DOI: 10.1534/genetics.107.080648

The Blast Resistance Gene *Pi37* Encodes a Nucleotide Binding Site–Leucine-Rich Repeat Protein and Is a Member of a Resistance Gene Cluster on Rice Chromosome 1

Fei Lin,* Shen Chen,*,† Zhiqun Que,* Ling Wang,* Xinqiong Liu*,‡ and Qinghua Pan*,[1]

*Laboratory of Plant Resistance and Genetics, College of Resources and Environmental Sciences, South China Agricultural University, Guangzhou, 510642, China, †Plant Protection Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510642, China and ‡College of Life Science, South-Central University for Nationalities, Wuhan, 430074, China

Manuscript received August 15, 2007
Accepted for publication September 4, 2007

ABSTRACT

The resistance (*R*) gene *Pi37*, present in the rice cultivar St. No. 1, was isolated by an *in silico* map-based cloning procedure. The equivalent genetic region in Nipponbare contains four nucleotide binding site–leucine-rich repeat (NBS–LRR) type loci. These four candidates for *Pi37* (*Pi37-1*, *-2*, *-3*, and *-4*) were amplified separately from St. No. 1 via long-range PCR, and cloned into a binary vector. Each construct was individually transformed into the highly blast susceptible cultivar Q1063. The subsequent complementation analysis revealed *Pi37-3* to be the functional gene, while *-1*, *-2*, and *-4* are probably pseudogenes. *Pi37* encodes a 1290 peptide NBS–LRR product, and the presence of substitutions at two sites in the NBS region (V239A and I247M) is associated with the resistance phenotype. Semiquantitative expression analysis showed that in St. No. 1, *Pi37* was constitutively expressed and only slightly induced by blast infection. Transient expression experiments indicated that the *Pi37* product is restricted to the cytoplasm. *Pi37-3* is thought to have evolved recently from *-2*, which in turn was derived from an ancestral *-1* sequence. *Pi37-4* is likely the most recently evolved member of the cluster and probably represents a duplication of *-3*. The four *Pi37* paralogs are more closely related to maize *rp1* than to any of the currently isolated rice blast *R* genes *Pita*, *Pib*, *Pi9*, *Pi2*, *Piz-t*, and *Pi36*.

BLAST, caused by the filamentous ascomycete *Mangnaporthe grisea* (Hebert) Barr, is one of the most devastating of rice diseases (Ou 1985). The rice/*M. grisea* combination has been developed into a well-established host–pathogen model (Velent 1990; Jia *et al.* 2000; Soderlund *et al.* 2006), particularly as many of the interactions between host resistance (*R*) and pathogen avirulence (*Avr*) genes can be satisfactorily explained by the classical gene-for-gene hypothesis (Flor 1971; Jia *et al.* 2000). Over 50 major rice blast *R* genes have been described in the literature (Chen *et al.* 2005; Liu *et al.* 2005), and 7 of these (*Pib*, *Pita*, *Pi9*, *Pid2*, *Pi2*, *Piz-t*, and *Pi36*) have now been isolated (Wang *et al.* 1999; Bryan *et al.* 2000; Chen *et al.* 2006; Qu *et al.* 2006; Liu *et al.* 2007). Six of the 7 belong to the nucleotide binding site–leucine-rich repeat (NBS–LRR) class of *R* gene, as they encode a protein carrying both a nucleotide binding site and a leucine-rich repeat domain. The exception, *Pid2*, encodes a receptor-like kinase protein.

NBS–LRR *R* genes are the commonest type of resistance gene (Hammond-Kosack and Jones 1997; Bai *et al.* 2002). The NBS domain contains three short amino-sequence motifs, a kinase-1a or P-loop (phosphate-binding loop), kinase-2, and kinase-3, and is thought to be involved in signal transduction (Traut 1994; Dangl and Jones 2001). The LRR region plays a critical role in the determination of resistance specificity (Parker *et al.* 1997; Meyers *et al.* 1998). The xxLxLxx motif within the LRR domain is predicted to form a β-strand/β-turn structure, allowing the variable residues to interact with the pathogen *Avr* gene product (Hammond-Kosack and Jones 1997; Jones and Jones 1997); it is these residues that are most subject to diversity selection (Parniske *et al.* 1997; Ellis *et al.* 2000; Sun *et al.* 2001; Zhou *et al.* 2006).

