A Chromosome Bin Map of 16,000 Expressed Sequence Tag Loci and Distribution of Genes Among the Three Genomes of Polyploid Wheat

L. L. Qi,* B. Echalier,* S. Chao,[†,1] G. R. Lazo,[‡] G. E. Butler,[§] O. D. Anderson,[‡] E. D. Akhunov,[¶]
J. Dvořák,[¶] A. M. Linkiewicz,[¶,2] A. Ratnasiri,[¶] J. Dubcovsky,[¶] C. E. Bermudez-Kandianis,[&]
R. A. Greene,[&] R. Kantety,[&] C. M. La Rota,[&] J. D. Munkvold,[&] S. F. Sorrells,[&]
M. E. Sorrells,[&] M. Dilbirligi,[**] D. Sidhu,[**] M. Erayman,[**] H. S. Randhawa,[**]
D. Sandhu,[**,3] S. N. Bondareva,[**] K. S. Gill,[**] A. A. Mahmoud,[††] X.-F. Ma,[††]
Miftahudin,[††] J. P. Gustafson,[‡‡] E. J. Conley,[§§] V. Nduati,[§§]
J. L. Gonzalez-Hernandez,[§§,4] J. A. Anderson,[§§] J. H. Peng,[¶¶]
N. L. V. Lapitan,[¶¶] K. G. Hossain,[&&] V. Kalavacharla,[&&]
S. F. Kianian,[&&] M. S. Pathan,[††] D. S. Zhang,[***,5]
H. T. Nguyen,[††] D.-W. Choi,[†††,6] R. D. Fenton,[†††]
T. J. Close,[†††] P. E. McGuire,[†] C. O. Qualset[†]
and B. S. Gill*,[7]

*Department of Plant Pathology, Wheat Genetics Resource Center, Kansas State University, Manhattan, Kansas 66506-5502, †Genetic Resources
Conservation Program, University of California, Davis, California 95616, ‡USDA-ARS Western Regional Research Center, Albany,
California 94710, §Arizona Genomics Institute, Department of Plant Sciences, University of Arizona, Tucson, Arizona 85721,
¶Department of Agronomy and Range Science, University of California, Davis, California 95616, &Department of Plant Breeding,
Cornell University, Ithaca, New York 14853, **Department of Crop and Soil Sciences, Washington State University,
Pullman, Washington 99164-6420, ††Department of Agronomy, University of Missouri, Columbia, Missouri 65211,
‡‡USDA-ARS Plant Genetics Research Unit, Department of Agronomy, University of Missouri, Columbia,
Missouri 65211, §§Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul,
Minnesota 55108, ¶¶Department of Soil and Crop Sciences, Colorado State University, Fort Collins,
Colorado 80523-1170, &&Department of Plant Sciences, North Dakota State University,
Fargo, North Dakota 58105-5051, ***Department of Plant and Soil Science,
Texas Tech University, Lubbock, Texas 79409 and †††Department of Botany
and Plant Sciences, University of California, Riverside, California 92521

ABSTRACT

Because of the huge size of the common wheat (*Triticum aestivum* L., $2n = 6x = 42$, AABBDD) genome of 17,300 Mb, sequencing and mapping of the expressed portion is a logical first step for gene discovery. Here we report mapping of 7104 expressed sequence tag (EST) unigenes by Southern hybridization into a chromosome bin map using a set of wheat aneuploids and deletion stocks. Each EST detected a mean of 4.8 restriction fragments and 2.8 loci. More loci were mapped in the B genome (5774) than in the A (5173) or D (5146) genomes. The EST density was significantly higher for the D genome than for the A or B. In general, EST density increased relative to the physical distance from the centromere. The majority of EST-dense regions are in the distal parts of chromosomes. Most of the agronomically important genes are located in EST-dense regions. The chromosome bin map of ESTs is a unique resource for SNP analysis, comparative mapping, structural and functional analysis, and polyploid evolution, as well as providing a framework for constructing a sequence-ready, BAC-contig map of the wheat genome.

THE three major cereal crops, common or bread wheat (*Triticum aestivum* L., $2n = 6x = 42$), rice (*Oryza sativa* L.), and maize (*Zea mays* L.), not only nourish humankind but also are plant genetic models

with varying degrees of genome complexity. Rice has a small genome (450 Mb; BENNETT and LEITCH 1995) that has been sequenced (GOFF *et al.* 2002; YU *et al.* 2002; http://rgp.dna.affre.go.jp). Maize is a paleopolyploid model with a genome of 2500 Mb (BENNETT and LEITCH

1995) and sequencing of gene-rich regions is under way (Peterson *et al.* 2002; Yuan *et al.* 2002; Lunde *et al.* 2003; http://www.maizegenome.org). Diploid and tetraploid wheats along with common wheat form a polyploid series with genome sizes of ~5000 (diploid), 13,000 (tetraploid), and 17,300 (hexaploid) Mb (Bennett and Leitch 1995). Wheat has emerged as a classic polyploid model. Polyploidy is a widespread evolutionary strategy in angiosperms, and research on wheat has greatly contributed to the understanding of this important phenomenon.

Because of triplication of genetic material, wheat can tolerate the loss of whole chromosomes, arms, and segments (Sears 1954, 1966; Sears and Sears 1978; Endo and Gill 1996). These cytogenetic stocks have been exploited for large-scale mapping of molecular markers into chromosomal "bins," regions delineated by neighboring deletion breakpoints (Werner *et al.* 1992; Gill *et al.* 1993, 1996a,b; Kota *et al.* 1993; Hohmann *et al.* 1994; Delaney *et al.* 1995a,b; Mickelson-Young *et al.* 1995; Faris *et al.* 2000; Weng *et al.* 2000; Qi and Gill 2001). Concurrently, RFLP-based genetic maps were developed for the 2*x*, 4*x*, and 6*x* wheats and a diploid ancestor (Chao *et al.* 1989; K. S. Gill *et al.* 1991; Devos and Gale 1993; Nelson *et al.* 1995a,b,c; Dubcovsky *et al.* 1996; Jia *et al.* 1996; Marino *et al.* 1996; Cadalen *et al.* 1997; Blanco *et al.* 1998; Boyko *et al.* 1999; Nachit *et al.* 2001). In spite of these genetic resources, map-based cloning of genes has been arduous in wheat because of polyploidy and a large genome mostly consisting of repeated DNA sequences; as a result only a few successes have been reported (Faris *et al.* 2003; Huang *et al.* 2003; Yan *et al.* 2003). At the same time, little DNA sequence information is available for wheat to take advantage of gene discoveries in model organisms.

It is in this context that expressed sequence tag (EST) analysis has opened exciting prospects for gene discovery in all organisms, irrespective of their genome size (Adams *et al.* 1991; Hillier *et al.* 1996; Covitz *et al.* 1998; Ewing *et al.* 1999; Fernandes *et al.* 2002; Shoemaker *et al.* 2002; Van der Hoeven *et al.* 2002). Large-scale mapping of EST unigenes can provide valuable insights into the organization of genomes and chromosomes. Tomato (*Lycopersicon esculentum* Mill.) chromosomes have proximal heterochromatic and distal euchromatic regions, and preliminary data have indicated that most ESTs map to the distal euchromatic regions (Van der Hoeven *et al.* 2002). Using EST data, gene density can be calculated for individual chromosomes and chromosome regions by dividing the number of observed loci by the number of expected loci (on the basis of size). In humans, gene density varied from 0.56 for chromosome X to 1.74 for chromosome 19 (Deloukas *et al.* 1998; http://www.ncbi.nlm.nih.gov/genome/guide/human). In cereals, inter- and/or intrachromosomal variation in gene density has been documented in wheat (Gill *et al.* 1996a,b), rice (Wu *et al.* 2002), and

maize (Davis *et al.* 1999). Obviously, EST distribution in relation to chromosome landmarks (short and long arms, euchromatin, heterochromatin, centromeres, and telomeres) and recombination is important in comparative analysis of chromosome structure and evolution, gene isolation, and targeted genome sequencing for a large genome species such as wheat. Such analysis of EST loci distribution for individual wheat chromosomes is presented in the accompanying articles in this issue (Conley *et al.* 2004; Hossain *et al.* 2004; Linkiewicz *et al.* 2004; Miftahudin *et al.* 2004; Munkvold *et al.* 2004; Peng *et al.* 2004; Randhawa *et al.* 2004). In this article, we report the genome-level comparative mapping analysis of 16,099 EST loci among the 21 chromosomes and variation in gene density among individual chromosomes, chromosomal regions, homoeologous groups, and A, B, and D genomes of hexaploid wheat. The implications of these results for polyploidy-driven evolution and especially for targeted sequencing of gene-rich regions of a polyploid genome are discussed.

