
    - To look for pedigree relations, GxE effects ….
- Key UK wheat population is Avalon/Cadenza
  - Many groups using same pop/genotype but for different traits
  - Should be simple to integrate data
  - No attempt to coordinate this – lots of excel files….
- AG work will present same type of problem

- Perhaps this could be anearly step in relating ontology to phenotype ie collecting the phenotype in a single location?

Conclusions

- Can standard Ontologies be mapped onto existing & accepted phenotypes (RL, Zadoks, Phenology)
- May be difficult for productivity traits directly
  - Need to agree sub-traits (& their naming!)
- Ontology is not the only issue eg QTL database for collation of phenotype