DOI: 10.1534/genetics.108.099614

Evolution of New Disease Specificity at a Simple Resistance Locus in a Crop–Weed Complex: Reconstitution of the *Lr21* Gene in Wheat

Li Huang,* Steven Brooks,† Wanlong Li,‡ John Fellers,§ James C. Nelson and Bikram Gill‡,1**

**Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, Montana 59717-3150, †U. S. Department of Agriculture–Agricultural Research Service, Dale Bumpers National Rice Research Center, Stuttgart, Arkansas 72160, ‡Department of Plant Pathology, Wheat Genetic and Genomic Resources Center, Kansas State University, Manhattan, Kansas 66506-5502, §U. S. Department of Agriculture–Agricultural Research Service–Plant Science and Entomology Research Unit, Department of Plant Pathology, Kansas State University, Manhattan, Kansas 66506-5502 and **Department of Plant Pathology, Kansas State University, Manhattan, Kansas 66506-5502*

Manuscript received December 11, 2008
Accepted for publication April 3, 2009

ABSTRACT

The wheat leaf-rust resistance gene *Lr21* was first identified in an Iranian accession of goatgrass, *Aegilops tauschii* Coss., the D-genome donor of hexaploid bread wheat, and was introgressed into modern wheat cultivars by breeding. To elucidate the origin of the gene, we analyzed sequences of *Lr21* and *lr21* alleles from 24 wheat cultivars and 25 accessions of *Ae. tauschii* collected along the Caspian Sea in Iran and Azerbaijan. Three basic nonfunctional *lr21* haplotypes, H1, H2, and H3, were identified. *Lr21* was found to be a chimera of H1 and H2, which were found only in wheat. We attempted to reconstitute a functional *Lr21* allele by crossing the cultivars Fielder (H1) and Wichita (H2). Rust inoculation of 5876 F_2 progeny revealed a single resistant plant that proved to carry the H1H2 haplotype, a result attributed to intragenic recombination. These findings reflect how plants balance the penalty and the necessity of a resistance gene and suggest that plants can reuse "dead" alleles to generate new disease-resistance specificity, leading to a "death–recycle" model of plant-resistance gene evolution at simple loci. We suggest that selection pressure in crop–weed complexes contributes to this process.

PLANTS possess large numbers of resistance genes (*R* gene) as a part of an elaborate plant defense system. In different plants, an *R*-gene locus may consist of a single-copy (simple) or of multiple copies of *R* genes (complex) in clusters as a result of gene duplication events. This duplication is considered as the birth of an *R* gene. *R* genes are necessary for plants to respond to pathogen attacks and to survive when pathogens are in the environment. Mutations, gene conversion, and recombination were found to be the means to create new specificities for various pathogens (for review, Leister 2004). However, an *R* gene could bring a penalty when the pathogen is absent (Stahl *et al.* 1999). In such a case, plants have better fitness when they get rid of the *R*-gene function. In nature, the presence of different pathogens maintains the diversity of *R*-gene specificities. So far, there is no report on the fates of nonfunctional *R* genes.

In native agricultural ecosystems, wild plants often grow as weeds intermixed with or adjacent to their crop relatives. Extensive gene flow occurs between wild and domesticated forms, spawning numerous crop landraces adapted to diverse environments and occasionally new species. Common (hexaploid or bread) wheat (*Triticum aestivum* L., $2n = 6x = 42$, genome formula AABBDD) arose from such a process by hybridization of domesticated tetraploid wheat (*T. turgidum* L., $2n = 4x = 28$, AABB) with goatgrass (*Aegilops tauschii* Coss., $2n = 2x = 14$, DD) growing as a weed in farmers' fields along the Caspian Sea in Iran *ca.* 8000 years ago (Kihara 1944; McFadden and Sears 1946; Nesbitt and Samuel 1998).

Because of the pivotal importance of *Ae. tauschii* in wheat evolution and crop improvement, Kihara *et al.* (1965) gathered extensive collections from Iran, Afghanistan, and adjacent regions. They suggested that Caspian Iran was the center of the genetic diversity of *Ae. tauschii*, a proposition later confirmed by molecular-marker analysis (Lubbers *et al.* 1991), as well as of resistance to leaf rust. Nine named and 12 new leaf-rust resistance genes have been documented in *Ae. tauschii*, and many more remain to be identified (Gill *et al.* 2008). Leaf rust, a scourge of wheat since before Roman times, is caused by the fungus *Puccinia triticina* (Eriks). It attacks mainly the leaf blade, producing small, elliptical, orange-red pustules on the upper surface, causing premature defoliation that results in as much as a 40% yield loss (McIntosh *et al.* 1995).

Supporting information is available online at http://www.genetics.org/cgi/content/full/genetics.108.099614/DC1.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. FJ876280-FJ876295.

[1]*Corresponding author:* Department of Plant Pathology, Wheat Genetic and Genomic Resources Center, Kansas State University, Manhattan, KS 66506-5502. E-mail: bsgill@ksu.edu

TABLE 1

Triticum and Aegilops accessions used for sequencing of *Lr21* and *lr21* alleles

Accession	Species	Known *Lr* genes	Collection site	Polymorphism pattern[a]
TA1599	*Ae. tauschii* spp. *tauschii*	*Lr21*	Ramsar, Iran	1
TA2527	*Ae. tauschii* spp. *tauschii*	*Lr21*	Ramsar, Iran	2
TA2528	*Ae. tauschii* spp. *tauschii*	*Lr21*	Ramsar, Iran	2
TA2529	*Ae. tauschii* spp. *tauschii*	*Lr21*	Ramsar, Iran	4
TA2530	*Ae. tauschii* spp. *tauschii*	*Lr21*	Ramsar, Iran	3
TA2472	*Ae. tauschii* spp. *tauschii*	*Lr21*	Ramsar, Iran	2
TA2378	*Ae. tauschii* spp. *tauschii*	*Lr21*	9 km from Ramsar, Iran	1
TA2476	*Ae. tauschii* spp. *tauschii*		13 km from Ramsar, Iran	9
TA2473	*Ae. tauschii* spp. *tauschii*		23 km from Ramsar, Iran	9
TA1649	*Ae. tauschii* spp. *tauschii*	*Lr21*	Bandar-e Anzali, Iran	1
TA2477	*Ae. tauschii* spp. *tauschii*		Bandar-e Anzali, Iran	6
TA2481	*Ae. tauschii* spp. *tauschii*		12 km from Bandar-e Anzali, Iran	7
TA2468	*Ae. tauschii* spp. *strangulata*	*Lr21*	51 km North of Babolsar, Iran	1
TA2469	*Ae. tauschii* spp. *tauschii*		51 km North of Babolsar, Iran	10
TA2470	*Ae. tauschii* spp. *strangulata*	*Lr21*	51 km North of Babolsar, Iran	3
TA2471	*Ae. tauschii* spp. *strangulata*		51 km North of Babolsar, Iran	7
TA2467	*Ae. tauschii* spp. *tauschii*	*Lr42*	51 km North of Babolsar, Iran	7
TA2450	*Ae. tauschii* spp. *strangulata*	*Lr39*	51 km North of Babolsar, Iran	6
TA1670	*Ae. tauschii* spp. *tauschii*	*Lr21*	Kutkashen, Azerbaijan	1
TA10110	*Ae. tauschii* spp. *tauschii*		East of Chrelet Kopetdag mountain range, Turkmenistan	5
TA2496	*Ae. tauschii* spp. *tauschii*		Tabriz, Iran	10
TA2495	*Ae. tauschii* spp. *tauschii*		Tabriz, Iran	10
TA1698	*Ae. tauschii* spp. *tauschii*		Dagestan, Russia	7
TA1699	*Ae. tauschii* spp. *tauschii*		Dagestan, Russia	7
TA1691	*Ae. tauschii* spp. *tauschii*	*Lr21*	Unknown	3
TA3009	*T. aestivum* cv. Wichita			9
TA3908	*T. aestivum* cv. Fielder			9

[a] Polymorphism patterns are based on KSUD14-STS marker analysis (Figure 2, A and B).

One *Ae. tauschii* accession, TA1599, collected in Caspian Iran, carries a gene named *Lr21* that confers resistance to all known *P. triticina* races. *Lr21*, transferred to wheat in the 1970s (Rowland and Kerber 1974; McIntosh *et al.* 1995), was recently cloned (Huang *et al.* 2003) and shown to be a simple (single-copy) locus encoding a nucleotide-binding site–leucine-rich repeats (NBS–LRR) protein of 1080 amino acids. Here we report how a simple locus such as *Lr21* evolved novel resistance specificities in a unique crop–weed system and how fragments of nonfunctional alleles could be reused in this process.

MATERIALS AND METHODS

Plant materials: Twenty-five accessions of *Ae. tauschii* were used for this study (Table 1). Of these, 12 are the *Lr21* carriers identified from the entire collection of 528 accessions of *Ae. tauschii* collected over a large geographic area representing its genetic diversity and maintained by the Wheat Genetic and Genomic Resources Center (WGGRC) at Manhattan, Kansas. Additional *Lr* genes are present in two accessions: *Lr39* in TA2450 and *Lr42* in TA2467. The remaining accessions carry no known *Lr* genes and are susceptible to leaf rust in the field. Among the 13 *lr21* accessions, 5 were sampled at the same collection sites as the *Lr21* accessions, 3 were collected along the Caspian Sea of Iran within 51 km of the *Lr21* accessions, and the remaining 5 were from places where no *Lr21* accessions were found.

Of 24 wheat cultivars (Table 1 and Table S1) with the *lr21* allele that were tested for polymorphism using the KSUD14-STS marker, a PCR-based molecular marker that distinguishes *Lr21* from *lr21* (Huang *et al.* 2003), "Fielder," a spring wheat, and "Wichita," a winter wheat, were chosen for this study. WGRC7 is a wheat germplasm developed by WGGRC (http://www.k-state.edu/wgrc/) by direct crossing Wichita with *Ae. tauschii* accession TA1649 and then backcrossing with Wichita twice (Raupp *et al.* 1983).

