Identification and High-Density Mapping of Gene-Rich Regions in Chromosome Group 5 of Wheat

Kulvinder S. Gill,* Bikram S. Gill,* Takashi R. Endo[†] and Elena V. Boyko*

*Wheat Genetics Resource Center and Department of Plant Pathology, Kansas State University, Manhattan, Kansas 66506 and [†]Laboratory of Genetics, Faculty of Agriculture, Kyoto University, Kyoto, 606 Japan

> Manuscript received August 17, 1995 Accepted for publication March 13, 1996

ABSTRACT

The distribution of genes and recombination in the wheat genome was studied by comparing physical maps with the genetic linkage maps. The physical maps were generated by mapping 80 DNA and two phenotypic markers on an array of 65 deletion lines for homoeologous group 5 chromosomes. The genetic maps were constructed for chromosome 5B in wheat and 5D in Triticum tauschii. No marker mapped in the proximal 20% chromosome region surrounding the centromere. More than 60% of the long arm markers were present in three major clusters that physically encompassed <18% of the arm. Because 48% of the markers were cDNA clones and the distributions of the cDNA and genomic clones were similar, the marker distribution may represent the distribution of genes. The gene clusters were identified and allocated to very small chromosome regions because of a higher number of deletions in their surrounding regions. The recombination was suppressed in the centromeric regions and mainly occurred in the gene-rich regions. The bp/cM estimates varied from 118 kb for gene-rich regions to 22 Mb for gene-poor regions. The wheat genes present in these clusters are, therefore, amenable to molecular manipulations parallel to the plants with smaller genomes like rice.

THE bread wheat (Triticum aestivum L. em. Thell., $1 2n = 6 \times = 42)$ possesses a large genome (16 billion bp per haploid genome), which is about six times the size of maize and 35 times that of rice (BENNETT and SMITH 1976). The three crop plants most probably originated from a common ancestor ~ 60 million years ago (BENNETZEN and FREELING 1993). Besides polyploidy in wheat, a key step in the evolution of these three crops was differential amplification of DNA, to a greater extent in wheat than in maize or rice. The amount of actively transcribing DNA is probably not much different among the three genomes. The genes in wheat may be present in uninterrupted clusters, individually interspersed by repetitive DNA blocks, or in a combination of the two arrangements. The distribution pattern of genes will greatly influence the choice and success of techniques for molecular manipulation of the genome. Map-based cloning of the genes individually interspersed by noncoding repetitive DNA would be difficult. Conversely, clustered genes will be amenable to such manipulations, parallel to the plants with smaller genomes like rice.

We proposed a mapping strategy to target gene-rich regions of the wheat genome (GILL and GILL 1994b). Briefly, single-break deletion lines are used to divide each wheat chromosome into small regions marked by chromosome bands, protein and DNA markers. The resultant physical maps are then compared with the corresponding genetic linkage maps to analyze the physical distribution of recombination and order of markers within each chromosome region. The genetic mapping may be performed in wheat or in any of its relative species. A physical map is compared with a genetic linkage map by drawing lines to join the common markers. The resultant composite map is called a cytogenetic ladder map (CLM) (GILL and GILL 1994b). The previously constructed CLMs have established that the distributions of markers and recombination are uneven along the wheat chromosomes (WERNER et al. 1992; GILL et al. 1993a; KOTA et al. 1993; HOHMANN et al. 1994; DELANEY et al. 1995a,b; MICKELSON-YOUNG et al. 1995). In the present study, we illustrate the effectiveness of the CLM strategy to identify and preferentially map gene-rich regions in homoeologous group 5 chromosomes of wheat.

MATERIALS AND METHODS

Genetic stocks: Genetic markers were physically mapped using 65 deletion lines for wheat group 5 chromosomes (5A, 5B, and 5D). Twenty of these deletion lines were for the short arms and 45 were for the long arms. There were eight deletion lines each for the short arm of chromosomes 5A and 5B, and four for 5D. For the long arm, there were 22, 14 and nine deletion lines for chromosomes 5A, 5B, and 5D, respectively. The deletion lines were generated using the gametocidal chromosome of *Aegilops cylindrica* (ENDO 1988; ENDO and GILL 1996). Nullisomic-tetrasomic (NT) lines (missing a pair of chromosomes, the deficiency of which is compensated by a pair of homoeologous chromosomes) and ditelosomic lines

Corresponding author: Kulvinder S. Gill, Wheat Genetics Resource Center, Department of Plant Pathology, 4307 Throckmorton Hall, Kansas State University, Manhattan, KS 66506. E-mail: ksg@ksu.ksu.edu

K. S. Gill et al.

TABLE 1

The clones used for genetic and physical mapping of wheat homoeologous group 5 chromosomes

