brassica.info

Bulletin

Jan 2011
(4th Edition)

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

For past editions, 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.
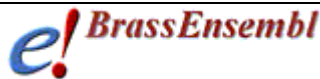
These are indicated by:

CS *: indicates curated within CropStore_DB system (http://www.cropstoredb.org/)*

Updates and information now available:

Ongoing *Brassica* sequencing projects	
http://www.brassica.info/resource/sequencing.php **Update on status of Genome Sequencing projects**	**Contact:** chris.love@bbsrc.ac.uk

2010 • Summary of all known ongoing *Brassica* sequencing efforts, including reference genotypes, and re-sequencing
 o Sequence data to download
  ▪ *B. rapa* R-o-18 re-sequencing from Rothamsted
  ▪ *B. oleracea* TO1000 from Warwick Uni.
 o Links to other downloadable genome sequences	

2011 • Updates on data in public domain as become available	*updated regularly*

Planned for 2011
New datasets, activities and services
2011 • SeqStore for *Brassica* online, includes concept of Registry 'Sequence Collections' • QTL repository • Ionomics datasets • Trait Data from other repositories • Genotype data • Links to transcriptomic datasets? • Any suggestions or datasets welcome.

Mailing list	
subscribe http://www.brassica.info/info/mailing.php • Currently ~510 members	
Geographical distribution of visitors (13 January 2011):	

BrassEnsembl	

<table>
<tr><td>

http://www.brassica.info/BrassEnsembl

2010 BrassEnsembl is the Ensembl based genome browser currently displaying *B. rapa* BAC sequence and annotation including:
- 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 *Arabidopsis* genes
- Blast interface for alignment against:
  - Gene and BAC sequence datasets
  - *B. rapa* tilling line R-o-18
- Additional tracks displaying alignments to:
  - 17,775 *Arabidopsis* coding sequences, with links to AtEnsembl
  - 6, 895 of the 95K *Brassica* unigene set
  - 681,689 *Brassica* ESTs 174, 907 *B. rapa* BAC end sequences
  - 135K *Brassica* unigene set (used in design of the Brassica Affymetrix Exon array)
  - 174, 907 *B. rapa* BAC end sequences

2011 Following release of the *B. rapa* genome paper, BrassEnsembl will also include:
- IVF-CAAS sequence scaffolds displayed as pseudochromosomes with gene features, as presented within *B. rapa* genome paper.
- Gene annotation derived through significant alignment to Arabidopsis genes
- Gene Ontology annotation for genes derived through alignment to SwissProt/Trembl
- Discussions with EBI are on-going regarding integration with EnsemblPlants and their Compara pipeline
- Integration of genomic sequence data with associated genetic maps and trait data via CropStoreDB, AlignStoreDB and SeqStoreDB. See InterStoreDB.

2012
- Scope for addition of *B. oleracea*, *B. napus* reference genomes as become available
- Integration with *B. rapa*, *B. oleracea*, *B. napus* re-sequencing data
- Incorporation of data from a standard MBGP annotation pipeline into BrassEnsembl
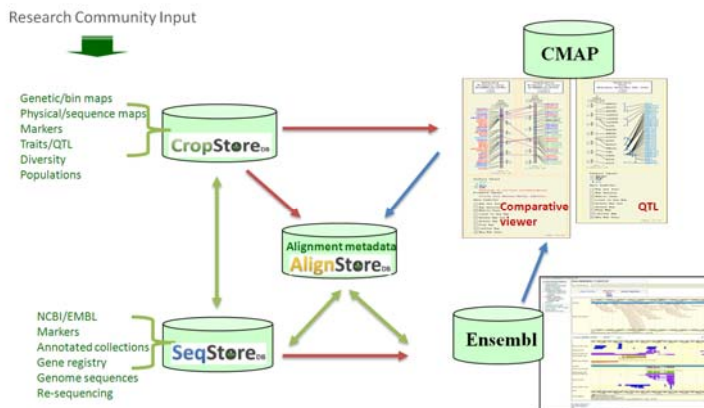
</td><td>

Contact:
chris.love@bbsrc.ac.uk

Developed as part of BBSRC funded UK contribution to the MBrSP and continuing at Rothamsted

</td></tr>
</table>

SeqStoreDB, AlignStoreDB	
2010 Development of InterStoreDB to link federated databases	
InterStoreDB:	

2011 Publicly accessible with regular updates, with web interfaces integrated with BrassEnsembl and CropStoreDB (Data structures and code managed under SVN)

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 http://www.brassica.info/resource/plants/mapping-populations.php • *B. napus* : 16 population (40 subsets) • *B. oleracea:* 9 population (19 subsets) • *B. rapa* : 4 populations (14 subsets) • *B. juncea* : 1 population Diversity Sets http://www.brassica.info/resource/plants/diversity_sets.php • *B. napus* : 1 population (3 subsets) • *B. oleracea:* 1 population (5 subsets) Genetic Resource Collections • European Brassica Database (*ex situ* genetic resources) **2011** • *B. rapa* Diversity Set (from Rothamsted) • USDA/ARS-GRIN Genetic Resource Collection	
Genetic maps	http://www.brassica.info/resource/maps.php Downloadable data + CMAP output *B. juncea* : 1 | *B. napus* : 16 | *B. oleracea* : 8 | *B. rapa* : 9 **2011** • Include consensus integrated maps • Links to BrassEnsembl • Genotype matrices made available for download • updates as become available	
Genetic markers	http://www.brassica.info/resource/markers.php • Marker sets - SSR and others updated in downloadable format **2011** • PFM marker set • SNP data as become available	
Trait data	• Web interface (http://www.brassica.info/CropStore/trait.php) • B. napus TNDH IMSORB data • EU-ResGen datasets **2011** • Ionomics data (mineral content - diversity & mapping populations) • Phosphate use efficiency • Fatty acid composition (diversity & mapping) • EU-ResGen datasets • GRIN datasets	
QTL repository	• Web interface (http://www.brassica.info/CropStore/qtl.php) • Qiu et al (2006) BnaTNDH • Seed oil content and composition **2011** • Mineral composition datasets • CMap and other graphic visualisation and navigation	

	Transcriptomics
	Updates on transcriptomic platforms
2010	Affymetrix Exon array probe design and 135k Unigene files
2011	Catalogue of available datasets (array and RNA-seq)

	Update of TILLING populations
	• TILLING screening service for *B. rapa* from RevGenUK
	• TILLING for *B. napus* available from CanTILL
2010	• *B. rapa* R-o-18 Illumina 12x sequence online
	http://www.brassica.info/datasets/Brassica_resequencing_data/
2011	• Additional R-o-18 re-sequencing data available

CropStoreDB

	Curation pipeline and Database	http://www.cropstoredb.org
2010	• Update to design for trait and trial data	
	• Simplified user input templates	
	• More generic for wide range of crops	
2011	• Additional interfaces and navigation to BrassEnsembl and CMAP	

	Publications
	Publications only available via www.brassica.info
2010	• Back catalogue of *Cruciferae Newsletter* + vol 29 (Mar 2010)
	• Posters from PAG (2010) and from Brassica2010 http://www.brassica.info/posters.php
	• *Also available*: Charles Darwin: from "The Variation of Animals & Plants Under Domestication", 2nd edtn, 1875 - section on brassicas
2011	• Please send us any publications you wish to be made available on www.brassica.info.