Chromosome Bin Map of Expressed Sequence Tags in Homoeologous Group 1 of Hexaploid Wheat and Homoeology With Rice and Arabidopsis

J. H. Peng,* H. Zadeh,* G. R. Lazo,[†] J. P. Gustafson,[‡] S. Chao,[§,1] O. D. Anderson,[†] L. L. Qi,[¶]
B. Echalier,[¶] B. S. Gill,[¶] M. Dilbirligi,[&] D. Sandhu,[&,2] K. S. Gill,[&] R. A. Greene,** M. E. Sorrells,**
E. D. Akhunov,[††] J. Dvořák,[††] A. M. Linkiewicz,[††,3] J. Dubcovsky,[††] K. G. Hossain,[‡‡]
V. Kalavacharla,[‡‡] S. F. Kianian,[‡‡] A. A. Mahmoud,[§§] Miftahudin,[§§] E. J. Conley,[¶¶]
J. A. Anderson,[¶¶] M. S. Pathan,[§§] H. T. Nguyen,[§§] P. E. McGuire,[§]
C. O. Qualset[§] and N. L. V. Lapitan*,[4]

*Department of Soil and Crop Sciences, Colorado State University, Fort Collins, Colorado 80523-1170, [†]U.S. Department of Agriculture
(USDA)-Agricultural Research Service (ARS), Western Regional Research Center, Albany, California 94710-1105, [‡]USDA-ARS Plant
Genetics Research Unit, Department of Agronomy, University of Missouri, Columbia, Missouri 65211, [§]Genetic Resources Conservation
Program, University of California, Davis, California 95616, [¶]Department of Plant Pathology, Wheat Genetics Resource Center,
Kansas State University, Manhattan, Kansas 66506-5502, [&]Department of Crop and Soil Sciences, Washington State
University, Pullman, Washington 99164-6420, **Department of Plant Breeding, Cornell University, Ithaca,
New York 14853, [††]Department of Agronomy and Range Science, University of California, Davis,
California 95616, [‡‡]Department of Plant Sciences, North Dakota State University, Fargo,
North Dakota 58105, [§§]Department of Agronomy, University of Missouri, Columbia,
Missouri 65211 and [¶¶]Department of Agronomy and Plant Genetics,
University of Minnesota, Saint Paul, Minnesota 55108

ABSTRACT

A total of 944 expressed sequence tags (ESTs) generated 2212 EST loci mapped to homoeologous group 1 chromosomes in hexaploid wheat (*Triticum aestivum* L.). EST deletion maps and the consensus map of group 1 chromosomes were constructed to show EST distribution. EST loci were unevenly distributed among chromosomes 1A, 1B, and 1D with 660, 826, and 726, respectively. The number of EST loci was greater on the long arms than on the short arms for all three chromosomes. The distribution of ESTs along chromosome arms was nonrandom with EST clusters occurring in the distal regions of short arms and middle regions of long arms. Duplications of group 1 ESTs in other homoeologous groups occurred at a rate of 35.5%. Seventy-five percent of wheat chromosome 1 ESTs had significant matches with rice sequences ($E \leq e^{-10}$), where large regions of conservation occurred between wheat consensus chromosome 1 and rice chromosome 5 and between the proximal portion of the long arm of wheat consensus chromosome 1 and rice chromosome 10. Only 9.5% of group 1 ESTs showed significant matches to Arabidopsis genome sequences. The results presented are useful for gene mapping and evolutionary and comparative genomics of grasses.

WHEATS are the universal cereals of Old World agriculture (HARLAN 1992) and belong to the world's foremost crop plants (FELDMAN *et al.* 1995; NEVO *et al.* 2002). Wheat is one of the most widely cultivated food crops and is the staple food in >40 countries and for over 35% of the global population (WILLIAMS 1993). Modern wheat cultivars belong primarily to two species,

i.e., hexaploid bread wheat (*Triticum aestivum* L., $2n = 6x = 42$) and tetraploid durum-type wheat (*T. turgidum* L., $2n = 4x = 28$). Bread wheat is most important and contains three genomes: A, B, and D (SEARS 1969).

The genome size of hexaploid wheat is the largest (16,979 Mbp) among all cereal crops, including oat (*Avena sativa* L.; 12,961 Mbp), maize (*Zea mays* L.; 2671 Mbp), sorghum (*Sorghum bicolor* L.; 735–1642 Mbp), and rice (*Oryza sativa* L; 490 Mbp; BENNETT and LEITCH 2003). The hexaploid wheat genome is also much larger than any of the current plant model species, ~35 times larger than rice and 99 times larger than *Arabidopsis thaliana* (172 Mbp) (BENNETT and LEITCH 2003).

In any genome-sequencing project, the central goal is the discovery of all the genes in the target organism and the establishment of their chromosomal location

(Bennetzen 2002). The completion of genome sequencing in rice (Goff *et al.* 2002; Yu *et al.* 2002; http://rgp.dna.affrc.go.jp/cgi-bin/statusdb/irgsp-status.cgi, http://www.gramene.org/) and Arabidopsis (Arabidopsis Genome Initiative 2000) opened the door to comparative plant genomics or comparative plant biology. Because of the large size of the hexaploid wheat genome, complete sequencing has not been feasible. Large-scale discovery and isolation of genes and deciphering of gene function in wheat and its relatives must rely on other, less direct methods.

An EST is a segment of a sequence from a cDNA clone complementary to an mRNA sequence. Thus ESTs are segments of expressed genes (Adams *et al.* 1991). The ESTs (mRNAs → cDNAs) can be isolated from multiple tissues under various treatments and used to identify as many genes as possible in an organism. The EST approach facilitates the tagging of genes in a relatively short time at a fraction of the cost of complete genome sequencing, provides new genetic markers, and serves as a resource in diverse biological research fields (Adams *et al.* 1991). This approach has provided a new resource for the analysis of chromosome sequences and gene discovery in many organisms, such as *Homo sapiens* (Adams *et al.* 1991, 1995; Hillier *et al.* 1996), *Mus musculus* (Marra *et al.* 1999), *Rattus norvegicus* (Scheetz *et al.* 2001), *Danio rerio* (Clark *et al.* 2001), *Medicago truncatula* (Covitz *et al.* 1998), maize (Fernandes *et al.* 2002), and rice (Ewing *et al.* 1999). Developing ESTs has become a top priority for crop genomics worldwide (Briggs 1998). Development and deletion mapping analysis of ESTs from hexaploid wheat were conceived by a group of U. S. researchers as a cost-effective approach and a short cut to gene discovery, comparative genomics, and evolutionary genomics.

Hexaploid wheat has seven homoeologous groups of chromosomes, each containing one A, B, and D chromosome from each one of the donor genomes (Sears 1969). The group 1 homoeologous chromosomes are the most studied and understood of the wheat chromosome groups, primarily because they house major clusters of agronomically important genes. The group 1 and the group 2 chromosomes have ~100 each of the 615 genes of known chromosomal location, more than any of the other groups (McIntosh *et al.* 2003). There are clusters of resistance genes in wheat chromosome 1B, including at least 22 genes and QTL conferring resistance (Peng 2000; Dilbirligi *et al.* 2004). Numerous genes and gene families expressed during seed development are located in group 1, such as high-molecular-weight glutenins, low-molecular-weight glutenins, γ- and ω-gliadins, the triplet protein, and several seed-specific globulins (Dubcovsky *et al.* 1997). Among seven domestication syndrome factors detected in wild emmer wheat [*T. turgidum* ssp. *dicoccoides* (Körn. ex Asch. & Graebn.) Thell.], two were located on chromosome 1B (Peng *et al.* 2003). In a single gene-rich region, the 1S0.8 region

of the short arm of chromosome 1 in Triticeae, as many as 75 genes were identified (Sandhu *et al.* 2001).

The goal of the U.S. wheat EST project was to establish the chromosomal location of genes in the hexaploid wheat genomes. ESTs representing wheat unigenes were physically mapped to individual chromosomes/chromosomal intervals using wheat nullisomic-tetrasomic and ditelosomic lines (Sears 1966) and deletion stocks (Endo 1988, 1990; Endo and Gill 1996). This article summarizes the EST mapping results for group 1 chromosomes of hexaploid wheat. Patterns of distribution and duplication of ESTs within and among the group 1 chromosomes of wheat and comparisons with rice and Arabidopsis genome sequences are discussed.

MATERIALS AND METHODS

EST clones: For the U.S. wheat EST project, 113,220 ESTs were produced from 41 libraries of wheat representing a wide range of tissues, developmental stages, and environmental stresses (Lazo *et al.* 2004; Zhang *et al.* 2004). The ESTs that mapped to wheat group 1 chromosomes were derived from 22 libraries. Amplified PCR products (inserts) were prepared and delivered by the U.S. Department of Agriculture Albany group to 10 mapping laboratories (http://wheat.pw.usda.gov/NSF) for Southern hybridization.

Plant materials: Hexaploid wheat has an extensive collection of aneuploid and deletion stocks in the Chinese Spring background that are ideal for chromosome mapping, allowing for virtually every DNA marker to be assigned to a specific chromosome/arm/interval without requiring intragenomic polymorphism (Sears 1954, 1966; Sears and Sears 1978; Endo 1988, 1990; Endo and Gill 1996). A total of 146 cytogenetic stocks including 21 nullisomic-tetrasomic, 24 ditelosomic, and 101 deletion lines were used. The genetic stocks for EST mapping were provided by the Wheat Genetics Resource Center, Kansas State University, and cytologically and/or molecularly verified by all 10 mapping laboratories (http://wheat.pw.usda.gov/NSF; Qi *et al.* 2003, 2004). DNA samples were isolated following protocols established in the individual laboratories.

Southern hybridization: The conventional Southern hybridization approach was adopted with some minor modifications (http://wheat.pw.usda.gov/NSF/project/mapping_data.html). About 20 µg of genomic DNA were digested with *Eco*RI enzyme. A mixture of digested products of λDNA using *Hin*dIII and *Bst*EII was used as the size ladder. The digested DNA samples were separated on a 1.0% agarose gel and blotted onto a Hybond N+ membrane (Amersham Biosciences, Buckinghamshire, UK). A set consisting of five Southern blots, each with 30 lanes (a size ladder was included for the first four blots), was hybridized with a single EST clone in each hybridization reaction. More details about Southern hybridization were described by Lazo *et al.* (2004) and Qi *et al.* (2004).