	oDNA /	Total no		No. of bands			Chromosome location	
Clone ^a	genomic	of bands	Enzyme	5A	5B	5D	T. tauschii	Wheat
þgsþ	С	3	EcoRI	1	1	1	55	
pTa71	С	8	EcoRI	0	0	1	5	1BS, 5DS
dhn2	С	12	EcoRI	2	1	1	5	5L
pHvabc309	С	5	HindIII	0	1	1	5	5L
pHvabg705	G	3	HindIII	1	1	1		5S
pHvcnlbcd157	С	4	HindIII	1	2	1	5	5L
pHvcnlbcd204	С	6	EcoRI	2	1	2	5	5L
pHvcnlbcd351	С	7	EcoRI	2	2	2	5	5L
pHbcnlbcd450	C	3	HindIII	1	1	1	5	5L
pHvcnlbcd508	C	10	HindIII	1	1	1	<i>I</i> , 5	5L, 1L
pHvcnlbcd1087	C	2	HindIII	0	0	1	2 76	2L 51 04 748 7DI 7DI
pAscnicao57	C C	1	Hindill	1	1	2	75 5	<i>JL, ZA, TAS, TBL, TDL</i>
pAscnica0215	C C	3 19	HindIII	1	1	2	19 21 21 49 5 6	JL 51 185 1D5 2D1
pAscrucu0300	C	15		1	1	9	13, 2L, JL, 43, J, 0 5	51 11
pAschicdo412	C	3	HindIII	1	1	1	5	51
hAscolicdo484	C	4	EcoRI	0	1	1)	51. 41. 21.
hAsculcdo677	Č	3	EcoRI	1	î	î	5	5S
pAscnlcdo687	č	8	HindIII	Ô	î	1	75	5S. 7L
pAscnlcdo786	Č	4	HindIII	Ŏ	0	1	75	5L
pAscnlcdo1049	Č	3	EcoRI	1	1	1	5	5L
pAscnlcdo1312	С	4	EcoRI	1	0	0	4, 5	5AL, 4BL, 4DL
pAscnlcdo1333	С	6	EcoRI	1	0	0	4, 5	5AL, 4BL, 4DL
pAscnlcdo1335	С	4	HindIII	1	1	1	5	55
pHvksu8	С	12	EcoRI	1	6	2	5	5L
pHvksu24	С	8	EcoRI	3	2	2	5	5L
pHvksu26	С	3	HindIII	1	1	1	5	5L
pHvksu58	С	5	HindIII	2	1	2	5	5L
pTtksua3	G	7	EcoRI	2	1	2	5	5L
pTtksud16	G	4	HindIII	1	1	1	5	5L
pTtksud30	G	7	EcoRI	2	2	2	5	5L
pTtksud42	G	3	HindIII	1	1	1		5L
pTtksuf1	G	5	EcoRI	1	I	1) 5 7	
pTtksug/	G	13	Hindill E. DI	1	0	2), / 5.71	JAL, JDL, IL 5DI 741
p1tksug12	G	0	ECOKI	1	1	1), /L 5	JDL, TAL 51
pItrsug14	C	3 9	LindIII	1	1	1	5	51
pTtRsug44	C	2 8	HindIII	1	1	1	5 2	51
pTiksug)7 hTtksug60	G	5	HindIII	9	1	Ô	<i>)</i> , <i>2</i> 5	55
hTtksub8	G	15	HindIII	ō	î	ŏ	2. 3. 5. 7	5BS. 7L
pTtksui26	Ğ	7	EcoRI	ľ	1	2	5	58
bTtksum2	Ğ	4	BamH1	1	1	2	5	5L
pTtksus1	G	4	HindIII	1	1	1	5	5L
pTamwg522	G	4	EcoRI	1	2	1	5	5L
pTamwg602	G	3	<i>Hin</i> dIII	0	1	1	5	5L
pTapsr79	С	5	EcoRI	1	1	1	5	5L
pTapsr115	С	7	EcoRI	0	2	1	5	4AL, 5BL, 5DL
pTapsr118	С	5	HindIII	1	1	1	5	5S
pTapsr120	C	8	EcoRI	2	2	3) F)L 51
pTapsr128	C	3	EcoRI	1	1	1	ן ד	フL 51
pTapsr145	C	12	E.coKI	3 1	4	4 0	ر ع	58 3I
pTapsr170	C C	8	ECOKI EcoDI	1 9	0	1	2 5	57, 5L
plapsr3/U		4	ECORI	4	1 1	1	5	4AL 5BL 5DL
praps 200 hTabs 628	C	5 4	EcoRI	1	1	1	5	5L
pTapsr020 hTapsr637	č	3	EcoRI	1	1	1	5	5L
pTag165	Ğ	3	HindIII	0	1	2	5	5

Continued									
Chromosome location									
Wheat									
2DL, 5BL									
5L									
5, 6A									
55									
4A, 4B, 5BL, 5DL									
5L, 6B, 6D									
5L									
5L.									
5, 7AS									
5L									
5, 7B									
4BL, 4DL, 5AL									
5S									
5L, 6L, 7L									
5L									
5L									
5L									
5L									
5L									
5L, 3L									
5L									
5L, 2BS, 7BS									

"Total no. of bands" and "No. of bands for 5A, 5B, and 5D" are from mapping on nulli-tetra lines of wheat cultivar Chinese Spring. The type of probe is indicated as C for the cDNA clones and G for the genomic. The restriction enzyme used for genomic DNA digestion of the aneuploid lines is mentioned under "Enzyme." The probes shown boldface were mapped on the 5B population of RSLs.

^a bcd, cdo, and wg were barley cDNA, oats cDNA, and wheat genomic clones, respectively, from Dr. MARK SORRELLS; mwg were barley genomic from Dr. A. GRANER; psr were wheat cDNA from Dr. MIKE GALE; Tag were wheat genomic from Dr. KOICHIRO TSUNEWAKI; Hvksu and Ttksu were barley cDNA and T. tauschii PstI genomic clones, respectively, from our laboratory.

(SEARS 1954) were used to assign DNA restriction fragment bands to their respective chromosome arms. The populations used for linkage analyses are described in the later sections. All the deletion lines, genetic and aneuploid stocks are maintained at the Wheat Genetics Resource Center (WGRC), Throckmorton Hall, Kansas State University, Manhattan, KS.

Arm ratio/arm fraction length (FL) measurements: The mean of 10 chromosomes was used for all measurements of the deletion lines. Standard errors were calculated for comparison of the measurements. Arm ratios were calculated by dividing the short arm measurements with that of the long arm. The breakpoint FL values of the short arm deletion lines were calculated by dividing the arm ratio of the deletion chromosome with that of the normal homologue. Fraction length values for the long arm deletions were calculated by dividing the arm ratio of the short ard errors for the the deletion chromosome. The standard errors for the FL measurements ranged from ± 0.005 to ± 0.021 .

DNA probes: Eighty wheat homoeologous group 5-specific probes were used for physical and genetic mapping. Chromosome locations, the numbers of DNA fragments detected, and the sources of these probes are given in Table 1. Thirty-eight (48%) probes were cDNA clones and 42 were genomic. Sixty (75%) probes ["bcd", "cdo", and "wg" (ANDERSON et al. 1992), "Tag" (NASUDA 1992), and "psr" probes (SHARP et al. 1989)]

were only interchromosomally mapped to group 5 chromosomes using NT lines. The "dhn2" probe is a barley dehydrin gene that mapped to chromosome 5H (TIM CLOSE, personal communication). The rest of the 20 probes were genetically mapped to chromosome 5D of T. tauschii (2n = 14, DD) by restriction fragment length polymorphism (RFLP) linkage analysis (GILL et al. 1991b).

