Project 7B – progress report.

Olsen, O.-A., Sandve, S.R, Kent, M., Belova,T., Rudi, H., Zhan, B. Lien, S. Norwegian University of Life Science, CIGENE/IPM, Norway.
Gullord, M. Graminor A/S, Norway.
Simkova, H., Kubalaskova, M, Dolezel, J. Institute of Experimental Botany, Laboratory of Molecular Cytogenetics and Cytometry, Czech Republic.
Cattonaro,F. Istituto di Genomica Applicata and IGA Technology Services, Italy
Kianian, S. Kumar, A. North Dakota State University, Department of Plant Sciences, USA.
Korol, A. University of Haifa, The Department of Evolutionary and Environmental Biology, Israel.
Li, L., Min,J., Chen, J., Yang,Y., Xu,X., Beijing Genomics Institute, China, Hu, J.

Talk outline.

- Chromosome sorting and BAC construction

1. 7B BAC fingerprinting and MTPs.

2. Integration of mate pair data in shotgun assemblies.

3. Development of Niblegen-chip for radiation hybrid and deletion mapping.

4. BAC-by-BAC sequencing.

5. Towards an endosperm cell/homeologue specific transcriptome.

1. 7B BAC fingerprinting and MTPs.

- 27648 BAC clones fingerprinted for 7BS:

  22000 BACs for MTP assembly, 12x coverage.

  Average insert size: 185kb.

- 60798 BAC clones fingerprinted for 7BL:

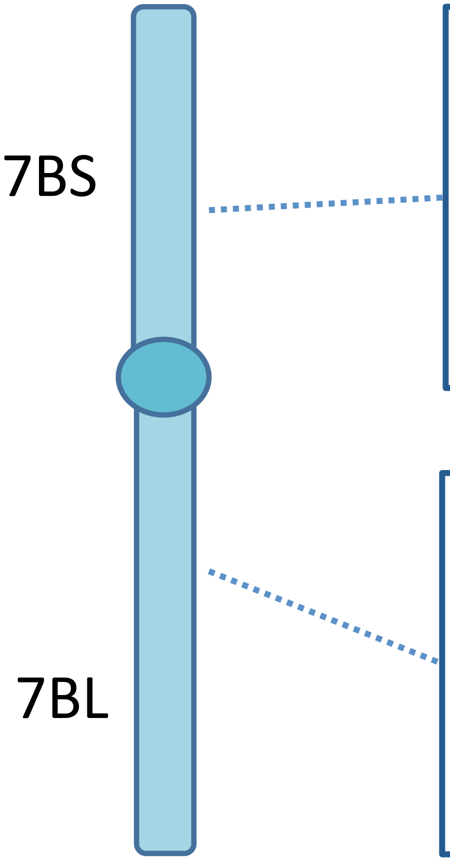
  44000 BACs for MTP assembly, 11x coverage.

  Average insert size: 120 kb.

Cattonaro,F. et al.

1. 7B BAC fingerprinting and MTPs cont´d.

MTP status



BAC’s in MTP: **3039**
BAC contigs: **255**
Mean BAC’s in contig: **12** (68-2)
Short arm coverage: **107%**

BAC’s in MTP: **5529**
BAC contigs: **286**
Mean BAC’s in contig: **18** (92-2)
Short arm coverage: **92%**

S.R.Sandve

A.Korol et al.

2. Integration of mate pair data in shotgun assemblies.

“Integration of mate pair sequences to improve shotgun assemblies of flow-sorted chromosome arms of hexaploid wheat” (MS ready for submission).

Tatiana Belova, Bujie Zhan, Jonathan Wright, Mario Caccamo, Torben Asp, Matthew Kent, Sigbjørn Lien, Frank Panitz, Christian Bendixen, Hanna Simkova, Jarozlav Dolezel, Odd-Arne Olsen and Simen R Sandve

-Paired-end seq.:	-Mate pair Illumina seq.:
350bp insert size	2,3,5Kb insert sizes
59x coverage 7BS	89x coverage 7BS
37x coverage 7BL	64x coverage 7BL

2. Integration of mate pair data in shotgun assemblies, cont´d.

Arm	Assemblies	Assembly size*	N50 (bp)	Mean contig size (bp)	Max length (bp)	Total (Mbp)
7BS	**Paired end**	178.789	2428	1152	50.938	206.1
	Mate pair	129.684	11033	1873	129.404	242.9
7BL	**Paired end**	328.725	1556	789	30.964	259.6
	Mate pair	260.821	8310	1201	122.119	313.3

- 5 fold increase in N50
- 50% increase in sequences with >2 FL genes
- 12% increase in FL genes

