Minutes
Peanut Genome Consortium Meeting, Zhengzhou China
6/17/13

PARTICIPANTS

Victor Nwosu
Lutz Froenicke
Peggy Ozias-Akins
Baozhu Guo
Corley Holbrook
Mark Burow
Richard F. Wilson
Steven Cannon
Richard Michelmore
David Bertioli
Charles Chen
Kim Moore
Howard Valentine
Tom Stalker
Noelle Barkley
Manish Pandey
Brian Scheffler
Xing Jun Wang
Hari Upadhyaya
Ran Hovav
Soraya Bertioli
Howard Shapiro
Dan O'Conner
Shyam Tallury
Shivali Sharma
Farid Waliyar
Rajeev Varshney
Boshou Liao
Scott Jackson
Xingyou Zhang
Guohao He

Jackson convened the meeting and welcomed those attending or via teleconference. A quorum was established. The agenda was approved. Minutes from 3/13/13 were approved. These documents are posted on *peanutbioscience.com.*

AAGB-2013 Update: Stalker (Secretariat) reported the technical program is composed of 40 oral presentations and 72 posters. 186 (excluding special guests) have registered, representing the U.S. (23), Australia (2), India (5), Brazil (2), Niger/Mali (1), Israel (1) and China (152). Speakers received information on uploading presentations, and were asked to provide Wilson with copies of their ppt files for preparation of stakeholder reports. Speaker ppt will not be posted on the web or used in other ways without their expressed consent and the direction of the PGC. The U.S. delegation was reminded of the group picture after the opening ceremony with Chairman Luo. Guo gave details and information regarding various social events and research facility tours. Wilson described the purpose of Session VIII Breakout sessions, and accepted the gracious offer of volunteers to facilitate the discussion groups. PROGRESS on GENOME SEQUENCING & ASSEMBLY. Group discussion considered best ways to move forward on sequencing & assembly of the tetraploid genome. Xun Xu and and Liu Xin proposed to explore a BACxBAC approach. Concern was expressed regarding the cost-effectiveness and information gained from BACxBAC. Merits and disadvantages of other approaches (Moleculo™, PacBio™, LFR, optical imaging) were discussed. BGI and Michelmore/ Froenicke will explore these options and share info with PGC members ASAP.

PHEN~~GENO~~TYPING: Holbrook updated group on activities with 16 RIL populations. Seed of 8 populations will be ~~are scheduled for seed~~ increased in 2013, the remainder in 2014. **Birdsong Inc has graciously stored all seed for genotyping in -18C freezers. A seed-inventory has been developed by Holbrook, and is posted on www.PeanutBioscience.com/.** The first 8 populations are being phenotyped for CBR, LLS, TSWV, WM (Holbrook & Ozias-Akins); ELS (Islieb); pod & seed characteristics (Hovav); Sclerotinia (Tallury). The second 8 populations will be phenotyped in 2014. **Discussions will be held with Sanders at APRES on screening for flavor traits.** PAC resistance will be screened in 2014 on the T-population RILs developed by Guo. An X-ray machine may be available from USDA at Dawson for determining grade~~root architecture~~. **Holbrook expressed concern that suitable plot combines are no longer commercially available**. This will force hand-harvesting and increase labor costs. ~~Burow reported that Michael Baring had constructed a combine;~~ Dan O'Connor also was building a machine in Australia. **Noelle Barkley will serve as the nexus for permits & distribution of seed upon request, plus information on all available genotyping resources. This includes information from Varshney on ICRISAT lines and synthetic amphidiploids from Bertioli; MTAs from Waliyar; and weather data at each location. All phenotyping data per se will be sent to Cannon and Holbrook, who will establish and disseminate a standard protocol/format.**

ULTRA-HIGH DENSITY DIPLOID MAPS. Michelmore and Froenicke reported that BGI has shared A- and B-genome specific sequence data for assembly. About 26 million SNPs from each genome were

filtered to 2 million good homogeneous data. 40% of each diploid genome is positioned on a high density genetic map. Work continues to achieve 60%. A pipeline has been established for *A. duranensis* RILs from interspecific crosses provided by Bertioli, and the ICRISAT diversity panel.

