

Multinational Brassica Genome Project (MBGP)

Meeting notes – PAG 2013 (January 13th 2013 @ 10am) by Pat Edger, edited by Chris Pires; distributed by email on January 18, 2013 additional edits from Boulos Chaloub, Isobel Parkin, Jinlin Meng, Rod Snowdon added by Rod Snowdon; redistributed by email on 9.12.2013 with invitation to next meeting at PAG 2014.

Note: ACTION items in bold and highlighted in yellow below

1. Introductions, Sign-up Sheet, Approval of Minutes

Rod Snowdon requested correction to minutes that there are 450 *Brassica napus* diversity lines available in public ERANET-ASSYST collection.

All minutes/notes will be posted on Brassica.info

2. Updates on Genetic resources (seeds and sequence).

German-Canadian consortium has developed 450 *Brassica napus* accessions (phenotyped and genotyped, metabolite data, RNAseq data for 100 lines available, future GWAS, etc.). Seeds are available free to universities and institutes as subsidized by charges to companies. Distributed currently to UK (Ian Bancroft), USA (Michael Gore) and others (China, US, Australia). See separate file attached to minutes for details (ASSYST Public *Brassica napus* Diversity set).

Pires read email from Graham King : “After a delay at Southern Cross University (SCU), we are now able to progress with selfing the *Brassica rapa* collection and this will be available in public domain once we get seed stocks.”

Guy Barker says *Brassica oleracea* and C genome diversity sets continue to advance at Warwick.

Ian Bancroft has sequenced 83 of 100 lines of *Brassica napus* for association studies, more planned.

3. Updates on genome sequencing activities

Pires read email from Graham King: “We have generated more R-o-18 B. *rapa* data (the line used for Lars Ostergaard’s EMS population) and will be putting the new assemblies for download on brassica.info website. We will continue to enhance this genome sequence as a reference for our and others work. (later this year we will also have a copy of the EMS population at SCU for collaboration with U. Nottingham and others if they wish).”

Boulos Chalhoub – *Brassica napus* AC genome should be available this year (chromosome scaffolding) withIN a consortium agreement between all three participating groups: French consortium (Boulos Chalhoub), Chinese consortium (Shengyi Liu) and Canadian consortium (Isobel parkin) to publish the French sequence of *Brassica napus* cv. Darmor together. The

participants are open to cooperating with other Brassica groups to make it an international *B. napus* initiative.

The C genome of *B. oleracea* was also discussed. It is expected that Shengyi Li's *B. oleracea* genome will be submitted for publication first (2013), and the aim is to submit the other genotype (TO1000: Isobel Parkin) when the *B. napus* genome publication is submitted.

Isobel Parkin and Chris Town: Resequencing of 3 morphotypes of each Brassica (cabbage, broccoli, etc) to approximately (~50X depth)

Rod Snowdon: German consortium resequencing 51 diverse *B. napus* (natural & resynthesized), data public from 2014

Pat Edger: Brassica rapa B3 and R500 – dataset on SRA and JGI on-going sequencing

No *Brassica rapa* morphotype resequencing projects were known of by anyone present.

Isobel Parkin: Several thousand more genes have been found in *Brassica rapa* (reported in Sicily meeting in late 2012) with Chifu version 2 coming soon (not yet public, only in China). As for B genomes: *Brassica nigra* genome coming with 2013 release anticipated; and *B. carinata* and *B. juncea* are being sequenced as well.

4. Updates on SNP discovery activities and assay resources

Isobel Parkin – Illumina Infinium 60K chip array is released for Brassica (consulting member access only); goes public in September? Chip was simple SNPs only (No Hemi-SNPs). Future – may have up to 90K (sensitivity limited)

5. Upcoming international meetings

Pires read email from Graham King: “I wondered if you knew who is organising the planned Brassica workshop at PAG Asia in Singapore in March? It may be worth raising question of how MBGP discussions are coordinated between the two conferences.”

PAG Asia Singapore – Jacqui Batley is organizing Brassica. Also an oilseed workshop (will include Brassica). #200-250 people attending. Graham Moore is organizing (supported by BGI).

Would it be useful to have this MBGP meeting (or similar) in Asia to discover availability of resources? General consensus – Yes. Italy meeting had no such meeting.

Canola meeting in Brazil in August 2013, Fede Iniguez-Luy on org. committee.

Crucifer Genetics Workshop (Brassica 2014) in Wuhan, China, March 31-April 2, 2014, hosted by Jinleng Meng.

International Rapeseed Congress in Saskatoon (end of June 2015; details coming from Isobel Parkin).