- FL genes coverage >70% of Brachy homologes

3. Development of Niblegen-chip for radiation hybrid and deletion mapping.

Radiation hybrid panel

Irradiate seeds at
25,35,45 krad

Langdon 7D(7B)
Chr 1-6: AABB
Chr 7: AADD

Langdon
Chr 1-6: AABB

RH panel for chromosome 7B

300 RH_1 lines selected with retention frequency 10-90%

T.Belova

S.Kiania. & A.Kumar

3. Development of Niblegen-chip for radiation hybrid and deletion mapping, cont´d.

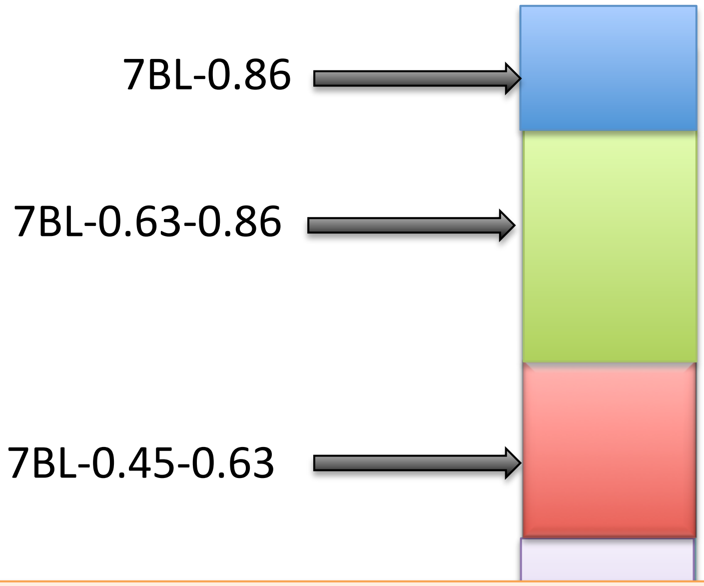
Probe selection for 7Bs and 7BL- Deletion BIN mapping

- Niblegen comparative genome hybridization chip:
  - 67500 probes, 4 replicates
    - 18000 control probes
    - 49500 7B specific probes
      - 19331 7BS
      - 30169 7BL
- Theoretical resolution 1 probe per 20Kb
  - 5-10 probes per BAC

T.Belova, S.R.Sandve, B.Zahn

N.Springer, D.Edvards

Probe verification for 7Bs and 7BL- BIN mapping


7BL-0.86

7BL-0.63-0.86

7BL-0.45-0.63

probes	contigs	scaffolds
2413	1211	339
2914	1465	472
1565	750	233

Radiation H_1-hybrid lines under characterization using these 10817 probes

T.Belova et al.

4. BAC-by-BAC sequencing

- Test run: 1000 BACs form the 7BL MTP

- 500bp Paired-end sequencing library
- Multiplexing on HiSeq2000 (clean data >92% Q20)
- Average sequencing depth: 30x

B.Zahn

Li, L., Min,J., Chen, J., Yang,Y., Xu,X.,

Analysis pipeline (BGI)

