




A brief update on data curation activities associated with the www.brassica.info website, managed in the UK at Rothamsted Research for the Multinational *Brassica* Genome Project (MBGP).

Editor: Graham King (graham.king@bbsrc.ac.uk)

Webmaster: Pierre Carion (pierre.carion@bbsrc.ac.uk)



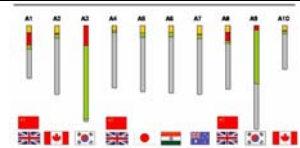
N.B. The value of the information made available via www.brassica.info is determined by the quality and coverage of the information made accessible by the research community. There are a number of areas we have identified that particularly require input from researchers elsewhere.

These are indicated by 

Updates and information now available:

BrGSP BAC Registry

<http://www.brassica.info/resource/sequencing.php>



Update on status of *B. rapa* Genome Sequencing project

2008

- which BACs are in pipeline (n=2,056)
- completed BACs, GenBank accession
- assignment to chromosomes, genetic map location, map graphics
- FISH assignment



2009

- Ensembl *Brassica*-centric viewer, to enable navigation and alignment to consensus genetic maps and conserved collinearity blocks

Gene Nomenclature

<http://www.brassica.info/info/reference/gene-nomenclature.php>

Update on status of *B. rapa* Genome Sequencing project

2008

- Agreed and published in Plant Methods

2009

- Online Gene Registry in progress





TILLING consortium

<http://www.brassica.info/research/activities/tilling.php>

Update of TILLING populations

- 2008**
- Established by Lars.Ostergaard@bbsrc.ac.uk
 - 17 populations collated, contact details etc
- 2009**
- Compile information on mutant load
 - Screening availability?



Genetic and genomic resources

Inventory and detailed curated datasets

Clone libraries

- <http://www.brassica.info/resource/clones.php>
- Provenance, status, ownership, etc
 - BACs
 - ESTs
 - GSS



Plant populations

- Mapping populations
[brassica.info/resource/plants/mapping-populations.php](http://www.brassica.info/resource/plants/mapping-populations.php)
- description, size, status, nomenclature
 - plant line names, subset populations used for different maps etc
 - seed and DNA availability

Diversity Sets

http://www.brassica.info/resource/plants/diversity_sets.php

- description, lines, numbers, status, provenance
- status of genetic fixing
- seed and DNA availability

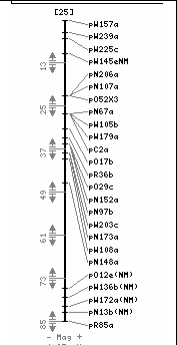




Genetic maps

<http://www.brassica.info/resource/maps.php>

- Nomenclature
- Assignment to populations and sub-populations
- Linkage groups and map positions
- Assignment of marker loci to marker_assays
- Downloadable data + initial CMAP output



2009

- Extend to include consensus integrated maps, links to genome sequence
- More sophisticated map interface

Genetic markers

<http://www.brassica.info/resource/markers.php>

- Assignment of marker_assays to 'canonical marker name' and hence to sequence assignment - for map integration and comparative analysis
- Marker sets - details of sequence-tagged genetic markers, with primer information

2009

- Extend to include MBrGSP SSRs that anchor genome sequence to genetic maps



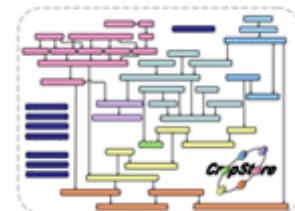
Explicit Nomenclature & Curation for Managing Integrated Genetic Datasets

Curation pipeline and Database

<http://www.brassica.info/resource/databases/cropstore.php>

2008

- Integrated schema to manage plant, map, marker, trait datasets
- Established nomenclature conventions and MBGP registries
- Input templates (MS Excel) available for download and data entry
- Web-based interface to download datasets in MS Excel, with detailed provenance data



2009

- Proposal and agreement on wider range of nomenclature conventions
- PC-based interface will be available for distribution; MySQL dumps to be available for local implementation



Reference Information

<http://www.brassica.info/information.php>

Look-up data for consistency

2008

Gene Nomenclature
Genome sizes
Linkage group assignments
Nomenclature conventions

2009



Publications

Publications only available via
www.brassica.info

2008

Cruciferae Newsletter

- vols 26-27 available for download
- vols 22-25 archive issues scanned (at INRA) and available

Cruciferae Trait Genetics (from AAFC)

- available for download

Guide to Wild Germplasm (from AAFC)

- available for download



2009

Cruciferae Newsletter

- vol 28 (Jan 2009) already available

Planned for 2009-

New datasets, activities and services

2009

- Map integration
- SeqStore for *Brassica*, to include concept of Registry 'Sequence Collections'
- QTL repository
- Ionomics datasets
- Links to transcriptomic datasets ?

