

## Identification of Wheat Chromosomal Regions Containing Expressed Resistance Genes

Muharrem Dilbirligi, Mustafa Erayman, Devinder Sandhu,<sup>1</sup> Deepak Sidhu and Kulvinder S. Gill<sup>2</sup>

Department of Crop and Soil Science, Washington State University, Pullman, Washington 99164 and Department of Agronomy and Horticulture, University of Nebraska, Lincoln, Nebraska 68583-0915

Manuscript received August 18, 2003  
Accepted for publication September 24, 2003

### ABSTRACT

The objectives of this study were to isolate and physically localize expressed resistance (*R*) genes on wheat chromosomes. Irrespective of the host or pest type, most of the 46 cloned *R* genes from 12 plant species share a strong sequence similarity, especially for protein domains and motifs. By utilizing this structural similarity to perform modified RNA fingerprinting and data mining, we identified 184 putative expressed *R* genes of wheat. These include 87 NB/LRR types, 16 receptor-like kinases, and 13 *Pto*-like kinases. The remaining were seven *Hm1* and two *Hs1<sup>pm-1</sup>* homologs, 17 pathogenicity related, and 42 unique NB/kinases. About 76% of the expressed *R*-gene candidates were rare transcripts, including 42 novel sequences. Physical mapping of 121 candidate *R*-gene sequences using 339 deletion lines localized 310 loci to 26 chromosomal regions encompassing ~16% of the wheat genome. Five major *R*-gene clusters that spanned only ~3% of the wheat genome but contained ~47% of the candidate *R* genes were observed. Comparative mapping localized 91% (82 of 90) of the phenotypically characterized *R* genes to 18 regions where 118 of the *R*-gene sequences mapped.

WHEAT crop is attacked by a variety of pests, including insects, pathogens, viruses, and nematodes. On average, these pests cause 20–37% yield loss worldwide, translating to ~\$70 billion/year (<http://pseru.ars.usda.gov>; PIMENTAL *et al.* 1997). Genetic manipulation of resistance (*R*) genes is an efficient, economical, and well-tested method of controlling wheat pests. About 269 *R* genes conferring resistance to various pests have been identified in wheat or its wild relatives (McINTOSH *et al.* 2000). Of the 110 that are well studied, 80 *R* genes show single-gene inheritance (3:1 ratio in F<sub>2</sub>) and seem to have the gene-for-gene interaction with the pests. Inheritance of the remaining 30 genes is complex and the resistance seems to be manifested in a nonspecific manner. The wheat genome is ~16 million kb/haploid cell (ARUMUGANATHAN and EARLE 1991). About 95–99% of the genome is nontranscribing (SANDHU and GILL 2003). Most of the wheat genes are present in clusters spanning physically small regions (gene-rich regions) of varying gene densities and numbers (GILL *et al.* 1996a,b; SANDHU and GILL 2002a; M. DILBIRLIGI, M. ERAYMAN, D. SANDHU and K. S. GILL, unpublished results). The gene-rich regions are spaced by variously sized blocks of nested retrotransposons and duplicated genes (SANMIGUEL *et al.* 1996; PANSTRUGA *et al.* 1998;

WENDEL 2000; WICKER *et al.* 2001). Although it is feasible to target the gene-rich regions, nontranscribing repeated DNA interspersing these regions poses a serious hindrance to genomic manipulation of genes (FEUILLET and KELLER 1999, 2002; STEIN *et al.* 2000; WICKER *et al.* 2001; SANMIGUEL *et al.* 2002; YAN *et al.* 2003). Furthermore, in addition to the repeated DNA, most wheat genes have multiple orthologs and in some cases paralogs. Nonfunctional or diverged-function gene copies may further complicate genome manipulation. As a result, very few genes with known function have been cloned in wheat. Therefore, approaches that specifically target an expressed part of the genome are particularly desirable for cloning wheat genes.

Thus far, 46 *R* genes conferring resistance against pathogens, insects, nematodes, and viruses attacking 12 plant species have been cloned (DILBIRLIGI and GILL 2004). Most of the cloned *R* genes are structurally conserved and can be grouped into four distinct classes on the basis of the presence of one or more of the nucleotide-binding (NB) sites, receptor-like transmembrane kinase (RLK), cytoplasmic protein kinase (PK), and leucine-rich repeat (LRR) domains (HAMMOND-KOSACK and JONES 1997; MEYERS *et al.* 1999; MONDRAGON-PALOMINO *et al.* 2002). The class I, containing NB and LRR (NB/LRR), is the largest with 31 members. This group of genes can have either a N-terminal coiled coil or a tool-interleukin receptor-like (TIR) domain (MEYERS *et al.* 1999; MONDRAGON-PALOMINO *et al.* 2002). The TIR domain is found mainly in the dicots. Class II *R* genes encode proteins that carry an extracellular N-terminal LRR region and a C-terminal transmembrane domain.

<sup>1</sup>Present address: G302 Agronomy Hall, Iowa State University, Ames, IA 50011-1010.

<sup>2</sup>Corresponding author: Department of Crop and Soil Sciences, 277 Johnson Hall, P.O. Box 646420, Pullman, WA 99164.  
E-mail: ksgill@wsu.edu

Class III genes carry an extracellular LRR and a cytosolic kinase domain that are connected by a single pass transmembrane domain. Class IV genes contain only the CPK domain. Significant structural and functional differences, however, may exist among *R* genes within each class. For instance, the LRR domain present in the *Hs1<sup>pro-1</sup>* gene is intercellular compared to *Cf2*, *Cf4*, *Cf5*, *Cf9*, *Xa21*, *Ve1*, and *Ve2*, where it is extracellular (DIXON *et al.* 1996; CAI *et al.* 1997; KAWCHUK *et al.* 2001).

The motifs present in the plant *R*-gene domains are highly conserved, and the sequence and order of the conserved amino acids is often unique for the *R* genes. The LRR domain of plant *R* genes contains 9–41 imperfect repeats, each ~25 amino acids long with a consensus amino acid sequence of xx(L)x(L)xxxx (COOLEY *et al.* 2000). The PK domain of both *Pto* and *Xa21* genes contains up to 25-amino-acid-long motifs where the first three (DFG) and the last two (PE) residues are highly conserved (LIU *et al.* 2002). Internal threonine (T) and serine (S) residues are essential for autophosphorylation and thus are conserved (LIU *et al.* 2002). The kinase-1a (p-loop), kinase-2, and putative kinase-3a motifs of the NB domain of plant *R* genes have a consensus sequence of GxxGxGK(T/S)T, LxxxDDVW, and Gxxxx TxR, respectively. These sequences are considerably different from that present in other NB-encoding proteins (HAMMOND-KOSACK and JONES 1997; MEYERS *et al.* 1999).

Genomic DNA sequence analyses revealed that there are 166 putative NB/LRR-containing genes in *Arabidopsis thaliana* and ~600 in rice (*Oryza sativa*; ARABIDOPSIS GENOME INITIATIVE 2000; GOFF *et al.* 2002; RICHLIY *et al.* 2002). The number of functional or expressed *R* genes, however, is not known. Analyses of NB/LRR-containing genomic sequences from various crop plants suggested that only a small fraction of the *R* genes may be functional (CHIN *et al.* 2001; SUN *et al.* 2001; SHEN *et al.* 2002). With the objective to clone resistance genes *en masse*, primers complementary to the conserved sequences have been used to amplify genomic DNA of various crop plants (LEISTER *et al.* 1996; YU *et al.* 1996; COLLINS *et al.* 1998). Expressed sequence tagged (EST) transcripts for only 9 of the 173 resistant gene analogs (RGAs) isolated from six major crop species have been observed (M. DILBIRLIGI and K. S. GILL, unpublished results; <http://www.ncbi.nlm.nih.gov/entrez>). The proportion of expressed *R* genes is expected to be even smaller in wheat because of the larger genome size.

It is well documented that recombination is highly uneven on the wheat and other eukaryotic chromosomes. Detailed analysis of wheat homeologous group I chromosomes showed that recombination occurs mainly in gene-rich regions encompassing ~13% of the genome (SANDHU and GILL 2002b). Distribution of recombination is very similar in other wheat chromosomes (M. DILBIRLIGI, M. ERAYMAN, D. SANDHU and K. S.

GILL, unpublished results). Most of the recombination occurs in the distal 50% of the chromosomes. There may be as much as a 20-fold difference in the rate of recombination even among various gene-rich regions (SANDHU and GILL 2002a; M. DILBIRLIGI, M. ERAYMAN, D. SANDHU and K. S. GILL, unpublished results). In many plants, *R* genes have been shown to cluster on the chromosomes usually in the highly recombinogenic subtelomeric ends (BOTELLA *et al.* 1997; MEYERS *et al.* 1999; SAKAMOTO *et al.* 1999; HALTERMAN *et al.* 2001; BRUEGGEMAN *et al.* 2002; ERNST *et al.* 2002; GOFF *et al.* 2002; RICHLIY *et al.* 2002; AKHUNOV *et al.* 2003). Recombination in regions immediately around the *R* genes, however, is usually suppressed (NOEL *et al.* 1999; CHIN *et al.* 2001; SUN *et al.* 2001; SHEN *et al.* 2002). Because of the massive differences in the extent of recombination particularly around the *R* genes, the linkage-based analysis may be misleading. It is, therefore, imperative to reveal the physical location of the *R* genes. Furthermore, it should be possible to use this unique structural and functional organization of *R*-gene-containing regions to verify the *R*-gene candidacy of any sequence.

The objectives of this study were: (i) to isolate and characterize the expressed fraction of the wheat *R* genes belonging to all known classes and (ii) to physically map both phenotypically characterized *R* genes and the cloned *R*-gene candidates via comparative analysis and deletion mapping to understand their distribution in the wheat genome.

## MATERIALS AND METHODS

**Plant material:** Fifty-four wheat lines possessing one or more *R* genes were used for the modified RNA fingerprinting method. Nineteen nullisomic-tetrasomic lines (missing a pair of chromosomes, the deficiency of which is compensated for by a pair of homeologous chromosomes) and 13 ditelosomic lines (missing a pair of chromosome arms; SEARS 1954) were used for arm location of the gene fragments. High-resolution physical mapping of the *R*-gene sequences was accomplished using 339 wheat deletion lines (ENDO and GILL 1996).

**RNA methods:** Total RNA was isolated from the pooled leaf tissue from 3- to 4-week-old seedlings of the 54 *R*-gene-containing wheat lines using the guanidinium thiocyanate-cesium chloride method (SAMBROOK *et al.* 1989) with a few modifications. The tissue was ground to a fine powder in liquid nitrogen and suspended in guanidinium thiocyanate buffer containing 1% mercaptoethanol and 0.5% sodium lauryl sarcosinate. After removing the cell debris by centrifugation, supernatant was layered on a 5.7 M CsCl/0.01 M EDTA solution in an ultracentrifuge tube and centrifuged in a swinging bucket rotor (SW55Ti) at 40,000 rpm for ~16 hr. The RNA pellet was washed with 70% ethanol and resuspended in TE containing 0.1% SDS. The poly(A)<sup>+</sup> RNA was isolated following the standard protocol (SAMBROOK *et al.* 1989).

**Modified RNA fingerprinting and cloning:** First-strand cDNA synthesis and the subsequent PCR amplification for the RNA fingerprinting reactions were carried out using the Delta RNA fingerprinting kit (CLONTECH, Palo Alto, CA) following the recommended protocols, except that <sup>35</sup>S isotope was used instead of <sup>33</sup>P. In addition to degenerate primers for the

TABLE 1  
Oligonucleotides used for RNA fingerprinting

Primer <sup>a</sup>	Sequence (5'-3') <sup>b</sup>	Degeneracy
p-loop	AAGAATTCGGNGTNGGNAACAAC	64
GLPLAL	AACTCGAGAGNGCNAGNGGNAGGCC	256
T1	CATTATGCTGAGTGATATCTTTTTTTTTAA	—
T2	CATTATGCTGAGTGATATCTTTTTTTTTAC	—
T3	CATTATGCTGAGTGATATCTTTTTTTTTAG	—
T4	CATTATGCTGAGTGATATCTTTTTTTTTCA	—
T5	CATTATGCTGAGTGATATCTTTTTTTTTCC	—
T6	CATTATGCTGAGTGATATCTTTTTTTTTCG	—
T7	CATTATGCTGAGTGATATCTTTTTTTTTGA	—
T8	CATTATGCTGAGTGATATCTTTTTTTTTGC	—
T9	CATTATGCTGAGTGATATCTTTTTTTTTGG	—

<sup>a</sup> Each T primer contains two underlined variant nucleotides at the 3' end.