MATERIALS AND METHODS

Genetic stocks: For the chromosome bin mapping of EST-specific fragments, cytogenetic stocks in the hexaploid cultivar Chinese Spring (*T. aestivum*) were used. The selected set includes 21 nullisomic-tetrasomic (NT) lines for mapping ESTs to individual chromosomes (Sears 1954, 1966), 24 ditelosomic (DT) lines for mapping ESTs to arms and centromeres (Sears and Sears 1978), and 101 deletion (del) lines with 119 chromosome-segment deletions for subarm mapping (Endo and Gill 1996). The fraction length (FL) value of each deletion identifies the position of the breakpoint from the centromere relative to the length of the complete arm. These stocks provide a complete coverage of the wheat genome, subdividing it into 159 chromosome bins. All the genetic stocks selected for EST mapping were cytologically and/or molecularly verified by C-banding and Southern hybridization with >500 EST clones (Qi *et al.* 2003). Seeds of the aneuploid and deletion stocks from the Wheat Genetics Resource Center (WGRC) were made available to 10 collaborating mapping labs (http://wheat.pw.usda.gov/NSF).

EST singleton clones and Southern hybridization: Under the auspices of the NSF wheat EST genomics project, 117,466 cDNA clones were sequenced from 43 libraries representing a wide range of wheat tissue and organs, developmental stages, and environmental conditions. The software program phrap was used to assemble the EST contigs with the parameters penalty, −5; minmatch, 50; minscore, 100 (Lazo *et al.* 2004). The ESTs from each contig and unassembled ESTs were considered as singleton clones and formed the wheat unigene set. A total of 7104 EST singleton clones (unigenes) were selected for chromosome bin mapping in the present study. The amplified PCR products of these clones were prepared at the U.S. Department of Agriculture (USDA)/Agricultural Research Service (ARS) Western Regional Research Center (Albany, CA) and delivered to 10 mapping labs (http://wheat.pw.usda.gov/NSF) for Southern hybridization.

Protocols for Southern hybridization, scoring, and data analysis were standardized as much as possible during a workshop of students and/or postdoctorals from participating labs at the WGRC, Kansas State University, Manhattan, Kansas, on

February 11–16, 2001. The detailed protocols are available on line at http://wheat.pw.usda.gov/NSF/project/mapping_data.html and were as described in QI *et al.* (2003). Five Southern blots (each with 30 lanes for a total of 150) with *Eco*RI-digested DNA of the cytogenetic stocks were hybridized with a single EST clone in one hybridization reaction. λDNA digested with *Hin*dIII and *Bst*EII was used as a size marker.

EST mapping, scoring, and data verification: An EST locus was assigned to a chromosome, an arm, or a bin according to whether its restriction fragment was present or absent in a given set of DNA lanes of a Southern blot (Figure 1). Some stocks have more than one deletion chromosome. In these cases, NT and DT lines were extremely helpful to assign a locus to a correct location. In other cases where a fragment was present in all deletion lines but absent in NT and DT lines, the fragment was assigned to a centromeric bin of the specific arm. Southern images were scored by at least two workers at one location, and all mapping data and images of autoradiographs can be accessed at the project website at http://wheat.pw.usda.gov/cgi-bin/westsql/map_locus.cgi. The coordinators of the seven homoeologous groups for the project (group 1, N. L. V. Lapitan; group 2, J. A. Anderson; group 3, M. E. Sorrells; group 4, J. P. Gustafson; group 5, J. Dubcovsky; group 6, K. S. Gill; and group 7, S. F. Kianian) further verified each set of data from all mapping labs. Only mapping data with identical scores by both the mapping labs and the coordinators were considered confirmed loci and were used in this analysis.

EST density: Chromosome size data on mitotic metaphase chromosomes of Chinese Spring wheat were taken from B. S. GILL *et al.* (1991). The distribution of EST loci was assumed to be uniform along the physical length of a chromosome. The mean number of loci per micrometer was calculated by dividing the total EST loci mapped by the total physical length (235.4 μm) of the wheat chromosome complement (B. S. GILL *et al.* 1991). The numbers of expected EST loci per chromosome, chromosome arm, and chromosome bin were calculated by multiplying the mean number by their physical length. The ratio of observed *vs.* expected EST loci was used to estimate EST (or gene) density. The chi-square goodness-of fit test was used to detect significant differences between observed and expected numbers of loci. The total number of mapped EST loci for a whole genome was directly taken from the project website and does not always correspond to the totals of the number of mapped ESTs reported for the each of the individual homoeologous groups (CONLEY *et al.* 2004; HOSSAIN *et al.* 2004; LINKIEWICZ *et al.* 2004; MIFTAHUDIN *et al.* 2004; MUNKVOLD *et al.* 2004; PENG *et al.* 2004; RANDHAWA *et al.* 2004). This is because analysis by genome differs from analysis by single bins within individual chromosomes in the various homoeologous groups with respect to several things. For example, at the genome level, if two ESTs have duplicate mapping patterns, one of them should be eliminated from the analysis; at the homoeologous-group level, both of them would be eliminated. Analysis at the genome level can include ESTs that map only to a chromosome, a chromosome centromere, or a chromosome arm. When the bins within an individual chromosome within a homeologous group were analyzed, only those EST loci that mapped to individual bins were included. Each analysis yields a different count of total ESTs mapped.

RESULTS

In hexaploid wheat, mapping of EST loci to each chromosome in a homoeologous group was based on interchromosomal polymorphism. An EST locus was as-signed to a chromosome bin according to presence or absence of a restriction fragment in a series of deletion lines. An example of localization of EST loci to individual bins of three chromosomes in a homoeologous group is shown in Figure 1. A total of 7104 EST unigenes were hybridized with the selected set of wheat aneuploid and deletion stocks. Of these, 741 EST probes did not provide useful information due to either high copy number or a lack of signal. For another 601 EST probes, Southern fragments could not be assigned to a chromosome or a chromosome bin for various reasons including comigrating fragments, missing DNA lanes, or poor blots. The remaining 5762 informative EST probes detected 27,658 restriction fragments. Of these, 17,761 fragments were assigned to either a chromosome or a chromosome bin by mapping laboratories. Each mapped fragment was considered a locus. The whole genome map contained 16,099 mapped EST loci that were verified by the coordinators of seven homoeologous groups and were analyzed for gene redundancy, gene density, and other structural features.

Gene redundancy: Assuming that each restriction fragment represented a single gene, the null hypothesis was that each EST unigene detected three restriction fragments, one in each genome of hexaploid wheat. However, this was not the case. The total number of restriction fragments in the three wheat genomes hybridizing with an EST unigene ranged from 1 to 24. Each EST detected a mean of 4.8 restriction fragments and 2.8 loci. About one-half of the EST probes (46%) detected 3 or 4 fragments. Twelve percent detected only 1 or 2 restriction fragments and 42% detected 5 or more restriction fragments. Some EST unigenes failed to display syntenic relationships and were chromosome or genome specific (Table 1). Eight of 14 EST unigenes that were found to be chromosome or genome specific involved B-genome chromosomes. As an example, view BE442924, a B-genome-specific EST unigene, at the web site (http://wheat.pw.usda.gov/cgi-bin/westsql/map_locus.cgi).