The analysis of *R* genes isolated from various host species has revealed that most are set within a complex locus composed of multiple copies of closely related genes. Prominent examples are the maize *rp1* cluster (Saxena and Hooker 1968; Collins *et al.* 1999), the barley *Mla* cluster (Wei *et al.* 2002), the wheat *Pm3* cluster (Yahiaoui *et al.* 2004), the flax *L* gene cluster (Shepherd and Mayo 1972; Islam *et al.* 1989), and the rice *Xa26* and *Pi9* clusters (Sun *et al.* 2004; Qu *et al.* 2006). The rice and *Arabidopsis thaliana* genome sequences have shown that the majority of NBS–LRR genes occur within

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession no. DQ923494.1.

[1]*Corresponding author:* Laboratory of Plant Resistance and Genetics, College of Resources and Environmental Sciences, South China Agricultural University, Guangzhou, 510642, China.
E-mail: panqh@scau.edu.cn

tandem arrays (Bai *et al.* 2002; Meyers *et al.* 2003). This characteristic clustering of *R* genes has been proposed to facilitate the evolution of novel resistance specificities via recombination or gene conversion (Hulbert 1997), with some well-characterized examples at the flax *L* locus and the maize *rp1* locus (Ellis *et al.* 1999; Smith and Hulbert 2005). The identification and isolation of both host *R* and pathogen *Avr* genes will serve to clarify many of the molecular mechanisms underlying specific host–pathogen recognition in plants, and a detailed understanding of gene organization within *R* gene clusters will help in the interpretation of the evolution of these complex loci.

The rice cultivar St. No. 1 confers partial resistance to Japanese and complete resistance to Chinese isolates of blast (Ezuka *et al.* 1969a,b; Yunoki *et al.* 1970; Chen *et al.* 2005). Much of this resistance is due to the presence of *Pi37* (Chen *et al.* 2005). In this article, we describe the *in silico* map-based cloning of this gene, which is located in a gene cluster set in a recombination-suppressed region.

MATERIALS AND METHODS

Candidate gene cloning: The gene annotation programs FGENESH (http://www.softberry.com) and RiceGAAS (http://ricegaas.dna.affrc.go.jp/rgadb/) were used to identify candidates for *Pi37* within the Nipponbare genomic sequence defined by the flanking markers RM543 and FPSM1 (Figure 1). These candidate sequences, including both their promoter and terminator, were amplified from the genomic DNA of St. No. 1 by the long-range PCR (LR-PCR) procedure described elsewhere (Liu *et al.* 2007). Necessary primer sequences and restriction enzymes are listed in Table 1. PCR products were purified by agarose gel electrophoresis and inserted into the *Sal*I site of the binary vector pCAMBIA1300 to form constructs R37L1CAM, R37L2CAM, R37L3CAM, and R37L4CAM. All clones were validated by sequencing.

Complementation analysis: Constructs containing a single candidate gene were transformed into *Agrobacterium tumefaciens* strain EHA105 by electroporation (GenePluser Xcell TM, Bio-Rad, Hercules, CA). Stability of the constructs was checked as previously described (Qu *et al.* 2003), and the constructs were then individually transformed into the highly blast-susceptible rice cultivar Q1063, following the methods elaborated by Hiei *et al.* (1994). The reaction to blast infection of the primary transgenics (T_0) and their progeny (T_1 and T_2) was tested by artificial inoculation with isolate CHL1159 (Pan *et al.* 2003; Chen *et al.* 2005). The *Pi37* donor St. No. 1 and the susceptible recipient Q1063 were used as controls for the efficacy of the pathological experiment. Transgene copy number in a number of blast resistant T_0 plants was assessed by Southern hybridization, as described previously (Liu *et al.* 2007). The presence of the transgene was also verified by cleaved amplified polymorphic sequence (CAPS) markers, using the primer pair 37CDSF and 37CDSR and digested by *Eco*RI. Patterns of transgene segregation and the association between the presence of the transgene and resistance to blast infection were studied in the T_1 and T_2 generations derived from single transgene copy T_0 individuals (supplemental Table S1 at http://www.genetics.org/supplemental/).