<sup>b</sup> Each underlined N within the p-loop and GLPLAL primers corresponds randomly to A, T, C, and G nucleotides.

p-loop (GVGKTT) and the hydrophobic (GLPLAL) domains of the cloned *R* genes, nine T primers were also used. The T primers had a common 19-bp sequence at the 5' end followed by nine thymidine and two variable bases at the 3' end in all possible pairwise combinations of A, C, and G (Table 1). The PCR reactions were performed in a total volume of 20  $\mu$ l containing 5  $\mu$ M each of dNTPs, 0.4  $\mu$ l of Advantage cDNA polymerase mix (CLONTECH), 20  $\mu$ M of each primer, 50 ng of the first-strand cDNA, 2  $\mu$ l of 10 $\times$  cDNA PCR reaction buffer, and 0.2  $\mu$ l of [<sup>35</sup>S]dATP. The PCR conditions were the following: one cycle of 5 min at 94 $^{\circ}$ , 5 min at 45 $^{\circ}$ , 5 min at 68 $^{\circ}$ ; 2 cycles of 2 min at 94 $^{\circ}$ , 5 min at 45 $^{\circ}$ , 5 min at 68 $^{\circ}$ ; 25 cycles of 1 min at 94 $^{\circ}$ , 1 min at 60 $^{\circ}$ , and 2 min at 68 $^{\circ}$ , followed by 7 min at 68 $^{\circ}$ . The amplification products were size separated on a 0.4-mm denaturing 5% polyacrylamide/8 M urea gel, following standard sequencing gel protocol (SAMBROOK *et al.* 1989). The gel containing 10  $\mu$ l of amplification product was run at 70 W for 3–4 hr. The gel was then rinsed in TE buffer for  $\sim$ 5 min, blotted onto a Whatman 3MM paper, and dried using a gel dryer (Bio-Rad, Richmond, CA) at 80 $^{\circ}$  for 2 hr. An X-ray film was placed on the dried gel and exposed for 3 to 4 days. The fragment bands were cut out of the gel and eluted by boiling in 50  $\mu$ l of TE for 5 min. The eluted DNA was reamplified using the corresponding primer pair, size separated on a 1% agarose gel, and purified using GENE-CLEAN III KIT (Q-BIOgene). The fragments were cloned using the pGEM-T easy cloning system I (Promega, Madison, WI). The DNA sequencing was done commercially (<http://geneseek.com>). The Sequencer 6.0 program and VecScreen (<http://www.ncbi.nlm.nih.gov/>) function were used to trim vector contamination. The trimmed sequences were grouped at 90% nucleotide identity level and the longest sequence from each group was used for further analysis. The TBLASTX (nucleotide query-translated database), BLASTX (nucleotide query-protein database), and BLASTN (nucleotide-nucleotide comparison) version 2.1.2 (<http://www.ncbi.nlm.nih.gov>) were used to assign putative functions to the sequences (ALTSCHUL *et al.* 1997). Numerical options were used at default values during the BLAST searches.

**R-gene identification by data mining:** For data mining, 22 putative protein sequences corresponding to four major *R*-gene classes were used to perform TBLASTN (protein query-translated database) searches at the National Center for Biotechnology Information (NCBI) GenBank EST database (non-mouse and nonhuman EST entries; [\[ncbi.nlm.nih.gov\]\(http://www.ncbi.nlm.nih.gov\); ALTSCHUL \*et al.\* 1997\). By comparing domain search, individual and multiple motif search, consensus sequence search, and individual full-length search, we observed that the individual full-length search is the most successful method for mining plant \*R\* genes \(M. DILBIRLIGI and K. S. GILL, unpublished results\). Therefore, the individual full-length BLAST search was performed for the putative protein sequences of 22 \*R\* genes \(Table 2\) belonging to all \*R\*-gene classes. Numerical options were left at the default values for all BLAST searches.](http://www.</a></p>
</div>
<div data-bbox=)

**Sequence analysis:** The BestFit [Genetic Computer Group (GCG); Wisconsin Package Version 10.1, Madison, WI] sequence comparison was performed to calculate the homology between the identified sequences and the known *R* genes (gap opening penalty of 3.0 and gap extension penalty of 1.0). The identified wheat sequences were compared with the known *R* genes to show their structural resemblance. Overall sequence similarity; presence and order of domains and motifs; sequence, location, and order of the conserved amino acids within the motifs; and size of the regions interspersing various motifs and domains were the criteria for structural resemblance comparisons. The MEME function of GCG and the CLUSTALW function of Vector NTI (default values) were used to identify the motifs. Additional details of the analysis of the identified sequences are given by M. DILBIRLIGI and K. S. GILL (unpublished results).

**Physical mapping of phenotypically characterized wheat *R* genes:** Physical localization of phenotypically characterized wheat *R* genes was accomplished by comparative mapping. There are 269 *R* genes, including quantitative trait loci (QTL), known in wheat or its wild relatives, of which 229 have been mapped on chromosomes, 147 have been mapped on chromosome arms, and 90 have been mapped relative to molecular markers (McINTOSH *et al.* 2000; <http://wheat.pw.usda.gov/ggpages/maps.shtml>). A comprehensive wheat consensus genetic linkage map containing 1380 commonly used DNA markers has been constructed by combining information from 137 published maps (M. DILBIRLIGI, M. ERAYMAN, D. SANDHU and K. S. GILL, unpublished results). To reveal the physical location of wheat *R* genes, 90 phenotypically characterized *R* genes were first placed on the consensus genetic linkage map by using the following criteria: First, the genes and markers present on multiple wheat genetic linkage maps were used as anchor markers for the construction of the consensus genetic maps. Second, the markers and *R* genes present between two

**TABLE 2**  
**Plant *R* genes used for data mining**

Plant	<i>R</i> genes <sup>a</sup>	Protein ID	Structure <sup>b</sup>	Reference/source
Wheat	<u><i>CRE3</i></u>	AAC05834	NB/LRR	LAGUDAH <i>et al.</i> (1997)
	<u><i>VRGA1</i></u>	AAF19148	NB/LRR	SEAH <i>et al.</i> (2000)
	<u><i>YR10</i></u>	AAG42168	NB/LRR	NCBI sequence no. AF149114
Barley	<i>MLA6</i>	CAC29241	NB/LRR	HALTERMAN <i>et al.</i> (2001)
Tomato	<i>PRF</i>	AAC49408	NB/LRR	SALMERON <i>et al.</i> (1996)
	<i>MI-1</i>	AAC67238	NB/LRR	MILLIGAN <i>et al.</i> (1998)
	<i>12C-1</i>	AAB63274	NB/LRR	ORI <i>et al.</i> (1997)
	<i>PTO</i>	AAF76313	PK	MARTIN <i>et al.</i> (1993)
Rice	<i>XA1</i>	BAA25068	NB/LRR	YOSHIMURA <i>et al.</i> (1998)
	<i>PIB</i>	BAA76282	NB/LRR	WANG <i>et al.</i> (1999)
	<i>RPR1</i>	BAA75812	NB/LRR	SAKAMOTO <i>et al.</i> (1999)
	<i>XA21</i>	T10725	PK/LRR (RLK)	SONG <i>et al.</i> (1995)
Maize	<i>RPI-D</i>	AAD47197	NB/LRR	COLLINS <i>et al.</i> (1999)
	<i>HM1</i>	T03970	Toxin reductase	JOHAL and BRIGGS (1992)
Arabidopsis	<i>RPP1</i>	AAC72977	NB/LRR	BOTELLA <i>et al.</i> (1998)
	<i>RPS2</i>	AAA21874	NB/LRR	BENT <i>et al.</i> (1994)
	<i>RPS5</i>	AAC26126	NB/LRR	WARREN <i>et al.</i> (1998)
	<i>RPP8</i>	AAC83165	NB/LRR	MCDOWELL <i>et al.</i> (1998)
	<i>RPP13</i>	AAF42832	NB/LRR	BITTNER-EDDY <i>et al.</i> (2000)
	<i>RPM1</i>	CAA61131	NB/LRR	GRANT <i>et al.</i> (1995)
Flax	<i>L6</i>	AAD25965	NB/LRR	LAWRENCE <i>et al.</i> (1995)
Sugarbeet	<i>HS1<sup>pro-1</sup></i>	AAB48305	LRR	CAI <i>et al.</i> (1997)

<sup>a</sup> Underlined *R* genes are functionally not proven.

<sup>b</sup> Only the common domains were considered.

anchor markers on multiple maps were integrated relative to the anchors. Genetic distances among the integrated markers were standardized relative to the flanking anchors. Third, markers and *R* genes with a consistent location on various maps were incorporated. Finally, the *R* genes present on only one or two genetic linkage maps were placed on the consensus genetic maps via common markers. If only one linked marker was available, the *R* gene was mapped in a window around the marker. Genetic distances used for the construction of the consensus genetic linkage map are relative.

The consensus genetic linkage map was then compared with the consensus physical maps to reveal the physical location of the phenotypically characterized *R* genes. First, flanking markers for each physical region were identified on the consensus genetic linkage map to localize encompassed *R* genes to the corresponding physical region. The *R* genes that were present very close to the flanking markers were also placed but were marked with an asterisk (\*). Additionally, the physical location of 17 group 5 *R* genes (marked “+”) was extrapolated on the basis of the location of other *R* genes present on the arm.

**Physical mapping of the *R*-gene candidates:** A two-step mapping strategy was followed to physically map the putative *R*-gene sequences. First the sequence clones were used as probes for gel-blot DNA hybridization analysis of 19 nullisomic-tetrasomic and 13 ditelosomic lines to reveal chromosome and arm location of the sequences. Either *EcoRI* or *HindIII* restriction enzymes were used for the analysis, except for BF201229 where *DraI* was used. The sequences were then analyzed on deletion lines for the corresponding chromosome. The names of the deletion lines along with the fraction length (FL) of the retained arm are given on the left of the chromosomes (Figure 2). Genomic DNA isolation and gel-blot analysis methods were as previously described (GILL *et*

*al.* 1993). The aneuploid stocks and the EST clones were kindly provided by Bikram S. Gill and Olin Anderson, respectively.

Each fragment band was mapped to a chromosomal region flanked by the breakpoints of the smallest deletion lacking the fragment band and the largest deletion possessing it. Multiple fragment bands mapping on a chromosome arm were scored to be nonallelic (shown by a, b, and c at the end of the probe name) if the bands mapped to different chromosomal regions. The physical mapping information from the three homeologues (*A*, *B*, and *D*) was combined to localize the known *R* genes as well as the *R*-gene candidates to the smallest possible physical regions. The consensus physical map for each of the seven wheat homeologous groups was constructed as previously described (GILL *et al.* 1996a).

## RESULTS

**RNA fingerprinting:** Size separation of the RNA fingerprinting PCR product showed 76 bands with the p-loop/GLPL primer combination. The p-loop/T primer combinations generated 80–100 bands each (Figure 1a). The size of the fragment bands ranged between ~150 and 1300 bp except for the p-loop/GLPL primer combination in which the size of the smallest band was ~300 bp. Many fragment bands amplified by different primer combinations appeared to be the same because of the size. About 900 fragment bands were observed on the fingerprinting gels. About 220 bands that appeared to be unique on the basis of size and intensity were excised from the gels, reamplified, and cloned (Figure 1b). Two

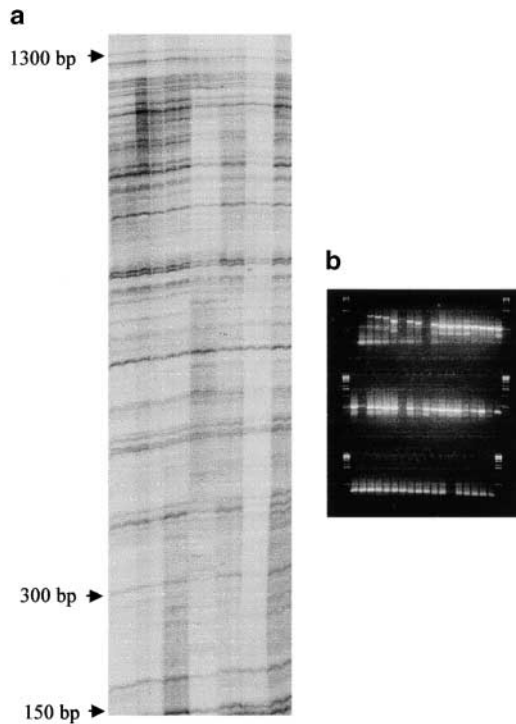


FIGURE 1.—Cloning of expressed candidate *R*-gene fragments using modified RNA fingerprinting. (a)  $^{35}\text{S}$ -labeled PCR product size separated on a polyacrylamide-urea gel. The product of each primer combination was loaded in two adjacent lanes. Approximate size of the fragments is shown on the left. (b) The excised bands were reamplified and size separated on a 1% agarose gel. The order of the PCR products on the agarose gel from left to right follows that of the polyacrylamide gel from top to bottom.

clones corresponding to each band were sequenced. Additional clones were sequenced if two clones from a sample were different. A total of 385 clones were sequenced and analyzed. ContigExpress analysis of the sequences resulted in 121 unique contigs. The longest sequence of each contig was used for further analysis after sequences were confirmed to contain a p-loop (kinase-1a) motif. Individual Gap (GCG) analysis showed that 121 sequences shared 21% (UNL115–UNL175) to 90% (UNL184–UNL185, UNL201–UNL202 and UNL208, and UNL139–UNL160) sequence similarity. Only ~27% of the sequences were closely related and sequence similarity for the remaining 73% was <50%. Putative protein sequences of the 121 clones were analyzed by BLASTX (ALTSCHUL *et al.* 1997; <http://www.ncbi.nlm.nih.gov/BLAST>). Only 48 of these sequences showed homology to the known or annotated genes and the remaining 73 were unique. Sequence similarity for 19 of the 48 was >80% ( $E \leq 10^{-40}$ ) and for the remaining was between 35 and 80% ( $E \leq 10^{-8}$ ). Twenty-six sequences were *R*-gene candidates because of their homology to NB/LRR, *Pto*, *Xa21*, or other types of *R* genes (Table 3 and Table 4). Twenty-two sequences were homologous (BLASTX  $E$  values were  $\leq 10^{-29}$  to  $\leq 10^{-57}$ ) to genes controlling cell

structural and metabolic activities and thus were not analyzed further.

**Motif analysis of the unique sequences:** The MotifSearch (GCG) was performed for the 73 unique sequences to identify putative functional motifs and domains. In addition to the conserved p-loop (kinase-1a), 38 also had the predicted kinase-2 and kinase-3a motifs (Tables 3 and 4). Of these, 17 contained an NB domain similar to that of the NB/LRR type of *R* genes and the remaining 21 were putative kinases. The consensus among the 17 NB/LRR types of sequences was GVGKTT for p-loop along with (L/V) (L/V) (L/V) (L/I/D) D (D/I/V/L) for kinase-2 and (E/G/F/V) (T/S/G/Q)  $\times$  (T/Y) (T/S) R for kinase-3a. The 21 putative kinases had an invariant aspartate (D) in kinase-2 and an arginine (R) in kinase-3a. Of the remaining 35 sequences, 21 had either a kinase-2 or a kinase-3a along with a p-loop motif as found in NB or kinase-encoding proteins. Fourteen sequences were not analyzed further as no motif other than a p-loop was observed.

