

Minutes v1

Peanut Genome Initiative Meeting (PGI) at Clearwater Beach, FL; 11 July 2016

PARTICIPANTS

Steve Brown	Jennifer Long	David Hoisington	Sanja Monteiro
Mark Burow	Rebecca Bennett	Carolina Chavarro	Charles Simpson
Kelly Chamberlin	Chris Liebold	Walid Korani	Dave MacKill
Charles Chen	David Bertioli	John Bennett	Jake Fountain
Jim Elder	Phat Dang	Daniel Whitehouse	Hannan Jones
Baozhu Guo	Carlos Ruiz	Naveen Puppala	Ethy Cannon
Corley Holbrook	Brian Scheffler	Mark Kline	Ze Peng
Scott Jackson	Barry Tillman	Kim Moore	DMarquine Wilson
Peggy Ozias-Akins	Lutz Froenicke	Doug Mann	Paul Styama
Howard Valentine	Soraya Bertioli	Sam Zhou	Gauphan Agapukal
Richard F. Wilson	Dan Ward	Youjia Chen	Victor Nwosu
Wesley Hancock	Darlene Cowart	Ruoyu Wen	Tom Stalker
Jamie Rhoads	Steven Cannon	Xue Qin	Graeme Wright

Representation in absentia via proxy to (name of voting member) or via teleconference:

Richard Michelmore	Jeremey Schmutz	George Birdsong,
Lutz Froenicke	Hari Upadhyaya (Varshney)	

Jackson convened the meeting. A quorum was established. Agenda and Minutes-v3 (7 December 2015) from Charleston SC were approved.

PGI Membership

There were no motions for changes in current PGI membership, nominations of new members, or changes in Policies & Procedures.

Discussion of the PI475845 *A. duranensis* genome assembly.

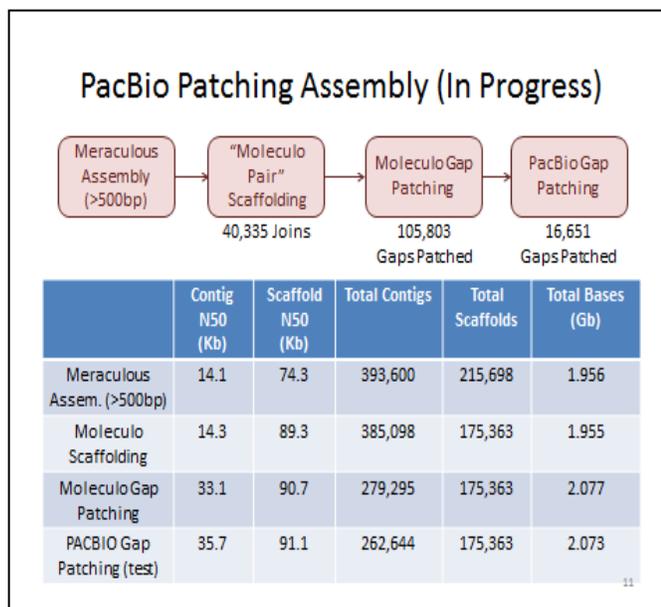
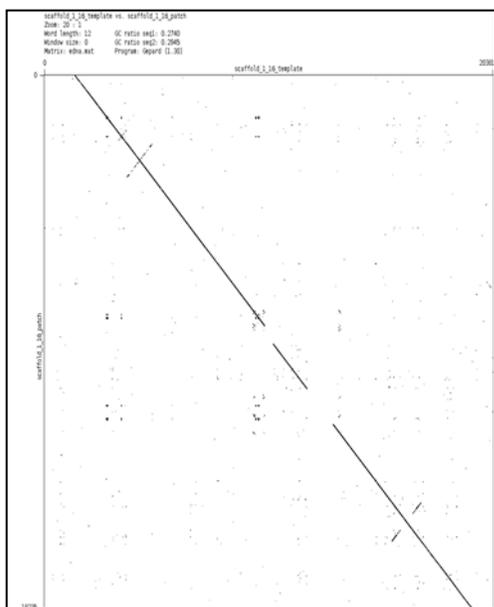
The paper entitled, "Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis and allergens," by Chen et al. was published on-line by PNAS (Proceedings National Academy of Science; doi:10.1073/pnA.1600899113). This work was conducted by the 'Parallel Workgroup in China' and led by PGC member Varshney. Prior discussion is presented for the record. The accession PI475845 (KGBSPCc30070) featured in the PNAS paper is associated with the 'Northern most collection of *A. duranensis* species, while V14167 (which is featured in the Nature Genetics paper) is associated with the Southern-most collection of the species. Both accessions are presumed to be highly diverged. Options were proposed for handling the PI475845 assembly: 1) create a hybrid assembly from both accessions, 2) include the PI475845 assembly in the V2 *A. duranensis* assembly being constructed at Hudson-Alpha, or 3) proceed with V2 without the PI475845 assembly. Action was delayed pending clarification of: 1) discrepancies noted between the PNAS manuscript & supplements; 2) evaluation of data from 454 and Illumina assemblies of the PI475845 genome, and 3) questions regarding how the PI475845 assembly was ordered.