Sequence analysis: Rapid amplification of cDNA ends (RACE) was conducted using the GeneRacer kit (Invitrogen, Groningen, The Netherlands), following the manufacturer's instructions. Total leaf RNA was extracted 24 hr after infection from both St. No. 1 and the highly blast-susceptible Lijiangxintuanheigu (LTH). Primers for the first round of amplification of the 5′ RACE were GS1 and the GeneRacer 5′ primer. A 50-fold dilution of the resulting PCR product served as template for the second round of amplification, using as primers GS2 and GeneRacer 5′. The 3′ RACE employed GS3 in combination with the GeneRacer 3′ primer. A mediate RT–PCR fragment was amplified by primers GS4 and GS5—this overlaps the 5′ RACE and 3′ RACE fragments (Figure 3A and supplemental Table S1 at http://www.genetics.org/supplemental/). The RACE products were inserted into pGEM-T (Promega, Madison, WI) for sequencing. Allelic variants for the coding sequences were derived from St. No. 1, LTH, Q1063, and Nipponbare (Table 2). Sequence similarities were calculated using the Matcher program (http://bioweb.pasteur.fr/seqanal/interfaces/matcher.html), while TSSP and POLYAH (http://www.softberry.com/berry.html) were used to identify the promoter and polyadenylation regions. Protein sequence homology was derived from a BLASTP analysis (Altschul *et al.* 1997). Multiple sequence alignments and phylogenetic analysis were conducted using MEGA 3.1 (http://www.megasoftware.net/mega.html). Theoretical isoelectric points (pI) and protein molecular weights were computed by the DNAstar software package.

Gene expression analysis: For semiquantitative RT–PCR analysis, total RNA was isolated with the TRIzol reagent (Invitrogen, Carlsbad, CA) from 250 mg of seedling (three- to four-leaf stage) leaf of St. No. 1 (*Pi37*) and LTH collected 0, 24, and 48 hr after inoculation with isolate CHL1159. RT–PCR was carried out in two steps: briefly, ~1 μg total RNA was reverse transcribed by SuperScript III RT (Invitrogen), and a 1-μl aliquot of the RT reaction used as template for the subsequent PCR. Primers GS4 and GS5 were used as *Pi37* gene-specific primers (supplemental Table S1 at http://www.genetics.org/supplemental/; Figure 3A). Primers Actin1F and Actin1R were used as an internal control (supplemental Table S1). The RT–PCR was initiated with 1 cycle at 94°/3 min, followed by cycling at 94°/30 sec, 62°/60 sec, and 72°/90 sec. A sample was removed from the thermocycler every 3 cycles between the 23rd and the 35th cycles. Equal volumes of these PCRs were electrophoresed through a 1.5% agarose gel for product quantification. The RT–PCR products were also ligated into pGEM-T for sequence validation.

Subcellular location of *Pi37*: The deduced *Pi37* peptide sequence was subjected to subcellular location prediction using WoLF POSRT (http://wolfpsort.org/). The domain containing a subcellular signal was amplified by primers Gfp371F and Gfp371R (supplemental Table S1 at http://www.genetics.org/supplemental/), containing the *Not*I and *Noc*I sites (underlined in supplemental Table S1). After digestion, the PCR fragment was ligated in frame to the C terminus of the eGFP coding region of pUC18 and expressed under the control of the CaMV 35S promoter. The constructs (1 μg) were coated on 1.1-mm diameter gold beads and shot into onion epidermal cell layers by a pneumatic particle gun (PDS-1000/He, Bio-Rad). Bombardment conditions were 128 in. Hg vacuum, 1300 psi He, target distance 6 cm). The cells were then cultured on MS medium for 24 hr at 22° and observed by a confocal microscopy (TCS SP2, Leica, Wetzlar, Germany) with a filter set providing 455–490 nm excitation and emission above 507 nm.