**Data mining:** The full-length sequence comparisons identified 344 sequences homologous to 22 known *R* genes with  $E$  values ranging from  $\leq 10^{-1}$  to  $\leq 10^{-117}$ . Detailed structural analysis of these sequences was reported by M. DILBIRLIGI and K. S. GILL (unpublished results). Briefly, comparison of the cloned *R* genes showed that the sequence, order, and gap length among various motifs and domains was unique to *R* genes even though the *R*-gene domains and motifs are also present in other proteins. Therefore, these structural features of the selected 344 sequences were compared with those of the cloned *R* genes. These analyses selected 176 sequences that were then used as query for BLASTX search in the nonredundant protein database. Only the sequences in which the best hit was a cloned *R* gene were selected. The selected 163 sequences were analyzed with VecScreen ORF Finder BLASTP programs (<http://www.ncbi.nlm.nih.gov>). These sequences assembled into 99 contigs representing 67 NB/LRR types, 12 *Xa21*-like (RLK), 11 *Pto*-like (PK), 7 *Hm1*-like, and 2 *Hs1<sup>pro1</sup>*-like sequences (Table 3). For physical mapping, the largest clone from each contig was used as a probe.

**Physical mapping of *R*-gene candidates:** A total of 166 best *R*-gene candidates, including 85 from RNA fingerprinting and 81 from data mining experiments, were physically mapped using wheat aneuploid stocks (Tables 3 and 4). These *R*-gene candidates belonged to all known classes. Forty-five of these sequences were unmappable because of either smear pattern or unresolved fragment bands on nullisomic-tetrasomic lines (Table 4). Physical location of 121 sequences was revealed using 335 deletion lines for all 21 wheat chromosomes (Figure 2). Four (4BL-6, 4BL-9, 5BL-11, and 6BS-9) of the 339 deletion lines were not used for the final analysis because of the discrepant results.

The 121 probes detected 664 fragment bands of which 371 were mapped on chromosomes using either *EcoRI*

**TABLE 3**  
**Summary of the type of *R*-gene sequences used for deletion mapping**

<i>R</i> -gene classes	RNAF	Data mining	Total sequence	Total attempted	Total mapped <sup>b</sup>
NB/LRR type	20	67	87	80	66 (82)
PK type	2	11	13	8	6 (9)
RLK type	4	12	16	12	5 (8)
<i>Hm1</i> type	—	7	7	5	3 (3)
<i>Hs1<sup>pro-1</sup></i> type	—	2	2	2	—
PR related	17	—	17	17	13 (14)
NB/kinase encoding <sup>a</sup>	42	—	42	42	28 (35)
Total	85	99	184	166	121 (151)

<sup>a</sup> Novel sequences annotated by MotifSearch analysis.

<sup>b</sup> Numbers within parentheses show the loci detected on the consensus physical maps by the sequences given outside the parentheses.

or *HindIII* enzymes (Table 4). These 371 fragment bands represented 310 loci. Of these, 94 mapped on the *A*-genome, 120 on the *B*-genome, and 96 on the *D*-genome chromosomes (Figure 2; Table 4). Forty-seven loci were detected for wheat homeologous group 1, 34 for 2, 45 for 3, 27 for 4, 67 for 5, 42 for 6, and 48 were for group 7 chromosomes. The highest number of loci was detected on homeologous group 5 (22%) and the lowest was detected on homeologous group 4 (9%). The difference was not significant among other homeologous groups. Of the 121 probes, 17 detected loci for only one, 21 for two, and the remaining 83 detected loci for all three homeologous chromosomes (Table 4).

Twenty-three probes detected paralogous loci. The probes UNL214, UNL216, BE426789, BF200008, and BE427790 detected loci on three different homeologous groups (Table 4). The number of bands detected by these probes ranged between 7 and 15. Eighteen probes detected two paralogous loci each. The paralogous loci detected by 6 of these probes were on the same chromosome (Table 4). The loci corresponding to these 23 probes were 25 on the *A*-, 34 on the *B*-, and 25 on the *D*-genome chromosomes (Figure 2; Table 4). These 23 probes included 11 NB/LRR, six unique NB/kinases, and three of the PK class of *R* genes. Some differences were observed among the various types of sequences for the number of bands detected. The average number of bands for the NB/LRR, RLK, and PK classes of *R* genes was 6.1, 6.6, and 7.3, respectively. The unique NB/kinases, *Hm1*, and the other types of *R* genes detected an average of ~4 bands.

The physical maps of the wheat chromosomes are shown in Figure 2. The 121 probes detected loci on all 21 wheat chromosomes. The highest number (67) of loci was detected on group 5 chromosomes and the lowest (27) on group 4. The number of loci for the remaining homeologous groups was between 35 and 48 (Figure 2). Most of the *R*-gene candidate loci mapped to the distal regions of the chromosomes. About 75%

of the loci (231 of 310) mapped to the distal 20% of the chromosomes. The only exception was *Xunl200*, which mapped in the proximal regions of group 7 chromosomes (Figure 2). The majority of the NB/LRR and NB/kinase types of *R*-gene candidates mapped to the distal regions of the chromosomes. The *Hm1* and PK-like gene sequences mapped in both distal and proximal regions (Figures 2 and 3). Also, the fragments belonging to the NB/LRR and NB/kinase classes of *R* genes mapped to all homeologous groups of wheat. The PR and RLK types of fragments mapped to five of the seven homeologous groups. The PR type did not map on groups 2 and 3, and the RLK type was absent on groups 1 and 3. The PK type was absent on homeologous groups 4, 6, and 7.

Various orthologous loci on the homeologues were in the syntenic regions; therefore, it was possible to combine the physical mapping information from the three homeologous chromosomes to generate a consensus physical map (Figure 3). About 43% of the loci mapped to the short arms and 57% to the long arms (Figure 3; Table 5). Most of the candidate *R*-gene loci localized to 26 chromosomal regions. These regions encompassed ~16% of the wheat genome. Major differences were observed among various regions for both the size and the number of the candidate *R*-gene loci. The size of the regions ranged from 1 to 11% of the respective chromosome and the number of loci ranged from 2 to 29. Some of these regions were very small in size but contained a large number of the candidate *R*-gene loci. Five of these regions (short arm of groups 1, 2, and 3 and long arm of groups 5 and 6) contained ~47% of the loci but encompassed only ~3.5% of the genome (Figure 3). These regions of higher loci density were present mainly at the distal ends of the chromosomes.

**Physical mapping of phenotypically characterized *R* genes:** Physical mapping results of phenotypically characterized wheat *R* genes are summarized in Figure 3 and Tables 5 and 6. Of the total 269 wheat *R* genes, 90 have been mapped relative to molecular markers and

TABLE 4  
Gel-blot DNA analysis of the putative *R* genes using aneuploid stocks

<i>R</i> -gene sequences <sup>a</sup>	Structure <sup>b,c</sup>	No. of fragments					Chromosome
		A	B	D	<i>Hind</i> III	<i>Eco</i> RI	
<i>unl135</i>	NB/kinase* <sup>Ψ</sup>	—	1	—		7	1
<i>unl143a</i>	NB/kinase* <sup>Ψ</sup>	—	—	1	7		1
<i>unl143b</i>	NB/kinase* <sup>Ψ</sup>	—	1	—	7		1
<i>unl220</i>	NB/kinase* <sup>Ψ</sup>	1	1	1	3		1
<i>unl128</i>	NB/LRR*	—	1	—		5	1
<i>BE406625</i>	NB/LRR**	—	1	1		4	1
<i>BE444145</i>	NB/LRR**	—	1	1		3	1
<i>BE498831</i>	NB/LRR**	—	1	—		11	1
<i>BE499561</i>	NB/LRR**	—	1	1		5	1
<i>BE605005</i>	NB/LRR**	1	1	1		3	1
<i>BF473313</i>	NB/LRR**	1	1	2		4	1
<i>BF474204</i>	NB/LRR**	1	1	—		2	1
<i>BF475048</i>	NB/LRR**	2	2	3	11		1
<i>BF483881</i>	NB/LRR**	2	2	—	12		1
<i>BE405778</i>	NB/LRR***	1	1	1		6	1
<i>BE518223</i>	NB/LRR***	1	1	—		5	1
<i>BG606866</i>	NB/LRR***	2	2	—	9		1
<i>BM135313</i>	NB/LRR***	1	1	—		5	1
<i>BE492937a<sup>d</sup></i>	PK**	1	1	1	8		1
<i>BE637867</i>	PK***	1	1	—		3	1
<i>unl111</i>	PR***	1	1	1	3		1
<i>unl207</i>	PR***	1	1	1	3		1
<i>BE518335</i>	Hm1 like***	—	1	1		3	2
<i>BE591764</i>	Hm1 like***	1	2	1		6	2
<i>unl110</i>	NB/kinase* <sup>Ψ</sup>	—	1	2		4	2
<i>unl136</i>	NB/kinase* <sup>Ψ</sup>	—	1	1	3		2
<i>unl149</i>	NB/kinase* <sup>Ψ</sup>	1	1	1	4		2
<i>unl145</i>	NB/LRR*	1	1	1	4		2
<i>unl147</i>	NB/LRR*	1	1	1		4	2
<i>BE404328</i>	NB/LRR**	—	1	—	7		2
<i>BE497251a</i>	NB/LRR**	—	1	1		7	2
<i>BE499523</i>	NB/LRR**	1	1	1		4	2
<i>BF200166</i>	NB/LRR**	1	2	—		5	2
<i>BF483292</i>	NB/LRR**	—	—	1		2	2
<i>BF484266</i>	NB/LRR**	1	3	1	7		2
<i>BE442802a<sup>d</sup></i>	PK**	—	2	—		14	2
<i>BE442854a</i>	PK**	—	2	—		14	2
<i>unl216a</i>	RLK**	—	3	1	15		2
<i>BE607045</i>	Hm1 like***	1	1	1		3	3
<i>unl130a</i>	NB/kinase* <sup>Ψ</sup>	—	2	—	8		3
<i>unl130b</i>	NB/kinase* <sup>Ψ</sup>	2	2	1	8		3
<i>unl132</i>	NB/LRR*	1	1	1		4	3
<i>unl150</i>	NB/LRR*	1	1	1		3	3
<i>unl153</i>	NB/LRR*	1	1	1		5	3
<i>unl167</i>	NB/LRR*	1	1	1	4		3
<i>BE405711</i>	NB/LRR**	1	1	1		3	3
<i>BE444148</i>	NB/LRR**	1	1	—		2	3
<i>BE499177</i>	NB/LRR**	1	1	1		3	3
<i>BE499910</i>	NB/LRR**	1	1	—		2	3
<i>BF200008a</i>	NB/LRR**	1	1	1		9	3
<i>BF482366</i>	NB/LRR**	1	1	1		5	3
<i>BF483203</i>	NB/LRR**	—	2	—		5	3
<i>BE591154</i>	NB/LRR***	1	1	—		2	3
<i>BE442854b</i>	PK**	1	1	1		14	3
<i>BE492937b<sup>d</sup></i>	PK**	1	1	1		8	3
<i>BE637850<sup>d</sup></i>	PK**	1	—	—		5	3

(continued)