One dicot species, *Arabidopsis thaliana*, and five cereal species, barley (*Hordeum vulgare*), oat (*Avena sativa*), rye (*Secale cereale*), maize (*Zea mays*), and rice (*Oryza sativa*) were chosen on the basis of the evolutionary time line to assess the approximate age of the *Lr21* locus.

DNA manipulation and sequence analysis: DNA isolation, digestion, blotting, and Southern hybridization followed the protocols described by Qi *et al.* (2004). Several pairs of primers were designed on the basis of coding regions and flanking sequences of *Lr21* from *Ae. tauschii* accession TA1649. Primers Sta (TTGTGATGGAGAAACGAGTGGCC) and Tor (CGGAC GAGTAGTTCTTTCAGGA) were designed to amplify the entire gene and 397-bp flanking regions (Figure 1). Each full-length allele was amplified by long-range PCR using Herculase-enhanced DNA polymerase (Stratagene, La Jolla, CA) from genomic DNA of each accession. The PCR products were then cloned directly using the pGEM-T easy system

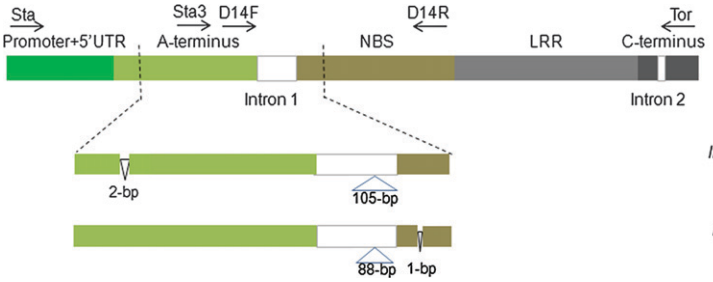

FIGURE 1.—Gene structure and primer location of the *Lr21* and *lr21* alleles. The enlarged regions show the major differences between the *lr21*-Wichita and the *lr21*-Fielder. A-terminus, amino terminus; NBS, nucleotide-binding site; LRR, leucine-rich repeats; and C terminus, carboxy terminus; 5′-UTR, 5′-untranslated region.

(Promega, Madison, WI). First, vector primers SP6 and T7 were used to sequence the ends of each clone. Internal primers were designed later on the basis of the sequences obtained from previous primers. At least three clones from each accession were sequenced from both directions. All the sequences were assembled using MacVector 6.5.3 (Oxford Molecular, Madison, WI). Primers Sta3 (TGGCTAATGCAGT GGGCACGG) and D14-R (GGACATTAGGCGATGCTTTGAA TTC) were used to amplify the NBS region of the alleles (Figure 1). The marker KSUD14-STS was designed on the basis of a 105- or 88-bp insertion/deletion (indel) in the first intron of the *Lr21* (Figure 1). A 1.36-kb fragment from this region is a signature of the *Lr21*, while a 1465-or 1448-bp fragment with a 105- or 88-bp insertion is a tag of *lr21*. Other sizes of fragments amplified with KSUD14-STS represent *Lr21* paralogs (Figure 2, A and B).

Gene expression study: Expression of the *Lr21* and *lr21* alleles was characterized by modified quantitative RT–PCR (KASHKUSH *et al.* 2003). The mRNA was isolated from leaf tissues with or without inoculation of the pathogen isolate PRTUS6 using MicroPoly(A) Pure (Ambion, Austin, TX). First-strand cDNAs were synthesized using oligo(dT) primer and second-strand cDNAs were amplified with gene-specific primers. *Lr21* or *lr21* were amplified with D14-F (CGAGAT TGGTCCTATGAGGTGGT) and D14-R (Figure 1). Actin gene expression was used for normalization for the expression study. Actin-F (GGTATCGTGAGCAACTGGGATG) and Actin-R (GTGAAGGAGTAACCTCTCTCGGTG) were used to amplify a 383-bp fragment. PCR was performed under the following conditions: 95° for 4 min and 12 cycles each with 95° for 30 sec, 60° for 30 sec, and 72° for 1.5 min. The amplicons were separated by electrophoreses on 1% agarose gels, transferred to Hybond-N$^+$ membranes (Amersham Biosciences, Piscataway, NJ), and probed with KSUD14 or the actin gene.

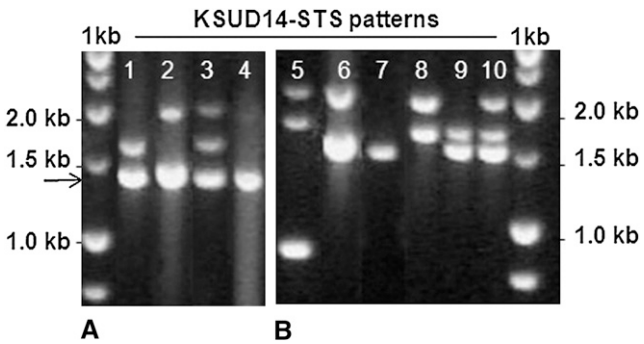

FIGURE 2.—Polymorphism survey of *Ae. tauschii* based on the KSUD14-STS marker. (A) Four patterns (1–4) were revealed among 12 *Lr21* accessions. The 1.36-kb fragment (indicated by an arrow) is a tag of *Lr21*, and other size fragments are *Lr21* paralogs. (B) Six patterns (5–10) were identified among the 13 *lr21* accessions, none of which has the 1.36-kb fragment.

RESULTS

The *Lr21* NBS–LRR family in cereal species: Hybridization of the *Lr21* NBS region to wheat, rye, barley, and oat revealed multiple bands (Figure 3), indicating that *lr21* homologs are present elsewhere in the genome and that the *Lr21* NBS–LRR family is shared by this group of grasses with a basic chromosome number of 7. No signal was detected in maize, rice, and Arabidopsis (Figure 3), indicating that *Lr21* homologs are absent in these

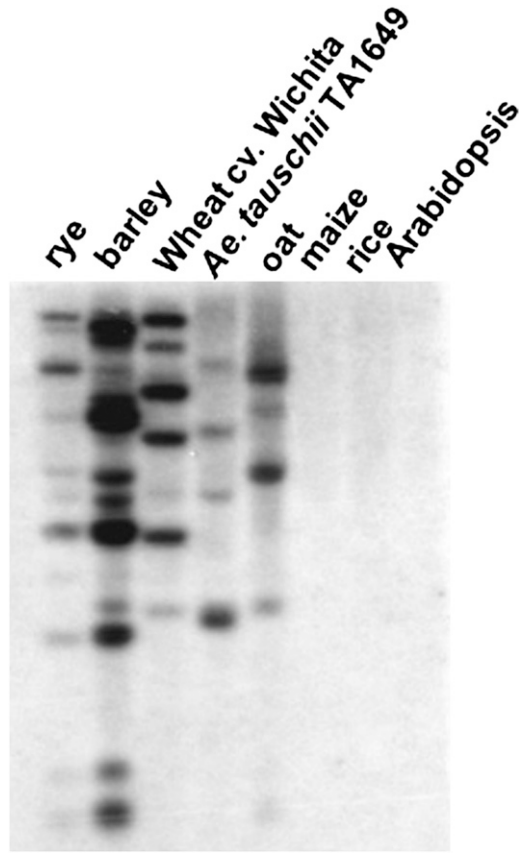

FIGURE 3.—Southern hybridization of genomic DNAs digested with restriction enzyme *Xba*I and probed with KSUD14 (NBS region of *Lr21*). The hybridization stringency amounted to 80% homology.

species. Primers designed on the basis of the *Lr21* NBS region were able to amplify fragments from rye, barley, and oat by PCR. Sequencing of the amplicons showed 95–98% similarity to the *Lr21* NBS sequence (File S1). BLASTn searches against the finished rice genome sequence revealed eight rice homologs with *E*-values > −20 and identity from 57 to 70%, consistent with the failure of hybridization. None of the homologs were detected on rice chromosome 5, which is homeologous to the wheat group 1 chromosomes including 1D, where the *Lr21* locus resides. Southern hybridization with probe KSUD14, a part of the *Lr21* gene, revealed two *Lr21* paralogs located on the short arm of chromosome 1D in the *Lr21* donor accession of *Ae. tauschii* (Huang *et al.* 2003). Two cosmid clones carrying the paralogs were identified. Full-length sequencing of the paralogs showed them to be NBS–LRR-like sequences with 80% identity to *Lr21* in the NBS region (File S2) and only ~50% identity in the rest of the gene.

Only one *Lr21* allele was identified: Of 528 accessions representing the geographic diversity of *Ae. tauschii* that were screened against a mixture of leaf-rust pathogens, 158 were found to be resistant at the seedling stage. The resistant accessions were tested with the *Lr21*-specific marker KSUD14-STS, and 12 were found to carry the *Lr21* allele. All 12 had been collected along the Caspian Sea in Iran and Azerbaijan within a range of 675 km (Figure 4 and Table 1). They shared an identical 1.36-kb fragment at the *Lr21* locus but were polymorphic at paralogous loci, with four patterns revealed by marker KSUD14-STS (Figure 2A). Amplification, cloning, and alignment of the *Lr21* sequences showed that all 12 accessions carry an identical allele at the *Lr21* locus.