Other ESTs detected intra- and interchromosomal duplications within a chromosome or among chromosomes within a genome (Table 2). Of the 5762 EST unigenes, 1086 (19%) hybridized to restriction fragments on chromosomes in different homoeologous groups. Within this group of 1086 EST unigenes, a majority (75%) detected loci in two homoeologous groups, 17% detected loci in three groups, and <10% detected loci in more than three groups. Some ESTs detected complicated patterns of duplications. For example, BE443755 hybridized to 17 fragments that could be assigned to 1A, 5A, all B-genome chromosomes, and 7D. Of those, nine loci were mapped on opposite arms within homoeologous groups on chromosomes 1AL-1BS and 5AS-5BL and on both arms of chromosomes 4B and 7B (Table 2; http://wheat.pw.usda.gov/cgi-bin/westsql/map_locus.cgi).
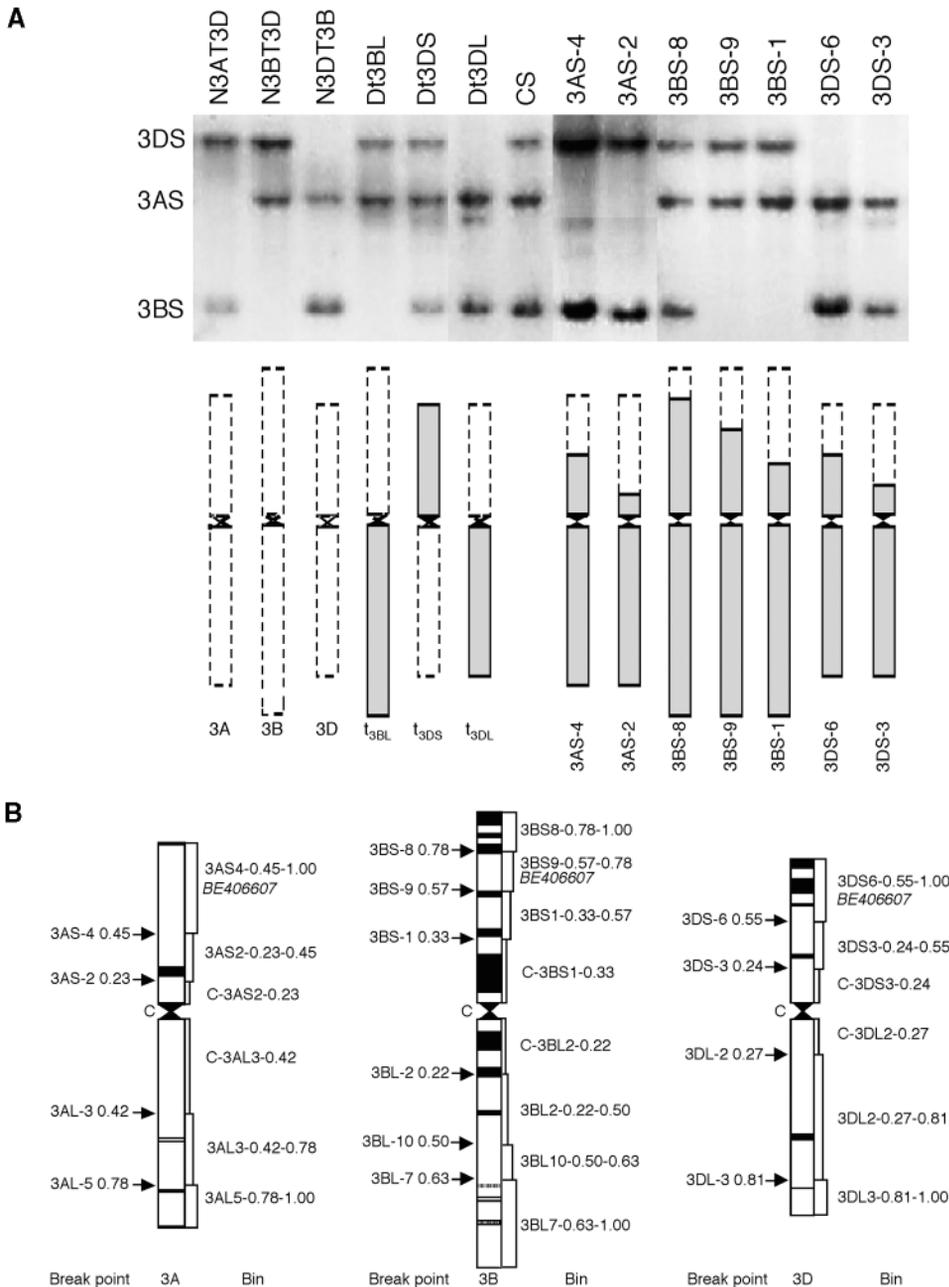
A

B

Figure 1.—An example of localization of EST loci to individual bins of three chromosomes in a homoeologous group. (A) An autoradiograph of Southern hybridization of genomic DNA of standard Chinese Spring (CS) and aneuploid and deletion stocks for group 3 chromosomes. Chromosome ideograms indicate the specific chromosome constitution in the corresponding genetic stocks. The dashed line means the chromosome, chromosome arm, or segment is missing. The genomic DNA of nullisomic-tetrasomic (NT), ditelosomic (DT), and deletion (del) lines was digested with *Eco*RI and hybridized with BE406607. The first and second fragments from the top detected by this EST were missing in all deletion lines of 3DS and 3AS arms, as well as corresponding NT and DT lines, and were mapped to the bins of 3DS6-0.55-1.00 and 3AS3-0.45-1.00, respectively. The third fragment was absent in N3BT3D, Dt3BL, del3BS-1, and del3BS-9, but present in del3BS-8, and was assigned to the bin 3BS9-0.57-0.78. (B) Depiction of chromosome bins and mapping information. Ideogram of C-banded chromosomes of group 3 is after B. S. Gill *et al.* (1991). The deletion names and breakpoints (indicated as fraction length from the centromere) are on the left and the bin assignments are on the right of each chromosome. The EST mapped to bins is in italics on the right of each chromosome.

Genome and chromosome distribution of EST loci: Of the 16,099 EST loci, 14,051 were mapped to chromosome bins and 2048 to only a chromosome or a chromosome arm. Eleven percent more loci were detected in the B genome (5774) than in the D (5179) and A (5146) genomes (Figure 2). The A and D genomes had almost the same number of loci. Analysis of the ratio of observed to expected loci in each genome (based on the physical length of the chromosomes in each genome) indicated a significantly higher gene density for the D genome and a significantly lower gene density for the A genome ($P < 0.001$). Among seven homoeologous groups, the numbers of EST loci observed were significantly higher than expected only for homoeologous group 2 and significantly lower than expected only for homoeologous group 6 ($P < 0.005$).

The 21 wheat chromosomes differ in absolute size and arm ratio (B. S. Gill *et al.* 1991). The expectation was that larger chromosomes would have a greater number of EST loci than the smaller chromosomes. Similarly, the long arms would have greater numbers of EST loci than the short arms within a chromosome. Both expectations were realized with some exceptions. Among the 21 chromosomes, 3B and 2B rank first and second in size and they also ranked first and second in number of EST loci, with 972 and 948 EST loci, respectively. Chromosome 1D is the smallest in size, yet chromosomes 6D (584 EST loci, ranked eighteenth on the basis of size) and 4B (612 EST loci, ranked eleventh on the basis of size) had the fewest number of EST loci. As a rule, the long arms had greater numbers of EST loci than the shorter arms (data not shown) and, among

<div style="text-align:center">

TABLE 1

Selected chromosome- and genome-specific ESTs of wheat

</div>

EST	No. of total fragments	Chromosome bin location	Specificity
BF474347	1	C-1BL6-0.32	1B
BM140322	8	2BS3-0.84-1.00	2B
BE445168	2	3AL5-0.78-1.00	3A
BE591233	1	1DS5-0.70-1.00	1D
BE517903	5	3BS8-0.78-1.00	3B
BG313712	1	6AL8-0.90-1.00	6A
BM135248	2	C-6BS5-0.75	6B
BE497897	1	6BL5-0.40-1.00	6B
BE443803	3	6BL5-0.40-1.00	6B
BM138667	2	6DL10-0.80-1.00	6D
BG607571	4	1AL1-0.17-0.61 C-4AL12-0.43 6A 7AL1-0.39-0.71	A genome
BE442924	7	C-1BL6-0.32 C-2BL2-0.36 C-3BL2-0.22 5BS4-0.43-0.56[a] 6BL3-0.36-0.40 7BS1-0.27-1.00	B genome
BM134384	3	C-2BS1-0.53 3BL7-0.63-1.00 7BL2-0.33-0.63	B genome
BE498837	6	2DL3-0.49-1.00 3D C-4DL9-0.31 5DL5-0.76-1.00 7DS4-0.61-1.00[a]	D genome

[a] Two fragments were mapped to this chromosome bin.

the long arms, 5BL is the longest and had the highest number of mapped EST loci (636).