**TABLE 4**  
(Continued)

R-gene sequences <sup>a</sup>	Structure <sup>b,c</sup>	No. of fragments					Chromosome
		A	B	D	<i>Hind</i> III	<i>Eco</i> RI	
<i>unl129</i>	NB/kinase* <sup>ψ</sup>	—	—	1		2	4
<i>unl133</i>	NB/kinase* <sup>ψ</sup>	2	1	1	5		4
<i>unl164a</i>	NB/kinase* <sup>ψ</sup>	1	1	2	9		4
<i>unl164b</i>	NB/kinase* <sup>ψ</sup>	—	—	1	9		4
<i>unl181</i>	NB/kinase* <sup>ψ</sup>	1	—	1	3		4
<i>unl198</i>	NB/kinase* <sup>ψ</sup>	1	—	1	7		4
<i>unl124</i>	NB/LRR*	—	2	—	9		4
<i>BE403251b</i>	NB/LRR**	1	—	—		17	4
<i>BE426789c</i>	NB/LRR**	1	—	—		12	4
<i>BE426790c</i>	NB/LRR**	1	—	—		12	4
<i>BE444404</i>	NB/LRR**	2	—	—	8		4
<i>BE445244a</i>	NB/LRR**	1	—	—	8		4
<i>BE497251b</i>	NB/LRR**	1	—	—	7		4
<i>BF201403</i>	NB/LRR**	2	—	—		8	4
<i>BF482358</i>	NB/LRR**	1	1	1	7		4
<i>unl163</i>	PR***	—	1	1	5		4
<i>unl216b</i>	RLK**	—	1	1	15		4
<i>unl109</i>	NB/kinase* <sup>ψ</sup>	1	1	1	4		5
<i>unl141a</i>	NB/kinase* <sup>ψ</sup>	1	1	1	15		5
<i>unl166</i>	NB/kinase* <sup>ψ</sup>	1	—	1	3		5
<i>unl169</i>	NB/kinase* <sup>ψ</sup>	1	1	1	3		5
<i>unl191</i>	NB/kinase* <sup>ψ</sup>	1	1	1	3		5
<i>unl214a</i>	NB/kinase* <sup>ψ</sup>	—	1	1		7	5
<i>unl217</i>	NB/kinase* <sup>ψ</sup>	2	1	2		6	5
<i>unl218</i>	NB/kinase* <sup>ψ</sup>	—	1	1	4		5
<i>unl117</i>	NB/LRR*	1	1	1	5		5
<i>unl119a</i>	NB/LRR*	—	1	1	11		5
<i>unl123</i>	NB/LRR*	1	1	1		3	5
<i>unl174</i>	NB/LRR*	1	2	—		3	5
<i>BE406845a</i>	NB/LRR**	1	—	—		11	5
<i>BE406845b</i>	NB/LRR**	—	1	1		11	5
<i>BE445244b</i>	NB/LRR**	—	1	—		8	5
<i>BE498327</i>	NB/LRR**	—	1	—		2	5
<i>BF482744</i>	NB/LRR**	1	1	—		2	5
<i>BF484437</i>	NB/LRR**	1	1	—		2	5
<i>BG604429</i>	NB/LRR**	2	—	—	3		5
<i>BM136556a</i>	NB/LRR**	—	1	—	6		5
<i>BM136556b</i>	NB/LRR**	1	2	1	10		5
<i>BM137909</i>	NB/LRR**	—	1	1		7	5
<i>BE442802b<sup>d</sup></i>	PK**	—	—	1		14	5
<i>BF201229</i>	PR***	—	1	1	2 <sup>e</sup>		5
<i>unl104</i>	PR**	1	1	1	3		5
<i>unl113</i>	PR***	1	1	1	3		5
<i>unl121</i>	PR***	1	1	1	3		5
<i>unl138</i>	PR***	—	1	1		2	5
<i>unl140</i>	PR***	—	1	1		2	5
<i>unl213a</i>	PR***	—	1	—	4		5
<i>unl216c</i>	RLK**	—	1	—	15		5
<i>unl152a</i>	RLK***	1	—	1	9		5
<i>unl141b</i>	NB/kinase* <sup>ψ</sup>	—	—	1	15		6
<i>unl214b</i>	NB/kinase* <sup>ψ</sup>	1	—	1		7	6
<i>unl119b</i>	NB/LRR*	—	1	—		11	6
<i>unl127</i>	NB/LRR*	2	1	2		5	6
<i>unl194a</i>	NB/LRR*	—	1	—	7		6
<i>BE403251a</i>	NB/LRR**	2	4	2		17	6
<i>BE405507b</i>	NB/LRR**	—	1	2		6	6
<i>BE426789a</i>	NB/LRR**	1	1	—		12	6

(continued)



TABLE 4  
(Continued)

<i>R</i> -gene sequences <sup>a</sup>	Structure <sup>b,c</sup>	No. of fragments					Chromosome
		A	B	D	<i>Hind</i> III	<i>Eco</i> RI	
<i>BE426790a</i>	NB/LRR**	1	1	—		12	6
<i>BE500158</i>	NB/LRR**	3	2	1		11	6
<i>BE638124</i>	NB/LRR**	1	—	—		5	6
<i>BF199778</i>	NB/LRR**	1	2	1		12	6
<i>BF200008b</i>	NB/LRR**	—	1	—		9	6
<i>BF478914</i>	NB/LRR**	—	2	1	7		6
<i>BG607637</i>	NB/LRR**	—	1	2		6	6
<i>BM137173</i>	NB/LRR**	—	1	1		2	6
<i>unl105</i>	PR**	1	1	1		3	6
<i>unl154</i>	PR**	—	1	2	4		6
<i>unl213b</i>	PR***	1	—	1	4		6
<i>unl122</i>	RLK**	1	1	1	4		6
<i>unl152b</i>	RLK***	—	—	1	9		6
<i>unl115</i>	NB/kinase* <sup>Ψ</sup>	—	—	1	3		7
<i>unl120</i>	NB/kinase* <sup>Ψ</sup>	—	—	1	3		7
<i>unl134</i>	NB/kinase* <sup>Ψ</sup>	1	1	1	3		7
<i>unl157</i>	NB/kinase* <sup>Ψ</sup>	1	1	1	3		7
<i>unl170</i>	NB/kinase* <sup>Ψ</sup>	—	2	—		3	7
<i>unl179</i>	NB/kinase* <sup>Ψ</sup>	—	1	1		3	7
<i>unl188</i>	NB/kinase* <sup>Ψ</sup>	1	1	1	4		7
<i>unl214c</i>	NB/kinase* <sup>Ψ</sup>	1	—	—		7	7
<i>unl219</i>	NB/kinase* <sup>Ψ</sup>	1	1	1	6		7
<i>unl173</i>	NB/LRR*	1	1	1	4		7
<i>unl194b</i>	NB/LRR*	1	—	—		7	7
<i>unl204</i>	NB/LRR*	—	2	—	7		7
<i>BE405003</i>	NB/LRR**	1	1	1		4	7
<i>BE405507a</i>	NB/LRR**	1	—	—		6	7
<i>BE426789b</i>	NB/LRR**	—	1	—		12	7
<i>BE426790b</i>	NB/LRR**	—	1	—		12	7
<i>BE518338a</i>	NB/LRR**		1			6	7
<i>BE518338b</i>	NB/LRR**	1	1	1		6	7
<i>BE591190</i>	NB/LRR**	1	1	1		7	7
<i>BF200008c</i>	NB/LRR**	1	1	—	9		7
<i>unl151</i>	PR***	1	2	1		6	7
<i>unl200</i>	PR***	1	1	1	3		7
<i>BE403270</i>	RLK**	1	—	—		2	7
<i>BE443579</i>	RLK**	1	1	1		3	7

BE415090, BE604247, BE213346, BE499505, BE403520, BF482232, BE405531, BE443325, BE445978, BE492332, BE518042, BF473126, BF473851, BE425456, BE429766, BM134978, BI479697, BE517538, BM138661, BF445532, UNL177, UNL215, UNL172, UNL159, UNL176, UNL203, UNL101, UNL103, UNL186, UNL102, UNL193, UNL205, UNL108, UNL114, UNL137, UNL180, UNL195, UNL131, UNL156, UNL158, UNL162, UNL165, UNL175, UNL178, UNL210, and UNL221 probes showed either smear pattern or unmappable bands. \*<sup>Ψ</sup> Novel sequences that show structural similarities to NB and kinase function rather than sequence homology.

<sup>a</sup> *a*, *b*, and *c* suffixes show multiple loci for a certain probe.

<sup>b</sup> Asterisks represent the identity level of sequences to the cloned *R* genes: \*, <40%; \*\*, 40–60%; \*\*\*, >60%.

<sup>c</sup> Sequences with NB/kinase structure are identified using MotifSearch as proposed by TRAUT (1994).

<sup>d</sup> Fragments contain >50% sequence homology to the kinase domain of both *Xa21* and *Pto* genes.

<sup>e</sup> *Dra*I enzyme was used.

147 have been localized to chromosomes (Table 5; McINTOSH *et al.* 2000). These wheat *R* genes confer resistance against 18 different wheat pests (see Table 5). The number of genes corresponding to the five major wheat pests *Lr*, *Sr*, *Yr*, *Pm*, and *H* are 45, 46, 47, 34, and 24, respectively. The number of resistance genes corresponding to the remaining 13 pests is 27 (McINTOSH *et al.* 2000).

An example of the comparison of the consensus physical with the genetic linkage maps is shown in online supplemental Figure 4 available at <http://www.genetics.org/supplemental/>. The 90 phenotypically characterized *R* genes were reliably mapped to small chromosomal regions by the comparative mapping approach. Seventeen additional *R* genes were tentatively localized to physical regions on the basis of the location of candi-

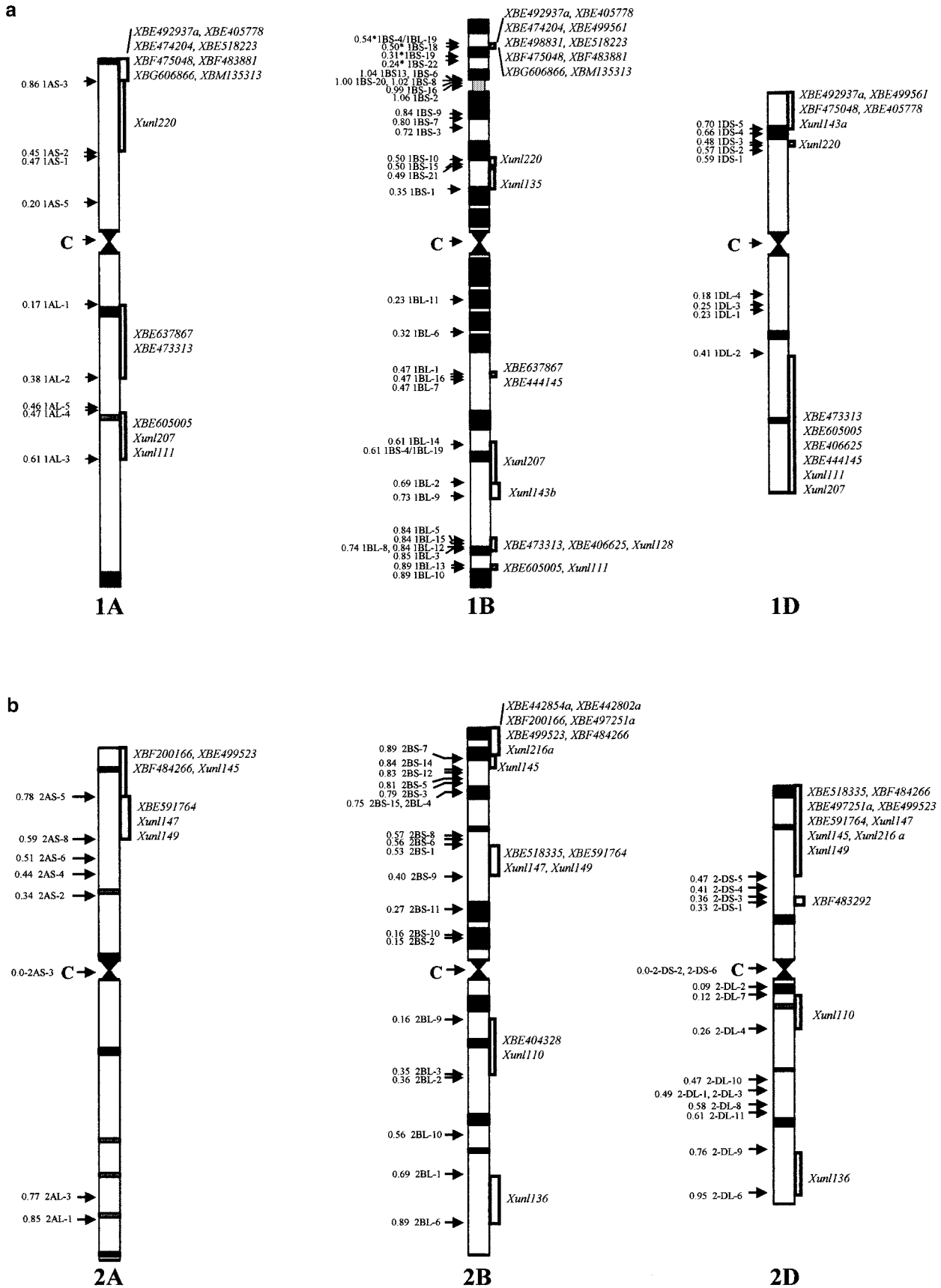
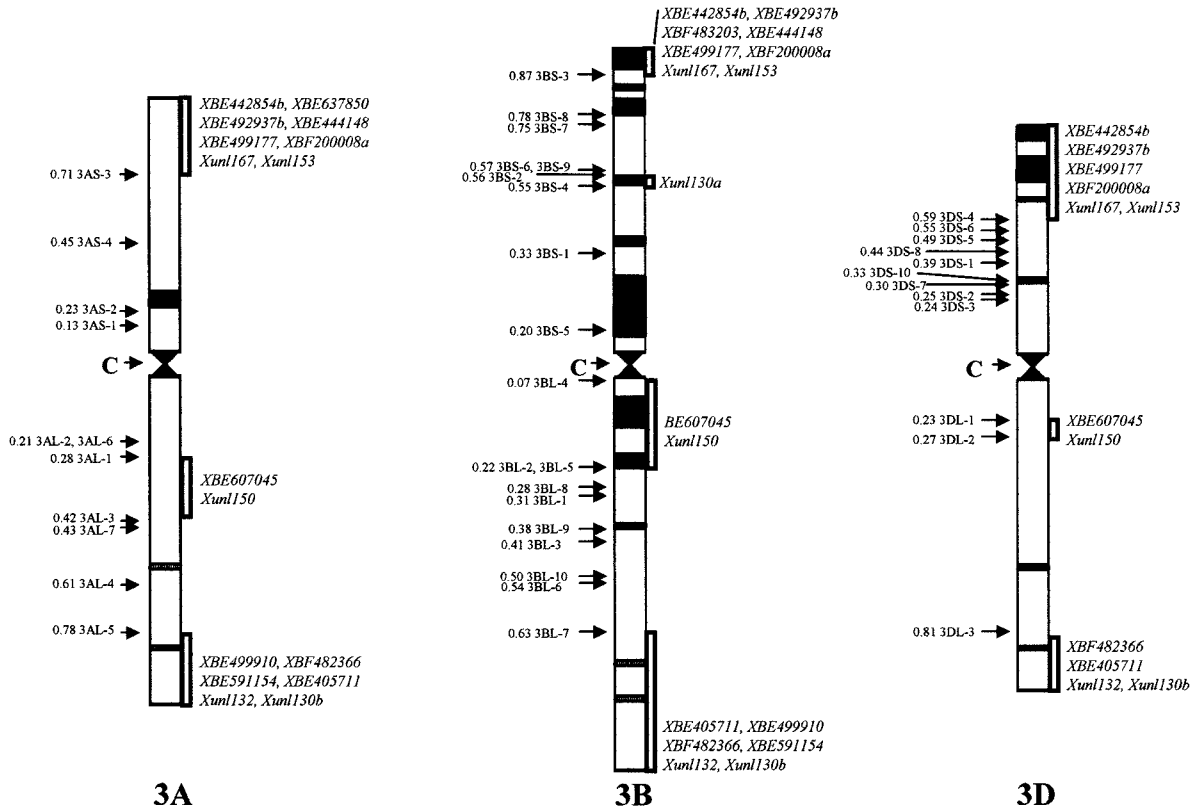


FIGURE 2.—Physical maps of *R*-gene candidates on the wheat chromosomes. Chromosome length, arm ratio, and the C-banding patterns were drawn to scale following the nomenclature for standard karyotypes (GILL *et al.* 1991). The deletion line breakpoints and FL of the retained arms are marked by arrows on the left of the chromosomes. Location of the candidate *R*-gene loci is given on the right.