The *lr21* alleles are pseudogenes: To elucidate the mechanisms underlying the origin of *Lr21* function, we characterized nonfunctional *lr21* alleles from 13 *Ae. tauschii* accessions (Figure 4 and Table 1) and two *Ae. tauschii*-derived *lr21* alleles present in the D genome of two hexaploid wheat cultivars (Table 1). On the basis of KSUD14-STS genotyping in *Ae. tauschii*, six different polymorphic patterns were observed, but none contained the 1.36-kb fragment (Figure 2B). On the basis of sequencing, 11 different *lr21* alleles were identified from the 13 *Ae. tauschii* accessions, of which TA2467, TA2473, and TA2481 carried the same allele (Figure 5, and Figure S1). The wheat D-genome alleles *lr21-f* (Fielder) and *lr21-w* (Wichita) were not detected in the sampled *Ae. tauschii* accessions. The sequences of the 11 *lr21 Ae. tauschii* and two wheat D-genome *lr21* alleles revealed a spectrum of the accumulated variation at this locus (Figure 5). There were 36 single-nucleotide polymorphisms (SNPs), none introducing a stop codon and only 11 representing nonsynonymous substitutions. There were also 15 indels, 3 of which caused frameshifts that introduced premature stop codons. With reference to the *Lr21* sequence, indel 3 (a 2-bp deletion at position +761) was present in all *lr21* alleles except *lr21-f* and

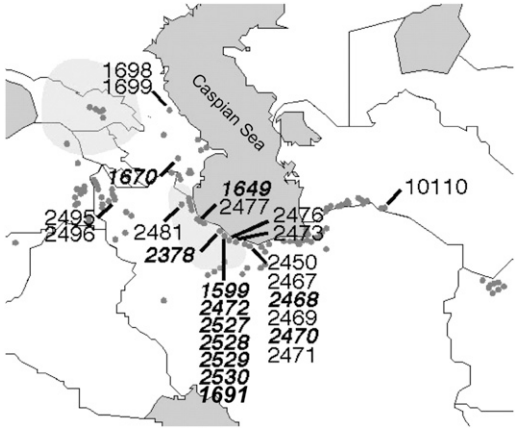

Figure 4.—The sampling sites and geographic distribution of the 25 *Ae. tauschii* accessions used in this study. Twelve accessions (in boldface italic type) are the *Lr21* carriers.

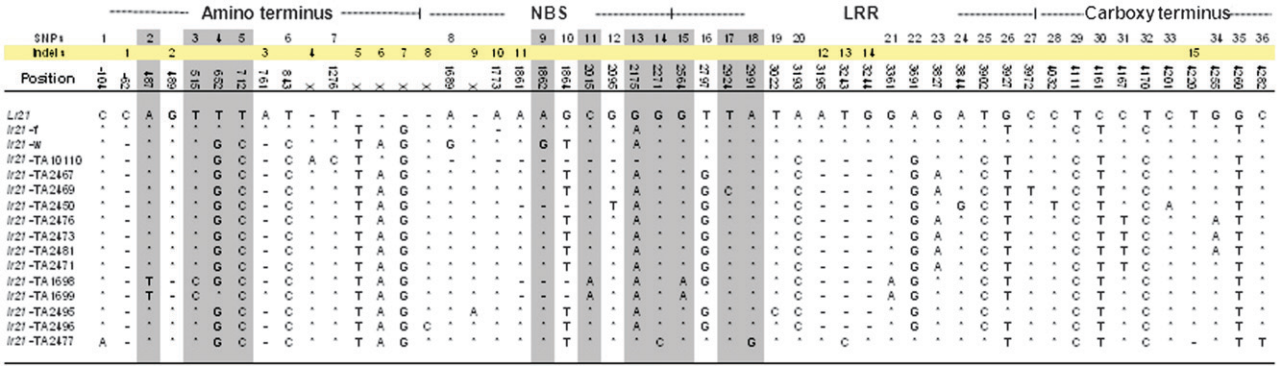

Figure 5.—Sequence variation among tested alleles of *Lr21* and *lr21*. Position numbers are based on the *Lr21* sequence and are shown as offsets from the transcription start site. *lr21-f* is the Fielder allele and *lr21-w* is the Wichita allele. SNP, single-nucleotide polymorphism; indel, insertion/deletion. An asterisk indicates that the sequence is identical to *Lr21*, and a hyphen notes a deletion. Shaded positions are those in which the nucleotide change resulted in an amino-acid change. The X's represent deletions in *Lr21*. Four functional domains are indicated above the sequence: the amino terminus, NBS (nucleotide-binding site), LRR (leucine-rich repeat), and the carboxy terminus.

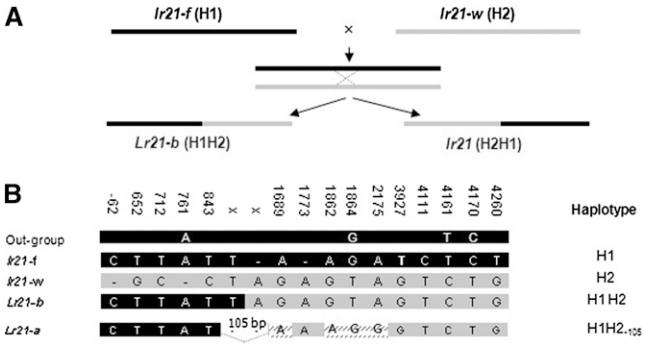

FIGURE 6.—(A) A crossing scheme for reconstituting a functional *Lr21* allele from two nonfunctional alleles. (B) The indel and SNPs that are shared among the three outgroup species rye, barley, and oat and the region in which intragenic recombination between H1 and H2 resulted in the chimerical allele *Lr21-b*. *Lr21-a* is the allele identified from *Ae. tauschii*. Position numbers are based on the *Lr21* sequence. In A and B, solid areas correspond to haplotype H1, representing allele *lr21-f* from cultivar Fielder, and shaded areas correspond to H2, representing *lr21-w* from Wichita. Hatching indicates SNPs that distinguish *Lr21-b* from *Lr21-a*.

lr21-TA1699. This deletion introduced an early stop codon, which resulted in a putative 151-aa peptide. Similarly, indel 10, a 1-bp deletion at position +1773, was identified only in the *lr21-f* allele, which encodes a putative 380-aa peptide. Indel 12 was a 4-bp deletion at +3195 identified in 12 *lr21* alleles, including the *lr21-TA1699* allele. Each *lr21* allele carried one or two of these three indels and was thus a nonfunctional pseudogene.

The two distinct *lr21* alleles found in the wheat D genome, *lr21-w* and *lr21-f*, were identified from chromosome 1D of Wichita and Fielder, respectively (Figure 1). The *lr21-w* and *lr21-f* alleles were distributed among the other 22 sampled wheat cultivars (Table S1) without bias as to winter or spring habit. Both encoded truncated proteins, resulting from a 2-bp indel at position +761 in *lr21-w* and a 1-bp indel at +1773 in *lr21-f* (Figure 5), and thus were pseudogenes.

The *Lr21* allele is a recombined allele of recent origin: Sequences from the NBS and part of the 3′ regions of the gene from rye, barley, and oat revealed that the 2-bp insertion at indel 3 (position +761) (Figure 6B) and nucleotides G at position +1864, T at position +4161, and C at position +4170 are shared among the three outgroup species. The presence in the *lr21-f* allele of this 2-bp insertion at indel 3 and the ancestral SNPs (Figure 6B) support its ancestral character. Sequence comparisons suggested that the 13 different *lr21* alleles were derived from three basic haplotypes and were subsequently modified by point mutation and insertion/deletion events (Figure 5 and Figure S1). We designated allele *lr21-f* as haplotype H1, *lr21-w* as haplotype H2, and *lr21-TA2467* as haplotype H3. The *Lr21* allele appears to be a chimera derived from intragenic recombination between H1 and H2 (Figure 5 and Figure S1), followed by the deletion of a 105-bp segment within the first intron, and was designated $H1H2_{-105}$. The putative crossover site lies between positions +762 and +1772. The recombined allele could encode a full 1080-aa protein and would be free of both the 2-bp (in H2) and the 1-bp (in H1) deletions. Another chimeric haplotype, H2H1, appeared in the *lr21-TA2477* allele, a putative product of reciprocal recombination carrying the 2-bp deletion (Figure 5 and Figure S1), and the lr21 protein product of this allele is truncated. The remaining *lr21* alleles are suggested to have been derived from the H3 haplotype as shown in the supplemental data (Figure S1). All of the *lr21* alleles are present as truncated pseudogenes but are transcribed as revealed by RT–PCR (Figure S2), suggesting that their promoters are still functioning.

A functional allele can be created from two dead alleles: To test the hypothesis that *Lr21* could originate from two dead (nonfunctional) alleles, we crossed wheat cultivars Wichita (H2) and Fielder (H1) (Figure 6A) and screened 5876 F_2 progeny in the greenhouse using the leaf-rust isolate PRTUS6. One plant, a putative recombinant designated as F/W-R, was identified on the basis of having an infection type lower than that of both parents. The critical region between +762 and +1772 was PCR amplified from the F/W-R plant, cloned, and sequenced. Six of nine clones had sequences identical to H2. The other three had sequences identical to H1 from the 5′-end to position +843 and thereafter were identical to H2 (Figure 6B). These data indicated that F/W-R was heterozygous at the *Lr21* locus for H2 and H1H2 haplotypes. The recombinant H1H2 was identical to the *Lr21* of the *Ae. tauschii* gene (*Lr21-a* in Figure 6B) except for a 105-bp insertion derived from H2 and four substitutions: G to A at positions +1689 and +1862, T to G at position +1864, and A to G at position +2175 (Figure 5 and Figure 6B). The presence of this insertion and the four SNPs showed that the reconstituted *Lr21-b* functional allele was created in this cross. The 105-bp sequence lies in the first intron of the gene, and the deletion or insertion of this fragment does not change the length of the peptide encoded by the gene. Two of four substitutions are synonymous, while the one at position +1862 changes a methionine to a valine (both neutral and nonpolar), and the second at +2175 changes the acidic polar aspartic acid to the neutral nonpolar glycine. However, the reconstituted allele *Lr21-b* still confers resistance to the same pathogen isolate, indicating that the amino-acid changes at these positions do not change the function of the protein. The F/W-R plant was selfed. Fifty progenies were tested for a reaction to leaf-rust isolate PRTUS6. Progeny testing revealed that resistance to the leaf-rust isolate was conferred by a single dominant gene (Figure 7). Ten resistance and 10 susceptible plants were selected for genotyping using the critical region of the *Lr21-b*. The result suggested that the resistance was conferred by the *Lr21-b* allele.