Because individual chromosomes within a homoeologous group were assumed to have similar gene content, we hypothesized that the number of EST loci should not differ significantly among individual chromosomes within a homoeologous group, irrespective of their size. However, this expectation was not realized because within six of the seven homoeologous groups, B-genome chromosomes had the highest number of mapped EST loci, followed by A- or D-genome chromosomes. The only exception was in group 4 where 4B had the lowest number of loci (4A is an exceptional case that is discussed later). In comparing A- and D-genome chromosomes, D-genome chromosomes had a greater number of EST loci except groups 5 and 6 where chromosomes 5A and 6A had a greater number of loci than 5D and 6D, respectively.

Gene density was analyzed among genomes, chromosomes, arms, and subarm regions. Assuming a uniform distribution of genes in the wheat genome, we expected 68 EST loci per micrometer for wheat chromosomes in this study. Gene density values (indicated by a ratio of numbers of EST loci observed/expected for a specified length) can be above or <1. There were differences in

gene density among A-, B-, and D-genome chromosomes. The gene density was less than one for A-genome chromosomes (except 4A and 6A) and more than one for D-genome chromosomes (except 5D and 6D). In the B-genome chromosomes, 1B, 2B, 3B, and 5B had average gene densities. Three chromosomes, 1A, 4B, and 6D, had significantly lower gene density ($P < 0.001$). Chromosomes 1D, 2D, 3D, and 7D had significantly higher gene densities ($P < 0.001$). The gene density in chromosome 2D was 1.6-fold higher than that in chromosome 4B with the lowest gene density.

In general, the short arms had lower gene density than the long arms. Among all chromosome arms, eight short arms, 2AS, 3AS, 4BS, 5AS, 5BS, 5DS, 6BS, and 6DS; and the long arm of 1AL; had significantly lower gene densities ($P < 0.001$). The significantly higher gene densities were observed on long arms 1DL, 2BL, 2DL, 3DL, 5AL, 5BL, and 7DL ($P < 0.001$, Figure 3). Chromosome arms 2DL and 1DL had the highest gene density. The 1DL arm accounts for the high gene density for chromosome 1D. Among the short arms 4BS had the lowest gene density.

Distribution of loci among chromosome bins: Gene density along the centromere-telomere axis of individual chromosome arms was revealed by analyzing the chromosome bin map (Figure 4). In general, gene density in each chromosome increased with increasing relative physical distance from the centromere. The Pearson correlation coefficient was $r = 0.57$ ($P < 0.0001$). Most gene-poor regions were located around the centromeres. The proximal region of the 1AS arm had the lowest gene density (0.05) and the telomeric bin of 6AL had the highest gene density (4.70, $P < 0.0001$), a 94-fold difference. Among 158 chromosome bins delineated by deletion breakpoints, 48 bins had significantly lower than expected numbers of EST loci observed and 50 had significantly greater than expected numbers of EST loci observed ($P < 0.001$). The overall breakdown of bins with varying gene densities is given in Table 3. Thirty-three chromosome bins (21%) had a very low gene density with a range of 0.05–0.49. Of these, 22 bins (67%) were located around the centromeric regions. None of these bins were located in the distal 20% of the arm. Another 27 of 41 chromosome bins (66%) with ratios between 0.50 and 0.99 were located in the proximal 60% of the arm length. All but one (2BL2-0.36-0.50) of the 23 chromosome bins with a twofold or greater gene density were located in the distal 40% of the arm; 16 of 23 bins (70%) were found in the distal 20% of the arm. The extremely high EST densities observed in chromosome bins 5BL9-0.75-0.79 (6.11), 5BL14-0.75-0.76 (11.0), and 6DS6-0.99-1.00 (17.00) may be due to errors in fraction length calculations of deletions delineating those bins.

Agronomic traits: A total of 67 previously mapped agronomic genes were assigned to chromosome bins on the basis of bin allocation of linked markers (Figure 4; GUPTA *et al.* 1999; PENG *et al.* 2000; LIU and ANDERSON

TABLE 2

Selected patterns of intra- and interchromosome duplications and the lack of homeology displayed by
multicopy ESTs of wheat

EST	No. of fragments	Chromosome arm	EST location in bin	Other mapped location
BE443775	17	1AL-1BS	1AL1-0.17-0.61	2BL, 3BL, 6BL
			1BS10-0.50-0.84	
		4BS-4BL	C-4BS4-0.37	
			C-4BL1-0.71	
		5AS-5BL	5AS1-0.40-0.75	
			5BL6-0.29-0.55	
		7BS-7BL	C-7BS1-0.27	
			C-7DS5-0.36	
			C-7BL2-0.33	
BE489313	15	1BS-1BL	1BS.sat18-0.50-1.00	7BS, 7DS
			1BL2-0.69-0.85	
		1DS-1DL	1DS5-0.70-1.00	
			1DL2-0.41-1.00	
BE500081	9	1AS-1BL	1AS1-0.47-0.86	7DS
			C-1BL6-0.32	
		2BS-2DL	2BS3-0.84-1.00	
			2DL3-0.49-0.76	
BE638132	4	2BS-2BL	C-2BS1-0.53	
			C-2BL2-0.36	
			2BL2-0.36-0.50	
BF474833	8	2BS-2BL	C-2BS1-0.53	1DS, 7AL
			2BL6-0.89-1.00	
BE489227	11	2BS-2DL	2BS3-0.84-1.00	5BL, 6BL, 7BS
			2DL9-0.76-1.00	
		3BS-3BL	C-3BS1-0.33	
			C-3BL7-0.63	
BE517312	8	3BS-3BL	3BS8-0.78-1.00	
			3BL7-0.67-1.00	
			3AL5-0.78-1.00	
BE489323	10	3BL-3DS	3BL2-0.22-0.50	1AS, 1BS, 2AL
			3DS3-0.24-0.55	6DL, 7DL
BE405234	12	3AL-3BS-3DS	3AL5-0.78-1.00	2AL, 2BL
			3BS1-0.33-0.57	
			3DS3-0.24-0.55	
BE499002	8	6AS-6AL	6AS1-0.35-0.65	
			6AL8-0.90-1.00	
			6BL5-0.40-1.00	
BE405852	14	6AS-6AL	C-6AL4-0.35	
			6AS5-0.65-1.00	
			6BS-Sat	
BE426715	7	7AS-7AL	7AS5-0.59-0.89	1AS, 2BL, 2DL
			7DS4-0.61-1.00	5AL
			7AL16-0.86-0.90	

2003; McINTOSH *et al.* 2003). Forty-one of the 67 genes were located in the distal 20% of individual chromosome arms. Most genes in distal bins relate to pest resistance. Only 2 of the 67 genes (3%) were mapped in the proximal 40% of the arm. The distribution of agronomic genes was consistent with the distribution pattern of EST loci; most colocalized in the EST-rich, distal chromosomal regions.

Chromosomal structural changes: In hexaploid wheat, chromosome and arm homoeologies within sets of tripli-

cated chromosomes are highly conserved with the exception of a translocation involving chromosomes 4A, 5A, and 7B, which is fixed in polyploid wheat (NARANJO *et al.* 1987; LIU *et al.* 1992; DEVOS *et al.* 1995; MICKELSON-YOUNG *et al.* 1995; NELSON *et al.* 1995a). In addition, a pericentric inversion occurred in chromosome 4A with arm homoeologies as 4AS = 4BL = 4DL and 4AL = 4BS = 4DS (MICKELSON-YOUNG *et al.* 1995). A pericentric inversion has been documented in 4B (ENDO and GILL 1984; B. S. GILL *et al.* 1991; FRIEBE and GILL 1994;
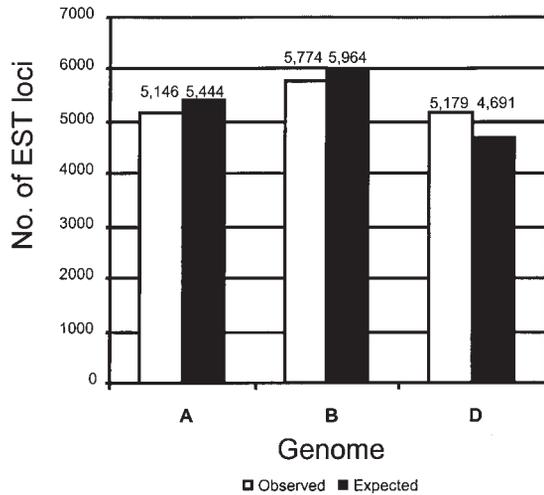
FIGURE 2.—Number of observed and expected EST loci for the A, B, and D genomes of wheat.
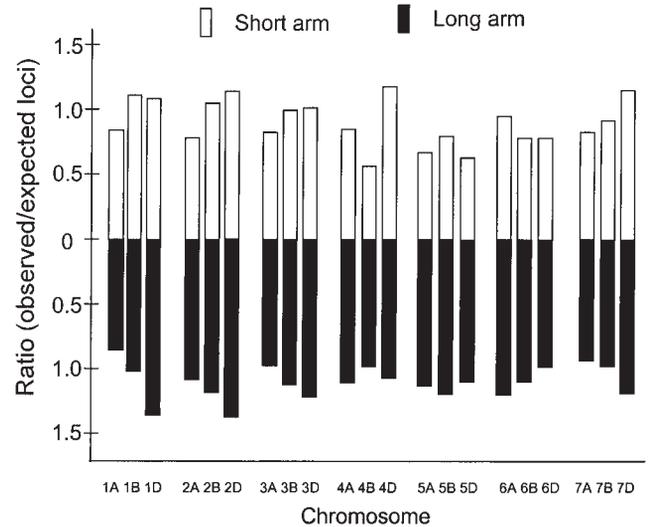