Localization of ESTs: Bread wheat is a hexaploid, and the allocation of homoeologous loci to specific chromosomes within a homoeologous group is based on interchromosomal polymorphism. According to presence or absence of the restriction fragments in a given set of DNA lanes of a Southern blot, EST loci were assigned to a specific chromosome, an arm, and/or a deletion bin (Sears 1954 1966; Endo 1988, 1990; Endo and Gill 1996). The EST loci were localized to individual bins in the homoeologous group 1 chromosomes as described by Akhunov *et al.* (2003) and Qi *et al.* (2003,

<div align="center">

TABLE 1

Distribution of EST probes and loci among the group 1 chromosomes in wheat

</div>

	All ESTs				ESTs with all bands mapped			
Chromosome	Probes	Loci	EST density[a]	Intrachromosome duplications[b]	Probes	Loci	EST density[a]	Intrachromosome duplications[b]
1A	597	660 (29.8%)	0.84	63 (20.2%)	252	280 (31.3%)	0.88	28 (24.6%)
1B	660	826 (37.3%)	0.99	166 (53.2%)	265	317 (35.5%)	0.94	52 (45.6%)
1D	643	726 (32.8%)	1.22	83 (26.6%)	263	297 (33.2%)	1.24	34 (29.8%)
Total	944	2212 (100%)		312 (100%)	326	894 (100%)		114 (100%)
χ^2	3.355	16.344		57.365	0.377	2.302		8.211
P	0.187	<0.001		<0.001	0.828	0.316		0.016

[a] EST density is the ratio of observed EST loci *vs.* expected EST loci based on chromosome size (GILL *et al.* 1991).

[b] The difference between the number of loci and the number of probes mapped in a particular chromosome was used to estimate the level of intrachromosome duplications. The χ^2 test was used to detect the probability of a departure from a hypothetically expected 1:1:1 ratio (even distribution) among the three chromosomes.

2004). Southern images were scored by at least two persons in each laboratory, and the edited images and mapping data were uploaded to the project website http://wheat.pw.usda. gov/NSF/project/. Each set of mapping data was further validated by corresponding coordinators of the seven homoeologous chromosome groups at the project website http://wheat.-pw.usda.gov/cgi-bin/westsql/map_locus_rev.cgi.

Data analysis: Only mapping data validated by three persons (http://wheat.pw.usda.gov/cgi-bin/westsql/map_locus_rev.cgi) were used for analyses. Mapping data were reverified by checking all the online images in the project database prior to the analyses. Any ambiguous data were excluded in the analyses. On the basis of the physical size (in micrometers) of chromosomes and chromosome arms (GILL *et al.* 1991) and the relative length of chromosome deletion bins (ENDO and GILL 1996), the expected number of ESTs/EST loci was estimated under the assumption of random distribution among and along chromosomes. But the expected numbers of ESTs/EST loci in 1A, 1B, and 1D were estimated on the basis of a hypothetical 1:1:1 distribution. EST densities were calculated as the ratio of observed *vs.* expected ESTs/EST loci for the individual chromosomes, chromosome arms, and chromosome bins. The χ^2 test was used to test for independence of distribution patterns of ESTs among and along chromosomes. Loci numbers and distributions of duplicated ESTs across the other six chromosome groups were also analyzed.

The method of constructing the consensus map was as described by GILL *et al.* (1996a,b). ESTs that were not mapped to specific bins or fell into two conflicting bins were assigned to consensus chromosome 1, to one chromosome arm, or to a larger combined bin encompassing the two conflicting bins.

To analyze the similarity of wheat consensus chromosome 1 with rice and Arabidopsis sequences, blastN searches of the ESTs—mapped to all three group 1 chromosomes and used to construct the consensus chromosome bin map against rice (http://rgp.dna.affrc.go.jp/cgi-bin/statusdb/irgsp-status.cgi) and Arabidopsis (http://www.arabidopsis.org/) sequence databases—were conducted. $E \leq e^{-10}$ was adopted as the standard to claim a significant match. BlastX was also used to retrieve the gene function of wheat chromosome 1 ESTs that matched to Arabidopsis. The binomial test was used to detect colinearity of consensus bins with rice chromosomes as described by LIN-KIEWICZ *et al.* (2004). A putative orthologous genetic map for each interval, significantly colinear with rice chromosomes 5 and/or 10, of wheat consensus chromosome 1 ESTs was constructed on the basis of the order and position of the corresponding P1-derived artificial chromosome (PAC)/bac-

terial artificial chromosome (BAC) clones in rice chromosomes 5 and 10. Details of wheat consensus chromosome 1 ESTs and the significantly matched PAC/BAC clones in rice chromosomes 5 and 10 can be retrieved as supplemental online material at http://wheat.pw.usda.gov/pubs/2004/Genetics/.

An anomaly was defined as an EST having loci mapped to nonoverlapping bins on homoeologous chromosomes (MUNK-VOLD *et al.* 2004). Mapping data and images of all the EST probes assigned to group 1 chromosomes were examined for such anomalies.

Types of EST duplications analyzed: Two types of duplications were distinguished, intrachromosome and interchromosome duplications. An intrachromosome duplication is inferred to have occurred when an EST generates two or more loci in one chromosome. The number of intrachromosome duplications was estimated from the difference between the number of loci and the number of EST probes mapped in a particular chromosome. An interchromosome duplication refers to the situation in which an EST maps to chromosomes other than the three homoeologous group 1 chromosomes (1A, 1B, 1D). In testing the observed number *vs.* the expected number of duplications, we assumed that duplications in other locations occurred randomly, so that the greater the length of a chromosome or bin, the greater the chance of a duplication occurring in that region. The expected number of duplications was therefore determined on the basis of the size of the chromosome or chromosome bin.

<div align="center">

RESULTS

</div>

ESTs mapped to homoeologous group 1 chromosomes and the consensus physical EST map: As of March 17, 2003, 944 ESTs were mapped to group 1 chromosomes. Of these, 597, 660, and 643 ESTs had loci in chromosomes 1A, 1B, and 1D, respectively (Table 1). A total of 367 (39%) of the 944 group 1 ESTs were mapped to each group 1 chromosome (1A, 1B, and 1D) and were used to construct the group 1 consensus deletion bin map. A list of these 367 ESTs has been deposited as supplemental online material at http://wheat. pw.usda.gov/pubs/2004/Genetics/. Of these 367 ESTs, 326 were allocated to specific bins and 41 were allocated
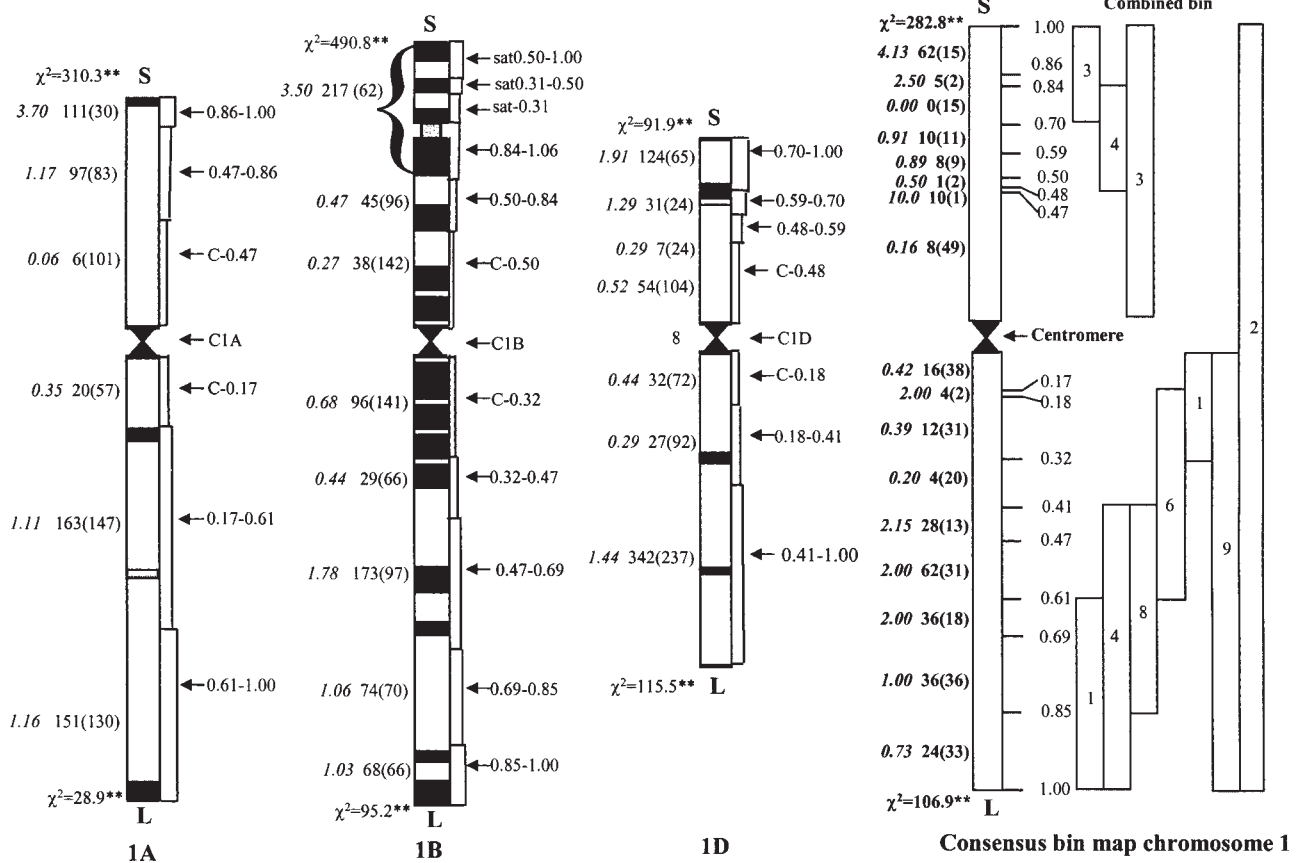
FIGURE 1.—EST deletion bin maps of wheat chromosomes 1A, 1B, and 1D and the consensus group 1 chromosome. Numbers of EST loci mapped only to chromosomes and chromosome arms were not shown in the individual chromosome bin maps. The numbers of mapped EST loci (expected numbers are in parentheses) and the gene density (italic type) are shown on the left of each chromosome and the intervals are indicated on the right. The 1BS satellite region is incorporated into bin 1BS-0.84-1.06. On the left of the consensus chromosome bin map are the numbers of ESTs (expected numbers are in parentheses) and gene density (italic type) and on the right are the fraction lengths. In the consensus chromosome bin map, the boxes on the right represent the combined bins and the number of ESTs is shown inside the boxes. χ^2 was used to test for random distribution of EST loci/probes on chromosome arms. * and **, significance of 0.05 and 0.01 probability, respectively.

to combined bins of the consensus chromosome 1. There were 251 and 114 ESTs located in the long and short arms, respectively, of the consensus chromosome bin map (Figure 1).