**c**



**d**

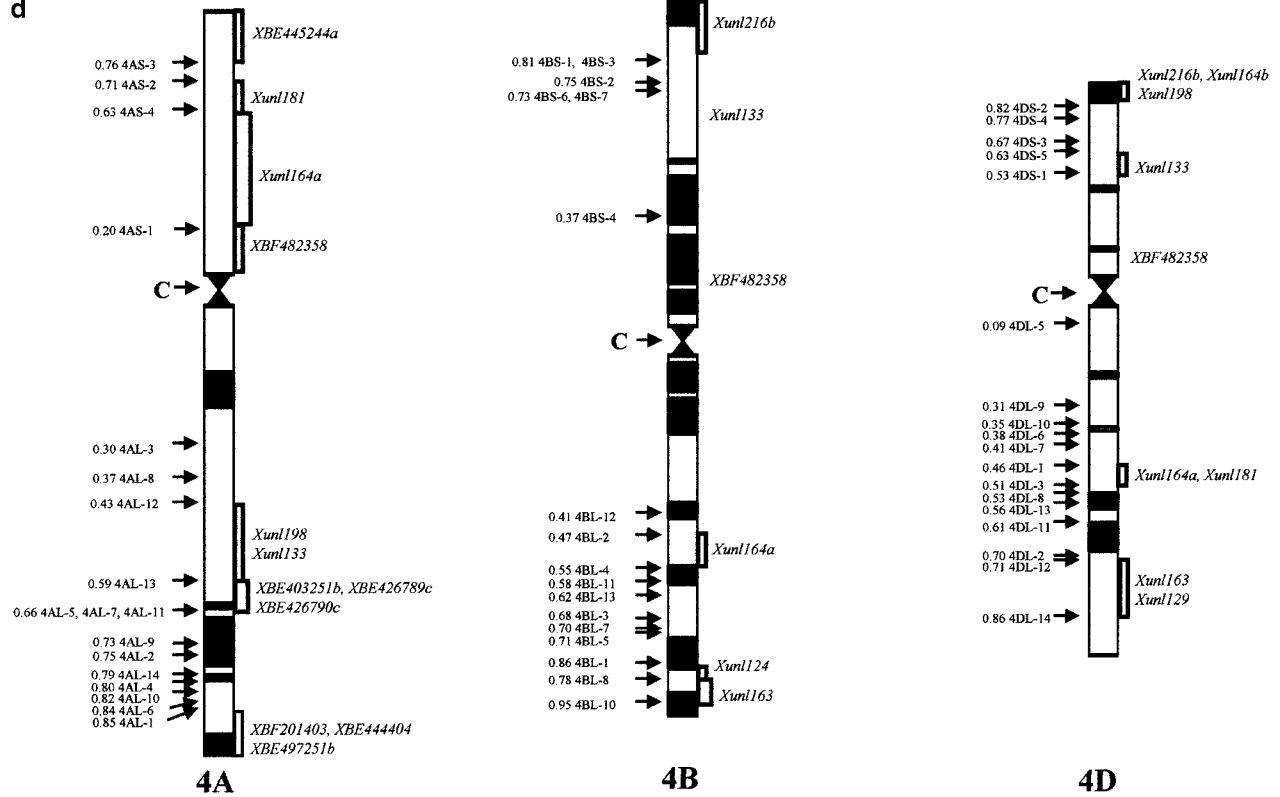


FIGURE 2.—Continued.

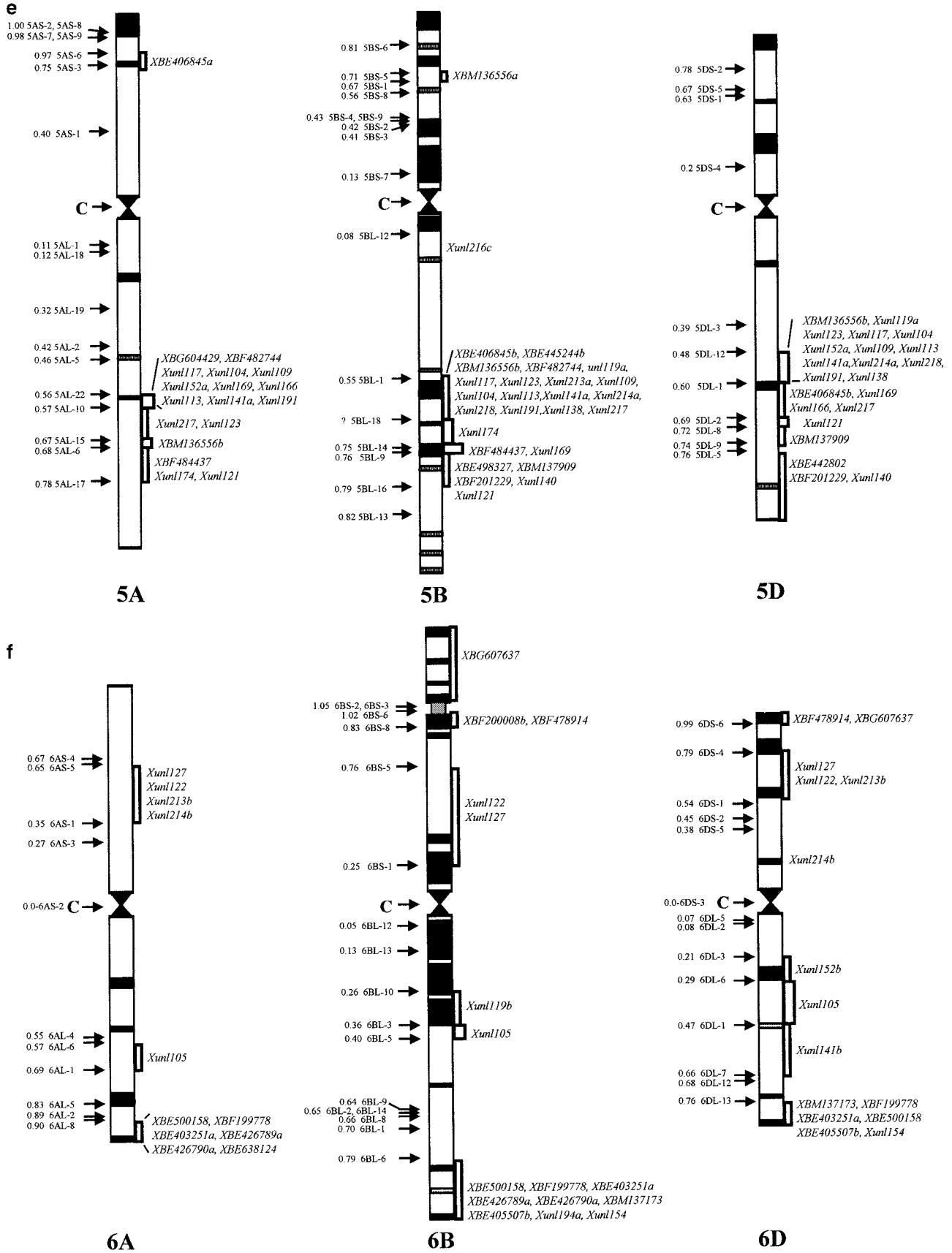


FIGURE 2.—Continued.

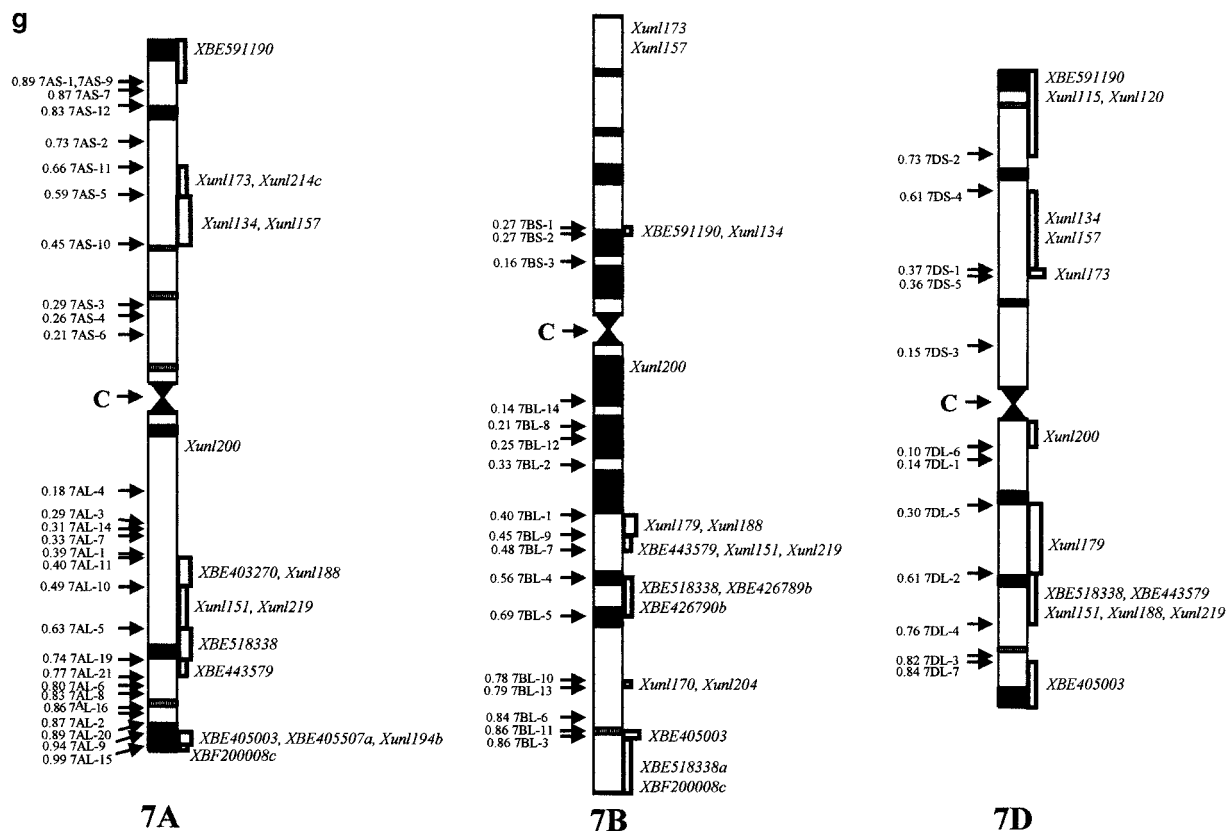


FIGURE 2.—Continued.

date *R* genes (marked ‘†’, Figure 3). These group 5 *R* genes (mainly *H*) were localized to a region where several candidate *R* genes mapped (Figure 3). These genes, however, were not included in any further analysis. Physically mapped wheat *R* genes included 31 *Lr*, 30 *Sr*, 8 *Yr*, 16 *Pm*, and an *H* resistance gene loci. The remaining 4 wheat *R*-gene loci were for *Crr*, *Kb*, and fusarium pathogens (Figure 3; Table 5). The short arms of the wheat chromosomes contained 50 of the 90 *R* genes and the long arms contained 40 (Figure 3; Tables 5 and 6). While wheat group 1 and 2 chromosomes contained 41% of the *R* genes, homeologous group 3 had the least (6%) number. The five major types of wheat *R* genes were present in all seven groups (Figure 3, Table 5). No specific pattern was observed for the type of wheat *R* genes.

Both the location and the distribution pattern between the phenotypically characterized wheat *R* genes and the candidate *R*-gene sequences were very similar. Of the 26 regions, 18 contained ~78% of the *R*-gene candidate loci and ~90% of the phenotypically characterized wheat *R* genes (Figure 3; Table 5). The proportional number in these regions was also similar. The exception was the long arm of group 2 chromosomes that contained eight of the known wheat *R* genes but lacked any of the candidate loci (Figure 3; Table 5).

The size of these regions totaled <16% of the wheat genome. About 50% of both the phenotypically characterized wheat *R* genes and the candidate *R*-gene loci mapped to five distal regions of chromosome groups 1, 2, and 3 short arms and 5 and 6 long arms encompassing <3% of the genome (Figure 3).

