5. SUMMARY

1. The present study on bread wheat (*Triticum aestivum* L. em Thell.) was undertaken essentially for physical mapping of additional SSR markers and for construction of an integrated physical map. Both wet-lab and *in silico* approaches were used for this purpose. Wet-lab approach involved use of cytogenetic stocks of Chinese Spring [including 42 nullisomic-tetrasomic (NT) lines, 24 ditelosomic (Dt) lines and 192 overlapping terminal deletion lines], while *in silico* approach involved SSR mining.

2. A total of 429 SSRs (including 154 gSSRs from chromosome 3B, 132 gSSRs from *Ae. tauschii,* and 47 gSSRs + 96 EST-SSRs from brachypodium) were tried for physical mapping using wet-lab approach. Of these, as many as 108 SSRs could be successfully mapped on 128 loci (76 wheat gSSR loci + 17 brachypodium gSSR loci and 35 brachypodium EST-SSR loci). Interestingly, mapping involving wheat SSRs and brachypodium SSRs did not differ. Proportions of EST-SSRs and gSSRs of brachypodium mapped were the same (36%) as those of wheat SSRs**.**

3. Cross-species amplification of brachypodium SSRs was also examined. A total of 89 (92.70%) out of 96 primer pairs of brachypodium EST-SSRs gave amplification in wheat (cv. Chinese Spring); as many as 72/89 (80.89%) of these gave products of expected size indicating high level of cross-species transferability.

4. An integrated physical map of SSRs comprising 1,903 SSR loci, developed earlier in our laboratory, was updated by combining the above 128 SSR loci mapped during the present study, so that an integrated physical map with 2,031 SSR loci became available. The distribution of mapped SSR loci on the three sub-genomes, seven homoeologous groups, 21 individual chromosomes and also on short and long arms of all individual chromosomes was non-random, when tested on the basis of DNA content and physical length of chromosomes. The distribution patterns of SSR loci within individual arms were also examined; the results differed for EST-SSRs and gSSRs, since more EST-SSRs were mapped in the distal 60% regions while more gSSRs were mapped in the proximal 40% regions.

5. In the integrated physical map, as many as 279 EST-SSRs were multilocus SSRs, which mapped on 677 loci. Similarly, among gSSRs, 169 gSSRs were multilocus, which mapped on 265 loci. More multilocus SSRs were found among EST-SSRs than among gSSRs.

6. Among 791 SSR loci that were earlier available on genetic maps, as many as 704 loci with their relative positions were linearly ordered in consensus genetic maps and integrated physical maps. As many as 87/791 (11%) common loci showed perturbed positions between genetic and physical maps; these were classified into the following two groups: (i) those positioned on the same arm (either short or long) of individual chromosome in genetic map as well as in physical map, and (ii) those positioned on different arms of individual chromosome in genetic and physical maps. These perturbed loci may be due to either differences in recombination rates along the chromosome length or due to structural changes in chromosomes.

7. Whole-genome relative genetic recombination frequencies have also been surveyed on the basis of based on the integrated physical and genetic maps. Using 704 loci that had no discrepancies between genetic and physical maps, the distribution of recombination was highly uneven along individual chromosome. Recombination was high in the distal region and was very low in centromeric regions except in the following 8 cases: C-1D, C-1AS, C-2AL, C-2DL, C-3BL, C-4DL, C-6DL and C-7A. Interestingly, slight differences were found between short and long arm of individual chromosomes in terms of both recombination frequencies and their distribution. In addition, recombination frequency also decreased in the most distal bins of the following chromosomes: 1A, 1B, 2A, 2D, 3D, 4A, 4D, 5B, 5D, 6A, 6D, 7A, 7B and 7D. These observation and the fact that the interstitial bins had an intermediate rates of recombination suggested that the recombination frequency (crossing over) is not increasing uniformly from the centromere to the distal ends.

8. Based on the observation concerning high transferability of brachypodium EST-SSRs to wheat, orthology was established among brachypodium, wheat and rice genomes. For this purpose, a total of 3,818 brachypodium EST contigs were employed for BLASTN analysis against the wheat EST contigs (containing bin-mapped wESTs) as well as against the rice whole genome sequences. The analysis revealed that as many as 449 and 743 brachypodium EST contigs were orthologous to sequences in wheat and rice genomes, respectively. The

observation of higher number of bEST contigs showing orthology with rice genome was mainly attributed to the fact that only a small fraction of wheat genome (0.02%) and almost complete rice genome (95%) were used for sequence comparison with brachypodium.

9. We used two new parameters for BLASTN analyses that take into account not only similarity of sequences but also the relative lengths of sequences: cumulative identity percentage (CIP) and cumulative alignment length percentage (CALP). These parameters were applied to all the BLAST alignments that were performed in the present study.

10. There were more orthologous loci in proximal regions than in the distal regions; this suggested that orthologs are more conserved in regions surrounding the centromere. It seems that higher degree of sequence conservation coincides with the low recombination proximal regions than with the high recombination distal regions.