FIGURE 3.—Gene density in the 42 chromosome arms of wheat. Gene density was calculated by the ratio of observed to expected loci (on the basis of chromosome length).

MICKELSON-YOUNG *et al.* 1995). With large-scale EST mapping, we have uncovered further structural changes in 4A and confirmed the 4B inversion and additional chromosomal structural changes in most of the B-genome chromosomes. These are described in detail in individual chromosome group articles (CONLEY *et al.* 2004; LINKIEWICZ *et al.* 2004; MIFTAHUDIN *et al.* 2004) and they account for some of the anomalies in gene distribution (see DISCUSSION).

DISCUSSION

ESTs have become an essential resource for gene discovery. Fifteen plant species had >50,000 ESTs each in public databases by January 30, 2004 (http://www.ncbi.nlm.nih.gov/dbEST), including wheat with the most entries (549,926). The second essential resource is the chromosome maps of ESTs where EST data can be related to genes controlling phenotypic traits (DAVIS *et al.* 1999; WU *et al.* 2002). Apart from Arabidopsis and rice, where a nearly complete genome sequence is available, common wheat now is one of the most densely mapped genomes among plants and highest among all polyploid organisms. Previously, 5537 RFLP loci, several hundred SSR loci, and 2049 protein loci and genes controlling phenotypic traits were mapped in wheat (MCINTOSH *et al.* 2003). The current chromosome bin map by the NSF EST project contains 16,099 EST loci. This vast EST resource and the chromosome bin maps (CONLEY *et al.* 2004; HOSSAIN *et al.* 2004; LINKIEWICZ *et al.* 2004; MIFTAHUDIN *et al.* 2004; MUNKVOLD *et al.* 2004; PENG *et al.* 2004; RANDHAWA *et al.* 2004) will hasten the pace of gene discovery in wheat. These data are also useful for analyzing general trends and peculiarities of gene distribution among chromosomes and genomes as influenced by features of karyotype, polyploidy, and evolutionary forces unique to polyploid organisms.

Mapping strategy: Several features of the wheat EST map distinguished it from other plant EST maps. First, Southern hybridizations were used, thus making it possible to locate diverged duplicated genes sharing 80% or more homology. This would not have been possible using a PCR-based approach. Second, aneuploid and deletion stocks were used for allocating restriction fragments to individual chromosomes, arms, and bins. The genome coverage was 100% because a set of nullisomic-tetrasomic and several telocentric stocks were included in the analysis. For many EST clones, all restriction fragments were mapped to specific chromosome bins. The choice of *Eco*RI as the restriction enzyme did not appear to have any effect on the number or distribution pattern of EST loci among chromosomes and genomes as judged from the analysis of previous deletion-mapping results using other restriction enzymes (DELANEY *et al.* 1995a,b; MICKELSON-YOUNG *et al.* 1995; GILL *et al.* 1996b; FARIS *et al.* 2000; QI and GILL 2001; SANDHU *et al.* 2001). One disadvantage of the chromosome bin map is that loci within chromosome bins cannot be ordered. However, this problem can potentially be overcome by *in silico* ordering of sequenced EST loci using the rice genome sequence as reported by SORRELLS *et al.* (2003) and as is reported in detail in the individual chromosome publications (CONLEY *et al.* 2004; HOSSAIN *et al.* 2004; LINKIEWICZ *et al.* 2004; MIFTAHUDIN *et al.* 2004; MUNKVOLD *et al.* 2004; PENG *et al.* 2004; RANDHAWA *et al.* 2004).

Why do wheat genomes have a variable number of EST loci? The A, B, and D genomes of the diploid progenitors of common wheat diverged from a common ancestor 1–5 million years ago (MYA; HUANG *et al.* 2002). Because of the recent evolutionary differentiation of the A, B, and D genomes from a common ances-
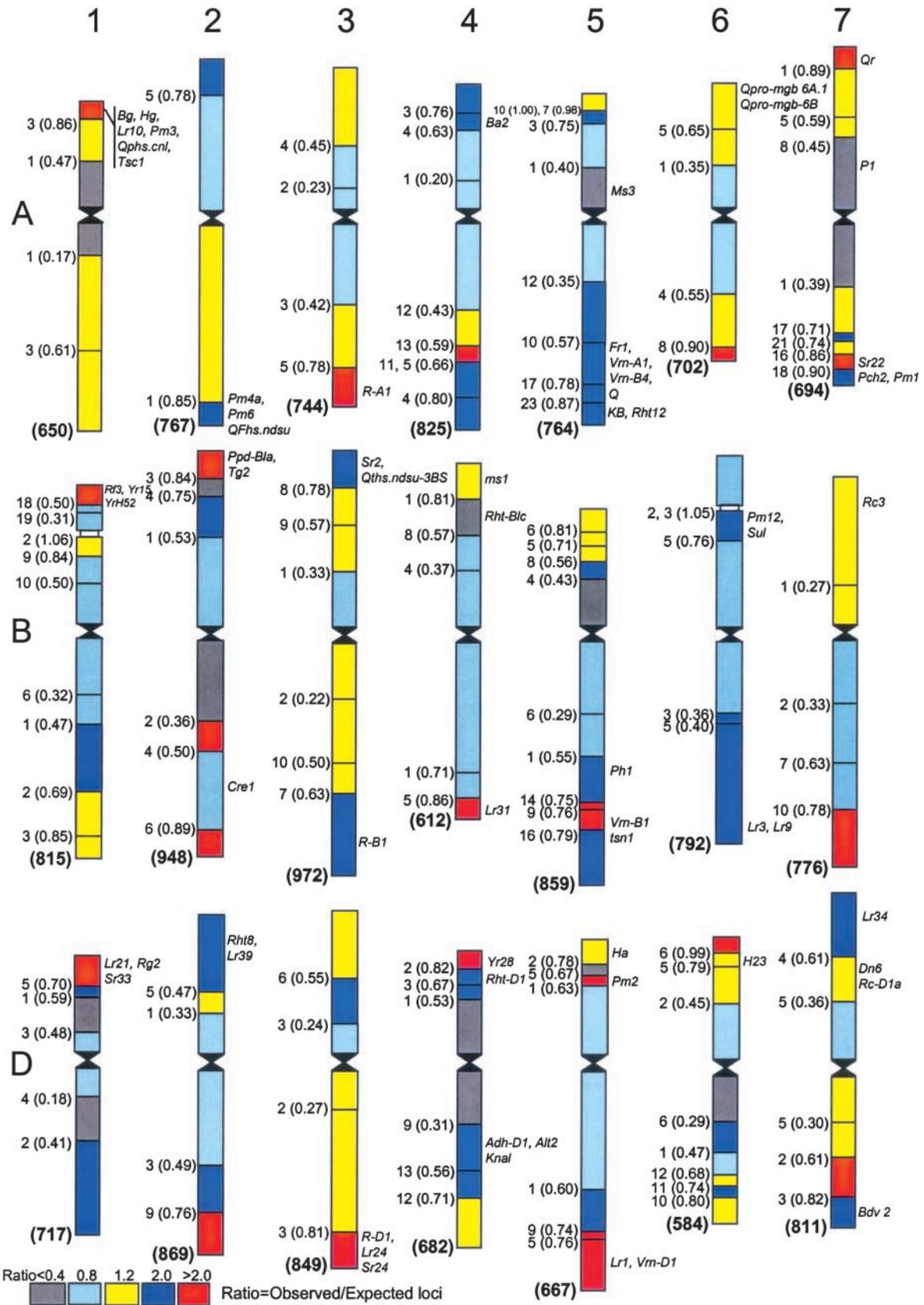
Ratio<0.4   0.8   1.2   2.0   >2.0   Ratio=Observed/Expected loci