Distribution of ESTs among homoeologous group 1 chromosomes: The 944 EST probes generated 2212 loci mapped to the group 1 chromosomes with 660, 826, and 726 loci located on chromosomes 1A, 1B, and 1D, respectively (Table 1). Among the three chromosomes, 1B had the highest number and 1A the lowest. The χ^2 test showed that deviation from the expected values based on a hypothetical ratio of 1:1:1 (even distribution) was highly significant ($P < 0.01$) for EST loci, but not significant for EST probes. Thus, it appeared that the mapped EST probes were randomly distributed among the three chromosomes, but the mapped EST loci were not. Chromosome 1B had a larger number and 1A had a smaller number than expected (Table 1).

As done by LINKIEWICZ *et al.* (2004) for homoeologous group 5, the previous analysis was repeated using the subset of EST probes for which all bands were

mapped (326 ESTs). Again the 1B chromosome had the largest proportion of probes (265) and loci (317). However, the numbers of EST probes and loci did not significantly deviate from the expected values based on random distribution (Table 1). It seems that this subset of mapped EST probes was randomly distributed among the three group 1 chromosomes.

Distribution pattern of EST loci on the chromosome arms: Of the 2212 EST loci (Table 1), 2076 were mapped to specific chromosome arms (Table 2) and the other 136 were mapped only to chromosomes or centromere bins. The numbers of EST loci mapped to the long arms of each of the three chromosomes were larger than those mapped to the short arms (Table 2). In comparison with the expected values based on arm length (GILL *et al.* 1991), the ESTs were randomly distributed between long and short arms for 1A ($P = 0.27$), nearly randomly for 1B ($P = 0.06$), and nonrandomly for 1D ($P = 0.02$) with more observed in the long arm and less than expected in the short arm (Table 2).

Chromosome bin mapping of 2212 EST loci (Table 1)

TABLE 2

Distribution of EST loci between chromosome arms

| Chromosome | Item | Chromosome arm | | Total[a] |
		Long	Short	
1A	Observed	371	215	586
	Expected[b]	383.80	202.20	586
	Deviation	−12.80	12.80	0
	χ^2	1.24 ($P = 0.265$)		
	Gene density	0.97	1.06	
1B	Observed	483	326	809
	Expected[b]	509.40	299.60	809
	Deviation	−26.40	26.40	0
	χ^2	3.69 ($P = 0.054$)		
	Gene density	0.95	1.09	
1D	Observed	459	222	681
	Expected[b]	428.90	252.10	681
	Deviation	30.10	−30.10	0
	χ^2	5.71 ($P = 0.016$)		
	Gene density	1.07	0.88	

[a] EST loci mapped only to chromosomes or centromere bins were excluded.
[b] The expected number is based on the arm ratio value (GILL *et al.* 1991).

yielded 1913 that were mapped to specific chromosome intervals/bins (Figure 1); the remaining 299 were mapped only to chromosomes or chromosome arms. Most of the EST loci were located in the distal regions of both long and short arms for all three chromosomes (Figure 1). This distribution of loci deviated significantly ($P = 0.01$) from the expected numbers based on interval lengths in each arm (Figure 1). The consensus chromosome deletion bin map had a higher resolution (the chromosome arms were divided into more and smaller intervals than any individual chromosome) and it also revealed a majority of the ESTs in the distal intervals of the chromosome arms (Figure 1). The number of ESTs located in distal regions (0.84–0.86 and 0.86–1.00) of the short arm was significantly larger than expected, but in the long arm this was true only for the middle intervals (0.41–0.47, 0.47–0.61, and 0.61–0.69); the reverse was found for the long-arm distal bin (0.85–1.00; Figure 1). In the 1A, 1B, and 1D long arms, ESTs were allocated to the most distal bins, but they appear to be located in the proximal regions of those bins and not at the ends of the chromosomes. The consensus map with a greater number of long-arm bins clearly points this out (Figure 1). Therefore, distribution patterns of mapped ESTs were different for the two arms of all group 1 chromosomes.

EST density: EST density is the ratio of the number of mapped ESTs to the expected value based on the length of chromosome/arm/interval. Chromosome 1D had relatively high EST density values, both when all ESTs were used (1.22) and when only the subset for which all loci were mapped was used (1.24); chromosome 1A had the lowest EST density with values 0.84 and 0.88; and chromosome 1B was in between with

values of 0.94 and 0.99 (Table 1). The short arms had slightly higher EST densities (>1) than the long arms (<1) for 1A and 1B, but the reverse was true for 1D where the long arm had a higher EST density (>1) than did the short arm (<1; Table 2).

Most of the intervals with EST density <1.0 were in proximal regions near the centromeres. Five intervals were EST rich, having high-EST density: 1AS-0.86-1.00 (3.70), 1BL-0.47-0.69 (1.78), 1DS-0.70-1.00 (1.91), 1DL-0.41-1.00 (1.44), and the 1BS satellite region (3.50; Figure 1). In the consensus chromosome bin map, the EST-rich regions in the short arm were located in the intervals 1S-0.47-0.48 (10.00), 1S-0.84-0.86 (2.50), and 1S-0.86-1.00, including the satellite region (4.13). The EST-rich regions in the long arm were all located in the middle intervals 1L-0.41-0.47 (2.15), 1L-0.47-0.61 (2.00), and 1L-0.61-0.69 (2.00; Figure 1).

Duplications of group 1 ESTs: EST intrachromosome duplications were not randomly distributed among the three chromosomes with ~50% of the duplications occurring in 1B (Table 1). Approximately 13% of the mapped EST unigenes contained at least one *Eco*RI restriction site, not including sites within introns (MUNKVOLD *et al.* 2004). Because internal *Eco*RI sites within the region of hybridization can create duplicated fragments within a bin, the rate of intrachromosome duplications might be overestimated in the present study. However, the multiple *Eco*RI restriction sites within ESTs would not affect the rates of interchromosome and intergenome duplications.

About 35% (335) of the ESTs mapped to group 1 chromosomes also had loci in one or more of the six other homoeologous groups (Table 3). The χ^2 test showed that deviation of the observed number of ESTs

EST probes mapped to group 1 chromosomes and to other homoeologous groups in wheat

	No. of duplicated ESTs		Percentage of duplicated ESTs[b]
Combination	Observed	Expected[a]	
Group 1 + group 2	158	135.0	16.74
Group 1 + group 3	132	136.9	13.98
Group 1 + group 4	113	122.8	11.97
Group 1 + group 5	126	129.3	13.35
Group 1 + group 6	118	123.2	12.50
Group 1 + group 7	129	128.9	13.67
χ^2	5.18 ($P = 0.39$)		

[a] The expected number is based on the chromosome length (GILL *et al.* 1991).
[b] The percentage is based on 944 mapped ESTs.

with duplications in other homoeologous groups from the expected number based on chromosome size was not significant, although group 2 had a relatively large number. It seems that none of the six homoeologous chromosome groups had a significantly closer homoeology with group 1 chromosomes.

The numbers of duplicated group 1 EST loci in the other homoeologous groups were evaluated for each of the group 1 chromosome bins (Table 4). The total numbers of duplications in the other six homoeologous groups were 430, 551, and 411 for 1A, 1B, and 1D, respectively. Therefore, the group 1 EST duplications in other homoeologous group chromosomes were not evenly distributed among the three group 1 chromosomes ($\chi^2 = 24.9$, $P < 0.01$), with a larger number than expected in 1B and a smaller number than expected in 1A and 1D. A similar result was observed for distribution of EST duplications along chromosome arms with clustering in the distal bins. In comparison with the expected number derived from chromosome size, deviations of EST duplication in the six other homoeologous groups for all three group 1 chromosomes were not significant.

Homoeology of wheat consensus chromosome 1 with the Arabidopsis genome: BlastN searches of the 367 ESTs involved in the consensus chromosome deletion bin map against sequence databases of the Arabidopsis genome (http://www.arabidopsis.org/) revealed that only 35 (9.5%) had significant ($E \leq e^{-10}$) matches with Arabidopsis sequences (Table 5). Of these, 27 (77%) were located on the long arm of wheat consensus chromosome 1 and the other 8 (23%) were on the short arm. These 35 that significantly matched wheat chromosome 1 ESTs were not randomly distributed ($P < 0.05$) among the five Arabidopsis chromosomes, with an apparently larger number located to chromosome 3 (37%) and fewer to chromosome 1 (11%). Therefore, homoeology of wheat consensus chromosome 1 (W1) ESTs with Arabidopsis was not high (<10%), and the relatively greatest homoeology exists with Arabidopsis chromosome 3.

Homoeology of wheat consensus chromosome 1 with rice genome: BlastN searches of 367 ESTs against the rice sequence database (http://rgp.dna.affrc.go.jp/cgi-bin/statusdb/irgsp-status.cgi) indicated that 274 (75%) had significant ($E \leq e^{-10}$) matches to rice sequences (Table 6). All but two of these were associated with specific rice chromosomes. Wheat group 1 ESTs that significantly matched to rice sequences were unevenly distributed ($P < 0.01$) among the 12 rice chromosomes with 49, 20, and 1–5% located on chromosome 5, chromosome 10, and the other 10 chromosomes, respectively. Therefore, the wheat group 1 ESTs analyzed have high homoeology ($\sim$75%) with rice, especially rice chromosomes 5 and 10.