RESULTS

Identification of candidate genes for *Pi37*: The location of *Pi37* has been defined by recombinational

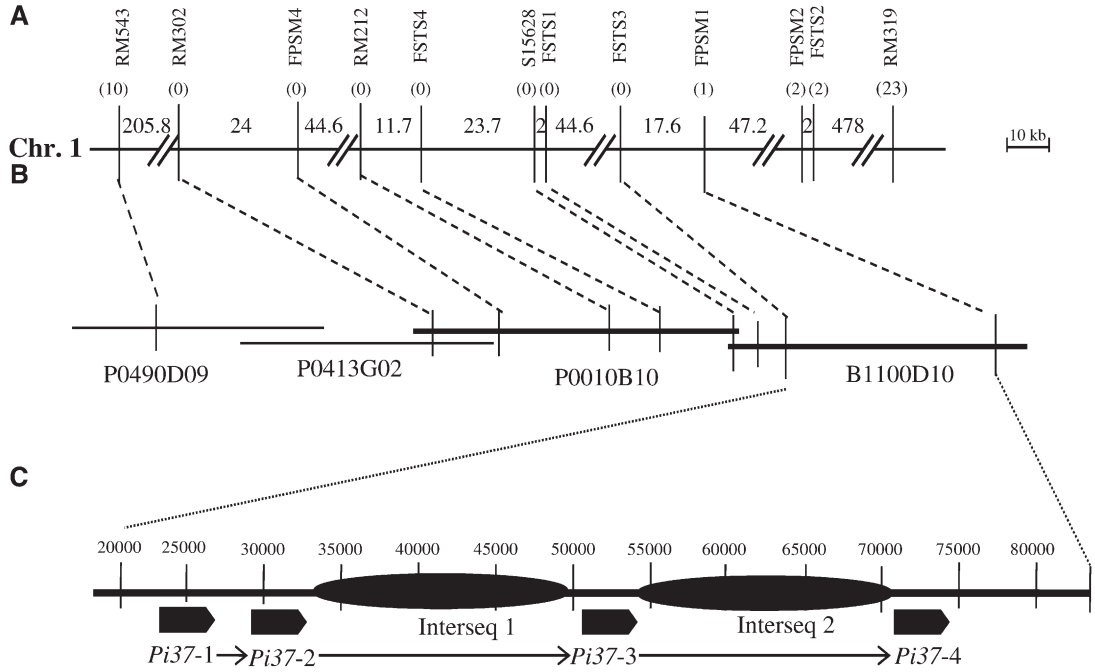

FIGURE 1.—Physical map of the *Pi37* cluster. (A) High-resolution physical map. The numbers above the map represent distances in kilobases, as derived from the Nipponbare genome sequence. Numbers in parentheses represent the number of mapping population recombinants (CHEN *et al.* 2005). (B) The Rice Genome Research Program bacterial artificial chromosome (RGP BAC) contig spanning the *Pi37* region. Individual BACs are shown in bold. (C) The *Pi37* NBS–LRR gene cluster. Candidate genes are indicated with rectangles with arrowheads and intergenic sequences with ellipses. The numbers above the map refer to positions within RGP BAC B1100D10. The thin arrows linking the four *Pi37* paralogs suggest how the gene cluster evolved.

analysis as lying between the two microsatellite loci RM543 and FPSM1 (separated by, respectively, 10 and 1 recombinants) and cosegregating with RM302, FPSM4, RM212, FSTS4, S15628, FSTS1, and FSTS3 (Figure 1A; CHEN *et al.* 2005). In Nipponbare, this 374-kb region is covered by the four BAC clones P0490D09, P0413G02, P0010B10, and B1100D10 (Figure 1B), and contains 118 predicted genes. BLAST analysis of these genes identified 4 as having an NBS–LRR structure, clustering within a 55-kb interval (from 22,313 to 75,167 bp) on B1100D10 (Figure 1C). On the basis that the majority of *R* genes are in the NBS–LRR class, these 4 genes, designated *Pi37-1*, *-2*, *-3*, and *-4* (Figure 1C), were considered to be the likeliest candidates for *Pi37*.

Isolation of candidate genes: Four primer pairs were designed from the Nipponbare sequence, including the necessary restriction sites and cloning protection base (Table 1). LR–PCR products of expected size 7.3, 9.3, 7.3, and 7.0 kb were successfully amplified, and these were ligated to form, respectively, the constructs R37L1CAM, R37L2CAM, R37L3CAM, and R37L4CAM. To guard against potential PCR-based artifacts, two independent LR–PCRs were conducted for each of the four candidate genes, and two to three clones per construct were sequenced. Sequence comparisons showed that *Pi37-1* and *-4* in St. No. 1 share identical sequence with their equivalents in Nipponbare, while St. No. 1 *Pi37-2* is 99.7 and *-3*, 99.8% homologous to their Nipponbare equivalents (supplemental Table S2 at http://www.genetics.org/supplemental/). Thus *Pi37-2* and *-3*, but not *-1* or *-4*, remained as potential candidates for *Pi37*.

