BMAP4 (*Brassicaceae* Map Alignment Project 4) Meeting Notes
Huazhong Agricultural University, Wuhan, China, 12-06-June (Notes by Dr. Yan Long; Edited by R. Wing and D. Weigel)

Attendees:

Name	Address	E-mail
Rod Wing	Arizona Genomics Institute, University of Arizona	rwing@ag.arizona.edu
Beom-Seok Park	Genomics Division, National Academy of Agricultural Science	pbeom@korea.kr
Yong Pyo Lim	Department of Horticulture, Chungnam National University, Republic of Korea	yplim@cnu.ac.kr
Xiaowu Wang	The Insititute of Vegetables and Flowers Chinese Academy of Agricultural Sciences	wangxw@mail.caas.net.cn
Shengyi Liu	Oil Crops Research Institute Chinese Academy of Agricultural Sciences	liusy@oilcrops.cn
Lixi Jiang	Zhejiang University	jianglx@zju.edu.cn
Maoteng Li	College of Life Science and Technology	limaoteng426@163.com
Jianjun Chen	Wuhan Botanical Garden, Chinese Academy of Sciences	jianjunchen@wbgcas.cn
Gong Xiao	Wuhan Botanical Garden, Chinese Academy of Sciences	gongxiaobio@gmail.com
Zhongsong Liu	Hunan Agricultural University	zsliu48@sohu.com
Jinling Meng	Huazhong Agricultural University	jmeng@mail.hzau.edu.cn
Meizhong Luo	Huazhong Agricultural University	mzluo@mail.hzau.edu.cn
Zaiyu Li	Huazhong Agricultural University	lizaiyun@mail.hzau.edu.cn
Kede Liu	Huazhong Agricultural University	kdliu@mail.hzau.edu.cn
Yan Long	Huazhong Agricultural University	longyan@mail.hzau.edu.cn
Ruiyuan Li	Huazhong Agricultural University	lry198010@webmail.hzau.edu.cn
Jing Wang	Huazhong Agricultural University	wjing@webmail.hzau.edu.cn

Brassicaceae Map Alignment Project(BMAP) IV Meeting
2010.06. wuhan

1. Self-introductions (13:00)

2. Introductions by Dr. Rod Wing with Crucifer evolution, *Brassicaceae* map alignment project (BMAP) and sequencing the complexity of *Oryza* (OMAP) project.

3.   Discussion about the one page summary submitted by attendees:
(1) Dr. Xiaowu Wang's said about his summary: My sequencing interests are in the Brassica vegetables crops, such as *B. rapa. B. oleracea*, and *B. juncea*. For sequencing the *B. juncea* genome, we will also do the B. *nigra* (B genome). We have finished sequencing the A and C genomes. For the A genome, we are writing a paper. For the C genome, we have assembled the genome, and are mapping the scaffolds to linkage groups, and are annotating. For the B genome sequencing project, we will collaborate with Dave in Austria and Isobel Parkin in Canada. In addition – other people in Chongqin, China, have shown interest in it.

In China, radish is also one of the most important crops, we will also sequence it.

I myself are interested in studying morphological variation in Brassica crops, both *B. rapa* and *B. oleracea*. We will compare those genes in morphological variation in the A and C genomes

Rod: What will be the first paper to be submitted?

Xiaowu: The A genome paper will be published first. We hope to publish it this year, and will open a Blast server to access the A genome sequence data by the end of June. Full service and downloading of all sequence data will be opened after the paper is published.

Rod: Are there competing groups doing A genome sequencing?

Xiaowu: I don't think there is any competing group. The Chiffu sequence will be the only RefSeq.

Lim: Some companies are doing sequencing, but are not competing since they won't publish their sequence.

Rod: if there is no completing it would make sense to let everyone have access to it.

Jinling: There is competition for C genome sequencing, one is Xiaowu's group in China, together with Dr. Shenyi Liu's group in oil crop's institute.

Xiaowu: Chris Pires is trying to sequence C genome in USA. At the moment, we are trying to find a way to collaborate.

Rod: Different cultivars and sequencing methods?

Xiaowu: Yes, different cultivars and methods. Regarding to the B genome, Isobel Parkin has sequenced over 70X genome coverage, but the assembly is still not good. We are planning to let

BGI get access to these data, so that BGI can help to improve the assembly. And we will start sequencing a new accession very soon.

Rod: How about a collaboration?

Xiaowu: We will try to organize the cooperation among all groups interested in the B genome.