Colinearity of wheat consensus chromosome 1 and rice chromosomes: The results shown in Table 6 indicate that bins in the regions of 0.47–1.00 on the short arm and of 0.41–1.00 on the long arm were significantly colinear with rice chromosome 5 (R5) and that bins in the region of C-0.47 on the long arm were significantly colinear with rice chromosome 10 (R10). The 1L-0.41-0.47 bin was significantly colinear with both R5 and R10. The 1L-0.61-0.69 bin was significantly colinear with both rice chromosomes 5 and 1. A large portion of the W1 short arm, especially the C-0.47 and 0.59-0.84 regions, was not significantly colinear with any of the 12 rice chromosomes.

Putative EST order in wheat consensus chromosome 1 based on rice sequences: Due to the high homoeology with wheat group 1 chromosomes, sequence orders of rice chromosomes R5 and R10 (http://rgp.dna.affrc.go.jp/cgi-bin/statusdb/irgsp-status.cgi) were used to construct an orthologous genetic consensus EST bin map of W1 as shown in Figure 2. Details of this orthologous genetic map are also presented as supplemental online material at http://wheat.pw.usda.gov/pubs/2004/Genetics/.

A total of 133 W1 ESTs corresponded to 93 R5 BAC/PAC clones spanning a genetic length of 122.3 cM. More than one-third (33) of the R5 sequences matched to two or more W1 ESTs in the same bins or different bins

TABLE 4

Distribution of group 1 ESTs having duplicated loci in other chromosome groups

Chromosome interval	Homoeologous group												Total		χ^2
	2		3		4		5		6		7				
	E	O	E	O	E	O	E	O	E	O	E	O	E	O	
1AS3-0.86-1.00	3	10	3	6	2	6	3	8	4	17	5	16	19	63	
1AS1-0.47-0.86	7	8	7	13	4	5	9	8	12	14	13	17	53	65	
C-1AS1-0.47	8	0	9	0	5	0	11	7	15	0	15	0	63	7	
χ^2	24.5**		17.1**		13.3**		9.9**		57.6**		40.4**		154.4**		
C1A		0		0		0		0		0		0		0	
C-1AL1-0.17	6	0	9	0	7	2	6	1	5	0	4	0	38	3	
1AL1-0.17-0.61	17	30	24	22	19	35	14	17	11	15	9	18	99	137	
1AL3-0.61-1.00	15	8	21	32	17	6	13	15	10	21	8	3	88	85	
χ^2	19.2**		14.9**		24.2**		5.1		18.6**		16.1**		46.9**		
Sum of 1A	75	69	76	88	68	63	72	65	68	80	71	65	464	430	6.05
1BS9-0.84-1.06	10	27	6	17	4	8	10	29	6	19	8	21	46	121	
1BS10-0.50-0.84	16	14	9	4	7	9	14	3	7	1	13	13	71	44	
C-1BS10-0.50	23	8	13	7	10	4	21	13	13	7	19	6	104	45	
χ^2	38.9**		25.7**		8.17*		47.8**		36.1**		30.0**		166.0**		
C1B		0		0		0		0		0		0		0	
C-1BL6-0.32	18	18	20	9	18	8	11	6	11	10	14	5	91	56	
1BL6-0.32-0.47	9	4	9	2	8	8	5	2	5	0	6	0	42	16	
1BL1-0.47-0.69	12	16	14	24	12	22	7	10	7	12	10	24	63	108	
1BL2-0.69-0.85	9	13	10	17	9	13	5	10	5	10	7	10	45	73	
1BL3-0.85-1.00	9	6	9	10	8	4	5	5	5	1	6	4	42	30	
χ^2	7.9		23.7**		17.7**		10.4*		16.9**		33.3**		82.6**		
Sum of 1B	96	118	97	99	87	84	92	82	87	73	92	95	464	551	8.62
1DS5-0.70-1.00	12	26	4	5	7	18	4	7	8	20	9	16	42	92	
1DS1-0.59-0.70	4	2	1.5	3	2	1	1.5	0	3	0	3	4	16	10	
1DS3-0.48-0.59	4	9	1.5	1	2	1	1.5	1	3	0	3	0	16	12	
C-1DS3-0.48	19	2	6	4	10	1	6	5	12	6	14	9	67	27	
χ^2	38.8**		2.6		26.4**		4.1		27.0**		10.6*		86.2**		
C1D		4		0		3		0		0		0		7	
C-1DL4-0.18	6	1	8	3	7	6	7	5	5	3	5	2	38	20	
1DL4-0.18-0.41	8	3	10	4	9	2	9	5	6	0	7	4	49	18	
1DL2-0.41-1.00	20	30	27	38	24	32	23	29	14	22	18	24	126	175	
χ^2	12.3**		11.2**		8.3*		3.9		11.4**		5.1		47.2**		
Sum of 1D	71	91	72	63	65	72	68	57	65	59	68	69	464	411	9.86
Overall χ^2 for the three chromosomes														24.9**	

E, expected values based on length of chromosome or interval except for the overall χ^2; O, observed value. * and **, significant at $P < 0.05$ and 0.01, respectively. Sum of each chromosome includes those group 1 ESTs not mapped to specific chromosome intervals. 1BS satellite region was combined into 1BS9-0.84-1.06.

in both short and long arms. There was arm correspondence between R5 and W1, with a small number (six) of exceptions where the linear order was not conserved (Figure 2).

A total of 53 W1 ESTs significantly matched to 37 R10 BAC/PAC clones spanning a genetic distance of 77.0 cM. About one-third (12) of the R10 sequences matched to two or more W1 ESTs in the proximal regions of the long arm of W1. Of the 37 R10 sequences, 33 matched to W1 ESTs on the long arm. Of the four exceptions, two were duplicated and two were specific to the short arm of W1 (Figure 2).

R5 sequence-matched ESTs apparently predominated over R10 sequence-matched ESTs for the bins with fraction length (FL) > 0.47 in both short and long arms of W1, the reverse was true for the bins with FL < 0.47 in the long arm of W1. There were five intervals without significant colinearity with R5 or R10, and four of them were in the short arm (C-1S-0.47, 1S-0.48-0.50, 1S-0.59-0.70, and 1S-0.70-0.84) of W1. Thus, either these five intervals have no similarity to R5 and R10 sequences or the similarity could not be detected with the existing data.

Anomaly: An anomaly refers to nonoverlapping map

TABLE 5

Wheat consensus chromosome 1 ESTs having significant matches with Arabidopsis sequences

GenBank accession no.	E-value	Arabidopsis chromosome	Gene function[a]	Wheat consensus chromosome bin
BE404660	6.00E-17	3	Ubiquitin-conjugating enzyme	1L-0.47-0.61
BE405167	8.00E-13	3	Ubiquitin-conjugating enzyme	1L-0.32-0.41
BE423193	7.00E-32	2	60S ribosomal protein	1L-0.47-0.61
BE426097	7.00E-54	5	Dim1 homolog, putative thioredoxin-like U5 small ribonucleoprotein particle protein	1L-0.41-0.85
BE426257	9.00E-16	3	Tetrahydrofolate dehydrogenase/cyclohydrolase, Methyltetrahydrofolate cyclohydrolase	1S-0.47-0.84
BE442716	7.00E-14	1	T-complex protein 1, ε-subunit/chaperonin	1L-0.41-0.47
BE442818	1.00E-11	3	Expressed protein	1L-0.47-0.61
BE443332	2.00E-11	4	Kelch repeat containing F-box protein family	1L-0.47-0.61
BE443378	2.00E-14	5	Microtubule binding protein D-CLIP-190 (GTP-binding protein)	1L-0.18-0.32
BE443531	2.00E-14	2	Neutral leucine aminopeptidase preprotein; metallo-exopeptidase; leucyl aminopeptidase	1S-0.47-0.48
BE444620	8.00E-16	3	Expressed protein	1L-0.61-0.69
BE446010	1.00E-11	3	Cyanate hydratase (cyanase), cyanate lyase	C-1L-0.32
BE446240	2.00E-13	2	Thiol methyltransferase, GDP dissociation inhibitor protein	1L-0.61-0.69
BE490592	2.00E-32	3	Mitochondrial NAD-dependent malate dehydrogenase	1L-0.85-1.00
BE490596	3.00E-37	5	H+-transporting ATP synthase β-chain (mitochondrial)	1L-0.41-0.85
BE494850	2.00E-20	3	20S proteasome β-subunit B (PBB1)	C-1S-0.48
BE495028	3.00E-55	3	Leucine-rich repeat transmembrane protein kinase, calmodulin	1L-0.47-0.61
BE497107	7.00E-14	3	Pseudogene, similar to En/Spm transposon protein; Ras family GTP-binding protein	C-1L-0.17
BE497584	6.00E-26	2	Mitochondrial F1-ATPase, mitochondrial precursor	C-1L-0.17
BE497808	7.00E-29	5	Glycosyl hydrolase family 17, histone H4 (TH091)	1L-0.61-0.69
BE500310	3.00E-15	4	Reverse transcriptase (RNA-dependent DNA polymerase) kinase-like protein	1L-0.17-0.61
BE500541	1.00E-14	1	GPI-anchor transamidase	1S-0.70-1.00
BE518048	5.00E-27	3	Mitogen-activated protein kinase	1L-0.85-1.00
BE518393	2.00E-34	1	60S ribosomal protein L23 (RPL23A)	C-1L-0.17
BE590822	1.00E-15	2	Glycosyltransferase family, UDP-glucoronosyl and UDP-glucosyl transferase, cyclophilin	1L-0.69-0.85
BE637867	6.00E-14	3	Crooked neck-related protein, receptor-protein kinase	1L-0.47-0.61
BF145399	1.00E-14	2	Homeodomain protein, vacuolar ATP synthase 16-kD proteolipid subunit	1S-0.86-1.00
BF200980	4.00E-15	5	Protein translation factor SUI1 homolog	1L-0.61-0.69
BF473056	3.00E-16	5	Phosphoinositide-specific phospholipase, actin-depolymerizing factor 1	1S-0.86-1.00
BF474139	1.00E-10	4	Fructose-6-phosphate-1 phosphotransferase	1L-0.61-0.69
BF483378	2.00E-14	2	DnaJ domain-containing protein, proteasome subunit α-type 3	1L-0.47-0.61
BF483456	6.00E-14	3	Nodulin/glutamate-ammonia ligase-like protein; pseudogene-gag-pol polyprotein; UDP-galactose transporter related protein 1; dTDP-glucose 4-6-dehydratase	1L-0.47-0.61
BF485305	2.00E-16	1	Unknown protein	1L-0.47-0.61
BG262410	7.00E-17	4	Transcriptional adaptor; Shaggy related protein kinase	1S-0.47-0.48
BM134392	1.00E-13	4	Cellulose synthase-1	1S-0.47-0.48