Raw data

Clean data

Split reads by index

Filter pollution reads

Assembly

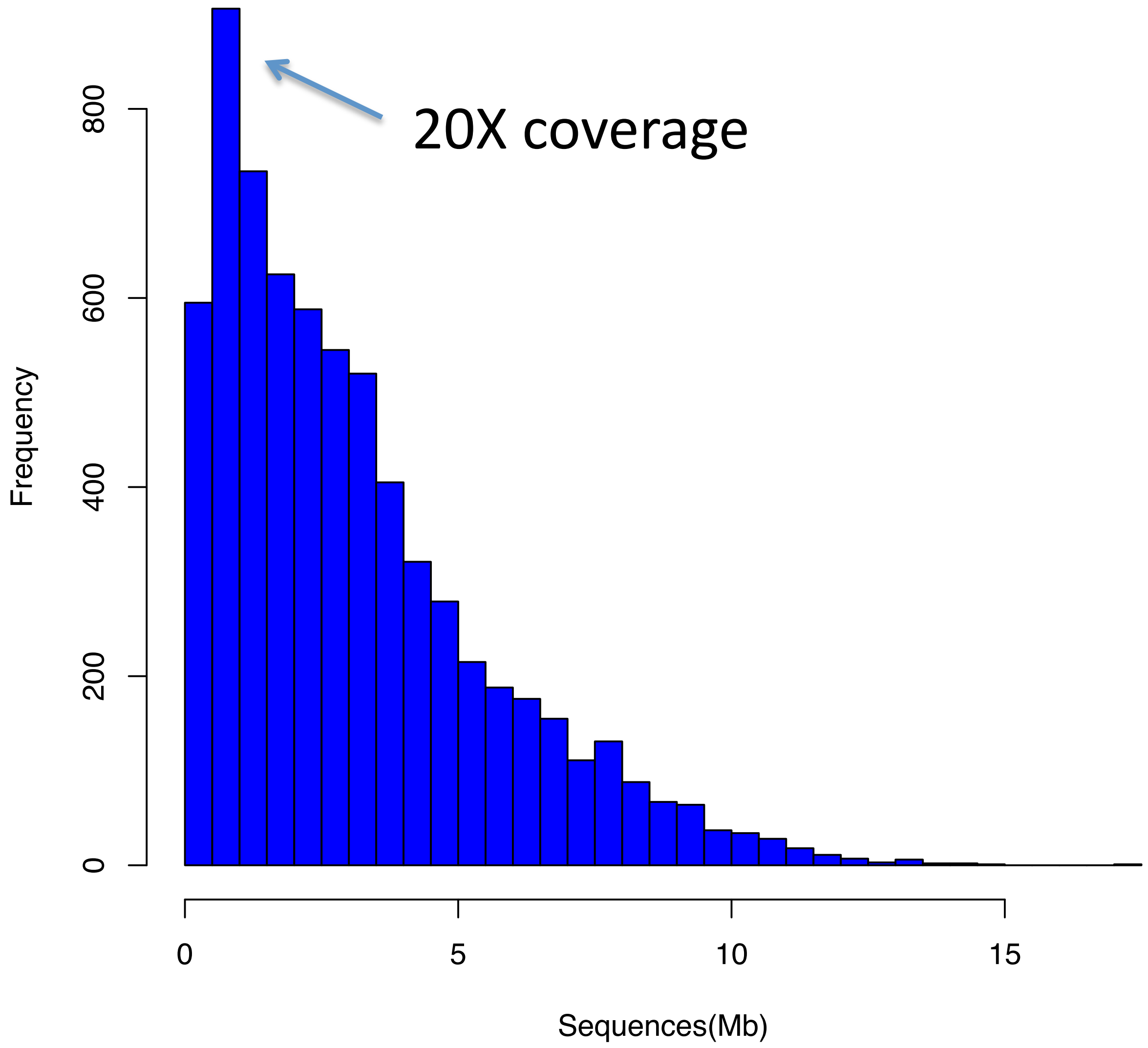
Filter pollution sequences

BGI analysis

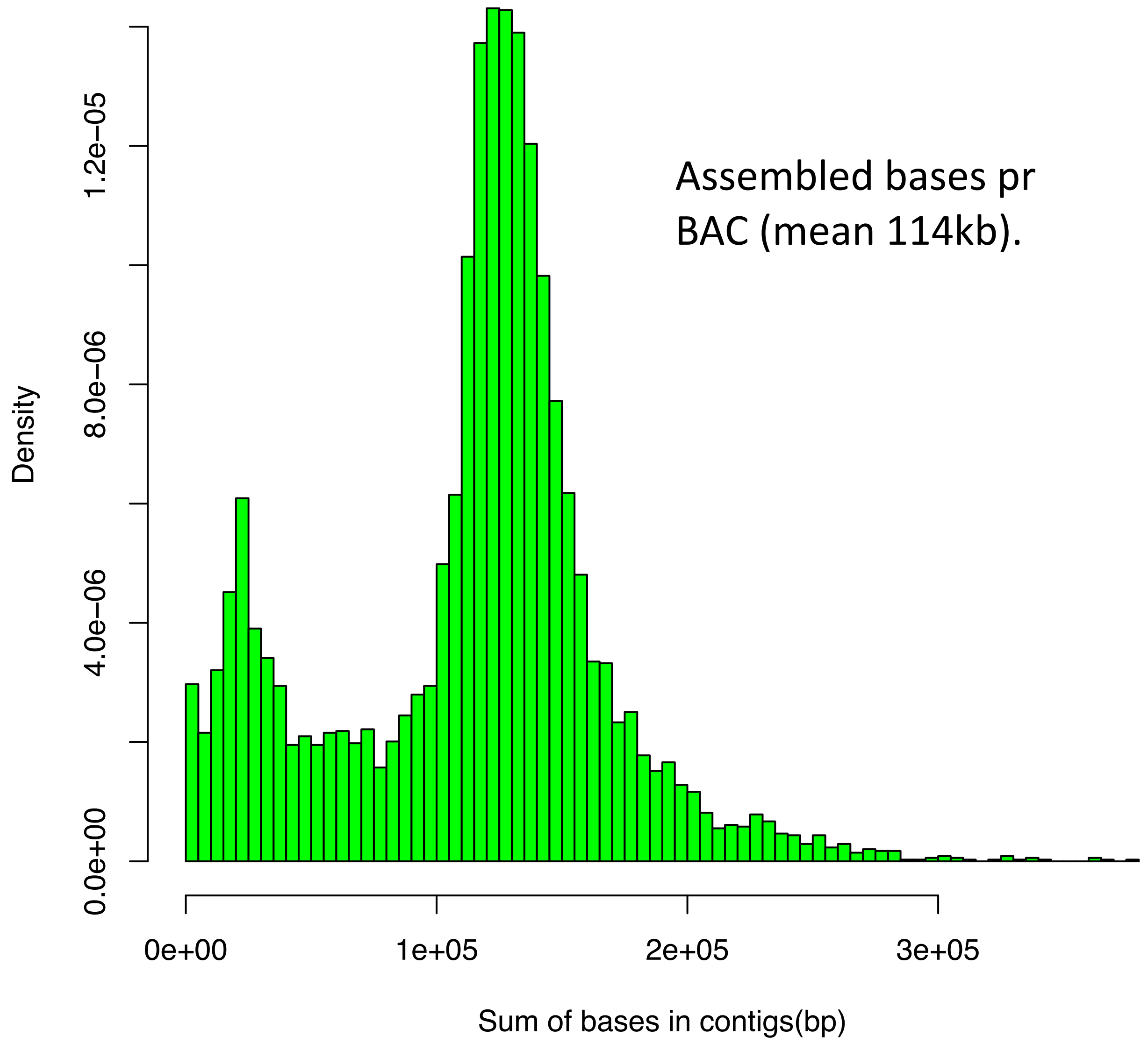
Final sequences

Norway analysis

