Physical map and sequencing of wheat chromosome arm 7DS

Helena Staňková[1], Ming-Cheng Luo[2], Paul Visendi[3], Zuzana Tulpová[1], Jacqueline Batley[4], Jan Bartoš[1], Jaroslav Doležel[1], David Edwards[3], Hana Šimková[1]

[1] Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Šlechtitelů 31, CZ- 783 71, Olomouc - Holice, Czech Republic
[2] Department of Plant Sciences, University of California, Davis, CA 95616
[3] School of Agriculture and Food Sciences and Australian Centre for Plant Functional Genomics, University of Queensland, Brisbane, QLD 4072, Australia
[4] Centre for Integrative Legume Research, University of Queensland, School of Agriculture and Food Sciences, Brisbane, QLD 4072, Australia

Background

Bread wheat (*Triticum aestivum*) is one of the most important crop species in the world providing staple food for 35% of world's population. Physical map construction and sequencing of wheat genome are hampered by huge genome size (1C = 17 Gb), presence of three homeologous subgenomes (A, B and D) and prevalence of repetitive sequences (>80%). The possibility to divide the wheat genome into individual chromosome arms by flow cytometric sorting enables to cope with polyploidy and to focus on the desired part of the genome.

The aim of this work was to construct physical contig map of 7DS chromosome arm and to sequence minimum tilling path (MTP). BAC clones from the 7DS-specific BAC library were fingerprinted using SNaPshot-based HICF technology and than automatically assembled into contigs using FPC software. Integration of the 7DS physical map with that of *Aegilops tauschii* (D genome ancestor) provided a clue for further merging of contigs. Reliability of the assembly was verified through LTC software. The physical map has been anchored to the genetic map applying both forward and reverse anchoring strategy. BAC clones representing 7DS MTP are being sequenced by Illumina platform and assembled into sequence contigs using Sassy software which was designed to assemble short complex sequences from Illumina paired read data. Anchored 7DS physical map and the sequence of the 7DS MTP will become a valuable tool for genetic mapping and positional cloning of genes located on 7DS chromosome arm of wheat.

Physical map construction

1) Fingerprinting

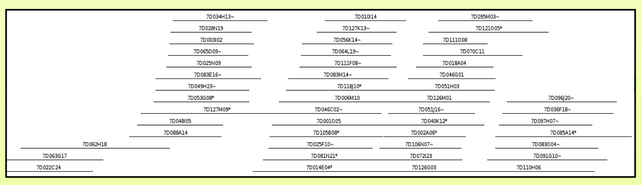
- All 49,157 BAC clones from 7DS-specific BAC library were fingerprinted with the SNaPshot high-information-content-fingerprinting (HICF) technology
- 39,765 fingerprints were useful for contig assembly

2) Contig assembly

- Clones were assembled into contigs based on fingerprint overlaps

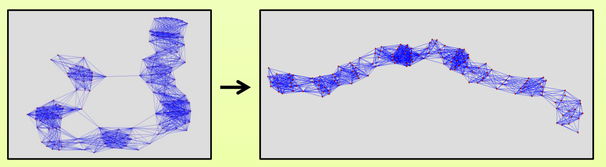
FPC software (FingerPrinted Contigs)

- Automatic assembly up to cut-off value $1e^{-45}$
- Manual end-merging of automatically assembled contigs
  → Based on integration of *Aegilops tauschii* physical contig map with physical contig map of 7DS chromosome arm



LTC software (Linear Topology Contigs)

- Three-dimensional view of contigs assembled by FPC software → verification of the assembly made by FPC
- Contigs were manually edited in FPC based on the 3-D view from LTC



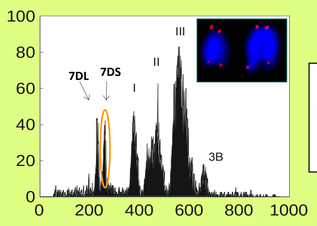
Selection of MTP BAC clones

- Selected from final assembly
- FPC software

7DS-specific BAC library

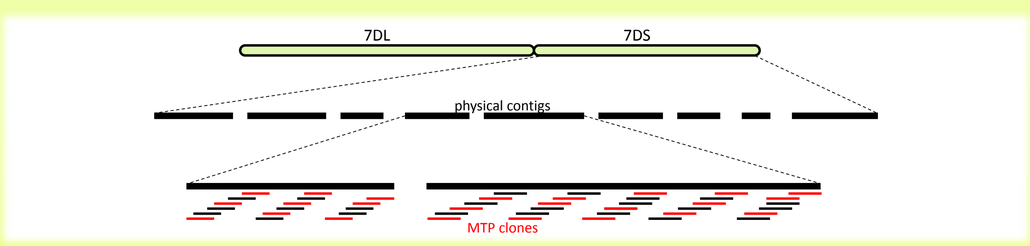
Size of the 7DS chromosome arm is 381 Mb (2,2% of the wheat genome)