[a] Gene function was retrieved via blastX.

positions for an EST in the chromosome bin maps of the homoeologues. Of the 944 EST probes mapped to group 1 chromosomes, 26 (2.8%) detected anomalies, as defined by MUNKVOLD *et al.* (2004), involving all three group 1 chromosomes and arms. The distribution of these anomalies among the three group 1 chromosomes is shown in Figure 3. For example, EST BE500081 in the bin C-1BL-0.32 was not mapped to the near-centro-

TABLE 6

Colinearity between wheat consensus chromosome 1 and rice chromosomes

Wheat consensus chromosome 1 bin	Rice chromosome[b]													No. with no match	Total matches	Probability	
	1[c]	2	3	4	5[c]	6	7	8	9	10[c]	11	12	0			1st	2nd
Chromosome 1[a]					1									1	1	0.0833	—
1S[a]					1	1								1	2	0.1597	—
1S-0.86-1.00	1	3	3	1	**16**	2	4	3	0	1	1	1	1	25	37	<0.0001	0.1022
1S-0.84-0.86					**3**									2	3	0.0006	—
1S-0.70-1.00[a]		1	1											1	2	0.1597	—
1S-0.70-0.84														0	0	—	—
1S-0.59-0.70			1		1		1			1				6	4	0.2939	—
1S-0.50-0.59					**7**	1								0	8	<0.0001	0.0909
1S-0.48-0.50														1	0	—	—
1S-0.47-0.84[a]				1						1				2	2	0.1597	—
1S-0.47-0.48	1	1			**8**									0	10	<0.0001	0.1736
C-1S-0.47					2	1	1			2				2	6	0.0831	0.0438
C-1L-0.17	1			2	1	1				*9*				2	14	<0.0001	0.0686
C-1L-0.32[a]							1							0	1	0.0833	—
1L-0.17-0.18									1	1				2	2	0.1597	—
1L-0.17-0.61[a]			1							*2*	1			2	4	0.0372	0.1736
1L-0.18-0.32			1				1			*8*				2	10	<0.0001	0.1736
1L-0.32-0.41					1					*3*				0	4	0.0022	0.0909
1L-0.41-0.47	1				**5**	1				*14*	1	2		4	24	<0.0001	0.0011
1L-0.41-0.85[a]			1		**2**			1		1				3	5	0.0586	0.2487
1L-0.41-1.00[a]		1			1	1								1	3	0.2297	—
1L-0.47-0.61	4	2	3		**29**	1	4	1		3			1	14	48	<0.0001	0.0744
1L-0.61-0.69	*4*		2	1	**22**						1			6	30	<0.0001	0.0035
1L-0.61-1.00[a]														1	0	—	—
1L-0.69-0.85				1	**19**	3		1	1	3	1			7	29	<0.0001	0.0554
1L-0.85-1.00	1	1		1	**13**				1	2				5	19	<0.0001	0.0968
1L[a]		1			1		1			*3*				3	6	0.0095	0.2487
Total	13	10	13	7	**133**	12	11	8	3	*53*	6	3	2	93	274	<0.0001	<0.0001

[a] Combined bins.

[b] A figure in the table with $P < 0.01$ is claimed as significant and is in boldface type for rice chromosome 5 (R5), in boldface and italic for R10, and in italic type for R1.

[c] Underlining denotes W1 bins with significant colinearity with rice chromosomes.

mere bin on the long arm of either 1A or 1D, but it did map to bin 1AS-0.47-0.86 on 1AS. A total of 13 anomalies were present among 18 of 21 noncentromere bins, and 4, 5, and 4 anomalies were present for 1A, 1B, and 1D chromosomes, respectively. But only 7 of the 13 anomalies were detected by two or more ESTs.

DISCUSSION

Large number of ESTs mapped to homoeologous group 1 chromosomes: In the catalog of wheat gene symbols (McIntosh *et al.* 2003), >1500 DNA marker loci including RFLPs and SSRs were documented for wheat group 1 chromosomes from ∼100 publications. With the results reported here, 2212 loci, and the previous 1500, >3700 loci have been identified for this group, making it particularly rich in DNA marker loci.

Gene-rich regions in wheat group 1 chromosomes: It is well established from this study and others that polymorphic DNA markers are not evenly distributed
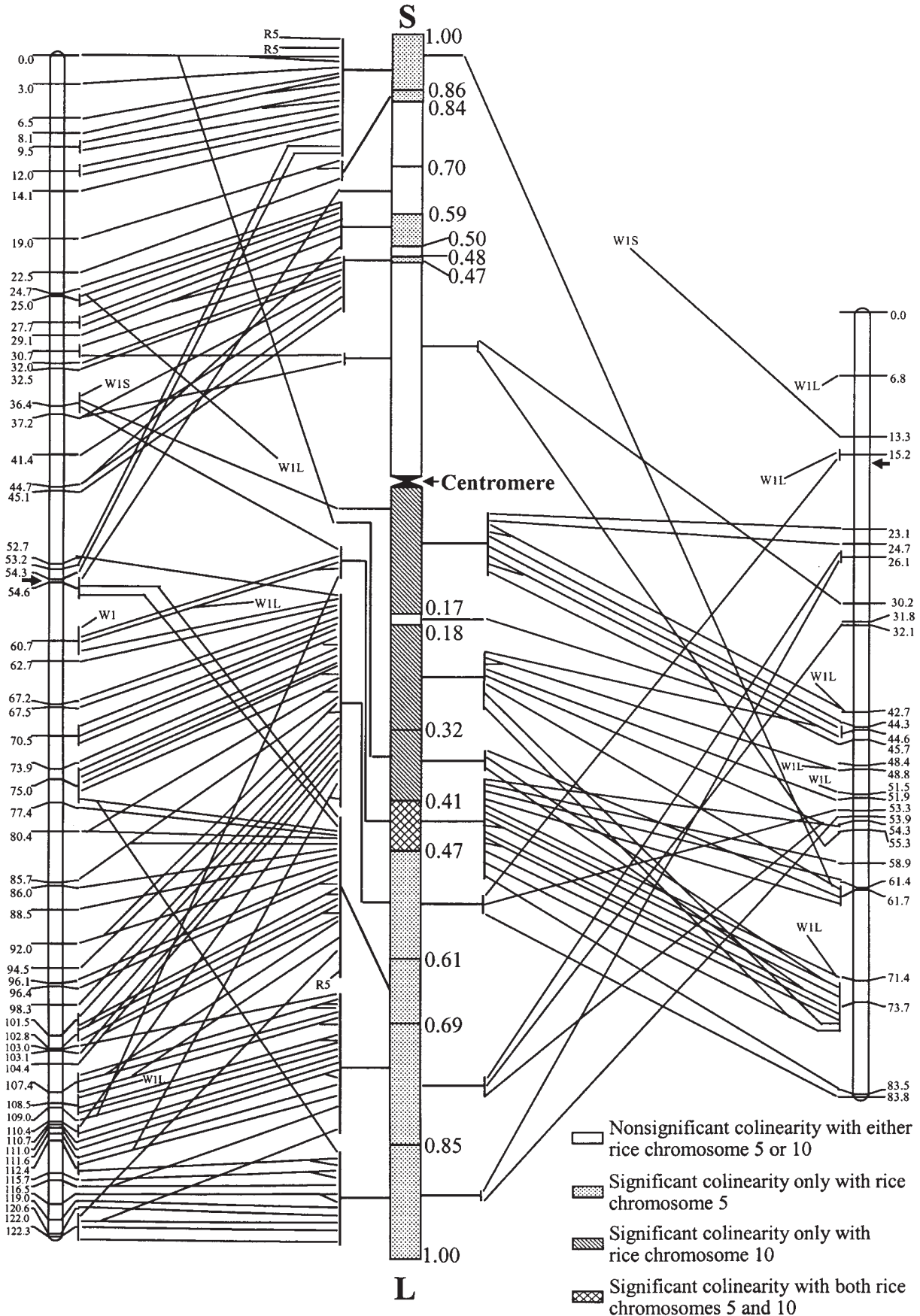
along the chromosomes in both genetic (Devey and Hart 1993; Dubcovsky *et al.* 1995; Gale *et al.* 1995; Van Deynze *et al.* 1995; Cadalen *et al.* 1997; Peng 2000; Peng *et al.* 2000) and physical maps (Delaney *et al.* 1995a,b; Gill *et al.* 1996a,b; Faris *et al.* 2000; Ma *et al.* 2001). Gill *et al.* (1996b) identified the 1S0.8 consensus region as gene rich. Sandhu *et al.* (2001) localized 75 useful genes into this region. We mapped >100 EST loci in regions of similar size in this study on each short arm of the group 1 chromosomes, accounting for >50% of those on the short arm for each chromosome. The short arm region of similar size in the wheat consensus chromosome 1 (16% of the arm) contained 67 ESTs (64.4% of the total in the consensus short-arm map; Figure 1; supplemental online material at http://wheat. pw.usda.gov/pubs/2004/Genetics/).