Uniformity of Multiplexing: only a small fraction of the BACs failed to produce sufficient coverage

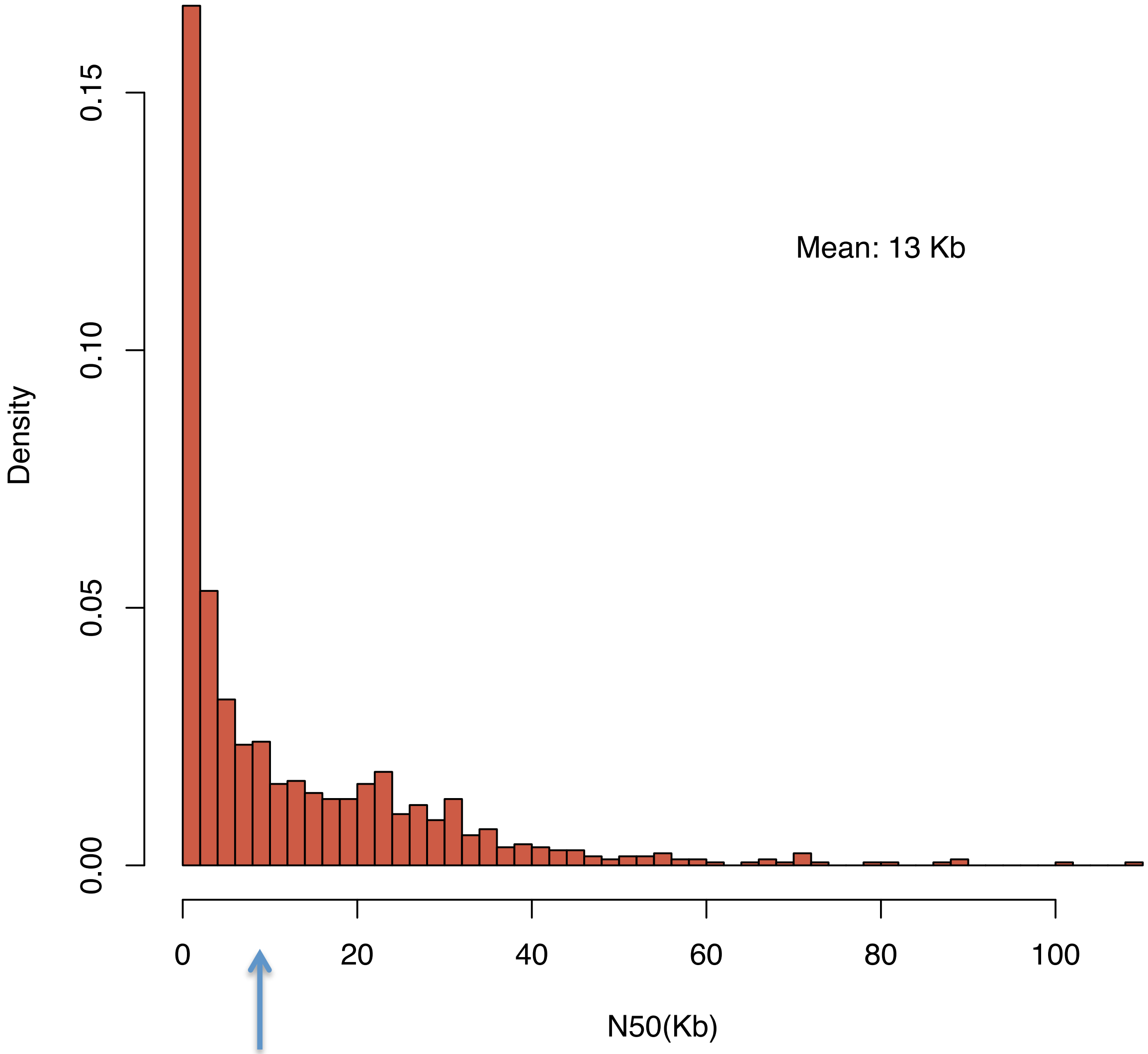


Assembly length of BAC clones: Most of the MTP BACs has an assembly size which correspond to the expected BAC size of 136Kb

