

IBSC business meeting

The annual IBSC business meeting was held Jan 15, 2011, in conjunction with the XIX PAG conference in San Diego.

Attendants: Alan Schulman, Robbie Waugh, Tim Close, Chengdao Li, Peter Langridge, Tsuyoshi Tanaka, Takashi Matsumoto, Gary Muehlbauer, Roger Wise, Kazuhiro Sato, Burkhard Steuernagel, Matthew Moscow, Maria Munoz, Peter Morrell, Jesse Poland, Klaus Mayer, Nils Stein

Agenda:

It is IBSC consent to elect the chair/co-chair of the consortium always for a period of three years.

Top1: election of IBSC chair / co-chair (unanimous)

- ➔ Nils Stein was re-elected as chair, Gary Muehlbauer as co-chair for another period of three years.

Top2: reports of ongoing projects (1h):

- **physical map**, BACs, BES, 3H (Nils). Nils provided more details on current status. BAC libraries show some differences—complementary. Assembly with 10% questionable clones, goes to only 4.4 Mb. 5% questionable goes to 5.3 Mb, closer to genome size. 1600 contigs do 50% of the assembly, the longer contigs of course are more genic. Anchoring info from various data sets.... total anchored contigs now 64% of the assembly / almost 90% of the cumulative contig length. 6k contigs are anchored by at least 1 anchor, 2.6 anchor points average per contig. 46k BACs anchored. Draft FPC version is accessible via <http://mips.helmholtz-muenchen.de/plant/triticeae/barleyDisclaimer.jsp> .
- **whole genome shotgun sequencing (Burkhard)** 28X paired-end (500 bp fragment size), read length 100, 25X mate-pair 3500, read length 100/150 on Illumina GA2. Estimate or real insert size – mapped reads against barley chloroplast. Actually 220-500 for paired-end for mate paired, much wider! CLC assembly cell worked much better than SOAP deNovo. Now 1.87 Gb of contigs, 2 million contigs. With f1cDNA, can find genes in the assembly! Read-based chromosome assignment (information from flow-sorted chromosome sequencing) of WGS assembly contigs is now 776 Mb – in other words 20% of all contigs can be anchored, covering half the complete dataset.

Roger enquires about the possibility to bring dataset in context of chip-based expression data.

Next steps: add 5X data from Udine, better assembly method. Anchoring of contigs to physical map in progress at MIPS.

- USDA sequencing project (Tim) USDA-AFRI project.** HighCot has been dropped in favor of shotgun sequencing to find the lower copy portion of the genome. Still doing hypomethylated genomic DNA at Bennetzen lab, expect to have data near end of February. HpaII and HpyCH4IV, Hpa+HPy, all with Illumina. Transcript sequencing was completed in Summer 2010 and the assembly is available online at www.harvest-web.org for BLAST and sequence retrieval. Only Morex was used for support of the USDA-NIFA project objectives, 88% of HarvEST:Barley assembly #35 unigenes whose members contain at least one Morex EST have high hits, ~25% of sequences have no significant hits to assembly #35 unigenes. Roger: there will be a lot of pathogen sequences there. 27% "new sequences". These transcript sequences were produced using Illumina GAII 2 x 78 from three cDNA libraries containing complex mixtures of RNA from developmental stages and after pathogen infection. Pathogen-infected RNA was provided by Brian Steffenson, Gary Muehlbauer and Roger Wise. The Close lab produced RNA from roots, shoots, shoot apices and young inflorescences. Draft genomic sequence assemblies are being released periodically for BLAST and sequence retrieval through www.harvest-web.org. The most recent release prior to PAG (from October 2010) was "Barley Genome 0.02", assembled using SOAPdenovo by undergraduate UC Riverside Computer Sciences student Matt Alpert in consultation with Close, Stefano Lonardi and others at UC Riverside. This assembly was made using Illumina 5.5 X coverage, 4.5X of which is paired end 70-100 bases, and 1X is paired-end of 2kb, 3kb, 5 kb "mate-pair" ~40 bases. Version 0.02 had high hits to 83% of known-genes ID'd with Bowtie. More lanes were running on HiSeq 2x125 during PAG, anticipating at least 10X coverage in the next version 0.03. BAC pool sequencing to assign genes to BACs using deconvolution of the shifted transversal design combinatorial is progressing. 14,674 MTP gene-bearing BACs selected from the assembly posted in 2008 on <http://phymap.ucdavis.edu/barley/> have been re-arrayed into 637 pools, and all pools have been applied to GoldenGate assays, relating 1967 SNP loci to BACs and validating the quality of the pools. The MTP BACs that have been directly related to SNP loci tie BAC contigs containing 14,699 BACs to the SBP genetic linkage map. Preliminary sequencing results from the BAC pools indicate that at least 20,000 genes can be related to specific BACs, if the USDA-NIFA project budget can afford to sequence all 637 pools. The project promised only to handle one set of 2197 BACs hoping to anchor ~2,000 to 3,000 genes to BACs, mainly on 4H, 5H and 6H, but it is much more efficient to tackle the entire MTP than to prioritize by chromosome or in response to specific requests from investigators.
- 5H sequencing (Chengdao):** Three activities ongoing: One with BAC set for MTP, to test seq method. 5H, 4 contigs, 25 BACs to be sequenced. Other 30k BACs for chrom 5H. Second activity RNAseq for developing and germinating seeds. 50k primary reads. Third activity WGS of 5H. Discussion on sequencing of pools from MTP rather than barcoding individual BACs.
- Haruna Nijo BACs (Kaz)** Pooled BAC clone sequencing of Haruna Nijo: 460 BACs from 3H and 140 from 5 HL. 5X depth of coverage of Haruna Nijo on 500 bp fragment, and also some 8 kb paired end and 20 kb paired end read by 454. 5X reads give 1.7M contigs. Combined with 8k PEs, gives 174k scaffolds, avg 2.9 kb. With 20k PEs, did not increase scaffold size.

Top3: new projects (30 min):

- **JGI barley WGS (Gary)** Rokhshar wanted to “see Community’s shotgun sequence” before deciding on whether to do more 100X of Morex vs doing something else. Tim sent 30 ug Morex, but JGI got only 14 they claim. They say that need 300 to sort out how to make the libraries. Consensus: Need to do more DHs, to get more coverage. Consensus: need to get JGI into working with the Community and adding to what it is. Discussion of genotyping-by-sequencing. → JGI will be updated by Gary. Need to deal with Dan and with Kerry Barry.
- **Tritex barley+ WGS (Nils) 2.8** MEURO for improving barley phys map and sequence, anchoring all barley genes by HT-sequencing, further integration of data. FPC map of wheat 6A.

Top4: next activities required (30 min):

- assembly
- **annotation and gene nomenclature (Klaus)** – Arabidopsis nomenclature, but not defacto standard. Has been used for rice etc. Then maize ID system. But doesn’t conform to other finished genomes. Should coordinate with other Triticeae communities! could be TrAes3bs123456UZHc1v2 c1 means coding seq 1 or p for protein or t for transcript. → Approved. → We come up with a list.

AOB

Scheduled meetings and conference calls:

April 15, 2011; 4 pm GMT+02:00: conference call via webex

Tentative: June 6-8, 2011: IBSC meeting in conjunction with the NA barley researchers conference