Physical mapping: Of the 65 deletion lines, 43 were homozygous (the deletion chromosome present as a pair), nine were hemizygous (only one deletion chromosome), and 13 were heterozygous (one deletion and one normal chromosome). Each group 5-specific DNA fragment band was mapped to a chromosome region flanked by breakpoints of the largest deletion possessing the fragment band and the smallest deletion lacking it. For homozygous and hemizygous deletion lines, the DNA probes were scored for the presence or absence of the DNA fragment(s). For the heterozygous deletion lines, however, the DNA fragment bands were scored for their intensities. The intensity of a DNA fragment detected in a deletion line was compared with that of Chinese Spring (CS) and/or other lines possessing the normal chromosome. For example, one or more of chromosome 5B or 5D deletion lines were used as a control for comparing intensities of 5Aspecific fragment bands. As an internal check, the intensities of noncritical fragment bands were also compared. For example, the intensities of 5A- and 5D-specific fragment bands were



chromosome in disomic condition, was selected and used for mapping.

FIGURE 1.—The crossing scheme used for generating the population of 5B RSLs. The two parents differed only for chromosome 5B and only one cycle of meiosis was allowed before the recombinant products were isolated individually as separate RSLs.

also compared while comparing the intensities of 5B-specific fragment bands. Ethidium bromide-stained picture of the gel was also used for intensity comparisons to make sure that an equal amount of DNA was loaded in the lanes being compared. If the DNA fragment band being compared was of approximately half the intensity of the CS band and the other homoeologous chromosome fragment band intensities were the same in the two lines, the probe was scored as "missing" in the deletion chromosome. For the probes that detected more than one DNA fragment band for a chromosome, the DNA fragments were scored to be nonallelic (different alleles shown by a letter at the end of probe names) if they map to different chromosome regions or if the probe detected more than two DNA fragments bigger than its own size.

Genetic mapping: The 5D RFLP genetic linkage map was developed using an F₂ population (56 plants) of a cross between two accessions of T. tauschii (TA1691 \times TA1704) (GILL et al. 1991b). All 80 probes were concurrently mapped physically and in T. tauschii, except for probes cdo484, ksud42, wg363, and wg889. These probes failed to detect polymorphism using restriction enzymes Dral, HindIII, EcoRI, and EcoRV. The mapping data were analyzed using the computer program MAPMAKER (LANDER et al. 1987). Twenty-seven of the 80 probes were genetically mapped in wheat using a chromosome 5B population of 136 recombinant substitution lines (RSLs) (marked bold in Table 1). The population was developed as described in Figure 1. The cultivar CS was crossed with a CS substitution line containing chromosome 5B from an emmer wheat (T. dicoccoides). The F_1 was crossed with mono telo 5BL to recover each recombinant chromosome individually as an RSL. For chromosome 5B, each RSL represents a gamete. The mapping data of 136 recombinant substitution lines were analyzed using MAPMAKER. The gel-blot DNA hybridization procedures were as previously described (GILL et al. 1993a).

RESULTS

Nature of chromosome breaks: The breakpoints of all 65 deletion lines are marked by arrows on the left of the chromosome drawings (Figure 2). To test the randomness of the deletion breakpoint distribution, the group 5 long arm was divided into 10 regions (FL 0.1, $0.2, \ldots, 1.0$). For each region, the number of deletion breakpoints in all three homoeologous chromosomes

were counted. Based on random distribution, 4.5 deletions are expected for each region (total 45 deletions). Eight deletions were observed in the proximal 30% of the long arm and no deletion was observed in the distal 10% (Figure 4). The expected numbers of deletions for these regions are 13.5 and 4.5, respectively. Conversely, five deletions were observed for a chromosome region from FL 0.3 to 0.4, and seven each for the FL 0.5–0.6, and FL 0.7–0.8 regions. For these three regions, deletions were observed in all three homoeologues. The three regions may be more susceptible to breaks and this property is conserved among homoeologues.

Physical maps: The physical maps of chromosomes 5A, 5B, and 5D are given in Figure 2. The 80 probes detected 209 loci on the three homoeologues (240 expected). Sixty-three loci were detected on chromosome 5A and 73 each on 5B and 5D. Fifty-one probes detected loci on all three chromosomes (marked bold in Figure 3), 18 detected loci on two, and 11 on only one chromosome. Several deletions that could not be ordered cytologically because of similar FL values were differentiated by DNA markers. For example, the FL value for deletions 5AL-10, -22, and -3 was 0.56. However, markers cdo388 and Tag644 mapped in the chromosome region distal to the breakpoints of deletions 5AL-10 and 5AL-22 but proximal to that of 5AL-3. Similarly, the submicroscopic chromosome region flanked by deletions 5AL-15 and 5AL-6 possessed nine markers. Significantly, the submicroscopic chromosome regions bracketed by deletions 5BL-14 and 5BL-9, and 5DL-8 and 5DL-9 were homoeologous to that flanked by deletions 5AL-15 and 5AL-6 (revealed by the presence of the same markers) (Figure 2).

Physical locations of two agronomically important genes were studied and both mapped to submicroscopic regions, each bracketed by two deletion breakpoints with the same FL value. The Q gene, which makes wheat heads square and free thrashing, mapped to 5AL chromosome region at FL 0.87 bracketed by deletions 5AL-7 and 5AL-23. Similarly, the Ph1, a chromosome-pairing regulator gene of wheat, mapped to a chromosome region at FL 0.55 bracketed by deletions 5BL-1 and ph1b.

