

Multinational Brassica Genome Project Steering Committee Meeting

Chantecleer room, Town & Country Resort - San Diego, California

Sunday 10th January 2010

Present:

Isobel Parkin (Chair)	Isobel.parkin@agr.gc.ca
Jacqui Batley	j.batley@uq.edu.au
Faouzi Bekkaoui	Faouzi.bekkaoui@nrc-cnrc.gc.ca
Dominique Brunel	brunel@versailles.inra.fr
Pierre Carion	pierre.carion@bbsrc.ac.uk
Boulos Chalhoub	Chalhoub@evry.inra.fr
Raju Datla	raju.datla@nrc-cnrc.gc.ca
Dave Edwards	dave.edwards@uq.edu.au
Pierre Fobert	pierre.fobert@nrc-cnrc.gc.ca
Erin Gilchrist	erin.gilchrist@ubc.ca
Jer�my Just	jeremy.just@rennes.inra.fr
Graham King	graham.king@bbsrc.ac.uk
Yong Pyo Lim	yplim@cnu.ac.kr
Kede Liu	kdlou@mail.hzau.edu.cn
Shengyi Liu	liusy@oilcrops.cn
Rob McClung	mcclung@dartmouth.edu
Dick McCombie	mccombie@cshl.edu
Jinling Meng	jmeng@mail.hzau.edu.cn
Lars �stergaard	lars.ostergaard@bbsrc.ac.uk
Jerome Pauquet	Jerome.Pauquet@biogemma.com
Chris Pires	piresjc@missouri.edu
Reno Pontarollo	rpontarollo@genomeprairie.ca
Carlos Quirios	Cfquiros@ucdavis.edu
Gijs Van Rooijen	vanrooijen@genomealberta.ca
Shusei Sato	ssato@kazusa.or.jp
Martin Trick	martin.trick@bbsrc.ac.uk
Richard Visser	Richard.visser@wur.nl
Jun Wang	jun.wang@bbsrc.ac.uk
Rod Wing	rwing@ag.arizona.edu
Tae-jin Yang	tjyang@snu.ac.kr

In attendance by phone: Chris Town (cdtown@jcvi.org)

1. B - genome nomenclature:

The issue of adopting a consistent nomenclature system for the description, assignment and orientation B genome linkage groups was raised. This arose due to the fact that the B genome groups described by Panjabi *et al.* (2008: BMC Genomics 9) for the B genome in *B. juncea* (the "BjuHVDH" map) conflicted in a number of respects with that of *B. nigra* described by Lagercrantz & Lydiate (Genetics 144: 1903-1910). It was agreed that although the logic of the Panjabi *et al.* assignments was based on similarity to linkage groups in other genomes, it would be appropriate to adopt the precedence of the nomenclature *sensu* Lagercrantz. It was therefore **agreed** that:

- The Lagercrantz B genome linkage group assignments would remain the reference point for subsequent publications.
- This would be re-iterated on the page defining linkage group assignments at <http://www.brassica.info/resource/maps/lg-assignments.php>
- Within the linkage map registry at www.brassica.info, for the BjuHVDH_01_2008a map, the linkage groups would be assigned to the appropriate groups *sensu* Lagercrantz.

- d) The record of these minutes would be cited if required to confirm that the Lagercrantz nomenclature will form the permanent assigned nomenclature for B genome linkage group names.

Action: [IP] agreed to contact the group in Delhi (Panjabi et al.) to alert them to the agreement above. Updates will be shown on www.brassica.info [PC].

2. Diversity sets:

An update on the progress and availability of the various *Brassica* diversity collections worldwide was given.

On-line information is available at http://www.brassica.info/resource/plants/diversity_sets.php

- | | |
|--------------------|--|
| <i>B. napus</i> | managed via the UK OREGIN project – details on status and availability are being maintained and are available online |
| <i>B. oleracea</i> | available from Warwick University (UK) - details available online and updated at least annually, currently ~65 DH lines fixed. |
| B. C genome | 'wild' species accessions - as for <i>B. oleracea</i> , less fixed to date as more challenging w.r.t. flowering, fixing. |
| <i>B. rapa</i> | work still ongoing at WAU (Wageningen, Guusje Bonnema) and IVF—CAAS (Beijing, Xiaowu Wang), aiming to produce ~250 accessions. <u>Unclear on release policy</u> . Additional lines (mutually exclusive from those being fixed by Bonnema/Wang are being generated by SSD in UK (King, Rothamsted), - status will be posted online as become available. |

Boulos Chalhoub commented on the *B. napus* collection being assembled in INRA (France) managed by Anne Laperche (anne.laperche@rennes.inra.fr) of ~600 accessions, of which a core collection is already public.