TABLE 3

Distribution of gene density in the wheat chromosome bins along the centromere-to-telomere axis

Observed/expected ratio	No. of total bins	Distribution[a]			
		C-0.40	0.40–0.60	0.60–0.80	0.80–1.00
0.05–0.49	33	22	5	6	—
0.50–0.99	41	18	9	12	2
1.00–1.49	37	2	10	12	13
1.50–1.99	25	—	4	12	9
>2.00	23	—	1	6	16
Total	159	42	29	48	40

[a] Relative physical length of arm from centromere. The relative physical length of the whole arm as 1 was divided into four chromosome bins, C-0.40, 0.40–0.60, 0.60–0.80, and 0.80–1.00.

tor, one might expect a similar gene content for these genomes. However, we have found this is not the case as the B genome had a significantly greater number of mapped EST loci than did the A and D genomes, which had similar numbers of loci ($P < 0.001$). In addition, the greater number of EST loci was not related to genome size as the A genome is larger than the B genome and the D genome is smaller than the A genome. Although a number of mechanisms may lead to this outcome, the most plausible explanation relates to the evolutionary history of each of the hexaploid wheat genomes. The two key parameters may be the breeding system (Akhunov *et al.* 2003) and age of the diploid donors. Both A and D genomes trace their origins to self-pollinating diploids *T. urartu* Thum. *ex* Gand. and *Aegilops tauschii* Coss., respectively. On the other hand, the B genome, having a greater number of EST loci, was derived from a species closely related to *Ae. speltoides* Tausch, a cross-pollinating diploid species. As a cross-pollinating species, *Ae. speltoides* is more diverse for RFLP variation than the self-pollinating *Ae. tauschii* (Dvořák *et al.* 1998). In situations of high polymorphism in a cross-pollinating species, it is usually difficult to distinguish between heterozygosity and a duplication event. However, on the basis of DNA sequence analyses, duplicate copies of genes *Acc1*, *Pgk1*, and *X2* were postulated in *Ae. speltoides* as compared to single copies in the self-pollinating A- and D-genome diploids (Huang *et al.* 2002; Li and Gill 2002). Therefore, part of the high polymorphism of the B genome may be due to gene duplications preexisting in the diploid donor.

The age of the lineage of the diploid donors at the time of the polyploidization events may be another contributing factor. Based on a molecular clock (Huang *et al.* 2002), the *Ae. speltoides* lineage is older (*ca.* 4 MYA) than *Ae. tauschii* (*ca.* 2.5 MYA), and the *T. urartu* lineage is the youngest (*ca.* 1 MYA). Older lineages would be expected to have accumulated more polymorphisms including gene duplications. A faster rate of B-genome evolution may also be indicated by the fact that most of the observed chromosomal inversions and translocations involved B genome chromosomes. Akhunov *et al.* (2003) analyzed a subset of the EST data and documented twice as many unique loci in the B genome as compared to the A and D genomes.

Why do wheat chromosomes within homoeologous sets vary in gene density? One factor leading to variation in gene density or number of loci is the evolutionary history of the genomes to which a specific chromosome belongs as discussed above. A second factor is individual chromosome size. The third contributing factor is structural aberrations that include unequal exchanges of genetic material within and among chromosomes. For short arms, the lowest gene density in the 4BS arm (Figure 3) is caused by an asymmetric pericentric inversion as documented by studies on chromosome morphology and C-banding patterns (Endo and Gill 1984) and confirmed by molecular mapping studies (Mickelson-Young *et al.* 1995). In the case of arms 4AL and 7BS, a reciprocal translocation event occurred, where unequal-size fragments were exchanged. Deletion mapping (Mickelson-Young *et al.* 1995) and FISH show that the 7BS segment on 4AL constitutes 25% of its arm length whereas the 5AL-specific segment on 7BS is not

Figure 4.—Gene densities (ratio of observed to expected loci) in individual chromosome bins of common wheat. Values <1 (gray and light blue) indicate low gene density and >1 denote high gene density (yellow, blue, and red). Note gene-dense regions are located toward the distal ends of chromosomes. The numbers on the left of a chromosome indicate the individual deletion lines and the numbers in parentheses are the FL values of each deletion. The number in boldface type in parentheses at the bottom of each chromosome is the total number of EST loci mapped in that chromosome. Previously mapped agronomic traits (indicated on the right in italics) were assigned to chromosome bins on the basis of the bin allocation of linked markers. Note that most traits colocalized to regions of high EST density.

visible at the resolution power of a light microscope (MUKAI *et al.* 1993; JIANG and GILL 1994).

There are additional significant differences in overall gene density in the absence of obvious structural changes. Chromosome 4B has the lowest gene density among the 21 chromosomes of wheat, which is all the more puzzling considering that B-genome chromosomes have ∼11% more loci overall than the A or D genome chromosomes do. One explanation is that 4B is a highly heterochromatic chromosome with proximal blocks of heterochromatin (B. S. GILL *et al.* 1991). Future work may shed light on the apparent slow pace of evolution of 4B, a member of the fast-evolving B genome.

Is polyploidy a trigger for fast genome evolution? It is now generally accepted that the genomes of most eukaryotic organisms have experienced repeated cycles of polyploidization and diploidization (ADAMS *et al.* 2003). Polyploidy has at least two evolutionary consequences. It provides a mechanism for fixing hybrid vigor, so there may be selection pressure for maintaining active/diverse alleles at homoeologous loci. On the other hand, gene redundancy triggers diploidization where duplicated genes may be eliminated, silenced, mutated to acquire new functions, or epigenetically regulated for tissue-specific expression (ADAMS *et al.* 2003). The data presented here on the mapping of gene motifs of a random set of ESTs provided an excellent opportunity for analyzing the role of gene elimination or amplification during polyploidy-driven evolution.

Gene duplications, in addition to those duplicated by polyploidy, were observed for 19% of the loci that were mapped on nonhomoeologous sets of chromosomes. Another 12% of EST clones detected only one or two fragments, indicating gene elimination. The actual deletion frequency may be lower due to comigrating fragments as AKHUNOV *et al.* (2003) estimated it to be 6%. We do not know if these duplication/deletion events occurred before or following polyploidization. AKHUNOV *et al.* (2003) reported that of the duplications events analyzed, 83% occurred at the diploid level and 17% at the polyploid level. Of deletion events, 60% occurred at the diploid level and 40% at the polyploid level. Previously, DUBCOVSKY *et al.* (1996) showed that ∼31% of the RFLP loci in the diploid *T. monococcum* L. map were present more than once in the genome. A similar level of duplications (30%) was observed in the barley (*Hordeum vulgare* L.) RFLP maps, suggesting that a large proportion of the duplications originated before the divergence of the Triticeae species.

Several patterns of gene duplication were observed. Many ESTs detecting more than seven restriction fragments showed both intra- and interchromosomal duplications. In several cases of intrachromosomal duplications, the duplicated fragments were detected in both arms. These events may represent peculiarities of chromosome behavior during which these duplications occurred. For example, opposite arms may be juxtaposed during prophase of meiosis, providing opportunities for gene-conversion type events. The other possibility is that these kinds of ESTs may be as yet unrecognized transposon-like elements that can move within and among different chromosomes.