## DISCUSSION

Resistance mechanism seems to be conserved among plants as most of the *R* genes share a strong structural similarity irrespective of the host or the pest type. With an objective to utilize this structural similarity to isolate *R* genes *en masse*, RGAs were amplified from genomic DNA using conserved motif primers (KANAZIN *et al.* 1996; LEISTER *et al.* 1996; SANMIGUEL *et al.* 1996; YU *et al.* 1996; AARTS *et al.* 1998; COLLINS *et al.* 1998; SEAH *et al.* 1998; SHEN *et al.* 1998). This approach was not very successful probably because of the presence of pseudogenes that have structural similarities to the functional *R* genes. Furthermore, most of the functional *R* genes have been shown to be single or few copies that will be out-competed by multiple-copy nonfunctional RGAs during random cloning and sequencing of the amplified genomic fragments. Analyses of ~800 RGAs from 20 plant species have identified only one functional

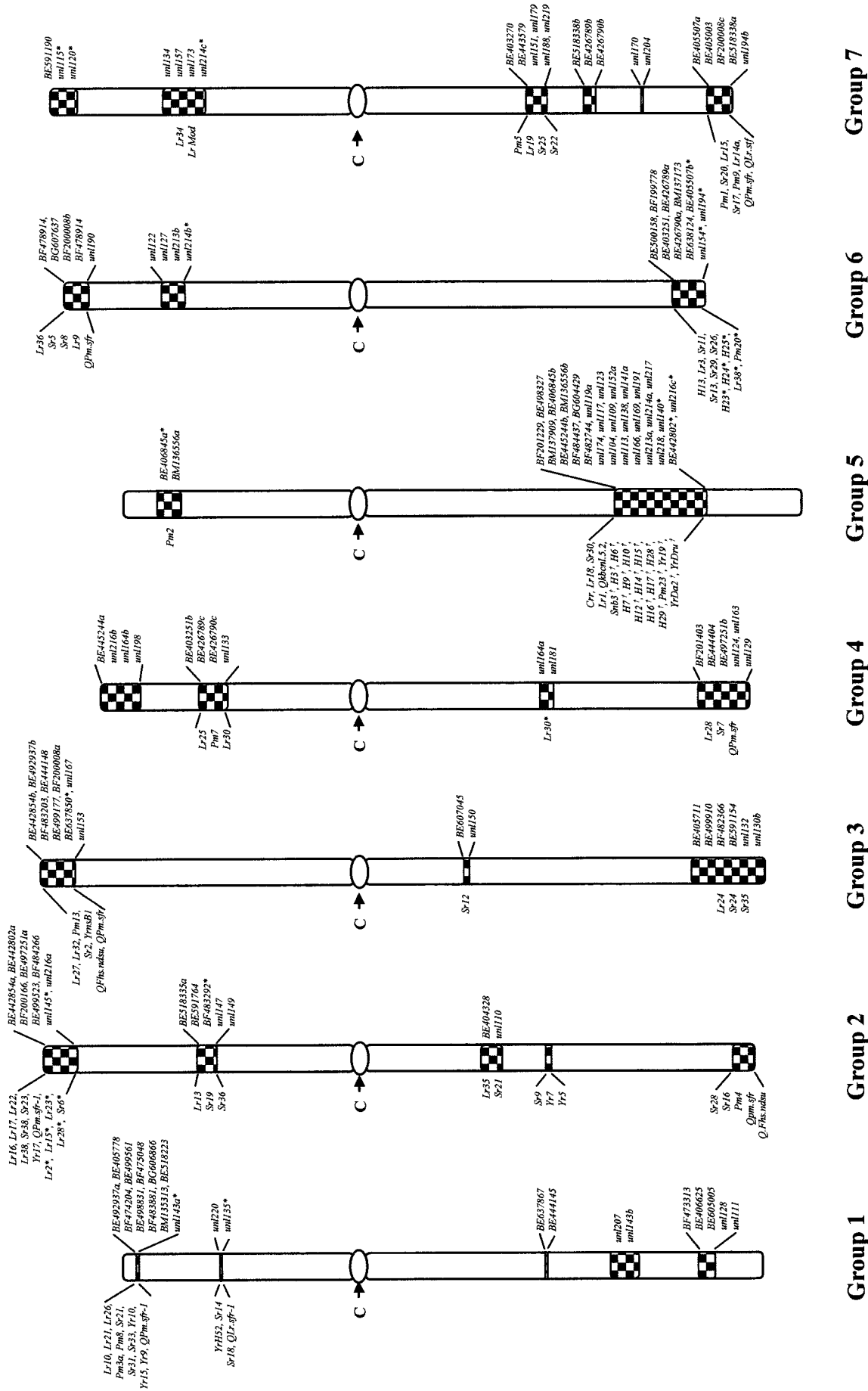


FIGURE 3.—Consensus physical maps showing distribution of candidate *R*-gene probes on wheat chromosome groups. Each arm of the consensus chromosome and the location and size of the *R*-gene-containing regions were drawn to scale on the basis of the average size of the three homologous chromosomes. Names of the deletion lines flanking the regions are given on the left side of the consensus chromosomes. A few loci, mapping in closer proximity to the flanking deletion lines of the *R*-gene clusters, were included in the regions and labeled by an asterisk (\*). Wheat *R* genes labeled as “+” on the long arm of group 5 were tentatively located since the majority of the candidate *R*-gene fragments mapped only to that region.

**TABLE 5**  
**Physical localization of phenotypically characterized wheat *R* loci based on mapping results of the putative *R*-gene sequences**

Unknown	Short/long arm <sup>a</sup>	Wheat <i>R</i> loci	Candidate loci <sup>b</sup>	DL and FL value
Group 1				
<i>Bt4</i>	<i>H5</i>	<i>Lr10</i>	<i>XBE492937a</i> (3)	1BS-4–1BS-19
<i>Bt5</i>	<i>Sr45</i>	<i>Lr21</i>	<i>XBE405778</i> (1)	1.47–1.54 (sat)
<i>Bt6</i>	<i>Yr24</i>	<i>Lr26</i>	<i>XBE474204</i> (1)	
<i>Gb2</i>	<i>QLr.sfr-1B</i>	<i>Pm3a</i>	<i>XBE499561</i> (1)	
<i>Gb6</i>		<i>Pm8</i>	<i>XBE498831</i> (1)	
<i>H11</i>		<i>Sr21<sup>c</sup></i>	<i>XBF475048</i> (1)	
<i>H22</i>		<i>Sr31</i>	<i>XBF483881</i> (1)	
<i>Lr1<sup>c</sup></i>		<i>Sr33</i>	<i>XBG606866</i> (1)	
<i>Lr2</i>		<i>Yr10</i>	<i>XBM135313</i> (1)	
<i>Lr24<sup>c</sup></i>		<i>Yr15</i>	<i>XBE518223</i> (1)	
<i>Lr42</i>		<i>QPm.sfr-1BD<sup>d</sup></i>	<i>Xunl143a<sup>e</sup></i> (5)	
<i>Lr44</i>		<i>Yr9</i>		
<i>Lr46</i>		<i>YrH52</i>	<i>Xunl220</i> (5)	1BS-21–1BS-10
<i>Pm10</i>		<i>Sr14</i>	<i>Xunl135<sup>f</sup></i> (5)	0.49–0.50
<i>Pm17</i>		<i>QLr.sfr-1B</i>		
<i>Pm22</i>		<i>Sr18</i>		
<i>Yr21</i>	<i>Dn4</i>		<i>XBE637867</i> (3)	1BL-1–1BL-7
<i>Yr25</i>	<i>Sr18</i>		<i>XBE444145</i> (1)	0.47–0.47
<i>Yr3</i>	<i>Lr38<sup>c</sup></i>		<i>Xunl207</i> (6)	1BL-14–1BL-9
<i>Yr8<sup>c</sup></i>	<i>Lr33</i>		<i>Xunl143b</i> (5)	0.61–0.73
	<i>Sr14</i>		<i>XBE473313</i> (1)	1BL-5–1BL-10
			<i>XBE406625</i> (1)	0.84–0.89
			<i>XBE605005</i> (1)	
			<i>Xunl111</i> (6)	
			<i>Xunl128</i> (1)	
Group 2				
<i>Bt1</i>	<i>Lr37</i>	<i>Lr16<sup>c</sup></i>	<i>XBE442854a</i> (2)	2BS-7 to telomere
<i>Bt7</i>	<i>Lr41</i>	<i>Lr17</i>	<i>XBE442802a</i> (2)	0.89–1.00
<i>Cre4</i>	<i>Sr40</i>	<i>Lr22</i>	<i>XBF200166</i> (1)	
<i>H20</i>	<i>Yr27</i>	<i>Lr38</i>	<i>XBE497251a</i> (1)	
<i>H21</i>		<i>Sr38</i>	<i>XBE499523</i> (1)	
<i>Lr11</i>		<i>Sr23</i>	<i>XBF484266</i> (1)	
<i>Lr45</i>		<i>QPm.sfr-A</i>	<i>Xunl216a</i> (2)	
<i>Pm6</i>		<i>Yr17</i>	<i>Xunl145</i> (1)	
<i>Sr32</i>		<i>Lr2</i>		2BS-3–2BS-14
<i>Sr34</i>		<i>Lr15<sup>c</sup></i>		0.79–0.84
<i>Sr39</i>		<i>Lr23</i>		
<i>Yr1</i>		<i>Lr28<sup>c</sup></i>		
<i>Yr16</i>		<i>Sr6</i>		
<i>Yr27</i>		<i>Lr13</i>	<i>XBE518335a</i> (4)	2BS-1–2DS-5
<i>Yr8<sup>c</sup></i>		<i>Sr19</i>	<i>XBE591764</i> (4)	0.47–0.53
<i>YrSte</i>		<i>Sr36</i>	<i>XBF483292<sup>e</sup></i> (1)	
<i>YrV23</i>			<i>Xunl147</i> (1)	
			<i>Xunl149</i> (5)	
	<i>Cre1</i>	<i>Lr35<sup>g</sup></i>	<i>XBE404328</i> (1)	2DL-4–2BL-3
	<i>Snb2</i>	<i>Sr21<sup>c</sup></i>	<i>Xunl110</i> (5)	0.26–0.35
	<i>Sr20</i>	<i>Sr9</i>		2DL-3–2DL-10
	<i>Ut-x</i>	<i>Yr7</i>		0.47–0.47
		<i>Yr5</i>		
		<i>Sr28</i>	<i>nd<sup>h</sup></i>	2DL-6 to telomere
		<i>Sr16</i>		0.95–1.00
		<i>Pm4</i>		
		<i>Qpm.sfr-D</i>		
		<i>Q.Fhs.ndsu-A</i>		

(continued)

**TABLE 5**  
(Continued)

Unknown	Short/long arm <sup>a</sup>	Wheat <i>R</i> loci	Candidate loci <sup>b</sup>	DL and FL value
Group 3				
<i>H24</i>	<i>Sr12</i>	<i>Lr27</i>	<i>XBE442854b</i> (3)	3BS-3 to telomere
<i>Rkn-mn1</i>	<i>Lr38<sup>c</sup></i>	<i>Lr32</i>	<i>XBE492937b</i> (3)	0.87–1.00
<i>Sr27</i>		<i>Pm13</i>	<i>XBF483203</i> (1)	
<i>YrS</i>		<i>Sr2</i>	<i>XBE444148</i> (1)	
<i>YrSte2</i> <i>YrTr2</i>		<i>YrnsB1</i>	<i>XBE499177</i> (1)	
		<i>QFhs.ndsu-B</i>	<i>XBF200008a</i> (1)	
		<i>Qpm.sfr-D</i>	<i>XBE637850<sup>e</sup></i> (3)	
			<i>Xunl167</i> (1)	
			<i>Xunl153</i> (1)	
	<i>Snb1</i>	<i>Sr12</i>	<i>XBE607075</i> (1)	3DL-1–3BL-2
	<i>SnbTM</i>		<i>Xunl150</i> (1)	0.23–0.27
		<i>Lr24<sup>†</sup></i>	<i>XBE405711</i> (1)	3AL-5 to telomere
		<i>Sr24</i>	<i>XBE499910</i> (1)	0.78–1.00
		<i>Sr35</i>	<i>Xunl130b</i> (5)	
			<i>XBF482366</i> (1)	
			<i>XBE591154</i> (1)	
			<i>Xunl132</i> (1)	
Group 4				
<i>H26</i>	<i>Yr28</i>		<i>XBE445244a</i> (1)	4DS-2 to telomere
<i>Lr12</i>			<i>Xunl216b</i> (2)	0.82–1.00
<i>Lr16<sup>c</sup></i>			<i>Xunl164b</i> (5)	
<i>Pm16</i>			<i>Xunl198b</i> (5)	
<i>Sr41</i>			<i>XBE403251b</i> (1)	4DS-1–4DS-3
<i>Yr22</i>			<i>XBE426789c</i> (1)	0.53 to 0.67
<i>YrCle</i>			<i>XBE426790c</i> (1)	
<i>YrHVIIYr</i>		<i>Lr25</i>	<i>Xunl198a</i> (5)	
<i>YrMor</i>		<i>Pm7</i>	<i>Xunl133</i> (5)	
<i>YrND</i>	<i>H25</i>	<i>Lr30</i>	<i>Xunl164a</i> (5)	4DL-1–4DL-3
<i>YrYam</i>	<i>Lr3</i>		<i>Xunl181</i> (5)	0.46–0.51
	<i>Lr30</i>	<i>Lr28<sup>c</sup></i>	<i>XBF201403</i> (1)	4BL-1 to telomere
	<i>Sr37</i>	<i>Sr7</i>	<i>XBE444404</i> (1)	0.86–1.00
	<i>Wsm1</i>	<i>Qpm.sfr-D</i>	<i>Xunl163</i> (6)	
			<i>XBE497251b</i> (1)	
			<i>Xunl124</i> (1)	
			<i>Xunl129</i> (5)	
Group 5				
<i>H3</i>	<i>Lr38<sup>c</sup></i>	<i>Pm2</i>	<i>XBE406845a<sup>f</sup></i> (1)	5DS-5–5AS-3
<i>H6</i>	<i>QPm.sfr-A</i>		<i>XBM136556a</i> (1)	0.67–0.75
<i>H7</i>	<i>Snb3</i>	<i>Crr</i>	<i>XBF201229</i> (3)	5BL-1–5BL-16
<i>H9</i>		<i>Lr18</i>	<i>XBE498327</i> (1)	0.55–0.78
<i>H10</i>		<i>Sr30<sup>g</sup></i>	<i>XBM137909</i> (1)	
<i>H12</i>		<i>Lr1<sup>g</sup></i>	<i>XBE406845b</i> (1)	
<i>H14</i>		<i>Qkb.cnl.5A.2</i>	<i>XBE445244b</i> (1)	
<i>H15</i>			<i>Xunl140<sup>e</sup></i> (3)	
<i>H16</i>			<i>XBM136556b</i> (1)	
<i>H17</i>			<i>XBF484437</i> (1)	
<i>H28</i>			<i>Xunl216c<sup>f</sup></i> (2)	
<i>H29</i>			<i>XBG604429</i> (1)	
<i>Pm23</i>			<i>XBF482744</i> (1)	
<i>Yr19</i>			<i>Xunl218</i> (5)	
<i>YrDa2</i>			<i>Xunl123</i> (1)	
<i>YrDru</i>			<i>Xunl119a</i> (1)	
			<i>Xunl174</i> (1)	
			<i>Xunl117</i> (1)	