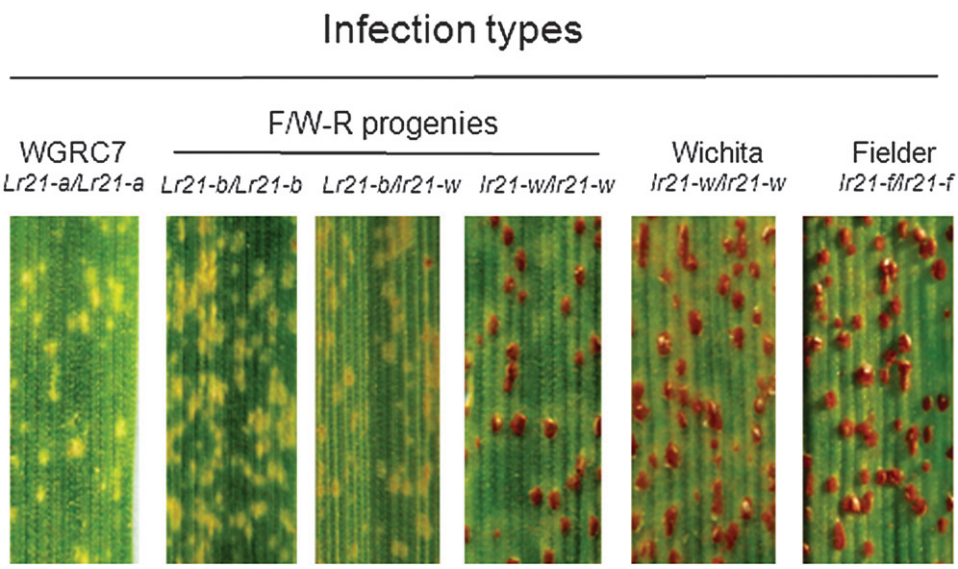

FIGURE 7.—Infection types of WGRC7 (*Lr21-a/Lr21-a*), progenies of the F/W recombinant (*Lr21-b/Lr21-b*; *Lr21-b/lr21-w*; *lr21-w/lr21-w*), Fielder (*lr21-f/lr21-f*) and Wichita (*lr21-w/lr21-w*) 9 days after inoculation with the leaf-rust isolate PRTUS6.

DISCUSSION

We selected the *Lr21* gene for map-based cloning >10 years ago because the gene was identified in a dozen *Ae. tauschii* accessions spread over a large area in Caspian Iran. Our hypothesis was that *Lr21* was probably a complex locus and was spawning new specificities as a result of unequal crossing over similar to the *Rp1* locus of maize (HULBERT *et al.* 2001). It was surprising when molecular cloning revealed *Lr21* to be a simple, single-copy locus (HUANG *et al.* 2003). After investigating the sequence variation at the *Lr21* locus in the 12 *Lr21*-carrier *tauschii* accessions and the 15 *lr21* alleles in a sample of *tauschii* and bread wheat accessions, our results showed an unexpected monomorphism at the *Lr21* locus among the 12 *Lr21*-carriers. The sequence data revealed *Lr21* to be a chimeric allele, providing possible clues to its recent origin through intragenic recombination. This hypothesis was experimentally verified by the recovery of a functional allele in the progeny of a cross between two susceptible parents. These results have important implications about the age of the *Lr21* locus, its origin, evolution, and fixation in the context of the crop–weed coevolutionary process as distinct from resistance evolution in wild populations; these aspects of this study are discussed below.

An ancient locus with a young allele of *Lr21* NBS–LRR: The evolutionary time line indicates that wheat diverged from rice and maize ~65 MYA, from barley ~14 MYA, and from rye ~7 MYA (HUANG *et al.* 2002). Compared to 80% identity with its paralogs (File S2), *Lr21* shared >95% identity with barley, oat, and rye homologs in the NBS region (File S1). Since, in general, orthologous genes in different species are more similar in sequence to one another than paralogous copies within a species (MICHELMORE and MYERS 1998; HULBERT *et al.* 2001), it is plausible that *Lr21* is an ancient locus shared by wheat, barley, and oat. However, its restricted geographic distribution and DNA-level monomorphism strongly suggest that the *Lr21* allele of *Ae. tauschii* originated more recently in a single event. In this scenario, it most likely spread by rare cross-pollination among *Ae. tauschii* populations, farming activity, and commerce of wheat grains contaminated with goatgrass seeds.

Chimeric origin of *Lr21*: *Ae. tauschii* is a self-pollinated species with an outcrossing rate of <5%. The presence in the wheat D genome of the H1 and H2 haplotypes indicates that these alleles were present in *Ae. tauschii* (donor of the wheat D genome), growing alongside domesticated tetraploid wheat. It thus appears that *Ae. tauschii* parents carrying H1 and H2 or similar haplotypes were involved in hybridization events leading to the origin of bread wheat *ca.* 8000 years ago and that bread wheat originated in at least two independent hybridization events (TALBERT *et al.* 1998). As H1 and H2 haplotypes were detected only in bread wheat, our failure to detect *Ae. tauschii* accessions carrying haplotypes H1 and H2 most probably was due to limited sampling. Alternatively, these haplotypes may be extinct in *Ae. tauschii* and preserved only in wheat.

The molecular mechanism underlying the origin of *Lr21* function may be associated with its location at the most distal point of the chromosome 1D short arm, a recombination hot spot (SPIELMEYER *et al.* 2000; QI *et al.* 2004). We previously reported one intragenic recombination event between positions −61 and +1354 involving alleles *Lr21* and *lr21-w* in a sample of 332 F_2 plants (HUANG *et al.* 2003). It involved a conversion tract of a minimum of 191 bp and a maximum of 1415 bp of DNA from *lr21-w* to *Lr21*, rendering the latter ineffective. We have now experimentally reconstituted *Lr21* through another intragenic recombination event. In addition, one H2H1 haplotype, an obvious product of intragenic recombination, was detected in a small sample of 13 *lr21* alleles. These findings suggest that recombination events

have occurred multiple times at the *Lr21* locus and that the evolutionary history of *Lr21* has been shaped by its location in a high-recombination region.

"Birth–recycle" at the *Lr21* locus: The presence of only one functional allele among an assortment of nonfunctional alleles at the *Lr21* locus suggests the cost of carrying the resistance allele in the absence of virulent pathogen strains. Two evolutionary classes of NBS–LRR genes have been characterized. One supports the so-called "arms race" model represented by the *L* locus of flax (Ellis *et al.* 1999) and the *RPP13* locus of Arabidopsis (Rose *et al.* 2004), which contain large numbers of different functional alleles and a high degree of variation in the regions responsible for specificity. The other class, consistent with a "trench warfare" model, is represented by the *RPM1* (Stahl *et al.* 1999) and *RPS2* (Mauricio *et al.* 2003) loci of Arabidopsis in which variation is low with no evidence of diversifying selection between functional and nonfunctional forms. The functional *RPM1* allele has been shown to impose a penalty in the absence of the pathogen. Complete deletion is one way to remove the deleterious effect of an allele such as *RPM1*. An alternative way is truncation, as seen with the *lr21* alleles.

Our discovery suggests a "death–recycle" model of plant-resistance gene evolution at simple loci. A "birth-and-death" process similar to that of the vertebrate major histocompatibility complex, T-cell receptor, and immunoglobulin genes has been proposed to explain the evolution of resistance genes at complex loci (Michelmore and Myers 1998). At a simple locus, there is no "birth" associated with gene duplication. A functional allele may become ineffective because of mutation or defeat by a new race of the pathogen and then may be reused in the creation of a new functional allele at that locus. Our results have confirmed that plants can reuse nonfunctional alleles to create new resistance specificity. The recycling of *lr21* hints at the potential usefulness of truncated alleles. New resistances similar to *Lr21* that occurred in nature may also arise in plant breeding programs more often than recognized because of extensive selection pressure for rare new disease specificities in segregating populations subjected to disease epidemics.

The wheat–goatgrass complex and the fixation of *Lr21*: *Ae. tauschii* accessions carrying the functional *Lr21* allele are predominant in regions where agriculture was practiced, to the point where at one location, Ramsar, Iran, all collected accessions carried this allele (Figure 4 and Table 1). This predominance would be expected if some evolutinary process in the crop–weed agroecosystem led to fixation of a new disease-resistance specificity created by a rare recombination event in *lr21*. Agricultural practice favors crop monoculture, or a single variety planted over large areas for long periods of time. This often leads to much higher rust-disease pressure in wheat fields than in wild *Ae. tauschii* populations, which harbor mixtures of different rust-resistance genes or even several leaf-rust resistance genes in a single accession (Gill *et al.* 2008). We propose that leaf-rust epidemics in a crop monoculture imposed selection pressure on goatgrass populations in or near wheat fields. A plant carrying the *Lr21* allele would have a fitness advantage in an environment with high leaf-rust inoculum. In this model, the crop–weed complex coevolutinary process was critical to the selection and retention of the *Lr21* gene in *Ae. tauschii* populations.

It appears that native agricultural ecosystems, located in Vavilovian world centers of crop plant origin, are virtual outdoor laboratories for the creation of genetic variation. Other "new" genes spawned by such ecosystems may well be of the same worldwide economic significance as *Lr21*. The example presented here argues for the careful preservation of the native agricultural ecosystems in the face of modern agricultural practices because the success of modern plant breeding hinges on the extensive use of genetic variation present in land races and in wild relatives of crop plants.

We are grateful to Scot Hulbert, Frank White, and Eduard Akhunov for critical review of the manuscript. This research was supported by the Kansas Wheat Commission and a U. S. Department of Agriculture–Cooperative State Research, Education, and Extension Service special grant to the Wheat Genetic and Genomic Resources Center. This is contribution 07-305-J from the Kansas Agricultural Experiment Station, Manhattan, Kansas.

