

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)

For 2009 edition, please see <http://www.brassica.info/info/bulletin.php>

Editor: Graham King (graham.king@bbsrc.ac.uk) - **Webmaster:** Pierre Carion (pierre.carion@bbsrc.ac.uk)

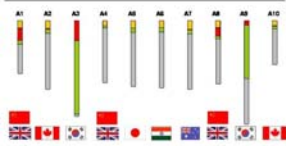


N.B. The value of the information made available via www.brassica.info is determined by the quality and coverage of the information made accessible by the research community. There are a number of areas we have identified that particularly require input from researchers elsewhere.

These are indicated by: 

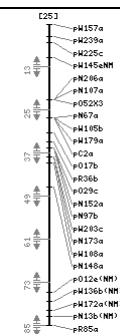
CS : indicates curated within CropStore_{DB} system (<http://www.cropstoredb.org/>)

Updates and information now available:

Ongoing <i>Brassica</i> sequencing projects	
http://www.brassica.info/resource/sequencing.php	
Contact: jun.wang@bbsrc.ac.uk	
Update on status of <i>B. rapa</i> Genome Sequencing project	
2009	<ul style="list-style-type: none"> • Last updated April 2009 <ul style="list-style-type: none"> ○ BACs in pipeline 2,426 ○ Sequenced = 1,819 ○ In GenBank = 968 ○ Sequence-tagged markers anchored to BACs (972)
2010	<ul style="list-style-type: none"> • Summary of all known ongoing <i>Brassica</i> sequencing efforts
 <i>updated regularly</i>	

BrassEnsembl	
http://www.brassica.info/BrassEnsembl/index.html	
Contact: chris.love@bbsrc.ac.uk	
2009	<p>Established to display the reference genome sequence and annotation, BrassEnsembl currently includes:</p> <ul style="list-style-type: none"> • 17, 512 gene models identified by PASA algorithm provided by JIC • 18, 554 gene models identified by SNAP algorithm • 77.9% (13, 642) with functional annotation from homology to <i>Arabidopsis</i> genes • Alignments to: <ul style="list-style-type: none"> ○ 17,775 <i>Arabidopsis</i> coding sequences, with links to AtEnsembl ○ 6, 895 of the 95K <i>Brassica</i> unigene set ○ 681,689 <i>Brassica</i> ESTs ○ 174, 907 <i>B. rapa</i> BAC end sequences
2010	<ul style="list-style-type: none"> • Work closely with EBI Plant Genomes and others to provide a focus for community-wide integration of information. • Alignments to new public access Affymetrix Exon array (~140k) <ul style="list-style-type: none"> • <i>B. rapa</i> sequence scaffolds (when released) • (P)re-sequencing data (<i>B. rapa</i>, <i>B. napus</i>, <i>B. oleracea</i>) • Orthologous and paralogous phylogeny data
<i>Developed as part of BBSRC funded UK contribution to the MBrSP</i>	