Except for the telomeric region of chromosome 5AL, the location and the order of the markers were the same among the three group 5 chromosomes, suggesting that the gene synteny is conserved among wheat homoeologues. The distal region of chromosome 5AL was involved in a complex translocation involving chromosomes 4A and 7B (NARANJO *et al.* 1987). A part of 5AL region distal to FL 0.87 is homoeologous to chromosomes 4B and 4D. The markers cdo1333, cdo1312, and wg114 are wheat homoeologous group 4 probes (Figure 3, see also chromosome 5A in Figure 2). Correspondingly, group 5 probes cdo484, psr115, and psr580 did not map on 5AL because the chromosome region possessing these probes is present on 4A (Figures 2 and 3).



breakpoints of various deletions, along with their FL values, are marked by arrows on the left of the chromosome drawings. The probe loci and chromosome regions FIGURE 2.—Physical maps of chromosomes 54, 5B, and 5D. C-bands on the chromosome drawings were drawn to the scale and according to their intensities. The paneled by the deletion breakpoints are given on the right. Each centromere is shown by a constriction.



FIGURE 3.—A CLM of wheat homoeologous group 5. Consensus physical map of wheat group 5 (center) was compared with the genetic linkage maps of 5D of T. tauschii (right) and 5B of wheat (left). The common markers across maps are joined by lines. The probes marked in bold on the consensus physical map detected loci on all three chromosomes of group 5. The region marked with dotted pattern at the tip of the long arm of the consensus physical map marks the region of chromosome 5A

Consensus physical map: Because of the conservation of gene synteny, the physical mapping information of the three homoeologues can be combined to generate a higher resolution consensus physical map. The group 5 consensus physical map was constructed as follows: a hypothetical chromosome was drawn that was divided into 67 chromosome intervals flanked by 65 deletion breakpoints (Figure 3). Each probe was placed on the consensus map to the shortest possible interval. For example, the probes ksua3, wg909, wg419, ksum2 and ksud42 mapped to the chromosome regions between FL 0.16 and 0.32 in 5A, and proximal to FL 0.29 in 5D. However, in chromosome 5B the probes mapped to the region spanned by FL 0.26 and 0.29. On the consensus map, the five markers were placed between FL 0.26 and 0.29. The FL values and the size of each marked region on the consensus map were the same as that of the chromosome where the size of the region was the smallest. For example, the probes ksus1 and ksu8 mapped to chromosome region between FL 0.46 and 0.56 in 5A and 0.48 and 0.60 in 5D. In chromosome 5B and the consensus map, the two probes mapped to a submicroscopic chromosome region at FL 0.55.

Although the relative order and the FL location of most of the markers were the same among three homoeologues, a few discrepancies were observed for FL location. The markers *wg1026*, *cdo400*, *psr637*, *ksu57*, *ksu24*, *psr120*, *Tag251*, and *wg583* mapped to submicroscopic regions at FL 0.75 in 5B and 5D. A submicroscopic region of 5A containing these markers was present at FL 0.67. On the consensus map, the region containing these markers was placed at FL 0.75.

Distribution of markers: The distribution of the markers was highly uneven along the chromosome length (Figures 3 and 4). Uneven distribution of the markers was more apparent from the consensus map than was from the individual physical maps. With the centromere as the reference point, more markers were present in the distal region compared to the proximal. No marker was observed in the chromosome region surrounding the centromere (20% of the chromosome). About 46% (32/69) of the long arm markers were present in the distal 25% of the arm. This general marker distribution pattern was superimposed by the presence of submicroscopic regions of high probe density. On the long arm, three marker clusters were obvious, surrounding FL 0.30, 0.55, and 0.75 (marked on Figure 3). Forty-two of the 69 (61%) long arm markers mapped to the three regions that encompass <18% of the arm. Ten markers mapped in the submicroscopic region at FL 0.75 and eight in the one at FL 0.55. Three other markers (cdo388a, Tag644, and bcd508a) also



FIGURE 4.—The distribution patterns of markers and deletion breakpoints on the long arm of wheat homoeologous group 5 chromosome. The long arm was divided into 20 regions, each equivalent to FL 0.05. The number of deletions, with their breakpoints in each region, were counted from chromosomes 5A, 5B, and 5D. The number of markers was counted from the consensus map. \blacksquare , number of markers; \blacksquare , the number of deletions.

mapped in submicroscopic chromosome regions. Thus, 30% of the long arm markers mapped in the submicroscopic chromosomal regions.

Genetic linkage maps: Figure 3 shows RFLP-based genetic linkage maps of chromosome 5D of T. tauschii (right) and 5B of wheat (left). All 80 markers (mapped physically) were also mapped in T. tauschii, except for cdo484, ksud42, wg363, and wg889, which did not detect polymorphism between the two parents of the population with restriction enzymes DraI, HindIII, EcoRI, and EcoRV. Fifty-six of the 80 probes were also tested for polymorphism on CS and the 5B substitution line (the two parents of the 5B mapping population). Thirtyseven (66%) probes detected polymorphic fragments between the two parents with one or more of the restriction enzymes DraI, HindIII, EcoRI, EcoRV, BamHI, KpnI, and, XhoI. The 5B genetic linkage map consists 27 of the 37 probes along with two C-band markers (Figure 3). The order of the probes is identical between the 5Band the 5D maps. Relative genetic distances among the markers were similar between the two maps, however, the overall recombination was reduced in the 5B map compared to the 5D map.

Cytogenetic ladder map (CLM): Both *5B* and *5D* genetic linkage maps gave a misleading appearance that

involved in translocation with chromosome 4A. The group 5 probes present on chromosomes 4A, 5B, and 5D are shown on the right of the region, whereas, the group 4 probes present on chromosome 5A, 4B, and 4D are shown on the left. The probes marked with * mapped to different chromosomal regions between wheat and T. tauschii. Centromeres are marked by C's. The two C-band markers on the 5B linkage map are 5BL1.42 and 5BL2.62.



FIGURE 5.—The distribution patterns of markers and recombination on the long arm of wheat homoeologous group 5 chromosome. The long arm was divided into 20 regions, each equivalent to FL 0.05. The number of markers were counted from the consensus map as in Figure 4. The recombination for each region was estimated from the 5D genetic linkage map. The recombination between markers mapping in adjoining regions was divided equally between the two regions. \blacksquare , number of markers; \blacksquare , recombination in 10 cM.

markers are concentrated around the centromeres. The physical maps clearly showed that centromeric regions are devoid of markers that are distally located in clusters. However, physical maps alone cannot reveal genetic distances and marker order within each region. It is only when information on the physical and genetic maps is combined in a CLM (GILL and GILL 1994b) that the precise distribution of recombination and genes in specific regions become apparent.