LITERATURE CITED

Ellis, J. G., G. J. Lawrence, J. E. Luck and P. N. Dodds, 1999 Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. Plant Cell **11:** 495–506.

Gill, B. S., L. Huang, V. Kuraparthy, W. J. Raupp, D. L. Wilson *et al.*, 2008 Alien genetic resources for wheat leaf rust resistance, cytogenetic transfer, and molecular analysis. Aust. J. Agric. Res. **59:** 197–205.

Huang, S., A. Sirikhachornkit, X. Su, J. Faris, B. S. Gill *et al.*, 2002 Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc. Natl. Acad. Sci. USA **99:** 8133–8138.

Huang, L., S. A. Brooks, W. Li, J. P. Fellers, H. N. Trick *et al.*, 2003 Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. Genetics **164:** 655–664.

Hulbert, S. H., C. A. Webb, S. M. Smith and Q. Sun, 2001 Resistance gene complexes: evolution and utilization. Annu. Rev. Phytopathol. **39:** 285–312.

Kashkush, K., M. Feldman and A. A. Levy, 2003 Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nat. Genet. **33:** 102–106.

Kihara, H., 1944 Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. Agric. Hortic. **19:** 13–14 (in Japanese).

Kihara, H., K. Yamashita, and M. Tanaka, 1965 *Results of the Kyoto University Scientific Expedition to the Karakoram and Hindukush*. Kyoto University, Kyoto, Japan.

Leister, D, 2004 Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet. **20:** 116–122.

Lubbers, E. L., K. S. Gill, T. S. Cox and B. S. Gill, 1991 Variation of molecular markers among geographically diverse accessions of *Triticum tauschii*. Genome **34:** 354–361.

Mauricio, R., E. A. Stahl, T. Korves, D. Tian, M. Kreitman *et al.*, 2003 Natural selection for polymorphism in the disease

resistance gene *Rps2* of *Arabidopsis thaliana*. Genetics **163:** 735–746.

McFadden, E. S., and E. R. Sears, 1946 The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J. Hered. **37:** 81–89, 107–116.

McIntosh, R. A., C. R. Wellings and R. F. Park, 1995 *Wheat Rusts: An Atlas of Resistance Genes*. Kluwer Academic Publishers, Boston.

Michelmore, R. W., and B. C. Meyers, 1998 Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. **8:** 1113–1130.

Nesbitt, M., and D. Samuel, 1998 Wheat domestication: archaeobotanical evidence. Science **279:** 1431.

Qi, L. L., B. Echalier, S. Chao, G. R. Lazo, G. E. Butler *et al.*, 2004 A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. Genetics **168:** 701–712.

Raupp, W. J., L. E. Browder and B. S. Gill, 1983 Leaf rust resistance in *Aegilops squarrosa*: its transfer and expression in common wheat (*Triticum aestivum* L.). Phytopathology **73:** 818.

Rose, L. E., P. D. Bittner Eddy, C. H. Langley, E. B. Holub, R. W. Michelmore *et al.*, 2004 The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. Genetics **166:** 1517–1527.

Rowland, G. G., and E. R. Kerber, 1974 Telocentric mapping in hexaploid wheat of genes for leaf rust resistance and other characters derived from *Aegilops* squarrosa. Can. J. Genet. Cytol. **16:** 137–144.

Spielmeyer, W., O. Moullet, A. Laroche and E. S. Lagudah, 2000 Highly recombinogenic regions at seed storage protein loci on chromosome 1DS of *Aegilops tauschii*, the D-genome donor of wheat. Genetics **155:** 361–367.

Stahl, E. A., G. Dwyer, R. Mauricio, M. Kreitman and J. Bergelson, 1999 Dynamics of disease resistance polymorphism at the *Rpm1* locus of Arabidopsis. Nature **400:** 667–671.

Talbert, L. E., L. Y. Smith and N. K. Blake, 1998 More than one origin of hexaploid wheat is indicated by sequence comparison of low-copy DNA. Genome **41:** 402–407.

Communicating editor: E. J. Richards

GENETICS

Supporting Information
http://www.genetics.org/cgi/content/full/genetics.108.099614/DC1

Evolution of New Disease Specificity at a Simple Resistance Locus in a Crop–Weed Complex: Reconstitution of the *Lr21* Gene in Wheat

Li Huang, Steven Brooks, Wanlong Li, John Fellers, James C. Nelson and Bikram Gill