6. Informatics

Pires read email from Graham King: “We have continued discussion with Rothamsted in moving www.brassica.info to SCU. I am pushing this now for next month or so. Chris Rawlings (head computational biology at Rothamsted Research) will be at PAG and should be at MBGP meeting. In 2012 we published the InterStoreDB paper (links CropStoreDB and BrassEnsembl, Journal of Integrative Biology 54: 345-355). In brief: plan is to move the links between CropStore and Ensembl, so that it picks up the B. rapa Ensembl managed at EBI (under management of Paul Kersey). In future, at SCU we would be interested in working to make links between genetic maps and other CropStoreDB resources of B. napus and/or B. oleracea and annotated genomes that can also appear in an EBI managed Ensembl genome resource. As you know I have had discussions with Chris Town re: AIP but not sure where they are with this at present. The common link that may work for us is via something called InterMine.”

Chris Town gave a description of the NSF call for an Arabidopsis Information Portal (AIP). Town was asked to submit a proposal to lead this large effort following a design workshop. NSF will no longer fund TAIR or individual portals (go to Monday PAG session on the International Arabidopsis Informatics Consortium for details). The main idea is to provide web services so proposed to not reinvent the wheel but to use InterMine already used by fly, yeast, mouse, zebrafish, and other communities. Web services will have three layers: presentation layer with GUI, webservice/transaction layer, and data layer. This will also house all the information currently in TAIR and will facilitate data distribution of Arabidopsis data (genomic, omic). Town plans to use Brassica and Medicago as “test case” to expand AIP to cross-species comparisons. Town is collaborating with iPLANT (computing performed with their infrastructure). No International efforts were addressed although several representatives from various countries were at the design workshops and are on the AIP advisory board.

Jim Beynon and Ian Bancroft and Chris Rawlings discussed the situation in the UK after departure of Graham King: TGAC transition was discussed. Brassica is key species for UK given agricultural importance. Sarah Ayling at TGAC is contact person, as well as Paul Kearsy at EBI for ENSEMBL.

Gene annotation uniformity across nations was discussed. Uniformly agreed that it was absolute importance to use common gene names to avoid previous problems experienced in rice and other communities. Each country / lab could have separate browsers, but keeping annotation standards is crucial.

ACTION: International annotation standards will be organized by Dave Edwards. Email Dave if you want to join/contribute to this important working subgroup.

Brassica MAKER pipeline (with set parameters) was discussed as another possible goal, which could expand eventually across the order Brassicales. Similar efforts across Solanaceae.

Initial volunteers from present meeting:

Australia – Dave Edwards

Canada – Isobel Parkin and Andy Sharpe (and contact Nick Provart/Stephen Wright?)

China – BRAD/Qteller – contact Feng and Xiaowu Wang?

France – Nathalie Nesi

Germany – Uwe Scholz, Birgit Samans

USA – Chris Town and Chris Pires

Others?

7. Future role of MBGP?

Continue to expanding meeting at PAG Asia; include more individuals / broader participation.

In addition to continuing website of brassica.info ; get into social media (Twitter? Facebook?).

Expand from genomics to Brassica metabolic network: Brassica cyc handed off to J. Moore?

8. Other MBGP business?

ACTION: Next organizer to replace Chris Pires is... Rod Snowdon!

9. Brassicales Map Alignment Project (BMAP)

Chris Pires and Pat Edger updated status of project, details to be posted on brassica.info

VEGI and BMAP genomes discussed. Aethionema arabicum, Sisymbrium irio, Leavenworthia uniflora, Cleome violaceae and Caulanthus amplexicaulis.

ASSYST Public *Brassica napus* Diversity Set

Approx. 450 homozygous inbred *Brassica napus* accessions covering a broad range of the genetic and geographic diversity present in the species (spring, winter and Asian oilseeds, kales, swedes). Being expanded to approx. 600 lines by Ian Bancroft.

Used since 2009 in an international consortium (DE, UK, CAN) performing associative expression analysis and GWAS to analyse and dissect complex traits.

Mission: Development of a public *B. napus* diversity collection and accumulation of publicly available genotype and phenotype data for multi-trait association mapping, systems biology, breeding, etc.

Genotype data:

- 6k Infinium SNP array (Isobel Parkin)
- part of collection with 60k Brassica SNP array
- approx. 20k RAD markers (Anja Bus/Benjamin Stich)
- SureSelect sequence capture of 30 flowering regulators (Rod Snowdon, in progress)
- RNAseq (seedlings) for 100 lines, to be expanded (Ian Bancroft)

Phenotype data:

- Automated phenotyping of seedling developmental traits
- Multiple years, multiple environments for seed quality traits (oil content, fatty acids, glucosinolate content, protein content, seed colour, seed fibre)
- Diverse environments for flowering time (Germany, China, Chile)

Distributed in part or whole to a number of research groups and seed companies in Germany, France, UK, China, Canada, US, Chile, Australia

Distribution to research groups so far free of charge, costs being covered by imparting a nominal charge per line for distribution to seed companies