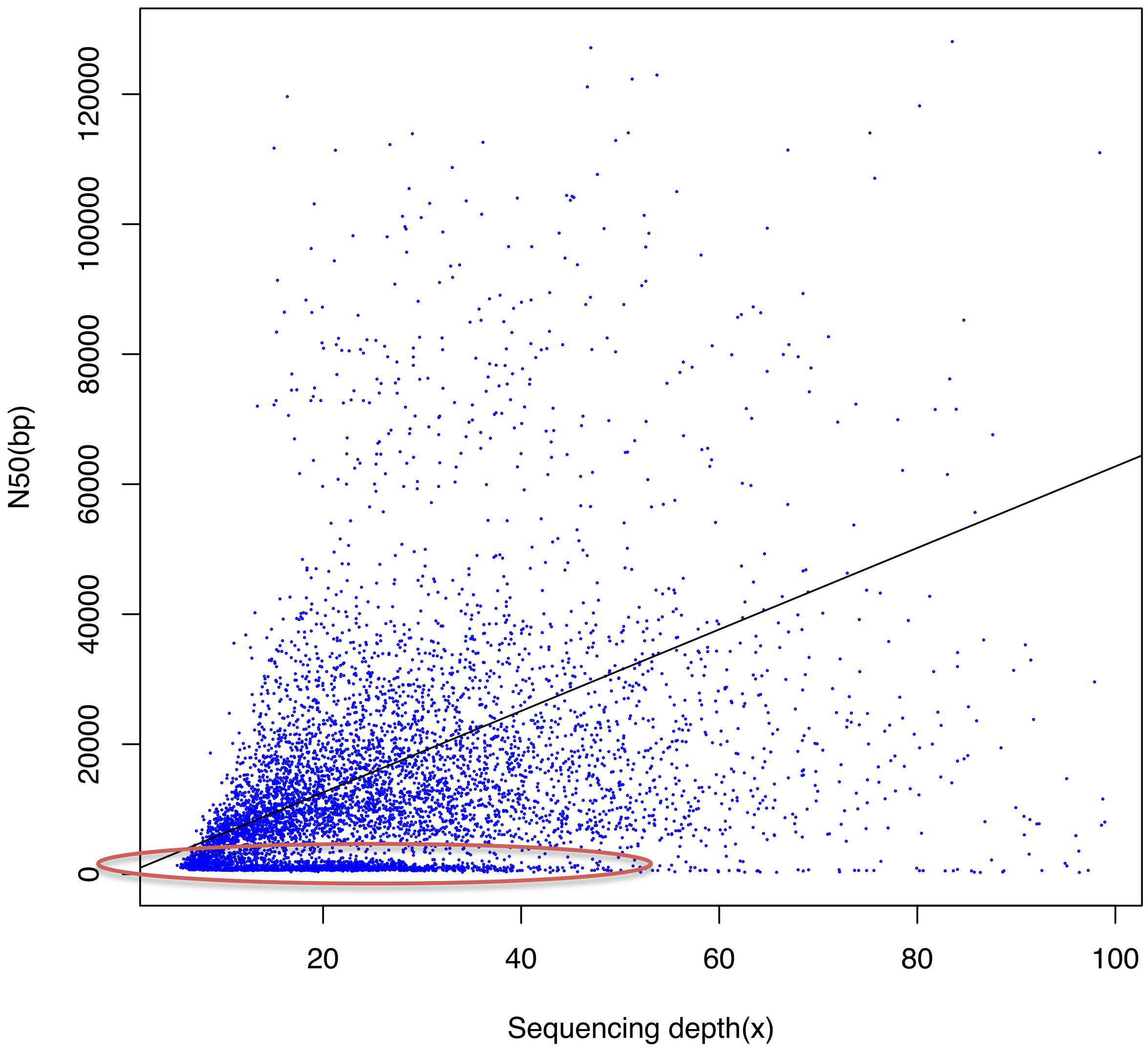


Assembled bases pr BAC (mean 114kb).