DOI: 10.1534/genetics.108.099614

FILE S1

Alignment of NBS regions among *Lr21* of *Ae. tauschii* and homologs of other species

```
Lr21_NBS_tauschii      GTTGGTGTCACAACTTATCAGGTACATCCAATTTCGTCGCCCACTCCTCATCTTATATAT
Lr21_NBS_rye           GTTGGTGTCACAACTTATCAGGTACATCCAATTTCGTCGCCCACTCCTCATCTTATATAT
Lr21_NBS_oat           GTTGGTGTCACAACTTATCAGGTACATCCAATTTCGTCGCCCACTCCTCATCTTATATAT
Lr21_NBS_barley        GGTGGTGTCACAACTTATCAGGTACATCCAATTTCGTCGCCCACTCCTCATCTTATATAT
                       * ************************************************************

Lr21_NBS_tauschii      ATGTACTCTATATACATGTAAAAGCTCTCCGTCCATGTGTTAAGAAAGATTCTTTTCTGA
Lr21_NBS_rye           ATGTACTCTATATACATGTAAAAGCTCTCCGTCCATGTGTTAAGAAAGATTCTTTTCTGA
Lr21_NBS_oat           ATGTACTCTATATACATGTAAAAGCTCTCCGTCCATGTGTTAAGAAAGATTCTTTTCTGA
Lr21_NBS_barley        ATGTACTCTATATACATGTAAAAGCTCTCCGTCCATGTGTTAAGAAAGATTCTTTTCTGA
                       ************************************************************

Lr21_NBS_tauschii      ACTTTATGTGTCATGCATCGTACTACAATATCTTTCTAATCCGTTACTATGATAATGAGA
Lr21_NBS_rye           ACTTTATGTGTCATGCATCGTACTACAATATCTTTCTAATCCGTTACTATGATAATGAGA
Lr21_NBS_oat           ACTTTATGTGTCATGCATCGTACTACAATATCTTTCTAATCCGTTACTATGATAATGAGA
Lr21_NBS_barley        ACTTTATGTGTCATGCATCGTACTACAATATCTTTCTAATCCGTTACTATGATAATGAGA
                       ************************************************************

Lr21_NBS_tauschii      GTAATACTAATTATCTTATTTGACATCAGTCATTGAGGTCAACAAAGAAAACCTGTATTT
Lr21_NBS_rye           GTAATACCAATTATCTTATTTGACATCAGTCATTGAGGTCAACAAAGAAAACCTGTATTT
Lr21_NBS_oat           GTAATACTAATTATCTTATTTGACATCAGTCATTGAGGTCAACAAAGAAAACCTGTATTT
Lr21_NBS_barley        GTAATACTAATTATCTTATTTGACATCAGTCATTGAGGTCAACAAAGAAAACCTGTATTT
                       ******* ****************************************************

Lr21_NBS_tauschii      GAAAAAAAAATGTGATCCATACGATAGATGGACAATATAGTTATCATTATCATATTTCCC
Lr21_NBS_rye           GAAAAAAAAATGTGATCCATACGATAGATGGACAATATAGTTATCATTATCATATTTCCC
Lr21_NBS_oat           GAAAAAAAAATGTGATCCATACGATAGATGGACAATATAGTTATCATTATCATATTTCCC
Lr21_NBS_barley        GAAAAAAAAATGTGATCCATACGATAGATGGACAATATAGTTATCATTATCATATTTCCC
                       ************************************************************

Lr21_NBS_tauschii      TTGTTTATTATTTACACTTTAATACTATTTCTAATAGATAGAATAGACATGTGACAGACA
Lr21_NBS_rye           TTGTTTATTATTTACACTTTAATACTATTTCTAATAGATAGAATAGACATGTGACAGACA
Lr21_NBS_oat           TTGTTTATTATTTACACTTTAATACTATTTCTAATAGATAGAATAGACATGTGACAGACA
Lr21_NBS_barley        TTGTTTATTATTTACACTTTAATACTATTTCTAATAGATAGAATAGACATGTGACAGACA
                       ************************************************************

Lr21_NBS_tauschii      TACCCATTTGAGATTTTGCATCTTGCAGGAGAATGCCACTGCATTGGATGTTGTCCTTAC
Lr21_NBS_rye           TACCCATTTGAGATTTTGCATCTTGCAGGAGAATGCCACTGCATTGGATGTTGTCCTTAC
Lr21_NBS_oat           TACCCATTTGAGATTTTGCATCTTGCAGGAGAATGCCACTGCATTGGATGTTGTCCTTAC
Lr21_NBS_barley        TACCCATTTGAGATTTTGCATCTTGCAGGAGAATGCCACTGCATTGGATGTTGTCCTTAC
                       ************************************************************

Lr21_NBS_tauschii      TGCTATCTCAAGATGGAACTTGAATAAAAGA----------ATAGAGAAGGTACAAAGTA
Lr21_NBS_rye           TGCTATCTCAAGATGGAACTTGAATAAAAGAATAGAGAAGAATAGAGAAGGTACAAAGTA
Lr21_NBS_oat           TGCTATCTCAAGATGGAACTTGAATAAAAGA----------ATAGAGAAGGTACAAAGTA
Lr21_NBS_barley        TGCTATCTCAAGATGGAACTTGAATAAAAGA----------ATAGAGAAGGTACAAAGTA
                       ******************************          ******************

Lr21_NBS_tauschii      CCATCAGCGAAGTGACGAAGTCACCGCTCTTGGGCACGGCAAGCAAGAGTGCACCAGACG
Lr21_NBS_rye           CCATCAGCGAAGTGACGAAGTCACCGCTCTTGGGCACGGCAAGCAAGAGTGCACCAGACG
Lr21_NBS_oat           CCATCAGCGAAGTGACGAAGTCACCGCTCTTGGGCACGGCAGGCAAGAGTGCACCAGACG
Lr21_NBS_barley        CCATCAGCGAAGTGACGAAGTCACCGCTCTTGGGCACGGCAAGCAAGAGTGAACCAGACG
                       ****************************************** ********* ********

Lr21_NBS_tauschii      ATATTGCTAACAAGAATAGGAGTAGAATTAGAACTGCTAGCAAGCGGAAGGTATTTGGTC
Lr21_NBS_rye           ATATTGCTAACAAGAATAGGAGTAGAATTAGAACTGCTAGCAAGCGGAAGGTATTTGGTC
Lr21_NBS_oat           ATATTGCTAACAAGAATAGGAGTAGAATTAGAACTGCTAGCAAGCGGAAGGTATTTGGTC
Lr21_NBS_barley        ATATTGCTAACAAGAATAGGAGTAGAATTAGAACTGCTAGCAAGCGGAAGGTATTTGGTC
                       ************************************************************

Lr21_NBS_tauschii      GAGAGGGGCTGCGTGATCATATCATGGTAAGGCTTCGTGAGATACCAGAGCATGA-TGCA
Lr21_NBS_rye           GAGAGGGGCTGCGTGATCATATCATGGTAAGGCTTCGTGAGATACCAGAGCATGA-TGCA
Lr21_NBS_oat           GAGAGGGGCTGCGTGATCATATCATGGTAAGGCTTCGTGAGATACCAGAGCATGA-TGCA
Lr21_NBS_barley        GAGAGGGGCTGCGTGATCATATCATGGTAAGGCTTCGTGAGATACCCGAGGATGAATGCA
```

```
                    ******************************************** *** **** ****

                                                          indel 3
Lr21_NBS_tauschii   GCAAGCTCAAGTGCTGATCCATGTTACTCAGTGATTGGTATATATGGTGTTGCTGGGTCT
Lr21_NBS_rye        GCAAGCTCAAGTGCTGATCCATGTTACTCAGTGATTGGTATATATGGTGTTGCTGGGTCT
Lr21_NBS_oat        GCAAGCTCAAGTGCTGATCCATGTTACTCAGTGATTGGTATATATGGTGTTGCTGGGTCT
Lr21_NBS_barley     GCAAGCTCAAGTGCTGATCCATGTTACTCAGTGATTGGTATATATGGTGTTGCTGGGTCT
                    ************************************************************

Lr21_NBS_tauschii   GGGAAGACCACGTTTGCAGGATATATTCAAGATTACATAAAGGAGGAATGCAAGGATGAG
Lr21_NBS_rye        GGGAAGACCACGTTTGCAGGATATATTCAAGATTACATAAAGGAGGAATGCAAGGATGAG
Lr21_NBS_oat        GGGAAGACCACGTTTGCAGGATATATTCAAGATTACATAAAGGAGGAATGCAAGGATGAG
Lr21_NBS_barley     GGGAAGACCACGTTTGCAGGATATATTCAAGATTACATAAAGGAGGAATGCAAGGATGAG
                    ************************************************************

Lr21_NBS_tauschii   AAACTTTTCGACACCATCATGTGCATTCATGTGACTGAAACTTTCAGTGTGGATGATATA
Lr21_NBS_rye        AAACTTTTCGACACCATCATGTGCATTCATGTGACTGAAACTTTCAGTGTGGATGATATA
Lr21_NBS_oat        AAACTTTTCGACACCATCATGTGCATTCATGTGACTGAAACTTTCAGTGTGGATGATATA
Lr21_NBS_barley     AAACTTTTCGACACCATCATGTGCATTCATGTGACTGAAACTTTCAGTGTGGATGATATA
                    ************************************************************
                              1864
Lr21_NBS_tauschii   TTT-CATGAAATGCTGAAGTATATTACCGGAGATAGTCACTCCAATATTTCAGATCGTGG
Lr21_NBS_rye        TTT-CATGAAATGCTGAAGTATATTACCGGAGATAGTCACTCCAATATTTCAGATCGTGG
Lr21_NBS_oat        TTT-CATGAAATGCTGAAGTATATTACCGGAGATAGTCACTCCAATATTTCAGATCGTGG
Lr21_NBS_barley     TTTTCATGAAATGCTGAAGTATATTACCGGAGATAGTCACTCCAATATTTCAGATCGTGG
                    *** ********************************************************

Lr21_NBS_tauschii   GGCTCTAGATAAGAAGTTGAAGGAAGCATTGTGTGGCAAACGTTTCTTCTTGATATTGGA
Lr21_NBS_rye        GGCTCTAGATAAGAAGTTGAAGGAAGCATTGTGTGGCAAACGTTTCTTCTTGATATTGGA
Lr21_NBS_oat        GGCTCTAGATAAGAAGTTGAAGGAAGCATTGTGTGGCAAACGTTTCTTCTTGATATTGGA
Lr21_NBS_barley     GGCTCTAGATAAGAAGTTGAAGGAAGCATTGTGTGGCAAACGTTTCTTCTTGATATTGGA
                    ************************************************************

Lr21_NBS_tauschii   TGATCTCTGGGTGAAAAACAAGAATGACCAACACCTAGAGGAGCTAATCTCTCCACTCAA
Lr21_NBS_rye        TGATCTCTGGGTGAAAAACAAGAATGACCAACACCTAGAGGAGCTAATCTCTCCACTCAA
Lr21_NBS_oat        TGATCTCTGGGTGAAAAACAAGAATGACCAACACCTAGAGGAGCTAATCTCTCCACTCAA
Lr21_NBS_barley     TGATCTCTGGGTGAAAAACAAGAATGACCAACACCTAGAGGAGCTAATCTCTCCACTCAA
                    ************************************************************

Lr21_NBS_tauschii   TGTTGGGCTGAAAGGAAGCAAAATCCTGGTGACGGCTCGAACAAAAGAAGCAGCTGGAGC
Lr21_NBS_rye        TGTTGGGCTGAAAGGAAGCAAAATCCTGGTGACGGCTCGAACAAAAGAAGCAGCTGGAGC
Lr21_NBS_oat        TGTTGGGCTGAAAGGAAGCAAAATCCTGGTGACGGCTCGAACAAAAGAAGCAGCTGGAGC
Lr21_NBS_barley     TGTTGGGCTGAAAGGAAGCAAAATCCTGGTGACGGCTCGAACAAAAGAAGCAGCTGGAGC
                    ************************************************************

Lr21_NBS_tauschii   TCTGGGTGCCGATAAATTTATTGAAATGCCTGATTTGGATGAGGATCAGTACTTGGCGAT
Lr21_NBS_rye        TCTGGGTGCCGATAAATTTATTGAAATGCCTGATTTGGATGAGGATCAGTACTTGGCGAT
Lr21_NBS_oat        TCTGGGTGCCGATAAATTTATTGAAATGCCTGATTTGGATGAGGATCAGTACTTGGCGAT
Lr21_NBS_barley     TCTGGGTGCCGATAAATTTATTGAAATGCCTGATTTGGATGAGGATCAGTACTTGGCGAT
                    ************************************************************

Lr21_NBS_tauschii   GTTTATGCATTATGCGCTAAGTGGTACAAGAGTTGCCCTTCAAGAATTTGAACAAGTTGG
Lr21_NBS_rye        GTTTATGCATTATGCGCTAAGTGGTACAAGAGTTGCCCTTCAAGAATTTGAACAAGTTGG
Lr21_NBS_oat        GTTTATGCATTATGCGCTAAGTGGTACAAGAGTTGCCCTTCAAGAATTTGAACAAGTTGG
Lr21_NBS_barley     GTTTATGCATTATGCGCTAAGTGGTACAAGAGTTGCCCTTCAAGAATTTGAACAAGTTGG
                    ************************************************************

Lr21_NBS_tauschii   GAGAGAGATTGCCAAAAAACTACACCGATCACCTATTGCAGCAGTAACAGTTGCAGGACG
Lr21_NBS_rye        GAGAGAGATTGCCAAAAAACTACACCGATCACCTATTGCAGCAGTAACAGTTGCAGGACG
Lr21_NBS_oat        GAGAGAGATTGCCAAAAAACTACACCGATCACCTATTGCAGCAGTAACAGTTGCAGGACG
Lr21_NBS_barley     GAGAGAGATTGCCAAAAAACTACACCGATCACCTATTGCAGCAGTAACAGTTGCAGGACG
                    ************************************************************

Lr21_NBS_tauschii   GCTTGGGGCAAACCCAAATATCAGTTTTTGGAAAAATGTTGCAAAGCTTGACATGTTGAA
Lr21_NBS_rye        GCTTGGGGCAAACCCAAATATCAGTTTTTGGAAAATTGTTGCAAAGCTTGACATGTTGAA
Lr21_NBS_oat        GCTTGGGGCAAACCCAAATATCAGTTTTTGGAAAAATGTTGCAAAGCTTGACATGTTGAA
Lr21_NBS_barley     GCTTGGGGCAAACCCAAATATCAGTTTTTGGAAAAATGTTGCAAAGCTTGACATGTTGAA
                    ********************************** ************************
```