(2) Dr. Jinling Meng's summary:

Jinling: We made a TN (Tapidor crossing Ningyou7) DH mapping population for *B. napus*. By cooperating with UK and Chinese scientists, about one thousand QTLs controlling 27 traits have been mapped. Some of the results have been published (Qiu et al., 2006; Long et al. 2007; shi et al., 2009). The map has been compared with the reference map of B. rapa, CK linkage map, based sequenced BACs of A genome. There is a BAC library of Tapidor available with sequenced BAC-ends, also an EMS-mutated population of Ningyou7 for TILLING. Now we are cooperating with JIC and BIG to *de novo* sequence the two parents using 454/Illumina sequencing technologies. We hope to obtain a high quality genome sequence as reference sequence for *B.napus* next year. Then, we plant to sequence another ten related cultivars of *B. napus* and *B. rapa*. At same time, Ian at JIC is doing transcriptome sequencing with 47 DH lines, plus two parents. This data will help to generate a high density gene-map.

B. carinata, originally from Africa, is another story. We hope to get some new traits and new genes from B and C genome into B. napus. We generated some DH lines of *B. carinata*, and then we have got a DH mapping population with 180 lines. If the community interested, we will provide the seeds (we promise to provide seeds to Isobel Parkin).

Rod: Is *B. carinata* a wild species or crop?

Jinling: A crop.

Lim: you gave me the seeds of DH lines. We also have inbred lines from India and they are different. We are doing DH population construction in India because flowering time in Korea is a problem.

Rod: Is it an important crop in Africa?

Jinling: Yes, also in parts of Asia and Canada. It grows in very dry places.

Park: Do you have any information regarding Canada's sequencing progress?

Jinling: They are sequencing a spring type of *B. napus*. France is sequencing *B. napus*. We have a friendly relationship with Canada for the sequencing, but we presently to not have a collaboration/cooperation with France.

Lim: A company (BGI) has already finished *B. napus* genome sequencing (Bayer and KeyGene).

Park: They worked together with BGI.

Xiaowu: A and C with BGI, and AC with KeyGene (physical map).

Rod: Will the company publish the results? Do they have plans to share the sequence?

Xiaowu: No. We release the data and the companies will use it.

(3) Dr. Yong Pyo Lim's summary:

Lim: We have a mapping population CKDH and CKRI of 200 lines. We tried to perform comparative mapping between the A and C genomes, between the A and AC genome (collaboration with Jinling), and the A and AB genome (collaboration with Deepak Pental, India). Additionally, in spite that we want to perform comparative mapping on the B genome, we do not have the B genome population. We have three population of radish for mapping.

Since B. rapa was mapped more than 1000 markers, recently we trie SSR markers in other Brassica species. When we checked 26 *Brassica* species with SSR markers originated from *B. rapa*, it was very successful. Also radish was successful. Since identification of these species by using SSR markers was successful, we can suggest to use this system to do diversity research.

For sequencing, we have 10 inbred lines for re-sequencing and some transcriptomics.
Additionally, I have two kinds of data. Six years ago, the Korean Brassica Ressource Center was established. We collected seeds and genetic materials from the world and from Korea. One list of handout indicated that *Brassica* species recorded in Korea were listed as 90 accessions. Other list showed the seeds and DNA of *Brassica* species in Korea, which the Brassica Resource Center has. If anybody wants to use them, we can provide it.

Jinling: Is the accession you mean still in Korea or in your stock?

Lim: I do not have all *Brassica* species recorded in Korea. Among those I have, some stocks were collected as seeds, some as DNA. Major are wild type species. This is natural ones, all accession are heterozygous type. For sequencing or genetic study, we should make the inbred lines by microspore culture. Until now, major is not doing. It is a problem.

Rod: Are these materials available?

Lim: Yes
Rod: Could you send the handout to Jinling?

Lim: Yes, of course.

(4) Dr. Xiaoming Wu's summary:

Xiaoming: We maintain the national collection from China of about 6000-7000 accessions in B. napus and *oleracea*, and core collection of *B. napus* and *juncea*. Our interest is in traits and extensive phenotyping. We want to do re-sequencing.

Xiaowu: Is there any *B. nigra* in your pool?

Xiaoming: yes.

Xiaowu: *B. nigra* is not a crop and we want to do segregation mapping

Xiaoming: For sequencing, wild species in China is important. *Sinapis arvensis* is located in the north west of China (Xingjiang province). It has important disease resistant traits, is native to China and is an ancestor of *B. nigra*.

Jiang: In 1994 a joint publication for Chinese germplasm catalogue was produced in Chinese only.

Rod: Is there any plan to make a catalogue in English?

Xiaoming: No (not yet).