Figure 3 shows a CLM of wheat group 5 chromosomes constructed by comparing the consensus physical map with the genetic linkage maps of 5B and 5D. The recombination distribution pattern coincided with that of the markers (Figure 5). The recombination was suppressed in the marker-poor chromosome region surrounding the centromere. Contrarily, 58% of the total long arm recombination occurred in the distal 25%. Most of the long arm recombination occurred in the regions around the three marker clusters (Figure 5). The genetic length of the submicroscopic region at FL 0.55 was 44 cM and at FL 0.75 was 114 cM. Similarly for the short arm, recombination was very low in the proximal 20% compared to the distal region where it was very high (Figures 3 and 5). The markers gsp, abg705, psr170, and Tag317 physically mapped in the distal 25% of the short arm. The four loci were not even linked to the 5D linkage map.

DISCUSSION

Uniqueness of the gametocidal system used: It is well established that some chromosomes from *Aegilops spel-*

toides, A. sharonensis, A. longissima, A. cylindrica, or Secale cereale, when present in wheat background, cause chromosome breakages (TSUJIMOTO and TSUNEWAKI 1985; ENDO 1988; KOTA and DVORAK 1988; LUKASZEWSKI 1995). For the purpose of deletion stock isolation, A. cylindrica-based system is the most ideal among them. Chromosome breaks occur in the gametes lacking the gametocidal chromosome of A. cylindrica, therefore, the deletions are recovered in a pure wheat background. In all the other Aegilops systems, breaks occur in the gametes carrying the gametocidal gene. Therefore, it is difficult to recover deletions in a pure wheat background. Second, the complex cytological aberrations or rearrangements that are frequent in the presence of A. speltoides or S. cereale chromosomes (KOTA and DVO-RAK 1988; LUKASZEWSKI 1995) were not observed in our deletion lines. More than 400 deletion lines for all 21 wheat chromosomes have been isolated and characterized using C-banding and RFLP markers (WERNER et al. 1992; GILL et al. 1993a,b; KOTA et al. 1993; MICKELSON-YOUNG et al. 1995; ENDO and GILL 1996). Almost all deletions resulted from single breaks followed by the healing of telomeres and subsequent loss of the regions distal to each break. It may partly be due to the fact that care was taken to select for deletions caused by single breaks (ENDO and GILL 1996). Third, the A. cylindrica-based system breaks all wheat chromosomes and arms compared to the other systems where breakages occur only at certain sites on a few specific wheat chromosomes (KOTA and DVORAK 1988; LUKASZEWSKI 1995). The deletions were equally frequent both in heterochromatin and in euchromatin (GILL et al. 1993a; ENDO and GILL 1996).

Sensitivity to breaks vs. marker density: In group 5 chromosomes, breakages were semi-random and the marker-rich regions were the preferred sites. As a result, marker clusters were identified and localized to submicroscopic chromosome regions bracketed by deletion breakpoints (Figure 4). This correlated pattern of deletion breakpoints distribution with that of markers was also observed for other wheat chromosomes (WERNER et al. 1992; GILL et al. 1993a,b; KOTA et al. 1993; MICKEL-SON-YOUNG et al. 1995). The deletions always occur in marker-rich regions, although, factors other than gene density also influence breakages. For chromosome 5A, six deletions were observed at FL 0.98 of the short arm and three at FL 0.11 of the long arm. The two regions were devoid of markers. An apparent exception was for the chromosome region distal to FL 0.87. No deletion was observed in any of the three chromosomes, although at least 10 markers mapped in the region (Figures 2 and 4). This is probably because of the lack of telomeric C-band in the long arm. The deletion chromosomes were identified by the C-banding analysis. Without a diagnostic telomeric C-band, loss of the distal 13% of the arm would not be easily detected. The six deletions at FL 0.98-1.0 of chromosome 5AS were identified because of the presence of the telomeric C-band.

Apparently, the susceptibility of the marker-rich chromosome regions to the gametocidal gene is conserved among homoeologues. Among the regions that possess breakpoints of more than one deletion at the same FL, the corresponding deletions in homoeologues were detected only for the marker-rich regions. The marker cluster at FL 0.75 is the best example for which deletions bracketing the submicroscopic region were recovered for all three homoeologues. One reason could be that loosely packed transcriptionally active chromosome regions are more accessible to breaks compared to tightly packed inactive regions. Moreover, gene-rich regions are also high in recombination (discussed later) that may also be related to high breakage frequency. Irrespective of the reason for the differential sensitivity to chromosome breaks, it would be interesting to see the extent of their conservation across the grass family.

Conversely, deletion clusters in marker-poor regions were not conserved across homoeologues. Three deletions were recovered at FL 0.11 of 5AL, but none was observed in the corresponding regions of 5BL or 5DL. The deletion cluster at FL 0.98 in 5AS is another example. No deletion was observed in the corresponding regions of 5BS or 5DS. Both these chromosome arms have telomeric C-bands. Deletions around FL 0.98, if occurred, would have been detected. Similarly, the breakage hotspot at FL 0.42 of 5BS was not observed at the corresponding regions of 5AS or 5DS. In these cases, marker density cannot be the cause of high breakages. Except for 5AS region at FL 0.98, other regions lie in recombination-poor areas and rule out the possible role of recombination in breakage frequency. Besides, the distribution pattern of recombination is conserved among homoeologues (discussed later). Thus, in certain cases chromosome-specific intrinsic elements may be related to high breakage frequency.