File S2

Alignment between NBS regions of *Lr21* and *Lr21* paralog-1. Identities = 1593/1965 (81%)

```
Lr21-paralog-1   1  ACACACTCATTTGAGATTTTGCTTCTTGCAGGAGCATGCCACTGCGTTTGATTTTGTACT  60
                    ||| || ||||||||||||||| |||||||||| |||||||||| || ||| |||| ||
Lr21_NBS         1  ACATACCCATTTGAGATTTTGCATCTTGCAGGAGAATGCCACTGCATTGGATGTTGTCCT  60

Lr21-paralog-1   61 TACTGCAATCCCAAGACGGAGTTTAAAGAAAAGAATTGAGAAGGTAGAAAGCACCATCAG  120
                    |||||| ||| ||||| |||  || || |||||||| ||||||||| |||| ||||||||
Lr21_NBS         61 TACTGCTATCTCAAGATGGAACTTGAATAAAAGAATAGAGAAGGTACAAAGTACCATCAG  120

Lr21-paralog-1  121 TGAAGTGAAGAAGTATCTGCTCTTAGGCACAGCAAGCAAGAGTGCGCCGAATGATATTGT  180
                     ||||||| |||||  | |||||| ||||| |||||||||||||| ||  | |||||||
Lr21_NBS        121 CGAAGTGACGAAGTCACCGCTCTTGGGCACGGCAAGCAAGAGTGCACCAGACGATATTGC  180

Lr21-paralog-1  181 CAACAAGAACAGGAGCAGAATCAKAACTGCTAGCAAGCGGAAGGTATTTGGCCGAGAGGC  240
                     |||||||| ||||| ||||| | ||||||||||||||||||||||||||| |||||||
Lr21_NBS        181 TAACAAGAATAGGAGTAGAATTAGAACTGCTAGCAAGCGGAAGGTATTTGGTCGAGAGGG  240

Lr21-paralog-1  241 GTTCCGCGATAGTATCATGGCAAAGCTCCGTGAGA---------CAT----CACCGAGCTC  288
                    | | || |||  |||||||| || ||| |||||||        |||    || | |||||
Lr21_NBS        241 GCTGCGTGATCATATCATGGTAAGGCTTCGTGAGATACCAGAGCATGATGCAGCAAGCTC  300

Lr21-paralog-1  289 GGGTACTGGTCCATGTTACTCGGTGATTGGCATATATGGTGTTGCAGGGTCTGGGAAGAC  348
                      || ||| |||||||||||| |||||||| |||||||||||||| ||||||||||||||
Lr21_NBS        301 AAGTGCTGATCCATGTTACTCAGTGATTGGTATATATGGTGTTGCTGGGTCTGGGAAGAC  360

Lr21-paralog-1  349 TACCTTTGCACGATACACCTGAGATTACATAGAGGAGGAATGCAAGGAGGAGAAACTTTT  408
                    || |||||| |||| |    |||||||||| ||||||||||||||| ||||||||||
Lr21_NBS        361 CACGTTTGCAGGATATATTCAAGATTACATAAAGGAGGAATGCAAGGATGAGAAACTTTT  420

Lr21-paralog-1   409 TGACACCACCATGTGCATTCATGTTTCGGAGACTTTCAGTGTCGATGATATATTTCATGA  468
                     ||||||| ||||||||||||||  | || |||||||||| ||||||||||||||||||
Lr21_NBS         421 CGACACCATCATGTGCATTCATGTGACTGAAACTTTCAGTGTGGATGATATATTTCATGA  480

Lr21-paralog-1  469 AATGCTGAAGGATATTACCGGAGATCGGCACTCCCATATTTCAGATCATGAGGAGCTTGA  528
                    |||||||||| |||||||||||||| | |||||| |||||||||||| || ||  || ||
Lr21_NBS        481 AATGCTGAAGTATATTACCGGAGATAGTCACTCCAATATTTCAGATCGTGGGGCTCTAGA  540

Lr21-paralog-1  529 AGAGAAGTTGAAGAAAGAATTGCATGGCAAACGTTTCTTTTTGATATTGGATGATCTCTG  588
                      |||||||||||| ||| ||||  |||||||||||||||| ||||||||||||||||||||
Lr21_NBS        541 TAAGAAGTTGAAGGAAGCATTGTGTGGCAAACGTTTCTTCTTGATATTGGATGATCTCTG  600

Lr21-paralog-1  589 GGTGAAGACCAAGAACGACCCACAACTGGAGGAACTAATCTCTCCACTTAATGTTGGGAT  648
                    |||||| | |||||| |||| ||| || ||||| |||||||||||||| ||||||||| |
Lr21_NBS        601 GGTGAAAAACAAGAATGACCAACACCTAGAGGAGCTAATCTCTCCACTCAATGTTGGGCT  660

Lr21-paralog-1  649 GACAGGAAGCAAAATCTTGGTAACGGCTCGAACAATAGTTGCAACTAGAGCTCTGTGTGA  708
                    || |||||||||||||| |||| |||||||||||||| ||  ||| || |||||||| |||
Lr21_NBS        661 GAAAGGAAGCAAAATCCTGGTGACGGCTCGAACAAAAGAAGCAGCTGGAGCTCTGGGTGC  720

Lr21-paralog-1  709 TGATGAACCTATTAAAATACCTGATTTGGACAAGGATTTGTACTTTTCGATGTTTATGGA  768
                     ||| ||  |||| |||| ||||||||||  |||||  ||||||  |||||||||| |
Lr21_NBS        721 CGATAAATTTATTGAAATGCCTGATTTGGATGAGGATCAGTACTTGGCGATGTTTATGCA  780
```

```
Lr21-paralog-1  769 TTATGCACTGGGCGGCACAAGCGTTGCTGATGAAAAAGAATTTATATGAGTTGGGCGGGT  828
                    ||||||  ||  | ||  |||||  |||||   |    ||||||||  |  |||||||  | |
Lr21_NBS        781 TTATGCGCTAAGTGGTACAAGAGTTGCCCTT---CAAGAATTTGAACAAGTTGGGAGAGA  837


Lr21-paralog-1  829 GATTGCAGAAAAGCTACACCAATCACCTATTGCAGCCGTAGTAGTGGCAGGACGGCTTGG  888
                    ||||||  ||||  |||||||  |||||||||||||||  |||  |||  ||||||||||||||
Lr21_NBS        838 GATTGCCAAAAAACTACACCGATCACCTATTGCAGCAGTAACAGTTGCAGGACGGCTTGG  897


Lr21-paralog-1  889 GACAAACCCAGATATCAAGTTTTGGAAAAATGCTGCAAACCATGAAATGTTGAATGACAC  948
                    | ||||||||  ||||||  |||||||||||||  ||||||  | |||  |||||||||  ||||
Lr21_NBS        898 GGCAAACCCAAATATCAGTTTTTGGAAAAATGTTGCAAAGCTTGACATGTTGAATTACAC  957


Lr21-paralog-1  949 CATGGATGCTCTTTGGTGGAGCTATCGGCCGCTCAATCCGGACATCAGGCGTTGCTTTGA 1008
                    ||||||||||||||||||||||||||  ||  |||  ||||||||||  |||||  ||||||||
Lr21_NBS        958 CATGGATGCTCTTTGGTGGAGCTATCAGCAGCTTGATCCGGACATTAGGCGATGCTTTGA 1017


Lr21-paralog-1 1009 ATTCTGCAATACATTCCCCCGAAGATCGAAGTTGAGAAGGGATGGGTTAATTCGTCTGTG 1068
                    ||||||||  ||  ||||||  |||  ||  ||  ||  ||  ||    |||  |  ||  |||||
Lr21_NBS       1018 ATTCTGCAGTATTTTCCCCAGAAAATTCAAACTGGAAAAAGACCAATTAGTCCGCCTGTG 1077


Lr21-paralog-1 1069 GATAGCGCAAGGGTTTGTAAAGAGCAGTTGTGCAACAGAGGACATGGAAGATGTAGCTGA 1128
                    ||  |||  |||||  ||||||||||  ||||||||||||  |||||  ||||||||||  ||  ||
Lr21_NBS       1078 GACAGCACAAGGATTTGTAAAGACCAGTTGTGCAACGGAGGAGATGGAAGATGTTGCCGA 1137


Lr21-paralog-1 1129 GGGCTACATTCAAGAGTTAATGTCATGCTCATTTATGCAACAAGAAGTAGATGATTCCTC 1188
                    ||||||||||||||||||  ||||||||||||||  |||||  ||  |||  |  ||  |   ||
Lr21_NBS       1138 GGGCTACATTCAAGAGTTAGTGTCATGCTCATTTCTGCAAGAAAAAGGAACTGGT---TC 1194


Lr21-paralog-1 1189 TGAAGAAGAATACTTCACAGTTCATGATCTGCTGCATGATTTATTAGCCAAGGTTGCTGG 1248
                    ||  |  ||  | ||  |||  ||||||||||  ||||||||||||||||  ||||||||||||
Lr21_NBS       1195 TGGCGTTGATCATTTTACAATTCATGATCTGTTGCATGATTTATTAGTCAAGGTTGCTGG 1254


Lr21-paralog-1 1249 TAGTGATTGCTTCAGAATTGAGAATACAAGGAGCCACAGAGGAGAAGGCTGGAAGGAAGA 1308
                    | ||||||||||||||  |||||||  | ||||  |   |||||||   ||  | |||||
Lr21_NBS       1255 AAATGATTGCTTCAGAATCGAGAATAGATGGAGTAAACAAGGAGAA---TGCCAAGAAGA 1311


Lr21-paralog-1 1309 AGTCCCTCGAGATGTCCGCCATCTTTTTGTTCAGAATTATGATGGAGAATTGATAACTAA 1368
                    ||||||||||||||||  ||||||||||||||||  ||||||||  ||||||||  || |
Lr21_NBS       1312 TGTCCCTCGAGATGTCCGTCATCTTTTTGTTCAGAAGTATGATGGGGAATTGATTACCAG 1371


Lr21-paralog-1 1369 GAAGATCCTTGGATTGGAAAATATACGCACTCTCATCATTGATGTTGTTCAAAGGGACAC 1428
                    ||||||||||||||||||||||  ||||||||||||  |||  ||||||  ||  | | |
Lr21_NBS       1372 GAAGATCCTTGGATTGGAAAATTTACGCACTCTCATTATTTGTGTTGTCGAAGAGAATAG 1431


Lr21-paralog-1 1429 ACCAGTTGAGGAAAAAGTCATTGAGAGTATATGCAAGAGGCTGCCAAAATTGCGGGTACT 1488
                    ||||||||||||||  ||||||||||||||  |||  ||||||||  ||  ||||||||||||
Lr21_NBS       1432 ACCAGTTGAGGAAAAGGTCATTGAGAGTATGTGCGTGAGGCTGCTGAAGTTGCGGGTACT 1491


Lr21-paralog-1 1489 AGCCGTTGCTTTCAGCAATAA---CCTTT----CTGGAATCTGGGAAATTAAGTTCTCGG 1541
                     |||  |||||||||  |||  ||   ||  ||    | |  ||||   |  |  ||||||||  ||
Lr21_NBS       1492 GGCCATTGCTTTCAACAAAAAATGCCGTTCAACCAGCAATCCCAGTGA-TAAGTTCTTGG 1550


Lr21-paralog-1 1542 TCCCAAAATCTATTGCTCAGTTAAAGCACCTACGTTATCTTGCTTTCAGGACATTTGGCT 1601
                    |||||  ||||  |||  ||||||||||||||||  ||  |||||  ||||||  |||  |  ||  |
Lr21_NBS       1551 TCCCAGAATCCATTACTCAGTTAAAGCACCTCCGCTATCTAGCTTTCCGGAAAAGTGACA 1610
```

```
Lr21-paralog-1 1602 CATGCTATGTAACTTTACCAAGCGCACTAACGAAACTTCACCATATCCAGTTGCTAGATT  1661
                      ||| | || | ||| ||||| |||||| | ||||||| |||||||||| |||||||||
Lr21_NBS       1611 ACTGCAAGGTTATTTTGCCAAGTGCACTAGCTAAACTTCTCCATATCCAGCTGCTAGATT  1670


Lr21-paralog-1  1662 TTTACTTGGCTAACATGTTGGAATTCCCCT---TTGTTGACCTTGTCAACTTGCGGCACA  1718
                     ||    | |  || || | ||| ||  |||    || ||||||| ||||||||||||||
Lr21_NBS        1671 TTGGTCTTGGCAAAATTTCGGATTTTACCTGTGCTGCTGACCTTATCAACTTGCGGCACA  1730


Lr21-paralog-1  1719 TGTTCTGCAGTGGAGACTGGAAATTCCCTAACTTGGGCAGGCTGATATCACTCCAAACGC  1778
                     | ||||||  ||| ||   ||  || |||||| | |||| |||||  ||||| ||||
Lr21_NBS        1731 TATTCTGC--TGG-GATGTGAGCTTTCCTAACATAGGCACGCTGAGCTCACTTCAAAGAA  1787


Lr21-paralog-1  1779 TACCAGGCTTCACAGTAAGCAATGAACAGGGTTATGAGATAAGGCAGCTGAGGGACCTAA  1838
                     ||||   |||||   |||| ||||||||||||||||||||||  || || |||||||| |
Lr21_NBS        1788 TACCCTTCTTCAGGCTAAGGAATGAACAGGGTTATGAGATAAAACAACTTAGGGACCTGA  1847


Lr21-paralog-1  1839 ACAAGCTTCGTGGCAGACTGTACATCGATGGCCTTGAAAATGTTAAAAGCAAGGAGGAAG  1898
                     ||||| |||||||||  |||    || ||||| ||||||||||||||||||||||||||
Lr21_NBS        1848 ACAAGATTCGTGGCATGCTGGTGGTCAATGGCTTTGAAAATGTTAAAAGCAAGGAGGAAG  1907


Lr21-paralog-1  1899 CTCTTGAAGCCAATCTAGCTGCCAAGGAACGGCTCACAGATCTGA  1943
                     |||||||||||||||||||||||||||| ||||||| || ||||
Lr21_NBS        1908 CTCTTGAAGCCAATCTAGCTGCCAAGGAAAGGCTCACGGAACTGA  1952
```