(5) Dr. Beom-Seok Park's summary:

Park: We are government institute. We have nearly finished *B. rapa* sequencing and are comparing linkage maps. Now we are going to sequence some *B. rapa* and *oleracea* genomes. We also have maps of radish.

Rod: Are you generating a reference genome and do you plan to publish?

Park: Yes, it's reference and will be published. We also have discussed to cooperate to sequence any Brassica crops.

Rod: Will the sequencing be performed at RDA? What is your sequencing capacity?

Park: We are going to have one Illumina machine.

Park: *De novo* sequencing is difficult for B. napus, even if we have good linkage maps. If we have a method analyze, we can choose some important species first.

Park: For BMAP how many species are you aiming for? For r sequencing, some countries have response, while some are late, progress is slow. We need to make a concrete plan? How many years to sequence the BMAP accessions?

Rod: Approximately 5 years.

Park: Sequencing is fast, while assembly is slow.

Xiaowu: One question is genome size, the other is the availability of a homozyous line. Most genomes are small, also some polyploids are difficult to get homozyous. What can we do? Selfing?

Another suggestion is one should have major partners to avoid slow progress.

Rod: I agree with you to have major partners. So genome size is a not problem, for polyploids, the idea is basic references for other species, so others can do resequencing.

Park: Please make announcement that people will finish their genomes.

Rod: The plan it to have a white paper by the end of August. This paper should be updated yearly to include progress on the international level. Funding is very important.

Xiaowu: It's important for funding from companies, they are interested in it.

Zhou: For OMAP how you manage the data?

Rod: I have a small analyzed team, and we co-operate with others. For us we release data immediately. More people can see the data.

Xiaoming: how to apply this information to others? how to use these data? We should have mechanism for people to use the data in different ways?

Ruiyuan: The huge data set proposed here will be different from that in OMAP. The question is how to provide a uniform platform to every person for data acquisition and retrieval with web service, which will provide API in REST or SOAP protocol? Cloud compuation maybe a good choice.

Rod: That should be major focus. What databases? How many? Sequence just began. The amount of data that will be released and is huge.

Xiaowu: This will attract more people to use it.

Lizhi: Some suggestions. Based on my genome sequencing experience, NGS goes very fast. We should put more attention to transcriptome sequencing, easy for assembling a transcriptome.

Rod: Transcriptome data are very very important, and should be coordinated.

Discussion of species (16:40)

(1) Priority accessions

Rod: Could everyone give me back in one week send e-mail, one for China and one for Korea. Add short reasons why choose the species? Accession Number 91? Should the priority be changed?

(2) Sequencing organizations:

Who will do the sequencing? Major organizations

ROD: who wants to do what? How to get funding? Everything is changing. Key for us is to produce high quality reference sequences. In Asia, BGI, BIG, Shanghai (NCGR) are the key players.

Lixi: Zhengjiang California International Nanosystems Institute

Xiaowu: Do they do assembly of large genome?

Lixi: They don't have experience for genome assembly, they are doing transcriptome analysis.

Lim: MACROGEN is a sequencing company in Korea. Also, the Korean government is planning to construct a genome center.

Rod: in the future, how about reference genomes? Key is who has populations, genome sequencing.

Rod: Future planning - what else is needed? Dedicacted computional infrastructure and databases? Single database? Or different?

Ruiyuan: The data produced by BMAP project will be huge, and the fee for the data storage and application maintenance may exceed the capacity of each individual institute or organization. My opinion is we can store the data in commercial data center which provide cloud computation servers such as Amazon AWS. The important thing is that it may be free, or much cheaper for us to store the BMAP data in a data center. The data storage provider will get money from the people who download the BMAP data from the data center, everyone can develop and deploy applications on the cloud computation platform and then the others can execute the application to utilize the data, the cloud computation servers provider will get money from the people who execute these application. This means that institute/organization will not need to dedicate and maintain a computational infrastructure.

Rod: Do you know of groups who have already have experience with such a model?

Ruiyuan: Yes, Steven Salzberg from Maryland (http://bowtie-bio.sourceforge.net/crossbow/index.shtml) has developed a cloud computational baseed application named "Crossbow" and used it to identified SNPs using the resequencing data of human genome in 38X coverage. It was reported that Ian from JIC will cooperate with a commercial company to deploy their transcriptome base SNP discovery software "TraitTag" on Amazon's EC2 cloud computation platform. A company named DNAnexus has successfully deployed some NGS analysis application on Amazon AWS and provide these NGS analysis servers to the customer at a very low price.

Xiaoming: We should set up working groups for different work, such as a biological materials group, sequencing group, analysis group, and an application group?

Rod: We can discussion this item further in Saskatoon this September.