Complementation analysis of the candidate genes: To carefully confirm the candidates deduced from the reference sequence information, all four constructs, which carry each candidate gene, were individually transformed into the highly susceptible cv. Q1063. A total of 6, 68, 132, and 39 independent primary transformants were generated using, respectively, R37L1CAM, R37L2CAM, R37L3CAM, and R37L4CAM. When infected with blast isolate CHL1159, all the transgenic plants involving R37L1CAM, R37L2CAM, or R37L4CAM were highly susceptible, but 24 out of 132 R37L3CAM transformants were resistant. Three of these (LZ75, LZ76, and LZ85) were genotyped by Southern blotting. One copy of *Pi37-3* was present in LZ76 and two in both LZ75 and LZ85 (supplemental Figure S1 at http://www.genetics.org/supplemental/). Monogenic inheritance of the *Pi37-3* transgene was displayed among the progeny of LZ76-4, which produced a segregation ratio between resistant and susceptible not significantly different from 3:1 (28 resistant and 6 susceptible progeny, $\chi^2 = 0.84$, $P < 0.30$). These progeny could also be used to demonstrate that perfect cosegegration between reaction to blast infection and the presence of the allele-specific marker generated by primer pairs 37CDSF and 37CDSR (Figure 2B). Thus *Pi37-3* is a strong candidate for *Pi37*.

TABLE 1

Long-range (LR)-PCR primers used to amplify *Pi37* candidates

Candidate gene	Primer	Sequence (5′–3′)[a]	Product size /kb	Restriction site	PCR conditions[b]	Vector
Pi37-1	R37L1F	*GTAGGTCGACTTCGAAGGGA AGGTCAAGGTGAGCGAGTGG*	7.3	*Sal*I	A	pCAMBIA1300
	R37L1R	*TTGCGTCGACAGCTCATTGCCGA TGGTCATGTCCTCCGTC*				
Pi37-2	R37L2F	*GTCAGTCGACGCTTTGTGTCTGTTCCAGCCTCTTGGTGTTC*	9.3	*Sal*I	B	pCAMBIA1300
	R37L2R	*TGGTGTCGACGTGCATACCTGATGCTGTGTTTCCGTGTC AG*				
Pi37-3	R37L3F	*GTACGTCGACCCTACCGAGTCCAGCAAAATCCAT*	7.3	*Sal*I	C	pCAMBIA1300
	R37L3R	*TCTAGTCGACGTTTCCCAATCTGGCACAGCAAGG*				
Pi37-4	R37L4F	*GTTCGTCGACCTGTGGCCTCGTCAATCTACATCAAG*	7.0	*Sal*I	A	pCAMBIA1300
	R37L4R	*CTCAGGTCGACCTGGCCCTTCCATTCAGATGACAGTGC*				

[a] Restriction recognition sites are underlined and protection bases are in italics.

[b] All PCRs began with a denaturation step of 94°/2 min, followed by 30 cycles of (A) 94°/30 sec, 70°/7.5 min; (B) 94°/30 sec, 70°/9.5 min; and (C) 94°/30 sec, 68°/7.5 min. The reactions were completed by a 10-min incubation at 72°.

Molecular characterization of *Pi37*: The full-length St. No. 1 *Pi37* cDNA was obtained by RT–PCR and RACE–PCR and was compared to the genomic sequence. The gene contains a 197-bp 5′ and a 603-bp 3′ untranslated region (UTR) and two introns. The first intron is 3943-bp long and is positioned within the 5′ UTR, ending 23 bp upstream of the ATG start codon. The second intron is 124-bp long, and is positioned within the 3′ UTR, starting 39 bp downstream of the TGA stop codon. The transcript length of *Pi37* is 3873 bp (Figure 3A). The deduced *Pi37* open reading frame encodes a 1291-residue polypeptide with an estimated molecular weight of 147 kDa, and a pI of 5.98. The N-terminal section contains three typical NBS family motifs (Van Der Biezen and Jones 1998), specifically GGAGKS (beginning at residue 222), LLVLDDV (beginning at residue 297), and GSRVLVTSRR (beginning at residue 327). These correspond to, respectively, the kinase 1a (P-loop), the kinase 2, and the kinase 3a consensus motifs (Traut 1994; Grant *et al.* 1995). The C-terminal region of the protein is composed of 25 irregular LRRs between residues 590 and 1290 (Figure 3B).