The distribution of ESTs can reveal the approximate expressed gene distribution patterns. We found that ESTs clustered in a few regions in the three group 1 chromosomes (Figure 1). It has been shown that centro-

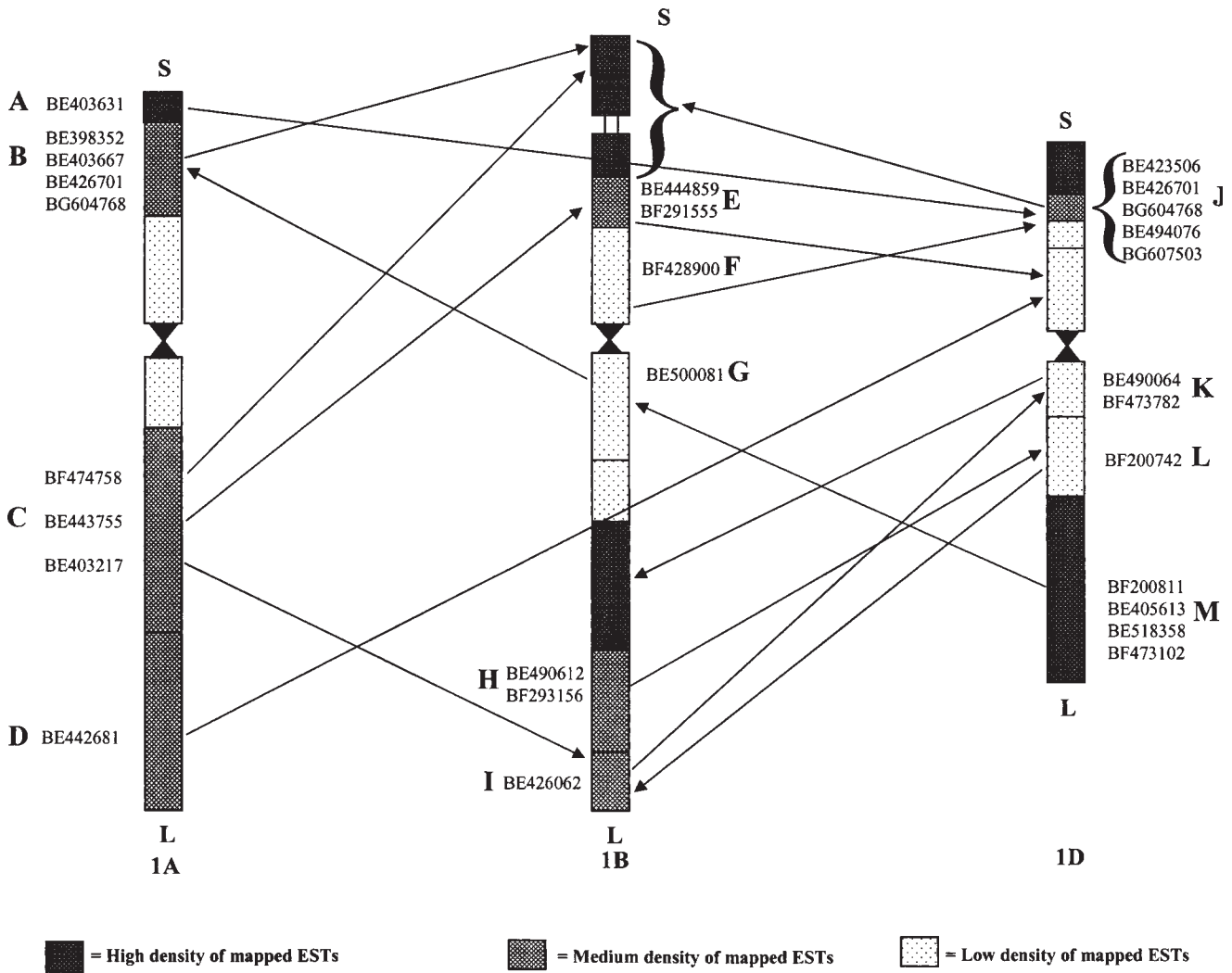**Rice chromosome 5**   **Wheat chromosome 1**   **Rice chromosome 10**

Nonsignificant colinearity with either rice chromosome 5 or 10

Significant colinearity only with rice chromosome 5

Significant colinearity only with rice chromosome 10

Significant colinearity with both rice chromosomes 5 and 10

= High density of mapped ESTs   = Medium density of mapped ESTs   = Low density of mapped ESTs

FIGURE 3.—Chromosome locations of mapped ESTs detecting anomalies. The GenBank accession numbers for the ESTs contained in the anomalies are included next to the appropriate chromosome intervals. A letter code was given to each anomaly for reference.

meric/proximal regions are relatively gene poor and distal/telomeric regions on the short arms are gene rich (GILL *et al.* 1996a,b; FARIS *et al.* 2000). Our results further showed that the telomeric regions on the long arms had a lower gene density than the middle portion of the arms, which contained EST-rich regions/clusters.

There are ESTs in the distal bins, but they are located in the proximal regions of those bins and are not at the ends of the chromosomes. The consensus map with a higher number of long-arm bins than the individual chromosomes clearly points this out (Figure 1).

We found that the number of ESTs mapped to the

FIGURE 2.—Schematic orthologous map of wheat consensus chromosome 1 ESTs with rice chromosomes 5 and 10. The wheat consensus chromosome 1 intervals are linked using lines with rice genetic maps in the locations where wheat ESTs are sequence matched with rice BAC/PAC clones. Order of the lines (ESTs) in wheat consensus chromosome 1 intervals was based on the order of homoeologous rice sequences (BAC/PAC clones). In the consensus bin map of wheat chromosome 1, the numbers and located intervals of wheat group 1 ESTs (lines) and the distance from the centromere (fraction length) are indicated. In the rice maps, the locations and numbers of BAC/PAC clones (lines linked with wheat map) and interval lengths (in centimorgans) are indicated. The arrows in rice chromosomes 5 and 10 indicate the approximate regions of centromere cited from http:// www.gramene.org/. For those wheat ESTs not mapped to specific intervals, the chromosome (W1) or chromosome arms (W1S for the short and W1L for the long arm) are indicated. The specific map locations of three rice BAC/PAC clones on chromosome 5 are not available, and the corresponding wheat ESTs are thus linked with a code R5 in the figure. The details including accession numbers of wheat ESTs and rice BAC/PAC clones and the orthologous map of wheat group 1 ESTs are also presented as supplemental online materials at http://wheat.pw.usda.gov/pubs/2004/Genetics/.

proximal parts of chromosome arms was significantly lower than that mapped to the distal parts (Figure 1). Comparisons of genetic linkage maps and physical maps have indicated that recombination is dramatically reduced in the centromeric regions of grass chromosomes. This recombination reduction is explained as the result of suppression around the centromere or by the centromere itself (Dvořák and Chen 1984; Lukaszewski and Curtis 1993; Van Deynze *et al.* 1995). As shown by Akhunov *et al.* (2003), EST density was positively related with recombination rate (*i.e.*, EST-rich regions have high recombination rates). But the reason for recombination reduction near the centromere may be the low EST (gene) density, rather than centromere suppression.

Duplication of wheat group 1 ESTs: If an EST probe had loci mapping to more than one chromosome in a genome, then those loci were considered duplicated (Akhunov *et al.* 2003). Anderson *et al.* (1992) reported that 40 of 210 DNA probes hybridized to fragments in more than one homoeologous group and that group 1 chromosomes were involved in the majority of these duplications. Van Deynze *et al.* (1995) found that many loci were duplicated between group 1 and group 7 chromosomes. Dubcovsky *et al.* (1996) pointed out that 30% of the loci mapped in *T. monococcum* were duplicated in other chromosomes. Akhunov *et al.* (2003) found that 21% of 730 ESTs had duplicated loci. In the present study, 335 (35%) of 944 ESTs analyzed were mapped to group 1 and at least one of the other six homoeologous groups. The number of group 1 duplications shared by each of the other six homoeologous groups ranged from 113 to 158 ESTs (Table 3). The EST duplication rate in this article seemed much higher than that found by Akhunov *et al.* (2003), and the difference may be derived from the data source (group 1, studied here, *vs.* all groups studied by Akhunov *et al.* 2003). Duplications of group 1 ESTs were randomly distributed among the other six homoeologous groups, but unevenly distributed among the three group 1 chromosomes with a larger number in 1B and a smaller number in 1A and 1D than expected and, along the chromosome arms, with a larger number in the distal and a smaller number in the proximal regions except 1AL and 1BL (Table 4). These results confirm the conclusion of Akhunov *et al.* (2003) who studied a smaller sample of the ESTs mapped in this project. The high rate of interchromosome duplication of ESTs also confirms that homoology exists among the seven groups of homoeologous chromosomes of wheat.

Anomalies in group 1 chromosomes: The rate of anomalies was much lower in group 1 than in group 3 (Munkvold *et al.* 2004). Anomalies may result from biological events such as chromosomal rearrangements, transposition, and gene duplication or be an artifact of technical errors. An anomaly evidenced by a single EST (6 of 13 for group 1 chromosomes) is questionable

and likely due to a technical error. However, anomalies involving two or more mapped ESTs with the same location pattern (7 of 13) are more likely to result from a biological event. Of the 7 multi-EST anomalies, three (B, J, and M, Figure 3) were supported by four to five ESTs. These three striking anomalies belong to an intra-arm anomaly, *i.e.*, long (or short) arm to long (or short) arm of two homoeoleogous chromosomes. The chromosome bins involved had different EST densities (Figure 3). These anomalies could have resulted from chromosomal rearrangements, possibly resulting from transposition and gene duplication mainly between chromosome intervals with different gene density.

Homoeologous relationship between wheat consensus chromosome 1 and the rice genome: Homoeology between wheat and rice genomes was first studied by Ahn *et al.* (1993) and followed by Kurata *et al.* (1994) and Van Deynze *et al.* (1995) at the macro level. Recently, Sorrells *et al.* (2003), studying a subset of project EST loci involving all homoeologous groups, compared rice and wheat genomes at the DNA sequence level. All studies indicated that rice chromosomes 5 and 10 were homoeologous with group 1 chromosomes of wheat. Sorrells *et al.* (2003) also showed that 81% of the rice BAC/PAC clones were matched by wheat ESTs. Our results (Figure 2; Table 6; supplemental online material at http://wheat.pw.usda.gov/pubs/2004/Genetics/) further corroborated the close syntenic or homoeologous relationship of wheat consensus chromosome 1 with rice chromosomes 5 and 10.