Action: [PC] to contact [DB] to obtain information to include in online information.

3. TILLING:

An update on the various *Brassica* TILLING platforms was provided:

Active services now available in public domain:

- | | |
|-----------------|--|
| <i>B. rapa</i> | R-o-18 EMS population generated at JIC by Lars Ostergaard is available via RevGenUK (http://revgenuk.jic.ac.uk/) and is being well used, with 20-30 external orders to date. A screen costs UK£1,500 for a single 1,000-1,200 bp amplicon |
| <i>B. napus</i> | CanTill project at UBC Canada (Erin Gilchrist, http://www.botany.ubc.ca/can-till/CanolaRequests.html) has 1,500 lines of <i>B. napus</i> DH12075. TILLING of a 1,500 bp amplicon is CA\$5000 |

Additional platforms for *B. napus*:

Jinling Meng (HAU, China);

Ian Bancroft has developed a *B. napus* TILLING population in collaboration with a private company. Ian can be contacted for access and availability - preliminary results indicate an excellent mutation rate, ~40% of EMS-sensitive bases hit somewhere in the ~6000-line population, based on the first ~1kb amplicon analysed in detail. Tracking back from DNA to seeds has been confirmed; seed viability is generally good.

No update on INRA (France) TILLING platform since last meeting.

Final version, posted 7th March 2010; minor edits 21st March 2010

Further information will be collated at <http://www.brassica.info/research/activities/tilling.php>

There was no update on any new BAC libraries

4. Transcriptomics:

Graham King confirmed the production of the Affymetrix GeneChip ® *Brassica* Exon 1.0 ST Array, which will be in the 2.4 million probe x 25mer format, and a catalogue item.

- Work carried out by Chris Love at Rothamsted extended the number of unigenes to ~135K. Annotation of arrays will be available via BrassEnsembl (<http://www.brassica.info/BrassEnsembl/index.html>) and standard downloads. The array was designed and initiated for a UK eQTL study focused on mineral nutrition (Nottingham, Warwick, Rothamsted).
- Will be in production January 2010, and should be available ~March - see <http://www.brassica.info/resource/transcriptomics.php> for details.
- The NASC Affymetrix Service (<http://affymetrix.arabidopsis.info/>) will offer this array as a service item, with costs comparable to existing advertised services.
- A discussion relating to the handling and storage of cumulative data arising from Brassica arrays (eg Agilent, Affymetrix) concluded that these should be made available through genome browsers and repositories that link to other tools, such as those available for Arabidopsis at NASC arrays (<http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl>).
- Integration of annotation associated with the different arrays will be required.

Arising from this discussion, it was mentioned that both NASC in the UK and TAIR in the US are having difficulty securing funding to maintain their services.

Action: [IP] agreed to contact TAIR, to see how the *Brassica* community could help by providing support to applications etc., as these resources are clearly invaluable to the *Brassica* research community. [GK] agreed to engage in discussions with NASC.

5. SNP / genotyping platforms for *Brassica*:

- The AAFC SNP discovery (Illumina Golden Gate Platform) will be completed in the next 6 months and data made available publicly.
- Other platforms of which the meeting participants were aware included additional Illumina arrays generated by: BioGemma, and Trait-Genetics in Germany (www.traitgenetics.de/).
- Additional platform: Bancroft/Trick SNP discovery by mRNA-seq (published: published (Trick et al., Plant Biotechnology J. 7:334-346, 2009). Conversion of this for high-throughput genotyping by barcoded PCR primers will be published. Both will be available shortly as commercial services; in the meantime if interested, people should contact Ian Bancroft directly by e-mail (ian.bancroft@bbsrc.ac.uk).
- The question of the requirement for a concerted effort to establish a SNP discovery platform was raised, as there are several institutes working independently using different platforms. It was agreed that if one was adopted, there would be considerable advantages in terms of economies of scale. However, this issue was not resolved at the meeting.

- There was some discussion about practical issues associated with problems in discovering SNPs that can reliably distinguish between homoeologues within A, B and C genomes.

Comment from IB re: establishing of a SNP Registry – this might be worthwhile setting up a simple Brassica registry (eg Excel format, available on brassica.info), but this mustn't become too burdensome. Would favour a simple nomenclature like [ref_seq]:[1-based coordinate]:[IUPAC code] e.g. JCVI_1234:123:Y . For this only those SNPs that relate to publicly retrievable reference sequences would be accepted.

Action: [IP] to send email to Brassica.info mailing list to assess progress and demand in more detail.