A full-length cDNA was also generated and sequenced from the susceptible LTH. Sequence comparison analysis revealed no structural differences between the alleles (data not shown), but there were nine single-nucleic-acid substitutions present, which generated six amino acid differences between the alleles (Table 2). From additional cDNA sequences obtained from Q1063 and Nipponbare, and from *Pi37-4* (see below), it was possible to conclude that all the susceptible genotype *Pi37* gene products share two residue differences relative to the sequence of the resistant-type product. These are V239A (a valine at position 239 replacing an alanine) and I247M (an isoleucine at position 247 replacing a methionine) (Table 2; Figure 3B).

Expression analysis of *Pi37*: After 23 RT–PCR cycles, the amplicon was barely detectable, but it became readily detectable from 26 cycles onward (Figure 4). The expression level of the resistant allele appears to be higher than that of the susceptible one, and there was some evidence for induction of expression during the 48 hr following inoculation. Overall, *Pi37* appears to be constitutively expressed, although its expression level was somewhat promoted in the presence of the pathogen. The C terminus of the predicted amino acid sequence of the *Pi37* product includes a cytoplasmic subcellular localization signal domain. Transient expression in onion epidermal cells was employed to confirm this prediction *in vivo*. This experiment showed that the eGFP–Pi37 C-terminus fusion protein was distributed throughout the cytoplasm, but not in the nucleus (Figure 5). (Note that the GFP-only control signal was evenly distributed throughout the cell.)

Evolutionary analysis of the *Pi37* gene cluster: A BLAST analysis of the Nipponbare sequence identified two duplications (33,911–54,539 and 54,540–75,167)