Comparative mapping based on cDNA clones indicated that rice chromosome 5 was largely conserved with wheat consensus chromosome 1 and that rice chromosome 10 was conserved with a portion of the long arm of wheat consensus chromosome 1. Our results, based on sequence comparison between wheat consensus chromosome 1 ESTs and rice BAC/PAC clones, further identified specific wheat consensus chromosome 1 bins with significant colinearity to rice chromosomes 5 and 10 (Figure 2; Table 6). Since distal segments of wheat chromosomes are gene-rich and recombination-rich regions (Delaney 1995a,b; Gill *et al.* 1996a,b; Faris *et al.* 2000), we may expect rice chromosome 5 to have close relationships to the gene clusters in the distal regions of wheat consensus chromosome 1. Rice chromosomes 5 and 10 and wheat chromosome 1 were probably differentiated by a chromosome fission/fusion after evolutionary divergence (Ahn *et al.* 1993). As shown in Figure 1, both the EST number and EST density in the C-1L-0.47 region significantly ($P < 0.01$) exceeded those in the C-1S-0.47 region of wheat consensus chromosome 1. Therefore, we also speculate that, in the process of divergence, the DNA segments corresponding to the current rice chromosome 10 might have been translocated to the proximal region of the long arm of rice chromosome 5. Because of this possible "translocation," the proximal part of the long arm of wheat chromosome

1 is not as gene poor as the counterpart of the short arm appears (Figures 1 and 2).

Kurata et al. (1994) provided evidence for the conservation of gene order between rice and wheat; i.e., many wheat chromosomes contain genes and genomic DNA fragments in a similar order to that found on rice chromosomes. Van Deynze et al. (1995) pointed out that although wheat orthologous loci span all of rice chromosome 5 and 10 linkage maps, loci from the distal portion of the linkage maps for the short arms of the Triticeae chromosomes are not represented in these rice chromosomes. The only two loci from the distal short arm of the linkage maps of the Triticeae that could be detected in rice did not map to homoeologous positions in rice chromosome 5. The linear organizations of genes in nine different genomes of grasses, including wheat, can be described in terms of only 25 "rice linkage blocks" based on genetic mapping using common DNA probes (Gale and Devos 1998). Our results, based on DNA sequence comparison and chromosome deletion bin mapping, corroborate, to some extent, the conservation of gene orders between rice and wheat. However, a number of genes not showing any colinearity were also identified. This noncolinearity may be due to chromosome structural changes, gene inversions, transposon-like movements, and inclusion of multi-copy probes. The ESTs in the orthologous regions of wheat consensus chromosome 1 may be ordered following the sequence (BAC/PAC clone) order in rice chromosome 5 or 10, as shown in Figure 2 and supplemental online material at http://wheat.pw.usda. gov/pubs/2004/Genetics/. The order of orthologous EST loci in the individual wheat consensus chromosome 1 deletion bins needs to be verified by genetic mapping or complete sequencing of the bins.

It is clear that the three chromosomes of homoeologous group 1 are not perfectly conserved. They are different in physical size and structure (Gill et al. 1991), the extent of polymorphism detected by both molecular markers (Peng et al. 2000) and QTL analysis (Peng et al. 2003), and gene content, density, and duplication (Table 1). Therefore, even though the consensus chromosome 1 is of considerable value in obtaining a general profile of EST distribution patterns of wheat group 1 chromosomes and their colinearity with rice chromosomes 5 and 10, there are limitations to its resolution and validity.

Homoeologous relationship of wheat consensus chromosome 1 with the Arabidopsis genome: Arabidopsis is the first plant with a completely sequenced genome (Arabidopsis Genome Initiative 2000). A. thaliana was estimated to contain 25,000 genes, but its genome is 99 times smaller than that of hexaploid wheat (Bennett and Leitch 2003). It would be greatly helpful for wheat genomics studies if Arabidopsis information could be transferable to wheat. Initial observations indicated that synteny might be extrapolated to monocot and dicot species (Paterson et al. 1996). Attempts to establish colinearity between the rice and Arabidopsis genomes suggest that colinearity cannot be detected in comparative genetic mapping studies (Devos et al. 1999). In the present study, only 35 (9.5%) of wheat homoeologous group 1 ESTs showed significant matches to the Arabidopsis genome sequences. No significant conservation of gene order was detected between wheat consensus chromosome 1 and Arabidopsis chromosomes (data not shown) with this small number of ESTs. These 35 ESTs represent <30 types of gene function (Table 5). The number of conserved genes shared by monocotyledonous and dicotyledonous plant species seems small. Therefore the DNA sequence data of Arabidopsis do not appear to be very useful in the study of genome organization of wheat or other grass species.

Arabidopsis chromosome 3 accounted for 13 (37.4%) of the DNA segments with significant matches to wheat group 1 ESTs and seems to have the closest relationship, among the five Arabidopsis chromosomes, with wheat consensus chromosome 1 (Table 5). Five (38.5%) of these 13 segments spanning 20.5 Mbp (2.6–23.1 Mbp) of Arabidopsis chromosome 3 corresponded to the ESTs located in the 1L-0.47-0.61 region of wheat consensus chromosome 1 (data not shown). However, as observed between rice and Arabidopsis (Devos et al. 1999), there is poor colinearity between wheat consensus chromosome 1 and Arabidopsis chromosome 3.

Interestingly, only 5 (14.3%) of the 35 group 1 ESTs with matches to Arabidopsis were located in the terminal regions of wheat consensus chromosome 1. The common genes shared by wheat consensus chromosome 1 and Arabidopsis appear located in the middle parts of the chromosome arms or the proximal region of the long arm (Table 5).

LITERATURE CITED

Adams, M. D., J. M. Kelley, J. D. Gocayne, M. Dubnick, M. H. Polymeropoulos et al., 1991 Complementary DNA sequencing: expressed sequence tags and human genome project. Science **252:** 1651–1656.

Adams, M. D., A. R. Kerlavage, R. D. Fleischmann, R. A. Fuldner, C. J. Bult et al., 1995 Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature **377** (Suppl.): 3–17.

Ahn, S. N., J. A. Anderson, M. E. Sorrells and S. D. Tanksley, 1993 Homoeologous relationships of rice, wheat and maize chromosomes. Mol. Gen. Genet. **241:** 483–490.

Akhunov, E. D., A. W. Goodyear, S. Geng, L. L. Qi, B. Echalier et al., 2003 The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. Genome Res. **13:** 753–763.

Anderson, J. A., Y. Ogihara, M. E. Sorrells and S. D. Tanksley,

1992 Development of a chromosomal arm map for wheat based on RFLP markers. Theor. Appl. Genet. **83:** 1035–1043.

Arabidopsis Genome Initiative, 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815.

Bennett, M. D., and I. J. Leitch, 2003 Plant DNA C-values database (release 2.0, Jan. 2003; http://www.rbgkew.org.uk/cval/home page.html).

Bennetzen, J., 2002 Opening the door to comparative plant biology. Science **296:** 60–63.

Briggs, S. P., 1998 Plant genomics: more than food for thought. Proc. Natl. Acad. Sci. USA **95:** 1986–1988.

Cadalen, T., C. Boeuf, S. Bernard and M. Bernard, 1997 An intervarietal molecular marker map in *Triticum aestivum* L. em. Thell. and comparison with a map from a wide cross. Theor. Appl. Genet. **94:** 367–377.

Clark, M. D., S. Hennig, R. Herwig, S. W. Clifton, M. A. Marra *et al.*, 2001 An oligonucleotide fingerprint normalized and expressed sequence tag characterized zebrafish cDNA library. Genome Res. **11:** 1594–1602.

Covitz, P. A., L. S. Smith and S. R. Long, 1998 Expressed sequence tags from a root-hair-enriched *Medicago truncatula* cDNA library. Plant Physiol. **117:** 1325–1332.

Delaney, D. E., S. Nasuda, T. R. Endo, B. S. Gill and S. H. Hulbert, 1995a Cytologically based physical maps of the group-2 chromosomes of wheat. Theor. Appl. Genet. **91:** 568–573.

Delaney, D. E., S. Nasuda, T. R. Endo, B. S. Gill and S. H. Hulbert, 1995b Cytologically based physical maps of the group 3 chromosomes of wheat. Theor. Appl. Genet. **91:** 780–782.

Devey, M. E., and G. E. Hart, 1993 Chromosomal localization of intergenomic RFLP loci in hexaploid wheat. Genome **36:** 913–918.

Devos, K. M., J. Beales, Y. Nagamura and T. Sasaki, 1999 Arabidopsis-rice: Will colinearity allow gene prediction across the eudicot-monocot divide? Genome Res. **9:** 825–829.

Dilbirligi, M., M. Erayman, D. Sandhu, D. Sidhu and K. S. Gill, 2004 Identification of wheat chromosomal regions containing expressed resistance genes. Genetics **166:** 461–481.

Dubcovsky, J., M. C. Luo and J. Dvořák, 1995 Differentiation between homoeologous chromosomes 1A of wheat and 1Am of *Triticum monococcum* and its recognition by the wheat *Ph1* locus. Proc. Natl. Acad. Sci. USA **92:** 6645–6649.

Dubcovsky, J., M. C. Luo, G. Y. Zhong, R. Bransteitter, A. Desai *et al.*, 1996 Genetic map of diploid wheat, *Triticum monococcum* L., and its comparison with maps of *Hordeum vulgare* L. Genetics **143:** 983–999.

Dubcovsky, J., M. Echeide, F. Giancola, M. Rousset, M. C. Luo *et al.*, 1997 Seed storage protein loci and RFLP maps of diploid, tetraploid, and hexaploid wheat. Theor. Appl. Genet. **95:** 1169–1180.