6. Systems Biology initiatives:

A number of projects have been initiated recently that will focus on collating Brassica 'omics' information for a range of systems biology/modeling exercises. These include:

- BBSRC project in UK (PI: Broadley, Nottingham) on biofortification
 - this will use the new Affymetrix Exon arrays for generating eQTL (in BralRRI population), with systems work at Rothamsted (incl. using ONDEX).
 - http://oasis.bbsrc.ac.uk/netans-bin/gate_exe?f=doc&state=r2f0li.1.1
 - <http://communications.nottingham.ac.uk/News/Article/Growing-greener-greens.html>
- ASSYST – ERA-PG project will be collating eQTL and other data, with systems analysis at JIC (Bancroft group)
 - <http://www.erapg.org/everyone/16790/18613/19533/19534>
- Lars Ostergaard: initiating project on distribution of hormones (hormonics) in organ development (silique), and development of mathematical models that can be tested experimentally.

Action: [PC] to add a “modeling” page to the www.brassica.info website.

7. Website update:

Pierre Carion presented the 2010 'brassica.info Bulletin', outlining the updates to the website, and data curation in the past year. It was emphasized that this is the community's site, and that any suggestions as to content should be made sent to Pierre.

Action: It was suggested that [PC] add a Jobs/Vacancies page, which would be a useful reference point.

8. BrassEnsembl:

Graham King announced the initial release of BrassEnsembl which has been established at Rothamsted by Chris Love. This includes the first set of *B. rapa* gene annotations from the publicly available BAC sequences.

- This initiative will continue, with collaboration from EBI, and any others who may wish to be involved in Europe or elsewhere.
- Once data from the complete *B. rapa* genome sequence is published these would be added, and used as a reference for re-sequencing data tracks.
- It was reiterated that any work on *Brassica* genome annotation and related information should ensure that the common GFF3 format was adopted. Data held in BrassEnsembl are available and can be imported in GFF3.

- A poster was presented at PAG http://www.intl-pag.org/18/abstracts/P05e_PAGXVIII_393.html and is also available at <http://www.brassica.info/posters.php>.

9. **Brassica map alignment project:**

Two aspects were discussed: Brassica specific and wider Brassicaceae

- Good progress has been made in collation curation and alignment of Brassica maps by Jun Wang, Pierre Carion and others (Rothamsted, via CropStoreDB) and a series of publications are in preparation from this work. The first will focus on 3 *B. napus* maps, and then extend to other *B. napus* maps, then A and C genomes diploid genomes.
- The map alignment project of Chris Pires, and Rod Wing was discussed. This involves identifying ~100 species within the Brassicaceae to be sequenced by a consortium, taking similar approach as for 12 *Drosophila* sequencing consortium and 1001 Arabidopsis genomes (<http://1001genomes.org/>). The aim is to select a set of representative species based on the phylogenetic relationships, and to sequence these first. These will be supplemented by more biology-based projects including study of comparative systems and genome evolution.
- The issue of suitable informatics and information sharing platform was discussion. The possibility of establishing a separate Brassicaceae database was raised. It was suggested that would be worthwhile to discuss with EBI in relation to their interests in plant genomes and comparative analysis pipelines.
- In the meantime, updates on this project may be included in www.brassica.info [CP] to contact [GK] and [PC] to discuss

10. **The following upcoming international meetings were mentioned.**

- 17th Crucifer Genetics workshop will be held in Saskatoon, Canada (5-8 September 2010).
- The next ISHS Brassica meeting, to be held in Sicily, with the year to be confirmed. 2011/2012.

11. **MBGP and BrGSP committees**

Due to the ongoing close integration of activities between the MBGP and Brassica Sequencing Projects, it was proposed that the relevant steering/management committees be merged.

- It was **agreed** unanimously that the MGBP, and BrGSP committees should be combined, as the first complete genome sequence is due to be released.
- It was **agreed** unanimously that Ian Bancroft should become Chair of this combined committee for the following year, as his current chairing of the BrGSP was due to end in 2011, and Isobel Parkin was due to complete her chairing of the MBGP at this 2010 meeting.

Note added after meeting: On behalf of the members of the Committee and wider community, I would like to thank Isobel Parkin for her excellent and hard work in maintaining the momentum of the MBGP over the past few years. [GK]

Brassica rapa Genome Sequencing Project Steering Committee Meeting

Chantecleer room, Town & Country Resort - San Diego, California

Sunday 10th January 2010

Present: see above.