(continued)



**TABLE 5**  
(Continued)

Unknown	Short/long arm <sup>a</sup>	Wheat <i>R</i> loci	Candidate loci <sup>b</sup>	DL and FL value
			<i>Xunl121</i> (6)	
			<i>Xunl104</i> (3)	
			<i>Xunl109</i> (5)	
			<i>Xunl152a</i> (2)	
			<i>Xunl113</i> (3)	
			<i>Xunl138</i> (6)	
			<i>Xunl141a</i> (5)	
			<i>Xunl166</i> (5)	
			<i>Xunl169</i> (5)	
			<i>Xunl191</i> (5)	
			<i>Xunl213a</i> (6)	
			<i>Xunl214a</i> (5)	
			<i>Xunl217</i> (5)	
			<i>XBE442802b<sup>c</sup></i> (3)	
		Group 6		
<i>Pm14</i>	<i>Cm1</i>	<i>Lr36</i>	<i>XBF478914</i> (1)	6BS-8 to telomere
<i>Pm24</i>	<i>Pm11</i>	<i>Sr5</i>	<i>XBG607637</i> (1)	0.90*–1.00
<i>Yr20</i>	<i>Pm12</i>	<i>Sr8</i>	<i>XBF200008b</i> (1)	
<i>Yr23</i>	<i>Pm21</i>	<i>Lr9</i>	<i>XBF478914</i> (1)	
<i>Yr4</i>	<i>Pm27</i>	<i>Qpm.sfr-B</i>	<i>Xunl190</i> (5)	
<i>YrD</i>	<i>Sr42</i>		<i>Xunl127</i> (1)	6DS-1–6AS-5
<i>YrDru</i>	<i>Yr26</i>		<i>Xunl122</i> (2)	0.54–0.65
<i>YrDru2Yr</i>			<i>Xunl213b</i> (6)	
<i>H46</i>			<i>Xunl214b<sup>f</sup></i> (5)	
<i>YrTr1</i>	<i>H23</i>	<i>H13</i>	<i>XBE500158</i> (1)	6AL-8 to telomere
<i>YrTy e</i>	<i>H24</i>	<i>Lr3</i>	<i>XBF199778</i> (1)	0.90–1.00
	<i>H25</i>	<i>Sr11</i>	<i>XBE403251a</i> (1)	
	<i>Lr38<sup>c</sup></i>	<i>Sr13</i>	<i>XBE426789a</i> (1)	
	<i>Pm20</i>	<i>Sr29</i>	<i>XBE426790a</i> (1)	
		<i>Sr26</i>	<i>XBM137173</i> (1)	
			<i>XBE405507<sup>c</sup></i> (1)	
			<i>Xunl154<sup>c</sup></i> (6)	
			<i>Xunl194<sup>c</sup></i> (1)	
		Group 7		
<i>Dn1</i>	<i>Bud1</i>		<i>XBE591190</i> (1)	7AS1 to telomere
<i>Gb3</i>	<i>Lr2</i>		<i>Xunl115</i> (5)	0.89–1.00
<i>Lr43</i>	<i>Pm15</i>		<i>Xunl120</i> (5)	
<i>Pm19</i>	<i>Pm18</i>	<i>Lr34<sup>g</sup></i>	<i>Xunl173</i> (1)	7AS-10–7DS-4
<i>Yr2</i>	<i>Sr44</i>	<i>Lr Mod</i>	<i>Xunl134</i> (5)	0.45–0.61
	<i>Yr18</i>		<i>Xunl157</i> (5)	
	<i>Yr6</i>		<i>Xunl214c<sup>c</sup></i> (5)	
	<i>Dn2</i>	<i>Pm5</i>	<i>XBE403270</i> (2)	7AL-11–7BL-7
	<i>Dn5</i>	<i>Lr19</i>	<i>BE443579</i> (2)	0.40–0.48
	<i>Lr20</i>	<i>Sr25</i>	<i>Xunl151</i> (6)	
	<i>LrVPM</i>	<i>Sr22</i>	<i>Xunl188</i> (5)	
	<i>Pch1</i>		<i>Xunl219</i> (5)	
	<i>Pch2</i>		<i>Xunl179</i> (5)	
	<i>Sr15</i>		<i>XBE518335b</i> (4)	7AL-5–7BL-5
	<i>Sr4</i>		<i>XBE426789b</i> (1)	0.63–0.69
			<i>XBE426790b</i> (1)	
			<i>Xunl170</i> (5)	7BL-10–7BL-13
			<i>Xunl204</i> (1)	0.78–0.79
		<i>Pm1</i>	<i>XBE405003</i> (1)	7BL-11 to telomere
		<i>Sr20</i>	<i>XBE405507a</i> (1)	0.86–1.00
		<i>Lr15<sup>c</sup></i>	<i>XBF200008c</i> (1)	

(continued)

**TABLE 5**  
(Continued)

Unknown	Short/long arm <sup>a</sup>	Wheat <i>R</i> loci	Candidate loci <sup>b</sup>	DL and FL value
		<i>Sr17</i>	<i>XBE518338a (1)</i>	
		<i>Pm9</i>	<i>Xunl194b (1)</i>	
		<i>Lr14a</i>		
		<i>QPm.sfr-B2</i>		
		<i>QLr.stf-B</i>		

<sup>a</sup> Bt, reaction to *Tilletia* spp.; Cm, reaction to *Eriophyes tulipae*; Cre, reaction to *Heterodera avenae*; Crr, reaction to *Cochliobolus sativus*; Dn, reaction to *Diuraphis noxia*; Gb, reaction to *Toxoptera graminum*; H, reaction to *Mayetiola destructor*; Kb, reaction to *Tilletia indica*; Lr, reaction to *Puccinia recondite*; Pch, reaction to *Pseudocercospora herpotrichoides*; Pm, reaction to *Erysiphe graminis*; Rkn, reaction to *Meloidogyne* spp.; Snb, reaction to *Phaeosphaeria nodorum*; Sr, reaction to *Puccinia graminis*; Ut, reaction to *Ustilago tritici*; Wsm, reaction to *Wheat streak mosaic virus*; Yr, reaction to *Puccinia striiformis*.

<sup>b</sup> The numbers in parentheses stand for the following: 1, NB/LRR type; 2, RLK type; 3, PK type; 4, *Hm1* type; 5, NB/kinase encoding; 6, PR type.

<sup>c</sup> The genes are reported to be mapped to more than one chromosome.

<sup>d</sup> QTL associated with markers with a <10-cM interval were included.

<sup>e</sup> Locus is placed in the regions; however, mapping information is lacking on A, B, or D genomes or on any two genomes.

<sup>f</sup> Locus is outside of the region in parentheses, although in very close proximity.

<sup>g</sup> *R*-gene loci may be outside of the region.

<sup>h</sup> Regions did not contain any candidate *R* gene.

<sup>i</sup> The FL values are given tentatively.

*R* gene (*DM3*; <http://www.ncbi.nlm.nih.gov/entrez>; SHEN *et al.* 2002). Our comparison for six major plants, each with an average of 200,000 ESTs, revealed that only ~5% of the RGAs are expressed (data not shown). The remaining are either nonfunctional or rare transcripts.

Pattern of gene expression is very different among the cloned *R* genes. About 81% of the *R* genes appear to be rare transcripts as no EST is present for 32% and only one or two for the remaining (data not shown). Expression of 19% of the cloned *R* genes is very high, making it very difficult to clone the rare transcripts by any random cloning and sequencing approaches. We

targeted the expressed *R* genes by using cDNA instead of the genomic DNA. We made our approach independent of expression by size separating on denaturing polyacrylamide gels. As a result, bands corresponding to abundant as well as rare transcripts were visible with similar intensity. Success of this approach can be realized from the fact that sequencing only 385 clones identified 121 unique sequences. Of these, 92 were rare transcripts, including 73 unique sequences for which no wheat homologs or ESTs were observed in the database. Only a subset of the identified sequences contained the GLPL motif, as observed among the cloned *R* genes.

**TABLE 6**  
Comparison of phenotypically characterized wheat *R* genes and mapped candidate *R*-gene sequences

Groups <sup>a</sup>	Morphologically confirmed <i>R</i> genes <sup>b</sup>			Candidate <i>R</i> -gene sequences <sup>c</sup>		
	Short	Long	Total	Short	Long	Total
1	16 (4)	— (5)	16 (44)	13	9	22
2	16 (4)	10 (4)	26 (51)	13	2	15
3	7 (2)	4 (2)	11 (17)	9	8	17
4	3 (1)	3 (5)	6 (24)	8	8	16
5	1 (2)	5 (1)	6 (25)	2	29	31
6	5 (7)	6 (5)	11 (34)	9	10	19
7	2 (7)	12 (8)	14 (34)	7	16	23
Total	50 (77)	40 (70)	90 (229)	61	82	143

<sup>a</sup> Each group represents a consensus physical chromosome in Figure 2.

<sup>b</sup> Numbers outside and inside the parentheses represent the number of wheat *R* genes located and not located, respectively, in any physical region in Table 4.

<sup>c</sup> Some sequences map to multiple loci.

These results suggest that using both GLPL and the poly(A)<sup>+</sup> tail primers was a worthwhile approach.

Various approaches to compare structural resemblance to the known *R* genes identified 99 *R*-gene candidates from the identified EST sequences. Including the 121 identified by RNA fingerprinting, all 220 expressed sequences showed varying levels of structural similarity to the known *R* genes. Detailed sequence analysis of these *R*-gene candidates is given by M. DILBIRLIGI and K. S. GILL (unpublished results). Of these *R*-gene candidates, 129 belong to the NB/LRR class, 38 to other types of *R* genes, and 17 are PR genes. The sequence similarity, presence, and location of the domains and motifs and other structural features of the 38 sequences matched with those of *Pto*, *Xa21*, *Hm1*, or *Hs1<sup>pm-1</sup>* *R* genes (Table 3). The remaining 36 of the 220 sequences had only a p-loop. No other *R*-gene motif or domain was observed, probably because of the shorter size of the sequences. Although many of these may structurally resemble known *R* genes, these sequences were not used for physical mapping.

A total of 229 genes conferring resistance to various wheat pests have been characterized and mapped on the chromosomes. Of these, 110 have been mapped relative to DNA markers whereas only arm location is known for 57. We were able to precisely localize 90 of these genes relative to deletion breakpoints (Figure 2; online supplemental Figure 4 at <http://www.genetics.org/supplemental/>; Table 5). It was not possible to physically map 20 *R* genes because the markers linked to these were mainly randomly amplified polymorphic DNA or amplified fragment length polymorphism types and thus were not present on the consensus genetic linkage map.

Although the phenotypically characterized *R* genes were present on all chromosomes, wheat homeologous groups 1 and 2 had the greatest number: 92 of the 229. Wheat groups 3 and 4 contained the lowest number: 44 of 229. Most of the wheat *R* genes are present in the telomeric or subtelomeric regions. About 75% of the wheat *R* genes mapped in the distal 20% of the chromosomes. All 90 *R* genes were clustered in ~12% of the wheat genome present as 20 small chromosomal regions. Five of these regions contained major *R*-gene clusters, accounting for ~50% of the *R* genes (Figure 3). Clustering of *R* genes has also been observed in other plant species such as *Arabidopsis*, soybean, maize, and lettuce (BOTELLA *et al.* 1997; MEYERS *et al.* 1999; CHIN *et al.* 2001; RICHLI *et al.* 2002). This clustering of *R* genes may have resulted from duplication events caused by unusual types of recombination that have been observed around *R* genes of various plants (CHIN *et al.* 2001; SUN *et al.* 2001; SHEN *et al.* 2002).

The distribution of the *R*-gene candidates on wheat chromosomes was very similar to that of the phenotypically characterized *R* genes. All 121 *R*-gene candidates mapped in 26 chromosomal regions, including the 20

in which phenotypically characterized wheat *R* genes mapped. The relative distribution of *R*-gene candidates among various regions roughly matched that of the known *R* genes (Figure 3; Table 5). The five major *R* gene cluster regions that account for ~50% of the known *R* genes contained 43 of the *R*-gene candidates (Figure 3; Table 5). Further characterization is necessary for finding a particular *R*-gene candidate corresponding to a particular disease.

We thank Olin Anderson for providing the EST clones and Bikram Gill for seed of the aneuploid stocks. We also thank Michael G. Hanly for critically reading the manuscript. This work was supported by a joint contribution of the Agricultural Research Center, Washington State University (journal series no. 207-03), and the University of Nebraska (journal series no. 14286). Financial support for the project also came from the Turkish government, the Nebraska Agricultural Experiment Station Hatch Funds, the Department of Agronomy and Horticulture, the University of Nebraska, the Vogel Endowment funds, and the National Science Foundation.