Alignment between NBS regions of *Lr21* and *Lr21* paralog-2. Identities = 115/145 (79%)

```
Lr21-paralog-2  AAGATGTTCCCAGAGACGTTCGCCATCTTTTTGTTCAGAGTTATGATGCAGCATTGATTA  158
                ||||||| ||  |||| || || ||||||||||||||||  |||||||  | ||||||||
Lr21_NBS        AAGATGTCCCTCGAGATGTCCGTCATCTTTTTGTTCAGAAGTATGATGGGGAATTGATTA  1367


Lr21-paralog-2  CAGGGAAGATTCTTGTATTGGAAAATTTACACACACTCGTCATTTATAGTGTTGGAGGGG  218
                |  ||||||| |||| |||||||||||||| ||| ||| | |||| |  ||| | || |
Lr21_NBS        CCAGGAAGATCCTTGGATTGGAAAATTTACGCACTCTCATTATTTGTGTTGTCGAAGAGA  1427

Lr21-paralog-2  ATACAACAGTTGAGGAAATAGTCAT  243
                ||| | ||||||||||||  |||||
Lr21_NBS        ATAGACCAGTTGAGGAAAAGGTCAT  1452


Identities = 49/59 (83%),

Lr21-paralog-2  AAAGTGTTGAAAGCAGAGAGGAAGCTCTTGCATTCGATCTAGCTGCCAAGAAACGGCTC  733
                ||| |||| ||||||  |||||||||||||| |  | |||||||||||||| || |||||
Lr21_NBS        AAAATGTTAAAAGCAAGGAGGAAGCTCTTGAAGCCAATCTAGCTGCCAAGGAAAGGCTC  1942
```

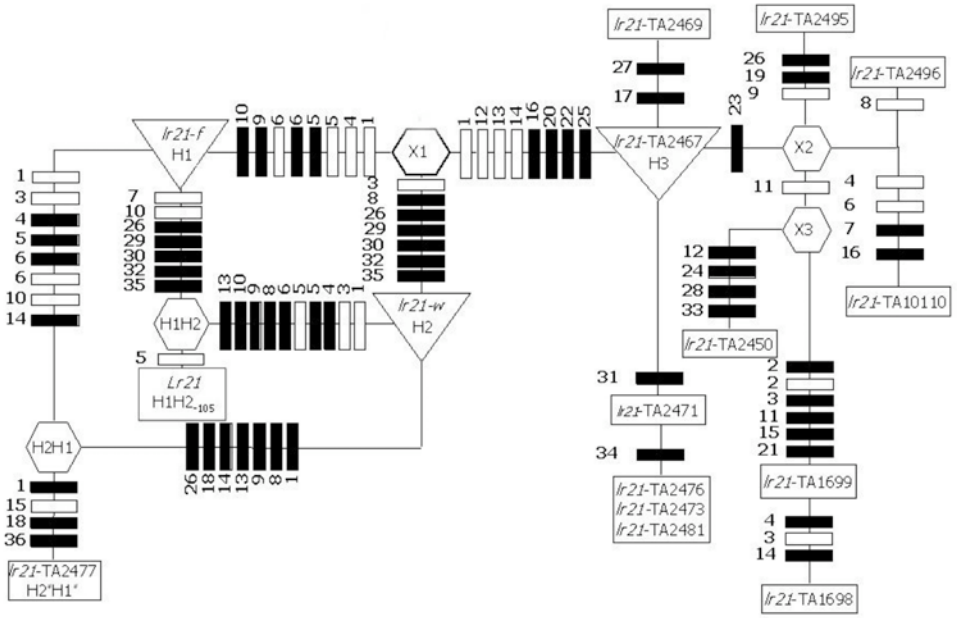

FIGURE S1.—A reconstruction of the haplotype structure at the *Lr21* locus based on the nucleotide variation corresponding to figure 4. The figure shows that three haplotypes H1, H2 and H3 (triangles) account for all the derived alleles but does not suggest their evolutionary order. A black bar indicates a SNP and a white bar an indel. Implied haplotypes X1, X2 and X3 were not observed in this study. Both *Lr21* and *lr21-TA2477* are recombinants of H1 and H2.

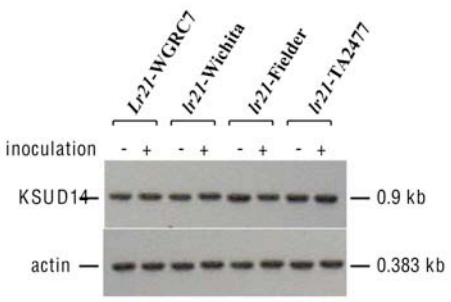

FIGURE S2.— Expression patterns of *Lr21* and *lr21* before and 24 hours after pathogen inoculation, as revealed by modified quantitative RT-PCR.

TABLE S1

Wheat cultivars tested for the *lr21* locus

	Cultivars	Polymorphism pattern
1	*T. aestivum* cv. Sunstar	10
2	*T. aestivum* cv. HD29	10
3	*T. aestivum* cv. Cheyenne	10
4	*T. aestivum* cv. Norin 61	10
5	*T. aestivum* cv. Sumai 3	10
6	*T. aestivum* cv. Katepwa	9
7	*T. aestivum* cv. Chihoku	10
8	*T. aestivum* cv. Thatcher	10
9	*T. aestivum* cv. AC Domain	10
10	*T. aestivum* cv. WL711	10
11	*T. aestivum* cv. Tam W101	10
12	*T. aestivum* cv. Opata	10
13	*T. aestivum* cv. Frontana	10
14	*T. aestivum* cv. Maringa	10
15	*T. aestivum* cv. Halberd	10
16	*T. aestivum* cv. Glenlea	10
17	*T. aestivum* cv. Tasman	10
18	*T. aestivum* cv. Egret	10
19	*T. aestivum* cv. Recital	9
20	*T. aestivum* cv. Cranbrook	10
21	*T. aestivum* cv. Renan	10
22	*T. aestivum* cv. Chinese Spring	9