Distribution of genes: Of the 80 markers used in the present study, 39 were cDNA clones, 29 were PstI genomic, and 12 were random genomic clones. No obvious difference was observed between the distribution patterns of cDNA and the genomic clones (Figure 6). Many PstI genomic clones probably also represent genes because PstI libraries are enriched with undermethylated actively transcribing regions of the genome (BURR et al. 1988, for review). The distribution pattern of the markers may, therefore, reflect that of wheat genes. It has been suggested that the highly conserved genes (housekeeping genes) may be present in recombination-poor regions of the genome (around centromeres) and less conserved genes map in the recombinationrich regions (AISSANI and BERNARDI 1991). If this hypothesis is true for wheat, the probes that do not detect polymorphism among various accession/parents should map in the centromeric regions, whereas, the probes that detect polymorphism should map toward telomeres. Sixty probes (75%) used in the present study were only interchromosomally mapped to group 5 us-



FIGURE 6.—Distribution of cDNA *vs.* genomic clones on the long arm of wheat homoeologous group *5*. The long arm was divided into 20 equal regions as in Figures 4 and 5. The number of cDNA or genomic clones was counted from the consensus map. ■, the cDNA clones; □, the genomic clones.

ing nulli-tetra lines and were not selected for their ability to detect polymorphism. The distribution pattern of *ksud42, bcd450, wg363,* and *wg889* that did not detect polymorphism in *T. tauschii* was not unusual; *ksud42* and *bcd450* mapped in the major gene clusters, *wg889* in an interstitial chromosome region and *wg363* in the short arm along with three other markers. The markers were from 10 different libraries from wheat, barley, oats, or *T.tauschii*. Therefore, the probes used in the present study represent a random sample of wheat genes.

The results from the present study lead to the conclusion that genes in wheat are present in clusters interspersed by blocks of repetitive sequences visualized as regions of low gene density. Three major gene clusters were observed in the long arm interspersed in genepoor regions (Figures 3 and 4). The precision with which the gene-rich and gene-poor blocks can be differentiated, depends upon the number of deletions for the chromosome. The uneven distribution of markers in the present study was not as obvious from the individual maps as it was from the consensus map where more deletion data points were available. These gene clusters may be localized to even smaller chromosome regions as more deletions become available. Eleven markers mapped in the region just distal to the submicroscopic region at FL 0.75 with 10 markers. Some of these adjacent markers may turn out to be part of the FL 0.75 cluster with the availability of fine physical mapping tools.

Gene distribution has been studied in animal systems and it seems that all animal genomes and chromosomes, to some degree, are divided into gene-rich and gene-poor compartments (SUMNER *et al.* 1993, for review). Gene distribution in the human genome was studied by several methods including *in situ* hybridization of a random pool of mRNA (YUNIS *et al.* 1977), *in situ* of isochore H3 (SACCONE *et al.* 1992), which comprises the highest gene concentration of any fraction of DNA (BERNARDI 1989; MOUCHIROUD *et al.* 1991), or by DNAase hypersensitivity (WEINTRAUB and GROUDINE 1976; ELGIN 1988). The unanimous conclusion was that the genes in the human genome are localized in Rbands and are more concentrated in the terminal Rbands called T-bands (SUMNER *et al.* 1993). In general, more R-bands are present toward the telomeric regions compared to the proximal regions. The gene distribution in wheat is similar to that in humans, although Rbands are not detected in wheat.

Differences in gene densities among chromosomal regions were observed in yeast, although, large regions devoid of genes were not detected (DUJON et al. 1994). Compared to that of wheat or humans, the yeast genome is very small. Correspondingly, the proportion of repetitive DNA in the yeast genome is also small. The differentiation of the genome into gene-rich and genepoor compartments is probably universal to all eukaryotes. The size of the gene-poor compartments and the extent of unevenness in the gene distribution depend upon the size of the genome. In smaller genomes of yeast or Arabidopsis, the gene distribution would appear more uniform compared to that of a large genome of wheat where it is highly uneven. The gene density in gene-rich compartments is probably comparable among various eukaryotes.

Distribution of recombination: Recombination in the telomeric region (distal 25%) of the group 5 long arm was more than 20 times the recombination in the centromeric region (proximal 25%) (Figures 3 and 5). These data support previous reports of uneven distribution of recombination in wheat (DVORAK and CHEN 1984; CURTIS and LUKASZEWSKI 1991; WERNER et al. 1992; GILL et al. 1993a; KOTA et al. 1993), animal systems (STEINMETZ et al. 1987; BOLLAG et al. 1989), and in other plants (RICK 1971; GANAL et al. 1989). Proximity to centromeres tends to suppress crossing over in other organisms also (GAUDET and FITZGERALD-HAYES 1990). In tomato, reduced recombination in the centromeric regions was attributed to the presence of heterochromatin (RICK 1971). In wheat, recombination in the proximal 20% of all wheat chromosomes is suppressed, even though these region may vary 10- to 20-fold in the amount of heterochromatin.

Earlier conclusions of suppressed recombination around centromeres and increased recombination toward telomeres in wheat were based on either a few Cbands (DVORAK and CHEN 1984; CURTIS and LUKASZEW-SKI 1991) or low density CLMs (WERNER *et al.* 1992; GILL *et al.* 1993a; KOTA *et al.* 1993). The high density maps presented here reveal that the distribution of recombination may depend upon marker density rather than on the chromosome location. Figure 5 shows the comparison of the distribution of markers with that of recombination. The recombination is the highest at the marker clusters. In fact, recombination around the marker clusters accounts for most of the long arm genetic length. Recombination in the largest marker cluster (at FL 0.75) was the highest among the three clusters.

It is apparent that the structure of each chromosome region determines the extent of recombination as the relative genetic distances between markers were similar in the genetic linkage maps of polyploid and diploid wheat (Figure 3). It implies that this chromosomal organization is conserved among homoeologues.