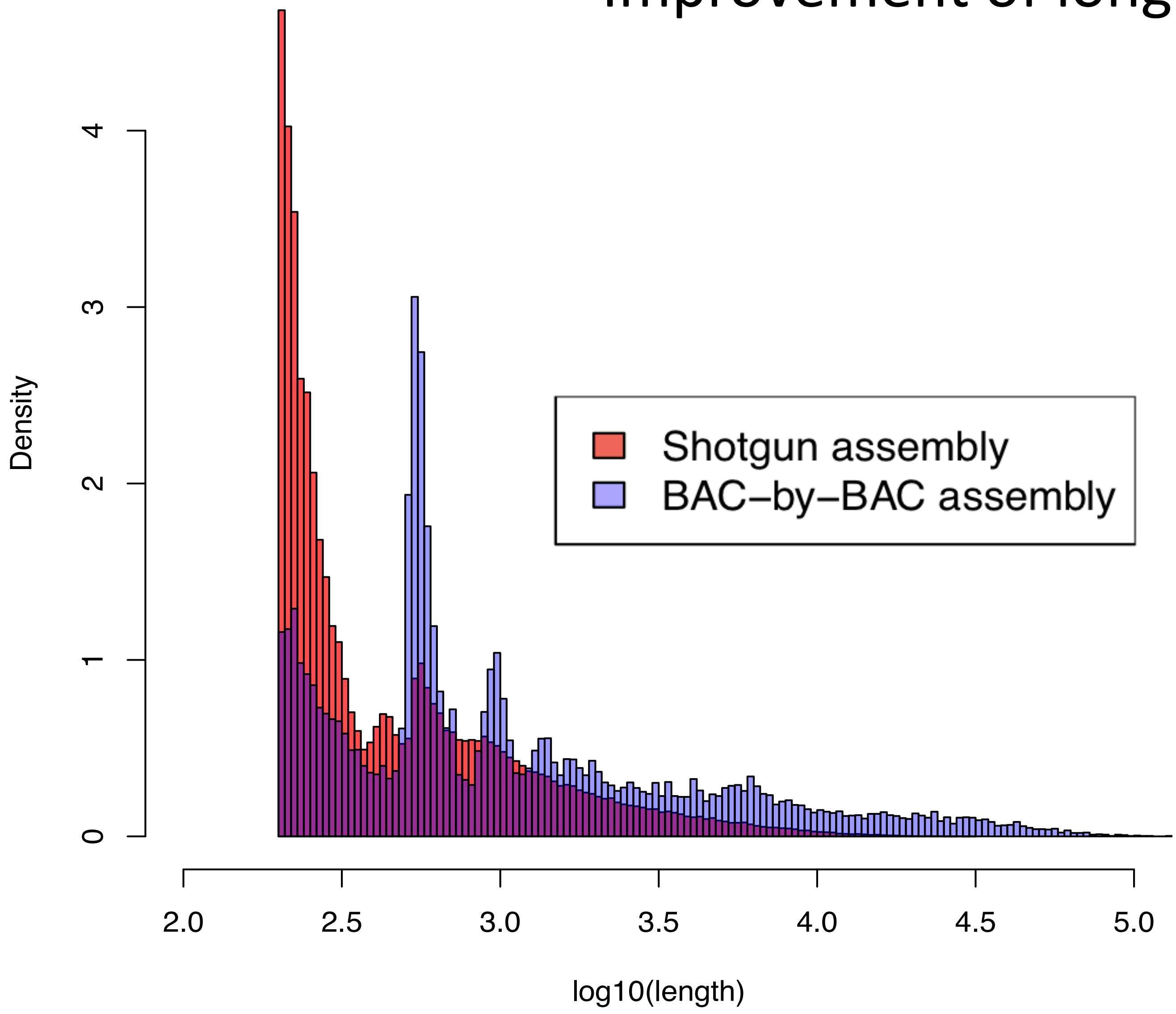
N50 of BAC clones: Mean is 13 kb: some BACs hard to assemble due to low coverage and/or repetitive sequences.