7DS arms were flow-sorted from 7DS/7DL double ditelosomic line of cv. Chinese Spring



Library characteristics

Number of clones	49,152
Average insert size	113 kb
Coverage	12.2x

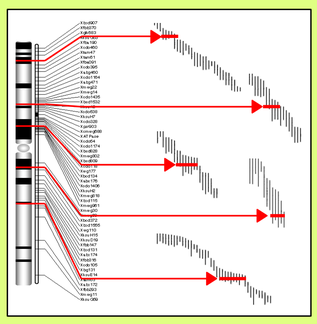


Parameters of the final assembly

Length of the assembly	362 Mbp (95% of the 7DS)
Number of clones	39,765
Number of contigs	931
Number of singletons	11,426
Number of MTP BAC clones	**4,608**

Physical map anchoring

Landing on the chromosome



Forward anchoring

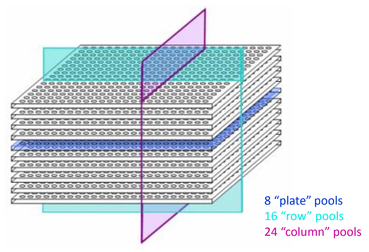
- **PCR screening of 3-D pools** from 7DS-specific BAC library with genetically mapped wheat SSR and STS markers
- ***In silico* anchoring** – integration of 7DS physical contig map with anchored physical contig map of *Aegilops tauschii*

- 31 markers anchored by PCR screening (20 SSR markers, 11 STS markers)
- 583 markers anchored *in silico*

BAC library screening

3-D pooling strategy
Organization of the three-dimensional pool for a set of eight 384-well plates

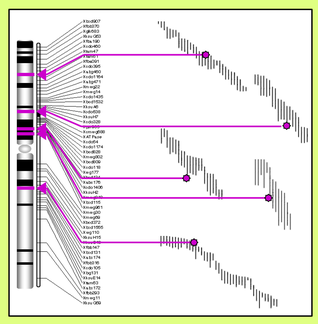
48 PCR reactions to screen 3072 samples



Reverse anchoring

- STS markers were developed based on sequences of MTP BAC clones
  - 9216 BAC-end sequences
  - Sequences of 7 completely sequenced and assembled BAC clones
- Markers were mapped on D-genome specific radiation hybrid (RH) panel (Kumar *et al*., 2012)

- 2 STS markers were mapped on D-genome specific RH panel



Sequencing of 7DS MTP

Sequencing of MTP BAC pools

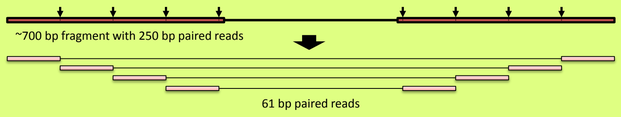
- 4 BAC clones were pooled into one pool
  → **1152 BAC pools** out of 4608 MTP BAC clones
- DNA sequencing libraries were constructed using TruSeq DNA PCR-Free Sample Preparation Kit
  - Size selection – fragments size ~550 bp
- Pools were sequenced on Illumina MiSeq sequencer
  → 250 bp pair-end reads

BAC-end sequencing

- All MTP BAC clones were sequenced by Sanger sequencing from both ends
  → **9216 BAC-end sequences**

Assembly of sequencing data

- Illumina pair-end reads were assembled into sequence contigs using **Sassy** software
- 250 bp paired reads from each BAC pool were cut into 61 bp offset paired reads and assembled



- BAC-end sequences were mapped to the sequence contigs
  → The ends of BAC assemblies and individual BACs in the pool were identified
- The assembly is supported by whole genome mate-pair data
- Sequence reads of individual BAC clones were assembled into 1 - 7 sequence contigs per BAC

Results and conclusions

- Physical map was anchored by 616 markers located in 309 contigs → **51%** of the assembly
- Wheat genome sequencing by proposed strategy is feasible
- Sequencing of all 7DS MTP BAC clones is in progress
- Sequences of BAC clones from the 7DS MTP will be used for further anchoring of the 7DS physical map

Reference

Kumar, A., Simons, K., Iqbal, M.J., Michalak de Jimenéz, M., Bassi, F.M., Ghavami, F., Al-Azzam, O., Drader, T., Wang, Y., Luo, M.-Ch., Gu, Y.Q., Denton, A., Lazo, G.R., Xu, S.S., Dvorak, J., Kianian, P.M., Kianian, S.F. (2012): Physical mapping resources for large plant genomes: radiation hybrids for wheat D-genome progenitor *Aegilops tauschii*. BMC Genomics 13: 597.

This work has been supported by the Czech Science Foundation (P501/12/2554), MSMT CR and EU (Operational Programme Research and Development for Innovations No. ED0007/01/01) and Internal Grant Agency PrF-2012-003.