Prospects for wheat gene discovery and genome analysis: This public resource offers information on sequence, map position, putative function, and often copy number of a given EST. Many EST-SSRs can be identified and mapped using this information. Similarly, single-copy EST loci will be good candidates for developing SNPs. Many agronomically useful genes have been localized to EST-rich regions in the distal ends of chromosomes and are targets for gene discovery (FARIS and GILL 2002; FARIS *et al.* 2003; HUANG *et al.* 2003; YAN *et al.* 2003). Once an EST tightly linked to a target gene has been identified, it can be the start of an *in silico* walk using the rice genome sequence contigs. The candidate clones gleaned from comparative wheat-rice maps are in turn mapped in the target gene population and the procedure is repeated until a cosegregating clone has been identified. Each mapped EST also has a unique size fragment that will be useful in allocating BAC contigs to individual wheat chromosomes. The chromosome bin map of EST loci, chromosome-specific BAC libraries (ŠAFÁŘ *et al.* 2004), and high throughput techniques for BAC fingerprinting (LUO *et al.* 2003) can now be used to make a sequence-ready global BAC-contig physical map of the wheat genome. The sequencing of BAC-contig maps spanning the gene-rich regions will be especially useful for accessing those genes that may be unique to wheat. This includes the *Ph1* gene controlling chromosome pairing in polyploids (RILEY and CHAPMAN 1958), genes affecting unique quality attributes such as bread-making properties (ANDERSON *et al.* 2002), durable resistance genes that underpin the green revolution such as *Lr34* (SINGH and HUERTA-ESPINO 2003), and many genes controlling domestication-driven traits (FARIS *et al.* 2003) that have made the wheat plant the most important food source for humanity.

LITERATURE CITED

ADAMS, K. L., R. CRONN, R. PERCIFIELD and J. F. WENDEL, 2003   Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc. Natl. Acad. Sci. USA **100:** 4649–4654.

ADAMS, M. D., J. M. KELLEY, J. D. GOCAYNE, M. DUBNICK, M. H. POLUMEROPOULOS *et al.*, 1991   Complementary DNA sequencing: expressed sequence tags and human genome project. Science **252:** 1651–1656.

AKHUNOV, E. D., A. R AKHUNOVA, A. M. LINKIEWICZ, J. DUBCOVSKY,

D. HUMMEL et al., 2003 Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates along chromosome arms. Proc. Natl. Acad. Sci. USA 100: 10836–10841.

ANDERSON, O. D., L. LARKS, M. J. CHRISTOFFERS, K. F. McCUE and J. P. GUSTAFSON, 2002 Comparison of orthologous and paralogous DNA flanking the wheat high molecular weight glutenin genes: sequence conservation and divergence, transposon distribution, and matrix-attachment regions. Genome 45: 367–380.

BENNETT, M. D., and I. J. LEITCH, 1995 Nuclear DNA amounts in angiosperms. Ann. Bot. 76: 113–176.

BLANCO, A., M. P. BELLOMO, A. CENCI, C. DE GIOVANNI, R. D'OVIDIO et al., 1998 A genetic linkage map of durum wheat. Theor. Appl. Genet. 97: 721–728.

BOYKO, E. V., K. S. GILL, L. MICKELSON-YOUNG, S. NASUDA, W. J. RAUPP et al., 1999 A high-density genetic linkage map of Aegilops tauschii, the D-genome progenitor of bread wheat. Theor. Appl. Genet. 99: 16–26.

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 wide cross. Theor. Appl. Genet. 94: 367–377.

CHAO, S., P. J. SHARP, A. J. WORLAND, E. J. WARHAM, R. M. D. KOEBNER et al., 1989 RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theor. Appl. Genet. 78: 495–504.

CONLEY, E. J., V. NDUATI, J. L. GONZALEZ-HERNANDEZ, A. MESFIN, M. TRUDEAU-SPANJERS et al., 2004 A 2600-locus chromosome bin map of wheat homoeologous group 2 reveals interstitial gene-rich islands and colinearity with rice. Genetics 168: 625–637.

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.

DAVIS, G. L., M. D. McMULLEN, C. BAYSDORFER, T. MUSKET, D. GRANT et al., 1999 A maize standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. Genetics 152: 1137–1172.

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.

DELOUKAS, P., G. D. SCHULER, G. GYAPAY, E. M. BEASLEY, C. SODERLUND et al., 1998 A physical map of 30,000 human genes. Science 282: 744–746.

DEVOS, K. M., and M. D. GALE, 1993 The genetic maps of wheat and their potential in plant breeding. Outlook Agric. 2: 93–99.

DEVOS, K. M., J. DUBCOVSKY, J. DVOŘÁK, C. N. CHINOY and M. D. GALE, 1995 Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor. Appl. Genet. 91: 282–288.

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.

DVOŘÁK, J., M. C. LUO and Z. L. YANG, 1998 Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing Aegilops species. Genetics 148: 423–434.

ENDO, T. R., and B. S. GILL, 1984 Somatic karyotype, heterochromatin distribution, and nature of chromosome differentiation in common wheat, Triticum aestivum L. em. Thell. Chromosoma 89: 361–369.

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., and B. S. GILL, 2002 Genomic targeting and high-resolution mapping of the domestication gene Q in wheat. Genome 45: 706–718.

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.

FARIS, J. D., J. P. FELLERS, S. A. BROOKS and B. S. GILL, 2003 A bacterial artificial chromosome contig spanning the major do-

mestication locus Q in wheat and identification of a candidate gene. Genetics 164: 311–321.

FERNANDES, J., V. BRENDEL, X. GAI, S. LAL, V. L CHANDLER et al., 2002 Comparison of RNA expression profile based on maize expressed sequence tag frequency analysis and micro-array hybridization. Plant Physiol. 128: 896–910.

FRIEBE, B., and B. S. GILL, 1994 C-band polymorphism and structural rearrangements detected in common wheat (Triticum aestivum). Euphytica 78: 1–5.

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., E. L. LUBBERS, W. J. RAUPP, T. S. COX and B. S. GILL, 1991 A genetic linkage map of Triticum tauschii (DD) and its relationship to the D genome of bread wheat (AABBDD). Genome 34: 362–374.

GILL, K. S., B. S. GILL and T. R. ENDO, 1993 A chromosome region specific mapping strategy reveals gene-rich telomeric ends in wheat. Chromosoma 102: 374–381.

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 group 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.

GUPTA, P. K., R. K. VARSHNEY, P. C. SHARMA and B. RAMESH, 1999 Molecular markers and their applications in wheat breeding. Plant Breed. 118: 369–390.

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.

HOHMANN, U., T. R. ENDO, K. S. GILL and B. S. GILL, 1994 Comparison of genetic and physical maps of group 7 chromosomes from Triticum aestivum L. Mol. Gen. Genet. 245: 644–653.

HOSSAIN, K. G., V. KALAVACHARLA, G. R. LAZO, J. HEGSTAD, M. J. WENTZ et al., 2004 A chromosome bin map of 2148 expressed sequence tag loci of wheat homoeologous group 7. Genetics 168: 687–699.

HUANG, L., S. A. BROOKS, W. LI, J. P. FELLERS, H. N. TRICK et al., 2003 Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploid genome of bread wheat. Genetics 164: 655–664.

HUANG, S. X., A. SIRIKHACHORNKIT, X. J. SU, J. D. FARIS, B. S. GILL et al., 2002 Genes encoding plastid acetyl-CoA carbxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and evolutionary history of polyploid wheat. Proc. Natl. Acad. Sci. USA 99: 8133–8138.

JIA, J., K. M. DEVOS, S. CHAO, T. E. MILLER, S. M. READER et al., 1996 RFLP-based maps of the homoeologous group-6 chromosomes of wheat and their application in the tagging of Pm12, a powdery mildew resistance gene transferred from Aegilops speltoides to wheat. Theor. Appl. Genet. 92: 559–565.

JIANG, J., and B. S. GILL, 1994 Different species-specific chromosome translocations in Triticum timopheevii and T. turgidum support the diphyletic origin of polyploid wheats. Chromosome Res. 2: 59–64.

KOTA, R. S., K. S. GILL, B. S. GILL and T. R. ENDO, 1993 A cytogenetically based physical map of chromosome 1B in common wheat. Genome 36: 548–554.

LAZO, G. R., S. CHAO, D. L. HUMMEL, H. EDWARDS, C. C. CROSSMAN et al., 2004 Development of an expressed sequence tag resource for wheat (Triticum aestivum L.): unigene analysis, probe selection and bioinformatics for a 16,000-locus bin map. Genetics 168: 585–593.

LI, W. L., and B. S. GILL, 2002 The colinearity of the Sh2/A1 orthologous region in rice, sorghum and maize is interrupted and accompanied by genome expansion in the Triticeae. Genetics 160: 1153–1162.