GENE EXPRESSION: Ozias-Akins reported on the status of gene expression analyses of the Tifrunner transcriptome. Twenty-four libraries (triplicate libraries from 8 tissues) have been sequenced by Scheffler (2 x 100 bp reads) and are being assembled via CLCBio™ or Trinity™ software. Libraries are almost complete. Froenike reported limited sequencing on diploid transcriptomes from mixed tissues by Mi-Seq. Deeper sequencing will be conducted via Hi-Seq plus PacBio for annotation of diploid and tetraploid sequences. Froenicke and Ozias-Akins will coordinate their efforts. Burow has produced 2x50 transcriptome reads from pod, leaf and root tissues of 22 cultivated and wild accessions in collaboration with Andrew Farmer (NCGR). Bertioli also has transcriptome data from diploid species. Xing Jun Wang offered to share data on 60,000 unigenes from two Chinese libraries from widely grown cultivars. Varshney will contribute transcriptome data from diploids sequenced at ICRISAT. All these data sets or assemblies will be used to analyze diploid and tetraploid genomes. Cannon suggested **Andrew Farmer would be the best to help integrate these data sets, to develop a gene expression atlas, and to develop a process for separating A and B-genomes. Cannon agreed to establish a common ftp site and to receive all these data.**

BIOINFORMATICS: Cannon reported progress on the development of **PeanutBase.org** for collecting published information on QTL, markers and maps to populate an informatics database (Peanut Breeders Toolbox). The Toolbox will be modeled after SoyBase and LIS. Stalker warned of significant genetic variation in *A. duranensis*, and recommended that the accession number name always be used to define the A-genome material used in sequence analysis. Bertioli also has noted from recombination studies that *A. ipanenis* may be very similar genetically to *A. magna*; therefore crosses between these species may not be truly interspecific. Bertioli also noted that even intraspecific crosses of cultivated peanut show high rates of distorted segreation. **Cannon agreed to set up a conference call on data set construction including: X.J. Wang, Varshney, Ozias-Akins, Burow, D. Bertioli, Scheffler, Froenicke, and Farmer.**

EXPERIMENTAL SEQUENCING TECHNOLOGIES: Froenicke reported extremely good progress with Moleculo™ technology on diploid genomes. A-genome analysis at 4X coverage produced high quality long reads (mean, 3.7 kb, longest, 22 kb); and B-genome analysis at 6X coverage yielded reads with mean length of 4.1 kb (longest, 20 kb). These long reads can differentiate A- and B-genomes with very low error. However, the peanut genome is AT-rich which causes assembly bias and gaps. The gaps are reproducible; so very high coverage highly filtered Illumina mate-pair data could be used to scaffold between long-read contigs. Unfortunately, Excelero™ assemblers do not work well with Illumina mate-pairs. It is possible to split the long-reads for assembly with short-read assemblers and then merge data onto existing scaffolds. Also possible to meld data from PacBio™ reads which are unbiased, but may have high error rates. Consensus was reached on Moleculo™ as the best protocol for Illumina mate-pairs. Therefore, a dedicated assembler is needed. Illumina has one under development, without a release date; possible cost for tetraploid assembly might be $60,000 to $120,000. **Shapiro agreed to approach IBM for help; Jackson agreed to draft a one-pager describing the status and assembly challenges posed by the tetraploid genome.**

PLANS and PUBLICATION: An ad hoc PGP Technical Committee (Jackson, Scheffler, Ozias-Akins, Froenicke, D. Bertioli, Cannon, Xin Liu, Xun Xu, Michelmore, Varshney, Hovav, X. Zhang, Nwosu. Shapiro, Valentine, Wilson, Guo) convened on July 20 at AAGB-2013 to develop a plan on next-steps and to discuss publication of peanut A- and B-genome data. All agreed that it was appropriate to proceed with the paper as soon as possible. **Michelmore, Froenicke and Shapiro will lead the effort, with Froenicke serving as 'whip'. Froenicke presented a rough draft tentative outline for consideration of collaborating author assignments. Froenicke agreed to provide a revised outline and convene a teleconference among authors to establish timelines and format requirements**. The group agreed on three additional actions to be completed **within the next two months: 1) Provide Shapiro with a justification for development of an dedicated assembler for Moleculo™ data (Jackson); 2) a full**

strength test of the BAC pool strategy (BGI); and 3) generating Moleculo™ data from the tetraploid genome (Michelmore/Froenicke).

OTHER BUSINESS:

- Wilson requested receipt of AAGB-2013 photos by email or at APRES.
- Valentine reported that David Hoisington will replace Tim Williams in USAID at UGA.
- Nwosu proposed future lab site-visits among PGP members to enhance coordination, cooperation.
- All agreed that AAGB-2014 should be held on a date between August and October 2014. Potential locations were: Saskatoon, Africa, Georgia, California. A suitable venue in California was selected as the most convenient for a majority of PGC members.
- Wilson polled the group on plans for meeting at PAG-2014 in San Diego

NEXT MEETING:

8 July 2013 (2:30 pm EDT), prior to American Peanut Research Education Society annual meeting at Brasstown Valley Resort, Young Harris, GA

ADJOURNED