The correlation between marker density and recombination suggests that recombination occurs only in the gene-rich regions. However, if recombination does occur in the gene-poor regions also, it would not be fully detected because of the lack of markers. A positive correlation was observed between the number of markers and observed recombination in wheat as well as in other crop plants (NILSSON et al. 1993, for review). Previously undetected recombination events were revealed in maize, barley (NILSSON et al. 1993), rice, and wheat by mapping more markers. In rice an RFLP genetic linkage map with 1383 markers (KURATA et al. 1994) is 30% longer than the map with 600 markers (TANKSLEY et al. 1993). Based on the data on two to four polymorphic C-bands, an average of about two crossovers/arm were observed for chromosome 5BL of wheat (DVORAK and CHEN 1984; CURTIS and LUKASZEWSKI 1991). However, when 36 markers were mapped on the arm, the average number of crossovers increased to about four (GILL and GILL 1994a; the present study). As many as five crossovers/arm were observed during a single cycle of meiosis. Double crossovers spanning <10% of the arm were frequent. The gene-poor region block may be as large as 25% of the arm. Double crossovers occurring in these blocks would not be detected.

Because of the nonrandom distribution of recombination along the chromosome length, bp/cM would differ among regions. The wheat CLMs reveal fairly precise estimates of bp/cM especially for the gene-rich regions. The current study shows that bp/cM estimates for wheat chromosomes may vary from 118 kb in generich regions to 22 Mb in gene-poor regions. Our preliminary results from the long-range mapping experiments indicate that one cM for the gene-rich cluster at FL 0.55 may translate into <54 kb (K. S. GILL and B. S. GILL, unpublished data). The above bp/cM estimates are comparable to that of similar regions in other crop plants. Based on the genome size and RFLP linkage map length, one cM in tomato is estimated to be 550 kb (GANAL et al. 1989). However, the estimates for regions around various disease-resistance genes ranged from 43 to 90 kb (GANAL et al. 1989; SEGAL et al. 1992; ZHANG et al. 1994). Similarly, rice contiguous maps have revealed that 1 cM may vary from 120 kb to 1 Mb compared to

the predicted 273 kb (UMEHARA *et al.* 1995). The bp/ cM estimates for the regions around the genes are comparable among various crop plants and may be anywhere from 40 to 120 kb/cM. However, the upper limit for such estimates depends upon the genome size. The upper limit for bp/cM estimates is determined by the size of gene-poor compartments. In rice the upper limit is 1 Mb compared to 22 Mb in wheat. The wheat genome is ~35 times larger than that of rice.

Conclusions: Because of its large size (16 billion bp), the wheat genome is commonly believed not to be suitable for molecular manipulations. Our results showed that a high proportion of actively transcribing wheat genes are present in clusters. The commonly used RFLP genetic analysis fails to detect these gene clusters; therefore, they cannot be efficiently used for molecular manipulations. The CLM strategy not only efficiently identifies the gene clusters but also preferentially maps them. The gametocidal system of A. cylindrica frequently causes breaks in the gene cluster regions and is therefore an aid in the identification and localization of gene clusters to small chromosomal regions. The gene clusters are also preferred sites for recombination and marker order can be resolved genetically. Furthermore, agronomically important genes are present in these clusters. For group 5, we mapped two such genes (Ph1 and Q) using the deletion lines and both mapped in the gene clusters. Two other genes conferring resistance to wheat against environmental stress (cold tolerance) and response to vernalization (Vrn1) (GALIBA et al. 1995) are flanked by two probes mapping in the gene cluster at FL 0.75. This appears to be a general pattern of distribution for most of the agronomically useful genes in wheat. Therefore, most wheat genes are amenable to map-based cloning similar to other plants with smaller genomes. Many wheat genes, especially the ones controlling the basic physiological functions, are probably similar to their counterparts in small genome crops like rice. However, some important genes are probably wheat-specific and must be cloned from wheat that is the only source of their genetic resources. According to our analysis, cloning would be feasible for the genes that are present in gene-cluster regions.

This is contribution no. 96-27-J from the Kansas Agricultural Experiments station, Kansas State University, Manhattan, KS 66506. The research was supported by U.S. Department of Agriculture-National Research Initiative Plant Genome Award and U.S.-Japan National Science Foundation-Japan Society of Plant Science collaborative research program. We thank Mr. DUANE WILSON for greenhouse help.

LITERATURE CITED

AISSANI, B., and G. BERNARDI, 1991 CpG islands: features and distribution in the genomes of vertebrates. Gene **106**: 173–183.

- 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.
- BENNETT, M. D., and J. B. SMITH, 1976 Nuclear DNA amounts in angiosperms. Philos. Trans. R. Soc. London, Ser. B 274: 227-274.

- BENNETZEN, J. L., and M. FREELING, 1993 Grasses as a single genetic system. Trends Genet. 9: 259-261.
- BERNARDI, G., 1989 The isochore organization of the human genome. Ann. Rev. Genet. 23: 637-661.
- BOLLAG, R. J., A. S. WALDMAN and R. M. LISKAY, 1989 Homologous recombination in mammalian cells. Annu. Rev. Genet. 23: 199-225.
- BURR, B., R. A. BURR, K. H. THOMPSON, M. C. ALBERTSEN and C. W. STUBER, 1988 Gene mapping with recombinant inbreads in maize. Genetics 118: 519–526.
- CURTIS, C. A., and A. J. LUKASZEWSKI, 1991 Genetic linkage between C-bands and storage protein genes in chromosome 1B of tetraploid wheat. Theor. Appl. Genet. 81: 245–252.
- DELANEY, D., 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., 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.
- DUJON, B., D. ALEXANDRAKI, B. ANDRE, W. ANSORGE, V. BALADRON et al., 1994 Complete DNA sequence of yeast chromosome XI. Nature 369: 371-378.
- DVORAK, 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.
- ELGIN, S. R., 1988 The formation and function of DNAase I hypersensitive sites in the process of gene activation. J. Biol. Chem. 263: 19259-19262.
- 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., and B. S. GILL, 1996 The deletion stocks of common wheat. J. Hered. (in press).
- GALIBA, G., S. A. QUARRIE, J. SUTKA, A. MORGUNOV and J. W. SNAPE, 1995 RFLP mapping of the vernalization (Vrn1) and frost resistance (Fr1) genes on chromosome 5A of wheat. Theor. Appl. Genet. 90: 1174–1179.
- GANAL, M. W., N. D. YOUNG and S. D. TANKSLEY, 1989 Pulsed field gel electrophoresis and physical mapping of large DNA fragments in the *Tm-2a* region of chromosome 9 in tomato. Mol. Gen. Genet. **215**: 395–400.
- GAUDET, A. M., and M. FITZGERALD-HAYES, 1990 The function of centromeres in chromosome segregation, pp. 118–131 in *The Eukaryotic Nucleus. Molecular Biochemistry and Macromolecular Assemblies*, edited by P. R. STRAUSS and S. H. WILSON. Telford Press, Caldwell, NJ.
- GILL, B. S., B. FRIEBE and T. R. ENDO, 1991a Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). Genome 34: 830–839.
- GILL, K. S., and B. S. GILL, 1994a Higher level of recombination in wheat is due to more cross-overs/arm and localized recombination hotspots, p. 64 in *Plant Genome II. The Second International Conference on the Plant Genome*, San Diego.
- GILL, K. S., and B. S. GILL, 1994b Mapping in the realm of polyploidy: the wheat model. BioEssays 16: 841-846.
- GILL, K. S., E. L. LUBBERS, B. S. GILL, W. J. RAUPP and T. S. COX, 1991b 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, 1993a A chromosome regionspecific mapping strategy reveals gene-rich telomeric ends in wheat. Chromosoma 102: 374–381.
- GILL, K. S., B. S. GILL, T. R. ENDO and Y. MUKAI, 1993b Fine physical mapping of *Ph1*, a chromosome pairing regulator gene in polyploid wheat. Genetics **134**: 1231–1236.
- 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.
- KOTA, R. S., and J. DVORAK, 1988 Genomic instability in wheat induced by chromosome 6Bs of *Triticum speltoides*. Genetics 120: 1085-1094.
- 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.