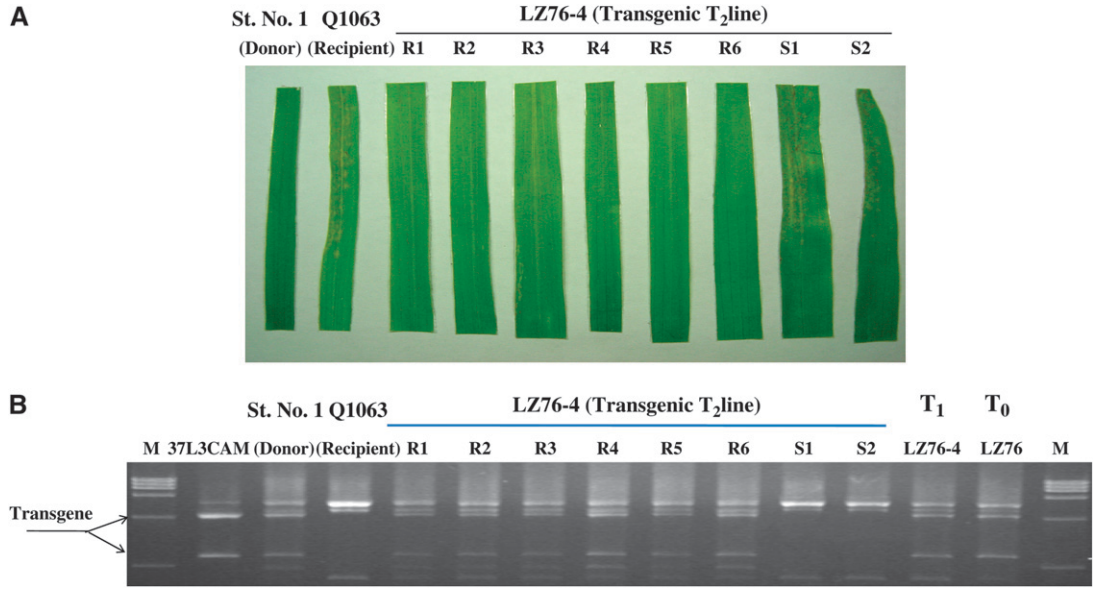

FIGURE 2.—Complementation test and the molecular analysis of transgenics. (A) Reaction to inoculation with blast isolate CHL1159 of the *Pi37* donor (St. No. 1), the susceptible recipient (Q1063), and a set of T_2 progeny derived from the T_1 plant LZ76-4. R, resistant; S, susceptible. (B) Cosegregation of the resistance phenotype with the presence of the *Pi37* transgene. The DNA fragment amplified by the primer pair 37CDSF and 37CDSR was digested with *Eco*RI and the digested product electrophoresed through a 1.2% agarose gel. M, standard molecular weight marker DL2000.

within the overall cluster (Figure 1C). The duplicated segment is composed of two elements—a gene member (*Pi37-3* or *-4*) and an intergenic sequence (Interseq1 or Interseq2), with a nucleotide identity, respectively, 98.8 and 99.1% (a part of these data is given in supplemental Table S2 at http://www.genetics.org/supplemental/). A YhyA-like transposon is present in each intergenic sequence (data not shown). The sequence identity between the four *Pi37* candidates ranged from 78.4 (*-1 vs. -2*) to 98.8% (*-3 vs. -4*) (supplemental Table S2). This suggests that *Pi37-3* was recently evolved from *-2*, which in turn was derived from *-1*, while *-4* is probably the youngest gene, emerging as a duplication of *-3*.

A wider BLAST search showed that the *Pi37* sequence shares 59% identity with *rp1*, which confers resistance to rust in maize (SUN *et al.* 2001). The other characterized rice blast *R* gene (*Pib*, *Pita*, *Pi9*, *Pi2*, *Piz-t*, and *Pi36*) product sequences were also included in the phylogenetic analysis. Altogether, these fifteen *R* proteins can be grouped into five clades: rp1-dp1/rp1-dp3/rp1-dp8/rp1-kp3/rp1-dp7/rp1-kp1 (I), Pi37-1/Pi37-2/Pi37-3/Pi37-4 (II), Pita/Pi36 (III), Pi9/Pi2/Piz-t (IV), and Pib (V) (Figure 6). This analysis proposes that the four *Pi37* paralogs belong to three clades and are more closely related to the maize *rp1* complex than to any of the currently characterized rice blast *R* genes.

DISCUSSION

The function of *Pi37*: Blast resistance in rice is commonly categorized into qualitative (complete) and quantitative (partial) (YUNOKI *et al.* 1970). The latter describes an incomplete form of generally race nonspecific resistance under multigenic control (SIMMONDS 1991), although race-specific effects are also known [*e.g.*, *Pif* (YUNOKI *et al.* 1970), *Pb1* (FUJII *et al.* 2000), *pi21* (FUKUOKA and OKUNO 2001), and *Pi34* (ZENBAYASHI-SAWATA *et al.* 2006)]. The cultivar St. No. 1 shows complete resistance to many Chinese blast isolates, and it is now understood that this resistance is based not only on *Pif*, but also on *Pi37* and perhaps other, as yet uncharacterized genes (CHEN *et al.* 2005). *Pi37* belongs to the NBS–LRR class of *R* genes, and its functionality depends on the identity of two residues (V239A and I247M) in the NBS region, located between the kinase 1a (P-loop) and kinase 2. The NBS region of certain *R* genes has been shown to be involved in intramolecular interactions with other domains of the protein, or in interactions with other proteins (MESTRE and BAULCOMBE 2006). Since the two critical *Pi37* substitutions ensure the complete resistance of St. No. 1, they are presumably involved in *AvrPi37* recognition, an avirulence factor carried by Chinese isolates, but presumably lacking in Japanese isolates. Further gene recombination and *in vitro* binding analysis will be necessary to provide a detailed molecular explanation for the major effect of such a small sequence difference (JIA *et al.* 2000; ELLIS *et al.* 2007).