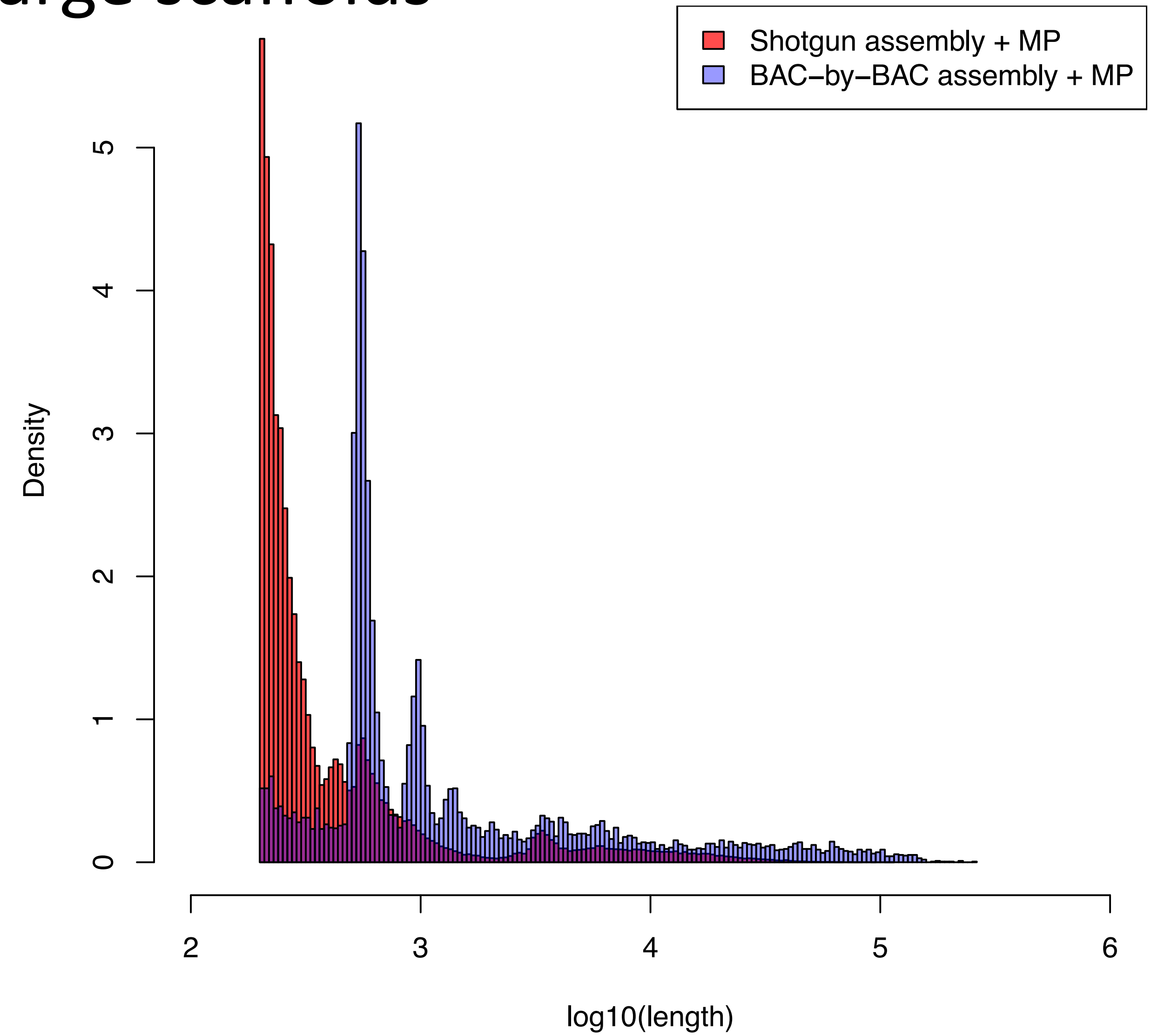
N50 vs Sequencing depth: generally a linear relationship, but some BACs are hard to assemble irrespective of depth of coverage.



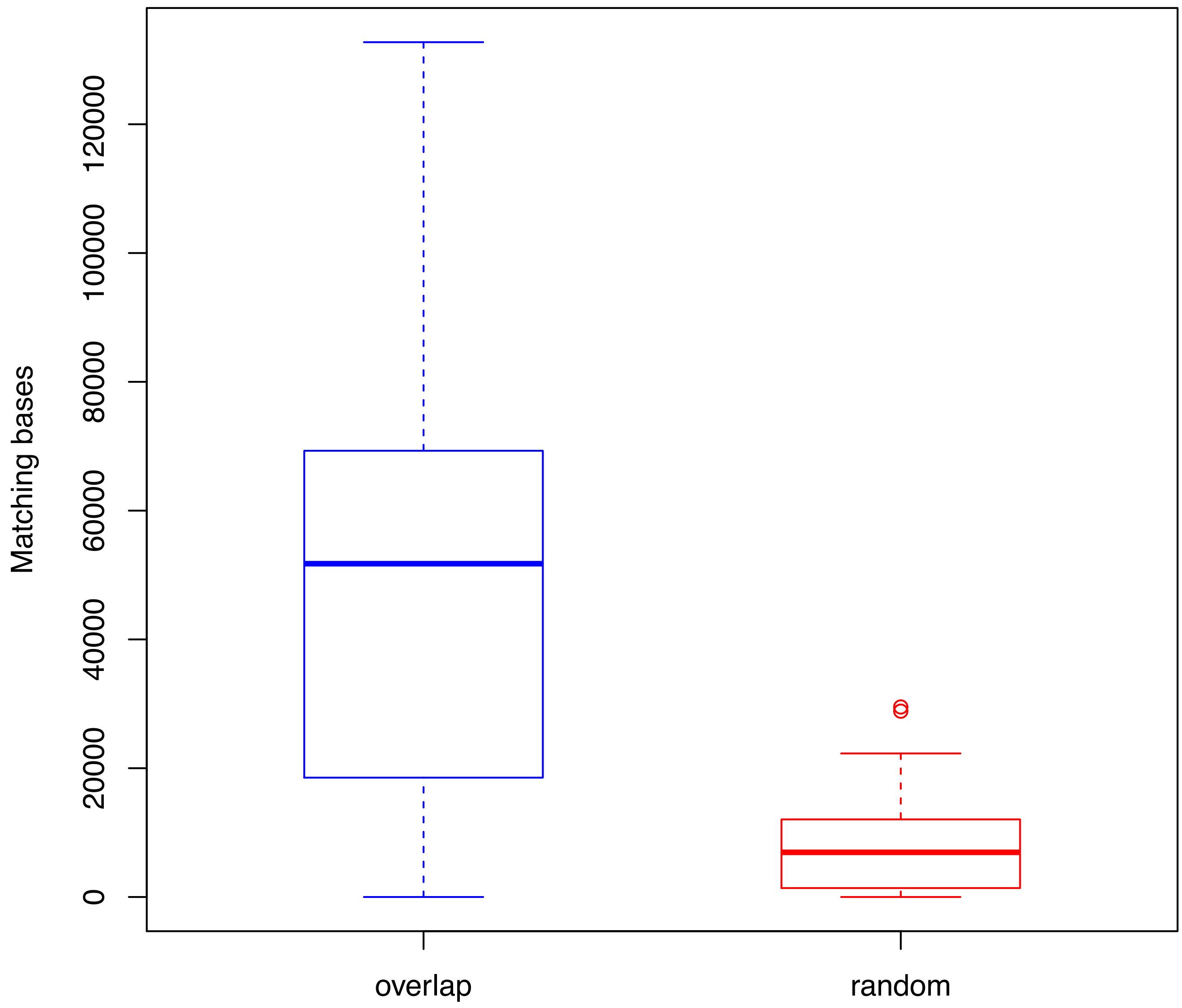
A comparison of the 7BL shotgun assembly (red) with the BAC by BAC assembly (blue/purple) : A substantial Improvement of long contigs