Dvořák, J., and K. C. Chen, 1984 Distribution of nonstructural variation between wheat cultivars along chromosome arm *6Bp*: evidence from the linkage map and physical map of the arm. Genetics **106:** 325–333.

Endo, T. R., 1988 Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L. in common wheat. J. Hered. **79:** 366–370.

Endo, T. R., 1990 Gametocidal chromosomes and their induction of chromosome mutations in wheat. Jpn. J. Genet. **65:** 135–152.

Endo, T. R., and B. S. Gill, 1996 The deletion stocks of common wheat. J. Hered. **87:** 295–307.

Ewing, R. M., A. B. Kahla, O. Poirot, F. Lopez, S. Audic *et al.*, 1999 Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. Genome Res. **9:** 950–959.

Faris, J. D., K. M. Haen and B. S. Gill, 2000 Saturation mapping of a gene-rich recombination hot spot region in wheat. Genetics **154:** 823–835.

Feldman, M., F. G. H. Lipton and T. E. Miller, 1995 Wheats, *Triticum* spp. (Gramineae-Triticinae), pp. 184–192 in *Evolution of Crop Plants*, edited by J. Smart and N. W. Simmonds. Longman Scientific & Technical Press, London.

Fernandes, J., V. Brendel, X. Gai, S. Lal and V. L. Chandler, 2002 Comparison of RNA expression profile based on maize expressed sequence tag frequency analysis and micro-array hybridization. Plant Physiol. **128:** 896–910.

Gale, M. D., and K. M. Devos, 1998 Comparative genetics in the grasses. Proc. Natl. Acad. Sci. USA **95:** 1971–1974.

Gale, M. D., M. D. Atkinson, C. N. Chinoy, R. L. Harcourt, J. Jia *et al.*, 1995 Genetic maps of hexaploid wheat, pp. 29–40 in *Proceedings of the 8th International Wheat Genetics Symposium*, edited by Z. S. Li and Z. Y. Xin. China Agricultural Scientech Press, Beijing.

Gill, B. S., B. Friebe and T. R. Endo, 1991 Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). Genome **34:** 830–839.

Gill, K. S., B. S. Gill, T. R. Endo and E. V. Boyko, 1996a Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. Genetics **143:** 1001–1012.

Gill, K. S., B. S. Gill, T. R. Endo and T. Taylor, 1996b Identification and high-density mapping of gene-rich regions in chromosome 1 of wheat. Genetics **144:** 1883–1891.

Goff, S. A., D. Ricke, T. H. Lan, G. Presting, R. Wang *et al.*, 2002 A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science **296:** 92–100.

Harlan, J. R., 1992 *Crops and Man*, Ed 2. American Society of Agronomy/Crop Science Society of America. Madison, WI.

Hillier, L., G. Lennon, M. Becker, M. F. Bonaldo, B. Chiapelli *et al.*, 1996 Generation and analysis of 280,000 human expressed sequence tags. Genome Res. **6:** 807–828.

Kurata, N., G. Moore, Y. Nagamura, T. Foote, M. Yano *et al.*, 1994 Conservation of genome structure between rice and wheat. Biotechnology **12:** 276–278.

Lazo, G. R., S. Chao, D. D. Hummel, H. Edwards, C. C. Crossman *et al.*, 2004 Development of an expressed sequence tag (EST) resource for wheat (*Triticum aestivum* L.): EST generation, unigene analysis, probe selection, and bioinformatics for a 16,000-locus bin-delineated map. Genetics **168:** 585–593.

Linkiewicz, A. M., L. L. Qi, B. S. Gill, B. Echalier, S. Chao *et al.*, 2004 A 2500-locus bin map of wheat homoeologous group 5 provides new insights on gene distribution and colinearity with rice. Genetics **168:** 665–676.

Lukaszewski, A. J., and C. A. Curtis, 1993 Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. Theor. Appl. Genet. **84:** 121–127.

Ma, X.-F., K. Ross and J. P. Gustafson, 2001 Physical mapping of restriction fragment length polymorphism (RFLP) markers in homeologous groups 1 and 3 chromosomes of wheat by *in situ* hybridization. Genome **44:** 401–412.

Marra, M., L. Hillier, T. Kucaba, M. M. Allen, R. Barstead *et al.*, 1999 An encyclopedia of mouse genes. Nat. Genet. **21:** 191–194.

McIntosh, R. A., Y. Yamazaki, K. M. Devos, J. Dubcovsky, W. J. Rogers *et al.*, 2003 Catalogue of gene symbols for wheat, pp. 1–34 in *Proceedings of the 10th International Wheat Genetics Symposium*, Vol. 4, edited by N. E. Pogna, M. Romano, E. Pogna and G. Galterio. Instituto Sperimentale per la Cerealicotura, Rome.

Munkvold, J. D., R. A. Greene, C. E. Bermudez-Kandianis, C. M. La Rota, H. Edwards *et al.*, 2004 Group 3 chromosome bin maps of wheat and their relationship to rice chromosome 1. Genetics **168:** 639–650.

Nevo, E., A. B. Korol, A. Beiles and T. Fahima, 2002 *Evolution of Wild Emmer and Wheat Improvement—Population Genetics, Genetic Resources, Genome Organization of Wheat's Progenitor, Triticum dicoccoides.* Springer-Verlag, Berlin/Heidelberg, Germany/New York.

Paterson, A. H., T. H. Lan, K. P. Reischmann, C. Chang, Y. R. Lin *et al.*, 1996 Toward a unified genetic map of higher plants, transcending the monocot-dicot divergence. Nat. Genet. **14:** 380–382.

Peng, J. H., 2000 Genomics of wild emmer wheat, *Triticum dicoccoides*: genetic maps, mapping of stripe rust resistance genes and QTLs for agronomic traits. Ph.D. Thesis, University of Haifa, Haifa, Israel.

Peng, J. H., A. B. Korol, T. Fahima, M. S. Röder, Y. I. Ronin *et al.*, 2000 Molecular genetic maps in wild emmer wheat, *Triticum dicoccoides*: genome-wide coverage, massive negative interference, and putative quasi-linkage. Genome Res. **10:** 1509–1531.

Peng, J. H., Y. I. Ronin, T. Fahima, M. S. Röder, Y. C. Li *et al.*, 2003 Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. Proc. Natl. Acad. Sci. USA **100:** 2489–2494.

Qi, L. L., B. Echalier, B. Friebe and B. S. Gill, 2003 Molecular characterization of a set of wheat deletion stocks for using in

chromosome bin mapping of ESTs. Funct. Integr. Genomics **3:** 39–55.

Qɪ, L. L., B. Eᴄʜᴀʟɪᴇʀ, S. Cʜᴀᴏ, G. R. Lᴀᴢᴏ, G. E. Bᴜᴛʟᴇʀ *et al.* 2004 A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. Genetics **168:** 701–712.

Sᴀɴᴅʜᴜ, D., J. A Cʜᴀᴍᴘᴏᴜx, S. N. Bᴏɴᴅᴀʀᴇᴠᴀ and K. S. Gɪʟʟ, 2001 Identification and physical location of useful genes and markers to a major gene-rich region on wheat group *1S* chromosomes. Genetics **157:** 1735–1747.

Sᴄʜᴇᴇᴛᴢ, T. E., M. R. Rᴀʏᴍᴏɴ, D. Y. Nɪsʜɪᴍᴜʀᴀ, A. McCʟᴀɪɴ, C. W. Rᴏʙᴇʀᴛs *et al.*, 2001 Generation of a high-density rat EST map. Genome Res. **11:** 497–502.

Sᴇᴀʀs, E. R., 1954 The aneuploids of common wheat. Univ. Mo. Agric. Exp. Stn. Bull. **572:** 1–58.

Sᴇᴀʀs, E. R., 1966 Nullisomic-tetrasomic combinations in hexaploid wheat, pp. 29–45 in *Chromosome Manipulations and Plant Genetics*, edited by R. Rɪʟᴇʏ and K. R. Lᴇᴡɪs. Oliver & Boyd, Edinburgh.

Sᴇᴀʀs, E. R., 1969 Wheat cytogenetics. Annu. Rev. Genet. **3:** 451–468.

Sᴇᴀʀs, E. R., and L. M. S. Sᴇᴀʀs, 1978 The telocentric chromosomes of common wheat, pp. 389–407 in *Proceedings of the 5th International Wheat Genetics Symposium*, edited by S. Rᴀᴍᴀɴᴜᴊᴀᴍ. Indian Society of Genetics and Plant Breeding, New Delhi.

Sᴏʀʀᴇʟʟs, M. E., M. Lᴀ Rᴏᴛᴀ, C. E. Bᴇʀᴍᴜᴅᴇᴢ-Kᴀɴᴅɪᴀɴɪs, R. A. Gʀᴇᴇɴᴇ, R. Kᴀɴᴛᴇᴛʏ *et al.*, 2003 Comparative DNA sequence analysis of wheat and rice genomes. Genome Res. **13:** 1818–1827.

Vᴀɴ Dᴇʏɴᴢᴇ, A. E., J. Dᴜʙᴄᴏᴠsᴋʏ, K. S. Gɪʟʟ, J. C. Nᴇʟsᴏɴ, M. E. Sᴏʀʀᴇʟʟs *et al.*, 1995 Molecular-genetic maps for group 1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. Genome **38:** 45–59.

Wɪʟʟɪᴀᴍs, P. C., 1993 The world of wheat, pp. 557–602 in *Grains and Oilseeds: Handling, Marketing, Processing*. Canadian International Grains Institute, Winnipeg, Manitoba, Canada.

Yᴜ, J., S. Hᴜ, J. Wᴀɴɢ, G. K. S. Wᴏɴɢ, S. Lɪ *et al.*, 2002 A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science **296:** 79–92.

Zʜᴀɴɢ, D., D. W. Cʜᴏɪ, S. Wᴀɴᴀᴍᴀᴋᴇʀ, R. D. Fᴇɴᴛᴏɴ, A. Cʜɪɴ *et al.*, 2004 Construction and evaluation of cDNA libraries for large-scale expressed sequence tag sequencing in wheat (*Triticum aestivum* L.). Genetics **168:** 595–608.

Communicating editor: J. P. Gᴜsᴛᴀғsᴏɴ