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 insights on gene distribution and colinearity with rice. Genetics 168: 665–676.

LIU, C. J., K. M. DEVOS, C. N. CHINOY, M. D. ATKINSON and M. D.

GALE, 1992 Homoeologous translocations between group 4, 5 and 7 chromosomes in wheat and rye. Theor. Appl. Genet. **83:** 305–312.

LIU, S. X., and J. A. ANDERSON, 2003 Targeted molecular mapping of a major wheat QTL for *Fusarium* head blight resistance using wheat ESTs and synteny with rice. Genome **46:** 817–823.

LUNDE, C. F., D. J. MORROW, L. M. ROY and V. WALBOT, 2003 Progress in maize gene discovery: a project update. Funct. Integr. Genomics **3:** 25–32.

LUO, M.-C., C. S. THOMAS, K. R. DEAL, F. M. YOU, O. D. ANDERSON *et al.*, 2003 Construction of contigs of *Aegilops tauschii* genomic DNA fragments cloned in BAC and BiBAC, Vol. 1, pp. 294–296 in *Proceedings of the 10th International Wheat Genetics Symposium*, edited by N. E. POGNA, M. ROMANO, E. POGNA and G. GALTERIO. Instituto Sperimentale per la Cerealicotura, Rome.

MARINO, C. L., J. C. NELSON, Y. H. LU, M. E. SORRELLS, P. LEROY *et al.*, 1996 Molecular genetic maps of the group 6 chromosomes of hexaploid wheat (*Triticum aestivum* L. em. Thell). Genome **39:** 359–366.

MCINTOSH, R. A., Y. YAMAZAKI, K. M. DEVOS, J. DUBCOVSKY, W. J. ROGERS *et al.*, 2003 Catalogue of gene symbols for wheat, Vol. 4, pp. 1–34 in *Proceedings of the 10th International Wheat Genetics Symposium*, edited by N. E. POGNA, M. ROMANO, E. POGNA and G. GALTERIO. Instituto Sperimentale per la Cerealicotura, Rome.

MICKELSON-YOUNG, L., T. R. ENDO and B. S. GILL, 1995 A cytogenetic ladder-map of the wheat homoeologous group-4 chromosomes. Theor. Appl. Genet. **90:** 1007–1011.

MIFTAHUDIN, K. ROSS, X.-F. MA, A. A. MAHMOUD, J. LAYTON *et al.*, 2004 Analysis of wheat expressed sequence tag loci on wheat chromosome group 4. Genetics **168:** 651–663.

MUKAI, Y., Y. NAKAHARA and M. YAMAMOTO, 1993 Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fluorescence *in situ* hybridization using total genomic and high repeated DNA probes. Genome **36:** 489–494.

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.

NACHIT, M. M., I. ELOUAFI, M. A. PAGNOTTA, A. EL SALEH, E. IACONO *et al.*, 2001 Molecular linkage map for an intraspecific recombinant inbred population of durum wheat (*Triticum turgidum* L. var. *durum*). Theor. Appl. Genet. **102:** 177–186.

NARANJO, T., A. ROCA, P. G. GOICOECHEA and R. GIRALDEZ, 1987 Arm homoeology of wheat and rye chromosomes. Genome **29:** 873–882.

NELSON, J. C., M. E. SORRELLS, A. E. VAN DEYNZE, Y. H. LU, M. ATKINSON *et al.*, 1995a Molecular mapping of wheat: major genes and rearrangements in homoeologous groups 4, 5, and 7. Genetics **141:** 721–731.

NELSON, J. C., A. E. VAN DEYNZE, E. AUTRIQUE, M. E. SORRELLS, Y. H. LU *et al.*, 1995b Molecular mapping of wheat. Homoeologous group 2. Genome **38:** 516–524.

NELSON, J. C., A. E. VAN DEYNZE, E. AUTRIQUE, M. E. SORRELLS, Y. H. LU *et al.*, 1995c Molecular mapping of wheat. Homoeologous group 3. Genome **38:** 525–533.

PENG, J. H., T. FAHIMA, M. S. RODER, Q. Y. HUANG, A. DAHAN *et al.*, 2000 High-density molecular map of chromosome region harboring stripe-rust resistance gene *YrH52* and *Yr15* derived from wild emmer wheat, *Triticum dicccoides*. Genetica **109:** 199–210.

PENG, J. H., H. ZADEH, G. LAZO, J. P. GUSTAFSON, S. CHAO *et al.*, 2004 Chromosome bin map of expressed sequence tags in homoeologous group 1 of hexaploid wheat and homoeology with rice and Arabidopsis. Genetics **168:** 609–623.

PETERSON, D. G., S. R. SCHULZE, E. B. SCIARA, S. A. LEE, J. E. BOWERS *et al.*, 2002 Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. Genome Res. **12:** 795–807.

QI, L. L., and B. S. GILL, 2001 High-density physical maps reveal that the dominant male-sterile gene *Ms3* is located in a genomic region of low recombination in wheat and is not amenable to map-based cloning. Theor. Appl. Genet. **103:** 998–1006.

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.

RANDHAWA, H. S., M. DILBIRLIGI, D. SIDHU, M. ERAYMAN, D. SANDHU *et al.*, 2004 Deletion mapping of homoeologous group 6-specific wheat expressed sequence tags. Genetics **168:** 677–686.

RILEY, R., and V. CHAPMAN, 1958 Genetic control of the cytologically diploid behavior of hexaploid wheat. Nature **182:** 713–715.

ŠAFÁR, J., J. BARTOŠ, J. JANDA, A. BELLEC, M. KUBALÁKOVÁ *et al.*, 2004 Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J. **39:** 960–968.

SANDHU, D., J. A. CHAMPOUX, S. N. BONDAREVA and K. S. GILL, 2001 Identification and physical localization of useful genes and markers to a major gene-rich region on wheat 1S chromosomes. Genetics **157:** 1735–1747.

SEARS, E. R., 1954 *The Aneuploids of Common Wheat.* Bull. 572, University of Missouri Agricultural Experiment Station, Columbia, MO.

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

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

SHOEMAKER, R., P. KEIM, L. VODKIN, E. RETZEL, S. W. CLIFTON *et al.*, 2002 A compilation of soybean ESTs: generation and analysis. Genome **45:** 329–338.

SINGH, R. P., and J. HUERTA-ESPINO, 2003 Effect of leaf rust resistance gene *Lr34* on components of slow rusting at seven stages in wheat. Euphytica **129:** 371–376.

SORRELLS, M. E., M. LA ROTA, C. E. BERMUDEZ-KANDIANIS, R. A. GREENE, R. KANTETY *et al.*, 2003 Comparative DNA sequence analysis of wheat and rice genomes. Genome Res. **13:** 1818–1827.

VAN DER HOEVEN, R., C. RONNING, J. GIOVANNONI, G. MARTIN and S. TANKSLEY, 2002 Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of large expressed sequence tag collection and selective genomic sequencing. Plant Cell **14:** 1441–1456.

WENG, Y., N. A. TULEEN and G. E. HART, 2000 Extended physical maps and a consensus physical map of the homoeologous group-6 chromosomes of wheat (*Triticum aestivum* L. em. Thell.). Theor. Appl. Genet. **100:** 519–527.

WERNER, J. E., T. R. ENDO and B. S. GILL, 1992 Toward a cytogenetically based physical map of the wheat genome. Proc. Natl. Acad. Sci. USA **89:** 11307–11311.

WU, J., T. MAEHARA, T. SHIMOKAWA, S. YAMAMOTO, C. HARADA *et al.*, 2002 A comprehensive rice transcript map containing 6591 expressed sequence tag sites. Plant Cell **14:** 525–535.

YAN, L., A. LOUKOIANOV, G. TRANQUILLI, M. HELGUERA, T. FAHIMA *et al.*, 2003 Positional cloning of the wheat vernalization gene *VRN1*. Proc. Natl. Acad. Sci. USA **100:** 6263–6268.

YU, J., S. HU, J. WANG, G. K. WONG, S. LI *et al.*, 2002 A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Indica*). Science **296:** 79–92.

YUAN, Y., P. J. SANMIGUEL and J. L. BENNETZEN, 2002 Methylation-spanning linker libraries link gene-rich regions and identify epigenetic boundaries in *Zea mays*. Genome Res. **13:** 1345–1349.

Communicating editor: J. P. GUSTAFSON