Xiaowu Wang (IVF-CAAS, project leader of *B. rapa* sequencing being carried out by BGI) sent his apologies, since unfortunately he was unable to attend the meeting in San Diego

1. Outcome of meeting in Saskatoon focused on *B. rapa* Chiifu-401 sequencing:

- A brief update was given from the last meeting held in Saskatoon in August 2009, <http://www.brassica.info/info/about-mbgp.php#meetings>. A presentation on the current status of *B. rapa* sequencing had been held the day before in *Brassica* workshop. http://www.intl-pag.org/18/abstracts/W14_PAGXVIII_104.html
- The meeting in Saskatoon discussed the available BGI scaffolds at that time and assessed their verity largely based on comparison to fully sequenced BAC clones. It was agreed that the BAC by BAC approach would be discontinued and efforts would be made to further verify the BGI scaffolds to aim for release of the *B. rapa* genome early in 2010. Three factors have delayed progress/release since the August meeting:
 - Boulos Chaloub is in the process of generating approximately 5X coverage of Chiifu with 454 sequencing. These data have now been completed and are being processed and integrated into the BGI scaffolds. These sequences will be released publicly by FTP by the end of April 2010.
 - Xiaowu Wang and the Korean group lead by Dr Park have agreed to combine their data sets, which has required additional analyses.
 - Additional anchoring of the BGI scaffolds to the genetic map using SNP analyses is on-going and will be completed in February 2010 [JB].

Timescales for release:

- The sequencing group in China is committed to releasing the sequence scaffolds once a publication describing the genome is accepted. Xiaowu Wang has indicated that a draft manuscript has been developed and he hopes to circulate this manuscript to relevant parties in mid-march with an intended submission date of mid-April. All being well it is hoped the sequence will be released in June 2010.

2. *B. oleracea* and *B. napus* sequencing:

It was clear that there were an increasing number of projects ongoing for generating 'reference' genome sequences, as well as other 're-sequencing' efforts. Below is a brief summary:

Species	Genotype	Aim	Strategy	Consortium	Contact
<i>B. oleracea</i>	cabbage O212	Reference	Illumina +454	CAAS, Wuhan	Shengyi Liu
<i>B. oleracea</i>		Re-sequence x5	Illumina	CAAS, Wuhan	Shengyi Liu
<i>B. oleracea</i>	TO1000	Reference	454+Illumina+BAC ends	Canada/USA/UK	Isobel Parkin
<i>B. oleracea</i>	Early Big	Re-sequence 50-70X	Illumina	USA	Dick McCombie
<i>B. napus</i>	Tapidor DH	Reference	454+Illumina+BAC ends	HAU(China)/UK	Jinling Meng

<i>B. napus</i>	Ningyou7	Reference/re-seq?	454+Illumina?	HAU(China)/UK	Jinling Meng
<i>B. napus</i>	?	Reference		CAAS, Wuhan	Shengyi Liu
<i>B. napus</i>	DH12075	Reference	454, BAC ends, Illumina	AAFC/NRC, Canada	Isobel Parkin/Andrew Sharpe
<i>B. napus</i>	Darmor DH	Reference	Illumina, 454, Sanger	INRA, France	Boulos Chaloub
<i>B. rapa</i>	Chiifu-401	Reference	BACs, BAC ends, 454+Illumina	CAAS/BrGSP	Xiouwu Wang
<i>B. rapa</i>	R-o-18	Re-seq. x12	Illumina 75bp paired	Rothamsted	Graham King
<i>B. rapa</i>	AcDsc	Reference	454, Illumina	AAFC/NRC, Canada	Isobel Parkin/Andrew Sharpe

More detailed updates were provided for *B. oleracea* from the groups in China (Shengyi Liu, OCRI, Wuhan) and North America (Isobel Parkin/Andrew Sharpe). More information will be provided on request.

N.B. Some additional *Brassica* re-sequencing datasets are available from Dave Edwards' group in Brisbane at <http://flora.acpfg.com.au/tagdb/cgi-bin/help>

Action: A new page will be established at www.brassica.info to summarize the status and details of ongoing *Brassica* sequencing efforts.

- It was also suggested that a *Brassica* Sequencing WIKI be set up by Chris T.
- In addition it was suggested that regular (every three months) meetings would be arranged among sequencing groups to discuss progress. Isobel Parkin volunteered to arrange the initial meeting sometime in the week of April 5th,

3. Annotation

- It was still early days for comprehensive concerted annotation. However, various pipelines are in place to enable this, and have been used successfully for e.g. KBr BACs (at JIC).
- Pipelines in place at BGI (China), ACPFG (Australia), JCVI (USA), JIC (UK). At JIC, the annotation pipeline has been used successfully to annotate IVF-BGI scaffolds as well as BACs.
- Agreed that at present the important issue is to ensure that annotations are available for comparison in common GFF3 format. In addition, it would be valuable to ensure that there was version control so that confusion was avoided at a later date.