Scaffolding with BAC-by-BAC vs shotgun assemblies with MP reads put lots of contigs into large scaffolds



Clones that are predicted to overlap in the MTP (fingerprinting) also have overlapping sequences



Gene coverage of assembly

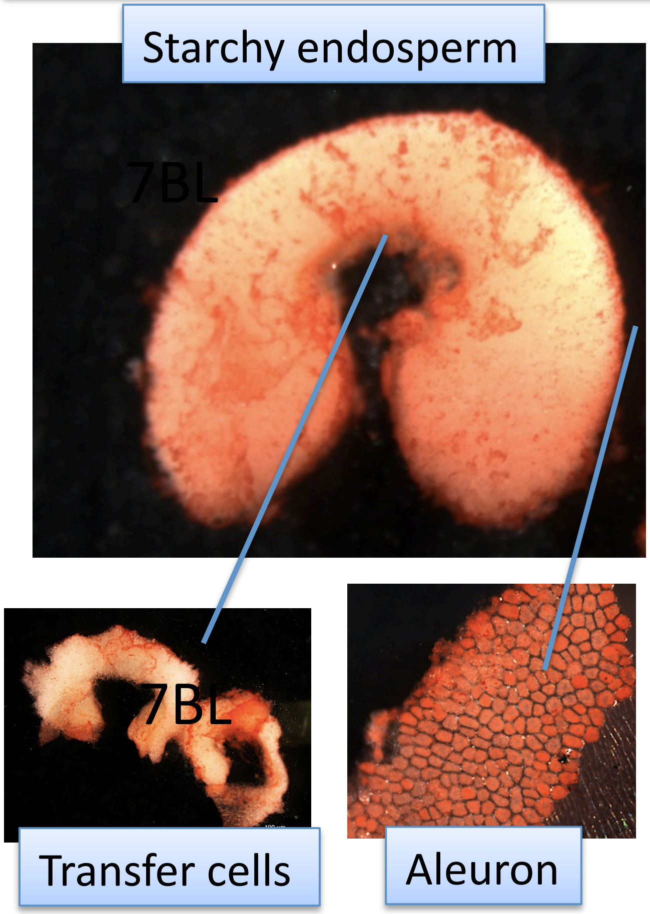
Data	Assembly	N50	Full length CDS	CDS fragmentation index	FL-CDS in multigene contigs
7BL-full	SSPACE	8.3 Kb	28%	17.4 %	21%
7BL MTP (10%)	SSPACE	50 Kb	42%	7.5 %	59%

Summary of 7BL assembly (BAC-by-BAC in MTP)

Contig#	20,766
Total bases	68,686,237
Minimum length	200
Maximum length	131,143
N50	16,279
Average	3307

Thanks BGI!
Li, L., Min,J., Chen, J., Yang,Y., Xu,X.

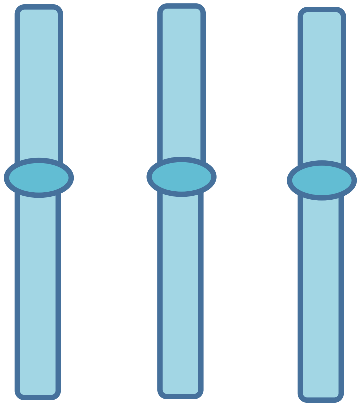
5. Towards endosperm cell/homeologue specific transcriptome for wheat baking quality



H,Rudi

RNAseq

Goal



+

-Aleurone
-Starchy endosperm
-Transfer cells

+

Developmental stages

Thanks to Kellye Eversole
Jaroslav Dolezel group

&

for your attention on a Sunday morning!