11. A total of 183 orthologous sequences were conserved among wheat, brachypodium and rice genomes. Functional annotation (BLASTP analysis) suggested that 137/183 (74.80%) brachypodium EST contigs also had conserved protein sequences that were involved in the metabolic pathways and assembly of structural proteins, thus suggesting that these 137 orthologs have remained stable during the course of divergence.

12. Above 183 brachypodium EST contigs were also used for mining SSRs; a total of 100 contigs contained 137 SSRs, of which 45 SSRs were conserved in wheat and 23 SSRs were conserved in rice. Results in the present study showed that the SSR motifs identified in brachypodium are more often conserved in wheat than in rice.

13. On the basis of above study on orthologous relationship and with the availability of whole genome sequences of brachypodium, we established syntenic and structural relationship between wheat and brachypodium genomes by using 8,210 mapped ESTs of wheat as query sequences and complete genome sequence of brachypodium as reference. The results of wheat-brachypodium synteny were also compared with

    wheat rice synteny reported earlier.

14. Out of 8,210 mapped wheat ESTs, 5,208 (63.41%) wheat ESTs showed significant hits against brachypodium chromosomes. As many as 4,804 (58.53%) wESTs showing

homology with brachypodium were mapped to specific bins of the wheat chromosomes belonging to 7 homoeologous groups of wheat.

15. On the basis of bin mapping information for each group of three homoeologs of wheat, a consensus map for all the 7 homoeologous groups (total 7 chromosomes; WC1-7) were prepared. A total of 1,388 wESTs or contigs showed homology with brachypodium chromosomes (Bd).

16. The wheat ESTs assigned to each wheat consensus chromosome (WC) was mostly shared by one or more of the five brachypodium chromosomes. For example, WC3 was largely syntenous with the brachypodium chromosome 2 (Bd2), WC4 was syntenic to Bd1 and WC6 was syntenic to with Bd3. Rest of the four wheat consensus chromosomes (WC1, WC2, WC5 and WC7) were each syntenic to two brachypodium chromosomes; WC1 = Bd2 and Bd3; WC2 = Bd1 and Bd5; WC5 = Bd1 and Bd4; WC7 = Bd1 and Bd3.

17. EST densities for all consensus bins of an individual consensus chromosome were determined on the basis of average physical length of each bin (Gill et al. 1991). Observations based on consensus chromosome maps suggested that small bins on wheat chromosomes had greater EST density. For example, bin (0.49-0.50) on the long arm of wheat consensus chromosome 2 contained alone 48 ESTs (36.6%) of the long arm ESTs that constituted only 1% physical length of the arm. These identified regions with high EST density and showing synteny with brachypodium are possibly conserved over time, and have evolutionary significance.

18. On the basis of colinear regions between wheat and brachypodium chromosomes, a total of 82 conserved syntenic blocks (represents 13 major blocks) could be identified in the present study. Colinearity was recorded on the basis of their relative best-hit order (sequence coordinates) on the brachypodium chromosomes. The above 82 syntenic blocks of wheat and brachypodium varied in size ranging from 318 kb to 58 Mb on corresponding brachypodium chromosomes. Two most dense syntenic block contained 32 and 34 wESTs that respectively spanned WC2L, 0.49-0.50 (1% of the arm) and WC5L, 0.55-0.57 (2% of the arm) on wheat chromosomes, which had orthologous counterparts on 10.9 Mb and 9.0 Mb regions of Bd5 and Bd4, respectively. These syntenic blocks provide a genome-wide

framework for understanding the genomic rearrangements between wheat and brachypodium.

19. Using 153 ESTs (wheat-brachypodium homologs), the ratios of nonsynonymous and synonymous substitutions (Ka/Ks) were worked out. As many as 98 homologs had Ka/Ks ratio <1 suggesting that these sequences evolved under purifying selection that did not alter the encoded amino acid sequence during speciation period. The remaining 55 homologs identified as fast evolving sequences (Ka/Ks ratio >1). These sequences are related to several molecular functions, including transcription factors, binding proteins, transport related-protein, fertility restoration, cold-induced proteins, disease resistance proteins, metal stress related proteins, etc. Therefore, we may conclude that these ESTs may be useful for identifying genes that perhaps evolved in response to the positive selection (Charlesworth et al. 2001) and might be responsible for speciation.

20. Of the 4,804 bin-mapped wESTs, as many as 793 (16.51%) wheat ESTs were mapped each to a single bin; these were considered as putative single copy ESTs/genes. Above single copy genes (single bin-mapped wESTs) of wheat were predominantly located each on one of the five brachypodium chromosomes, identified as synteny, indicating a common origin of the specific wheat and brachypodium chromosomes. Such an identification of single copy genes (ESTs) is important for a number of reasons, particularly for their unique status in genomes that is full of paralogs like hexaploid wheat.

21. The more exciting features of synteny analysis was that, using the same data set of wheat, patterns of syntenic relationship between wheat-brachypodium (reported in this study) and between wheat-rice (reported earlier studies) was almost comparable. For example, Bd2/R5 and Bd3/R10 had the similar conservation of syntenies over the whole length of wheat homoeologous group 1.