Genetic and genomic resources	
Inventory and detailed curated datasets	
Clone libraries	<p style="text-align: right;">http://www.brassica.info/resource/clones.php</p> <p>Provenance, status, ownership, etc</p> <ul style="list-style-type: none"> • BACs, ESTs, GSS
Plant populations	<p>Mapping populations http://www.brassica.info/resource/plants/mapping-populations.php</p> <ul style="list-style-type: none"> • <i>B. napus</i> : 16 population (40 subsets) • <i>B. oleracea</i>: 9 population (19 subsets) • <i>B. rapa</i> : 4 populations (14 subsets) • <i>B. juncea</i> : 1 population <p>Diversity Sets http://www.brassica.info/resource/plants/diversity_sets.php</p> <ul style="list-style-type: none"> • <i>B. napus</i> : 1 population (3 subsets) • <i>B. oleracea</i>: 1 population (5 subsets)
Genetic maps	<p style="text-align: right;">http://www.brassica.info/resource/maps.php</p> <p>Assignment of marker loci to marker assays, sequence id, Arabidopsis and <i>Brassica</i> BLAST hits</p> <p>2009</p> <ul style="list-style-type: none"> • Downloadable data + CMAP output <ul style="list-style-type: none"> • <i>B. juncea</i> : 1 • <i>B. napus</i> : 16 • <i>B. oleracea</i> : 8 • <i>B. rapa</i> : 9 • Data used for CMap3D at ACPFG http://acpfg.imb.uq.edu.au/cmap3d.php <p>2010</p> <ul style="list-style-type: none"> • Include consensus integrated maps • Links to BrassEnsembl • Genotype matrices available for download
Genetic markers	<p style="text-align: right;">http://www.brassica.info/resource/markers.php</p> <p>2009</p> <ul style="list-style-type: none"> • Marker sets - SSR and others updated in downloadable format • BAC anchor markers <ul style="list-style-type: none"> ◦ BraCKDH_01_2009b (185) ◦ BnaTNDH_05_2008b (80)
Trait data	<p>2009/10</p> <ul style="list-style-type: none"> • Web interface (http://www.brassica.info/CropStore/traits.php) • <i>B. napus</i> TNDH 'IMSORB' data • Prioritisation of datasets from OREGIN project (http://www.oregin.info/resources/traits.php) <p>2010</p> <ul style="list-style-type: none"> • Ionomics data (mineral content - diversity & mapping populations) • Phosphate use efficiency • Fatty acid composition (diversity & mapping)
QTL repository	<p>2009/10</p> <ul style="list-style-type: none"> • Web interface (http://www.brassica.info/CropStore/qtl.php) • Qiu et al (2006) BnaTNDH • Seed oil content and composition <p>2010/11</p> <ul style="list-style-type: none"> • Additional datasets • CMap and other graphic visualisation and navigation





<http://www.cropstoredb.org/>

Curation pipeline and Database

- 2009**
 - Simplified management of marker data
 - Trials and QTL schema refined
 - Input templates 'road tested'

- 2010**
 - Paper describing CropStore_{DB} and nomenclature conventions
 - Additional interfaces and navigation to BrassEnsembl



SeqStore_{DB}, AlignStore_{DB}

- 2009** Databases developed to manage *Brassica* sequence data and alignments as reference resource to assist in map integration- retaining versioning and explicit sequence collections (e.g. library sets, marker sets)
- 2010** Publicly accessible with regular updates, with web interfaces integrated with BrassEnsembl and CropStoreDB



Transcriptomics

- 2009** Updates on transcriptomic platforms
- 2010** Affymetrix Exon array and service

Update of TILLING populations

- 2009**
 - TILLING screening service from RevGenUK
- 2010**
 - *B. rapa* R-o-18 Illumina 12x sequence online
 - http://www.brassica.info/datasets/Brassica_resequencing_data/



Publications

Publications only available via www.brassica.info

- 2009**
- Cruciferae Newsletter**
 - vol 28 (Jan 2009)
 - Cruciferae Trait Genetics** (from AAFC)
 - available for download
 - Guide to Wild Germplasm** (from AAFC)
 - 3rd edition (published 2009) available for download
 - Brassica crosses**, by A.W. Sutton, published 1908 in J. Linnean Society, with photographs
- 2010**
- Please send us any publications you wish to be made available on www.brassica.info.
 - Back catalogue of **Cruciferae Newsletter** + vol 29 (Mar 2010)
 - We are also making space for relevant posters and presentations.
<http://www.brassica.info/posters.php>

**Planned for 2010****New datasets, activities and services**

- 2010**
- Map integration
 - SeqStore for *Brassica*, includes concept of Registry 'Sequence Collections'
 - QTL repository
 - Ionomics datasets
 - Genotype data
 - Links to transcriptomic datasets?
 - Any suggestions or datasets welcome.

Mailing list

subscribe <http://www.brassica.info/info/mailing.php>

- Currently ~290 members



Geographical distribution of the www.brassica.info visitors (4 Jan 2010)