- KURATA, N., Y. NAGAMURA, K. YAMAMOTO, Y. HARUSHIMA, N. SUE et al., 1994 A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genet. 8: 365-372.
- LANDER, E. S., P. GREEN, J. ABRAHAMSON, A. BARLOW, M. J. DALY et al., 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181.
- LUKASZEWSKI, A. J., 1995 Chromatid and chromosome type breakage-fusion-bridge cycles in wheat (*Triticum aestivum* L.). Genetics 140: 1069-1085.
- 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.
- MOUCHIROUD, D., G. D'ONOFRIO, B. AISSANI, G. MACAYA, C. GAUTIER et al. 1991 The distribution of genes in the human genome. Gene 100: 181–187.
- NARANJO, T., A. ROCA, P. G. GOICOECHEA and R. GIRALDEZ, 1987 Arm homoeology of wheat and rye chromosomes. Genome **29**: 873–882.
- NASUDA, S., 1992 Molecular genetical studies of common wheat and its relatives—telosomic analysis of the RFLP loci in common wheat and a study on chloroplast genome diversity in *Aegilops caudata* L., M.Sc. Thesis. Kyoto University, Japan.
- NILSSON, N.-O., T. SALL and B. O. BENGTSSON, 1993 Chiasma and recombination data in plants: are they compatible. Trends Genet. 9: 344-348.
- RICK, C. M., 1971 Some cytological features of the genome in diploid plant species. Stadler Genet. Symp. 1,2: 153-174.
- SACCONE, S., A. D. SARIO, G. D. VALLE and G. BERNARDI, 1992 The highest gene concentrations in the human genome are in telomere bands of metaphase chromosomes. Proc. Natl. Acad. Sci. USA 89: 4913-4917.
- SEARS, E. R., 1954 The aneuploids of common wheat. Research Bulletin 572 University of Missouri Agricultural Experiment Station.
- SEGAL, G., M. SARFATTI, M. A. SCHAFFER, N. ORI, D. ZAMIR et al., 1992 Correlation of genetic and physical structure in the region sur-

rounding the 12 Fusarium oxysporum resistance locus in tomato. Mol. Gen. Genet. 231: 179-185.

- SHARP, P. J., S. CHAO, S. DESAI and M. D. GALE, 1989 The isolation, characterization and application in the Triticeae of a set of wheat RFLP probes identifying each homoeologous chromosome arm. Theor. Appl. Genet. 78: 342–348.
- STEINMETZ, M., Y. UEMATSU and L. K. FISCHER, 1987 Hotspots of homologous recombination in mammalian genomes. Trends Genet. 3: 7-10.
- SUMNER, A. T., J. D. L. TORRE and L. STUPPIA, 1993 The distribution of genes on chromosomes: a cytological approach. J. Mol. Evol. 37: 117-122.
- TANKSLEY, S. D., T. M. FULTON and S. R. MCCOUCH, 1993 Linkage map of rice (*Oryza sativa*) (2n = 24), pp. 6.61–6.76 in *Genetic Maps*, edited by S. J. O'BRIEN. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- TSUJIMOTO, H., and K. TSUNEWAKI, 1985 Hybrid dysgenesis in common wheat caused by gametocidal genes. Jpn. J. Genet. 60: 565-578.
- UMEHARA, Y., A. INAGAKI, H. TANOUE, Y. YASUKOCHI, Y. NAGAMURA *et al.*, 1995 Construction and characterization of rice YAC library for physical mapping. Mol. Breed. *1*: 79–89.
- WEINTRAUB, H., and M. GROUDINE, 1976 Chromosomal subunits in active genes have an altered conformation. Science 193: 848-856.
- WERNER, J. E., T. R. ENDO and B. S. GILL, 1992 Towards a cytogenetically based physical map of the wheat genome. Proc. Natl. Acad. Sci. USA 89: 11307–11311.
- YUNIS, J. J., M. T. KUO and G. F. SAUNDERS, 1977 Localization of sequences specifying messenger RNA to light-staining G-bands of human chromosomes. Chromosoma 61: 335–344.
- ZHANG, H. B., G. B. MARTIN, S. D. TANKSLEY and R. A. WING, 1994 Map-based cloning in crop plants: tomato as a model system. II. Isolation and characterization of a set of overlapping yeast artificial chromosomes encompassing the jointless locus. Mol. Gen. Genet. 244: 613–621.

Communicating editor: J. A. BIRCHLER