Peanut Genome Project Milestones & Action Items

Status of the DoveTail project: Michelmore presented data that demonstrated substantially improved contiguity of the lettuce genome using the DoveTail 'Hi-Rise' assembly. DoveTail Hi-Rise reduced the scaffold N50 from 1446 with minimum length of 0.476 Mb to 103 scaffolds with minimum length of 6.44 Mb. By prior agreement, Michelmore ceded coordination of future work with DoveTail to Jeremey Schmutz at Hudson-Alpha (in conjunction with Jackson, Ozias-Akins, D. Bertioli at UGA). A contract is in place between DoveTail and Hudson-Alpha for work on the tetraploid peanut genome. Michelmore strongly recommended similar work on both *A. duranensis* (V14167) and *A. ipaensis* diploid genomes. Schmutz received a quote of \$11,800 for the *A. ipaensis* genome. **A motion passed to accept the new quote and to include high molecular weight DNA of both progenitor diploid accessions in new contracts with DoveTail. D. Bertioli agreed to grow the diploid species and provide suitable quantities of tissue to DoveTail. DNA extraction and library construction will be done by DoveTail. Subsequent sequencing will be done or coordinated by Schmutz.**

Progress Report on the Exome Capture Project: Michelmore/Froenicke reported results of work to evaluate exome capture as a useful alternative for assaying genetic variation in peanut. Exome capture targets 1% of the genome that contains transcribed genes that account for a majority of genes annotated in the peanut genome for traits of greatest interest (disease resistance, stress resistance, oil metabolism, nodulation, flowering, nitrogen assimilation). In addition, the exome capture bait-set contained sequences matching about 7,000 previously characterized SNPs. However, exome capture also should reveal variants that are not represented on the 58K SNP chip. Exomes of 28 genotypes have been sequenced to date (11, *A. duranensis*; 1, *A. ipaensis*; 16, *A. hypogaea*-which include the 10 parents used to construct the NAM (CAP) mapping populations). Results suggested that 20X coverage could capture 80% of the total genetic variation in the peanut genome. Priorities for future work were not established among proposed accessions of germplasm collections. **Additional evidence was requested to compare results from exome capture with the 58K SNP chip before recommending additional investment. Clevenger, Ozias-Akins and D. Bertioli will compare post-SWEEP SNP analysis of 10 parents of the NAM populations with variants found in exomes from those lines, and develop recommendations.**

Progress on the Hudson-Alpha project: Schmutz recapped results of the Meraculous assembly of the tetraploid genome. Meraculous generated 1,028,755 scaffolds in a 2.098 Gb assembly with 12.8 Kb contig N50 and 68.4 Kb scaffold N50. These scaffolds were integrated with tetraploid Moleculo scaffolds. Moleculo Gap patching filled about 105,000 gaps and improved contig N50 to 33.1 kb and scaffold N50 to 90.7 kb. The Moleculo Patched assembly will be released to PGC members soon. In addition, high molecular weight DNA from Tifrunner was sequenced by PacBio technology (12 kb mean read length) to finish scaffolding and patching in the next version (V2) of the Arachis tetraploid genome assembly. Interim results showed that PacBio Gap patching filled another 16,000+ gaps. PacBio data collection at Hudson-Alpha and USDA-ARS at Stoneville MS is expected to be completed in 3 to 4 months. Chromosome scale models produced by DoveTail will be linked with the PacBio denovo assembly if possible. About 80% of the tetraploid genome assembly should be completed early in 2017. Collection of PacBio sequence data on 100 clones of Tifrunner BAC libraries (from BIOCCOMPARE) has not been initiated. Examples of PacBio patching and assembly progress are shown below.



Improvements in SNP Discovery and Genotyping Tools. Ozias-Akins reported that the efficiency of SNP-calling based on data generated with the 58K SNP chip may be only about 10% unless tools such as SWEEP (developed by Clevenger) are used to eliminate false positives. Typically, about 40% of SNPs discovered in resequenced tetraploid genomes are true positive. This condition poses a substantial restraint on the implementation of GBS as a useful breeding strategy. However, work at Tifton may present a remedy to this problem. For example, machine-learning technology (developed by Korani) is being applied to SNP-calling. Several models were compared. ‘Neural Net’ and ‘Treebagger’ models

performed best, generating about 80% true positive genomic SNPs in blind validated supervised comparisons. Preliminary results also show this approach may improve the efficiency of true positive SNP-calling in RNA-seq data from 67% to 90%. Work also is underway to create second generation arrays with about 6000 validated SNPs from runner and Virginia market-types.

IPGI 2015-2016 Research Accomplishment Report: Brown requested input on the current draft of the subject report. A final draft will be submitted to TPF Board by July 31, 2016.

IPGI Strategic Plan 2017-2021: A final draft plan was accepted by The Peanut Foundation Board during the World Peanut Congress in Charleston SC in June 2016. The new plan is now in effect.

AAGB Update: Stalker reported on logistics and issued the first-announcement flyer for AAGB-2017 in Cordoba, Argentina on March 14-17, 2017. Wilson will post AAGB-2017 meeting information on PeanutBioscienc.com as materials become available. Brown reported that a proposal to host a future AAGB will be submitted by members in China.

Next Meetings:

TPF Review of 2016 projects will be scheduled during the APC Winter meeting (date & location TBD)

An informal PGI meeting will be held during PAG XXV in San Diego CA (January 14-18, 2017). Wilson will survey PGC members for a convenient time/date and make arrangements.

IPGI during AAGB-2017 in Cordoba, Argentina: March 2017

ADJOURNED