#### LITERATURE CITED

- AARTS, M. G., B. TE LINTEL HEKKERT, E. B. HOLUB, J. L. BEYNON, W. J. STIEKEMA *et al.*, 1998 Identification of *R*-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. *Mol. Plant-Microbe Interact.* **11**: 251–258.
- AKHUNOV, E. D., A. W. GOODYEAR, S. GENG, L. L. QI, B. ECHALIER *et al.*, 2003 The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res.* **13**: 753–763.
- ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHAEFFER, J. ZHANG, Z. ZHANG *et al.*, 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389–3402.
- ARABIDOPSIS GENOME INITIATIVE, 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796–815.
- ARUMUGANATHAN, K., and E. D. EARLE, 1991 Estimation of nuclear DNA content of plants by flow cytometry. *Plant Mol. Biol. Rep.* **9**: 229–233.
- BENT, A. F., B. N. KUNKEL, D. DAHLBECK, K. L. BROWN, R. SCHMIDT *et al.*, 1994 RPS2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. *Science* **265**: 1856–1860.
- BITTNER-EDDY, P. D., I. R. CRUTE, E. B. HOLUB and J. L. BEYNON, 2000 RPP13 is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. *Plant J.* **21**: 177–188.
- BOTELLA, M. A., M. J. COLEMAN, D. E. HUGHES, M. T. NISHIMURA, J. D. JONES *et al.*, 1997 Map positions of 47 *Arabidopsis* sequences with sequence similarity to disease resistance genes. *Plant J.* **12**: 1197–1211.
- BOTELLA, M. A., J. E. PARKER, L. N. FROST, P. D. BITTNER-EDDY, J. L. BEYNON *et al.*, 1998 Three genes of the *Arabidopsis* RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. *Plant Cell* **10**: 1847–1860.
- BRUEGGEMAN, R., N. ROSTOKS, D. KUDRNA, A. KILIAN, F. HAN *et al.*, 2002 The barley stem rust-resistance gene Rpg1 is a novel disease-resistance gene with homology to receptor kinases. *Proc. Natl. Acad. Sci. USA* **20**: 20.
- CAI, D., M. KLEINE, S. KIFLE, H. J. HARLOFF, N. N. SANDAL *et al.*, 1997 Positional cloning of a gene for nematode resistance in sugar beet. *Science* **275**: 832–834.
- CHIN, D. B., R. ARROYA-GARCIA, O. E. OCHOA, R. V. KESSELLI, D. O. LAVELLE *et al.*, 2001 Recombination and spontaneous mutation at the major cluster of resistance genes in lettuce (*Lactuca sativa*). *Genetics* **157**: 831–849.
- COLLINS, N. C., C. A. WEBB, S. SEAH, J. G. ELLIS, S. H. HULBERT *et al.*, 1998 The isolation and mapping of disease resistance gene analogs in maize. *Mol. Plant-Microbe Interact.* **11**: 968–978.

- COLLINS, N., J. DRAKE, M. AYLIFFE, Q. SUN, J. ELLIS *et al.*, 1999 Molecular characterization of the maize Rpl-D rust resistance haplotype and its mutants. *Plant Cell* **11**: 1365–1376.
- COOLEY, M. B., S. PATHIRANA, H. J. WU, P. KACHROO and D. F. KLESSIG, 2000 Members of the Arabidopsis HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell* **12**: 663–676.
- DILBIRLIGI, M., and K. S. GILL, 2004 Identification and analysis of expressed resistance gene sequences in wheat. *Plant Mol. Biol.* (in press).
- DIXON, M. S., D. A. JONES, J. S. KEDDIE, C. M. THOMAS, K. HARRISON *et al.*, 1996 The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* **84**: 451–459.
- ENDO, T. R., and B. S. GILL, 1996 The deletion stocks of common wheat. *J. Hered.* **87**: 295–307.
- ERNST, K., A. KUMAR, D. KRISELEIT, D. U. KLOOS, M. S. PHILLIPS *et al.*, 2002 The broad-spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. *Plant J.* **31**: 127–136.
- FEUILLET, C., and B. KELLER, 1999 High gene density is conserved at syntenic loci of small and large grass genomes. *Proc. Natl. Acad. Sci. USA* **96**: 8265–8270.
- FEUILLET, C., and B. KELLER, 2002 Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. *Ann. Bot.* **89**: 3–10.
- GILL, B. S., B. FRIEBE and T. R. ENDO, 1991 Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome* **34**: 830–839.
- GILL, K. S., B. S. GILL 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.
- GRANT, M. R., L. GODIARD, E. STRAUBE, T. ASHFIELD, J. LEWALD *et al.*, 1995 Structure of the Arabidopsis RPM1 gene enabling dual specificity disease resistance. *Science* **269**: 843–846.
- HALTERMAN, D., F. ZHOU, F. WEI, R. P. WISE and P. SCHULZE-LEFERT, 2001 The MLA6 coiled-coil, NBS-LRR protein confers AvtMla6-dependent resistance specificity to *Blumeria graminis* f. sp. hordei in barley and wheat. *Plant J.* **25**: 335–348.
- HAMMOND-KOSACK, K., and J. JONES, 1997 Plant disease resistance genes. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**: 575–608.
- JOHAL, G. S., and S. P. BRIGGS, 1992 Reductase activity encoded by the HM1 disease resistance gene in maize. *Science* **258**: 985–987.
- KANAZIN, V., L. F. MAREK and R. C. SHOEMAKER, 1996 Resistance gene analogs are conserved and clustered in soybean. *Proc. Natl. Acad. Sci. USA* **93**: 11746–11750.
- KAWCHUK, L. M., J. HACHEY, D. R. LYNCH, F. KULCSAR, G. VAN ROOIJEN *et al.*, 2001 Tomato Ve disease resistance genes encode cell surface-like receptors. *Proc. Natl. Acad. Sci. USA* **98**: 6511–6515.
- LAGUDAH, E. S., O. MOULLET and R. APPELS, 1997 Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the Cre3 nematode resistance locus of wheat. *Genome* **40**: 659–665.
- LAWRENCE, G. J., E. J. FINNEGAN, M. A. AYLIFFE and J. G. ELLIS, 1995 The L6 gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene RPS2 and the tobacco viral resistance gene N. *Plant Cell* **7**: 1195–1206.
- LEISTER, D., A. BALLVORA, F. SALAMINI and C. GEBHARDT, 1996 A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nat. Genet.* **14**: 421–429.
- LIU, G. Z., L. Y. PI, J. C. WALKER, P. C. RONALD and W. Y. SONG, 2002 Biochemical characterization of the kinase domain of the rice disease resistance receptor-like kinase XA21. *J. Biol. Chem.* **277**: 20264–20269.
- MARTIN, G. B., S. H. BROMMONSCHENKEL, J. CHUNWONGSE, A. FRARY, M. W. GANAL *et al.*, 1993 Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**: 1432–1436.
- MCDOWELL, J. M., M. DHANDAYDHAM, T. A. LONG, M. G. AARTS, S. GOFF *et al.*, 1998 Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *Plant Cell* **10**: 1861–1874.
- MCINTOSH, R. A., G. E. HART, K. M. DEVOS, M. D. GALE and W. J. ROGERS, 2000 Catalogue of gene symbols for wheat supplemental, p. 235 in *Proceedings of the 9th International Wheat Genetics Symposium*, edited by A. E. SLINKARD. University Extension Press, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.
- MEYERS, B. C., A. W. DICKERMAN, R. W. MICHELMORE, S. SIVARAMAKRISHNAN, B. W. SOBRAL *et al.*, 1999 Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.* **20**: 317–332.
- MILLIGAN, S. B., J. BODEAU, J. YAGHOUBI, I. KALOSHIAN, P. ZABEL *et al.*, 1998 The root knot nematode resistance gene Mi from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. *Plant Cell* **10**: 1307–1319.
- MONDRAGON-PALOMINO, M., B. C. MEYERS, R. W. MICHELMORE and B. S. GAUT, 2002 Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana. *Genome Res.* **12**: 1305–1315.
- NOEL, L., T. L. MOORES, E. A. VAN DER BIEZEN, M. PARNISKE, M. J. DANIELS *et al.*, 1999 Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of Arabidopsis. *Plant Cell* **11**: 2099–2112.
- ORI, N., Y. ESHED, I. PARAN, G. PRESTING, D. AVIV *et al.*, 1997 The I2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell* **9**: 521–532.
- PANSTRUGA, R., R. BUSCHGES, P. PIFFANELLI and P. SCHULZE-LEFERT, 1998 A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. *Nucleic Acids Res.* **26**: 1056–1062.
- PIMENTAL, D., C. WILSON, C. MCCALLUM, R. HUANG, P. DWEN *et al.*, 1997 Ecological and environmental benefits of biodiversity. *Bioscience* **47**: 747–757.
- RICHLY, E., J. KURTH and D. LEISTER, 2002 Mode of amplification and reorganization of resistance genes during recent Arabidopsis thaliana evolution. *Mol. Biol. Evol.* **19**: 76–84.
- SAKAMOTO, K., Y. TADA, Y. YOKOZAKI, H. ARAGI, N. HAYASHI *et al.*, 1999 Chemical induction of disease resistance in rice is correlated with the expression of a gene encoding a nucleotide binding site and leucine-rich repeats. *Plant Mol. Biol.* **40**: 847–855.
- SALMERON, J. M., G. E. OLDROYD, C. M. ROMMENS, S. R. SCOFFIELD, H. S. KIM *et al.*, 1996 Tomato Prf is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the Pto kinase gene cluster. *Cell* **86**: 123–133.
- SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- SANDHU, D., and K. S. GILL, 2002a Gene-containing regions of wheat and the other grass genomes. *Plant Physiol.* **128**: 803–811.
- SANDHU, D., and K. S. GILL, 2002b Structural and functional organization of the '1S0.8 gene-rich region' in the Triticeae. *Plant Mol. Biol.* **48**: 791–804.
- SANMIGUEL, P., A. TIKHONOV, Y. K. JIN, N. MOTCHOULSKAIA, D. ZAKHAROV *et al.*, 1996 Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**: 765–768.
- SANMIGUEL, P. J., W. RAMAKRISHNA, J. L. BENNETZEN, C. S. BUSO and J. DUBCOVSKY, 2002 Transposable elements, genes and recombination in a 215-kb contig from wheat chromosome 5A(m). *Funct. Integr. Genomics* **2**: 70–80.
- SEAH, S., K. SIVASITHAMPARAM, A. KARAKOUSIS and E. S. LAGUDAH, 1998 Cloning and characterization of a family of disease resistance gene analogs from wheat and barley. *Theor. Appl. Genet.* **97**: 937–945.
- SEAH, S., W. SPIELMEYER, J. JAHIER, K. SIVASITHAMPARAM and E. S. LAGUDAH, 2000 Resistance gene analogs within an introgressed chromosomal segment derived from *Triticum ventricosum* that confers resistance to nematode and rust pathogens in wheat. *Mol. Plant-Microbe Interact.* **13**: 334–341.

- SEARS, E. R., 1954 The aneuploids of common wheat. *Univ. Mo. Agric. Exp. Stn.* **572**: 1–58.
- SHEN, K. A., B. C. MEYERS, M. N. ISLAM-FARIDI, D. B. CHIN, D. M. STELLY *et al.*, 1998 Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. *Mol. Plant-Microbe Interact.* **11**: 815–823.
- SHEN, K. A., D. B. CHIN, R. ARROYO-GARCIA, O. E. OCHOA, D. O. LAVELLE *et al.*, 2002 Dm3 is one member of a large constitutively expressed family of nucleotide binding site-leucine-rich repeat encoding genes. *Mol. Plant-Microbe Interact.* **15**: 251–261.
- STEIN, N., C. FEUILLET, T. WICKER, E. SCHLAGENHAUF and B. KELLER, 2000 Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L. spans the Lr10 resistance locus in hexaploid wheat (*Triticum aestivum* L.). *Proc. Natl. Acad. Sci. USA* **97**: 13436–13441.
- SONG, W. Y., G. L. WANG, L. L. CHEN, H. S. KIM, L. Y. PI *et al.*, 1995 A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science* **270**: 1804–1806.
- SUN, Q., N. C. COLLINS, M. AYLIFFE, S. M. SIMITH, J. DRAKE *et al.*, 2001 Recombination between paralogues at the Rp1 rust resistance locus in maize. *Genetics* **158**: 423–438.
- TRAUT, T. W., 1994 The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *Eur. J. Biochem.* **15**: 9–19.
- WANG, Z. X., M. YANO, U. YAMANOUCHI, M. IWAMOTO, L. MONNA *et al.*, 1999 The Pib gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. *Plant J.* **19**: 55–64.
- WARREN, R. F., A. HENK, P. MOWERY, E. HOLUB and R. W. INNES, 1998 A mutation within the leucine-rich repeat domain of the *Arabidopsis* disease resistance gene RPS5 partially suppresses multiple bacterial and downy mildew resistance genes. *Plant Cell* **10**: 1439–1452.
- WENDEL, J. F., 2000 Genome evolution in polyploids. *Plant Mol. Biol.* **42**: 225–249.
- WICKER, T., N. STEIN, L. ALBAR, C. FEUILLET, E. SCHLAGENHAUF *et al.*, 2001 Analysis of a contiguous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. *Plant J.* **26**: 307–316.
- 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.
- YOSHIMURA, S., U. YAMANOUCHI, Y. KATAYOSE, S. TOKI, Z. X. WANG *et al.*, 1998 Expression of Xa1, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. *Proc. Natl. Acad. Sci. USA* **95**: 1663–1668.
- YU, Y. G., G. R. BUSS and M. A. MAROOF, 1996 Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. *Proc. Natl. Acad. Sci. USA* **93**: 11751–11756.

Communicating editor: J. A. BIRCHLER