The Pi37 protein shares more sequence homology with the products derived from the *rp1* complex than with any of the other *R* genes characterized to date. This maize complex gene confers race-specific resistance to *Puccinia sorghi* (SAXENA and HOOKER 1968; COLLINS *et al.* 1999). The order and arrangement of alleles at this

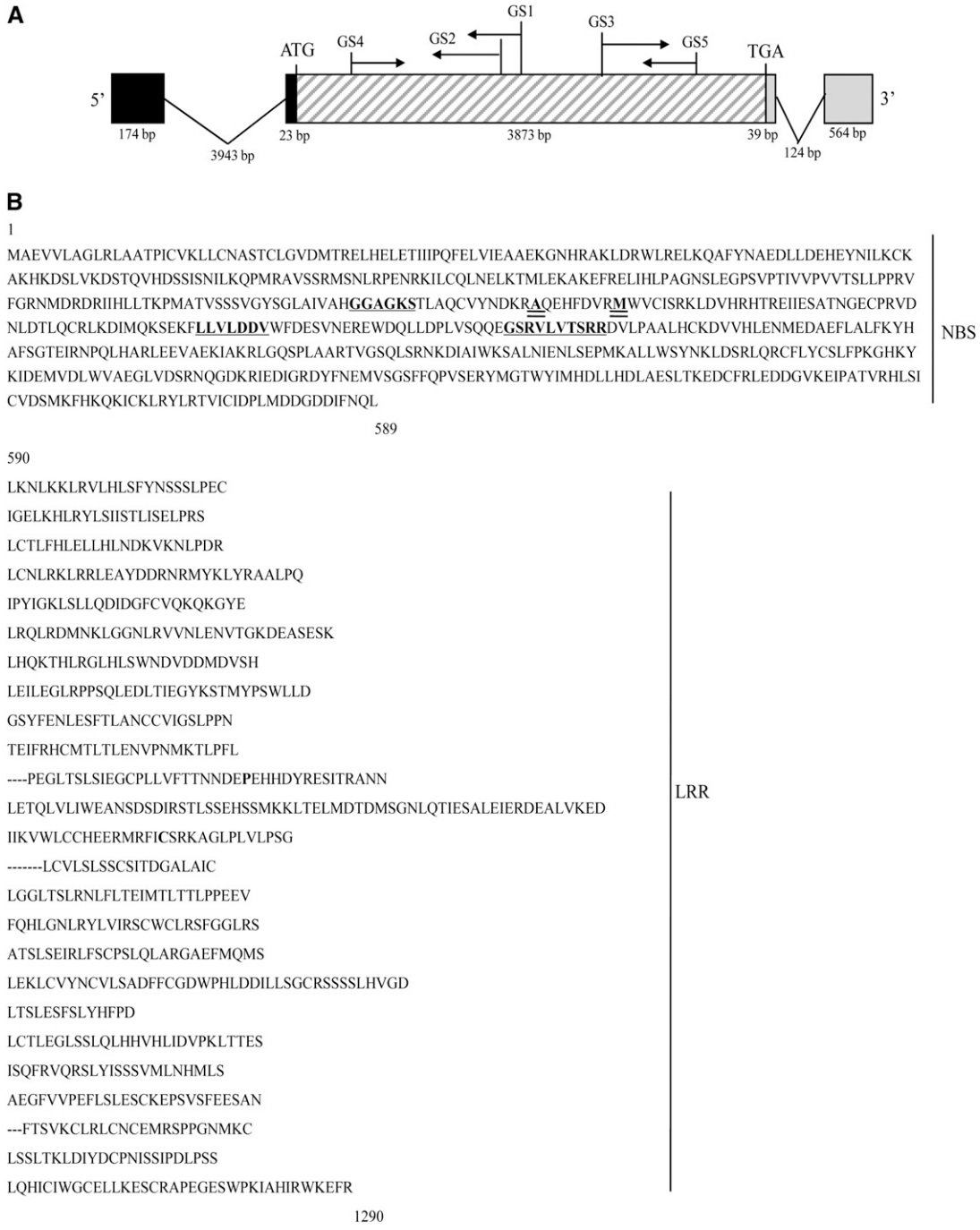

FIGURE 3.—The structure of *Pi37* and its gene product. (A) The structure as determined by 5′ and 3′ RACE. Hatched box, coding region; black box, 5′-UTR; gray box, 3′-UTR. The translation start codon (ATG), translation stop codon (TGA), 5′ and 3′ intron, and RACE primers are also indicated. (B) Deduced peptide sequence of the *Pi37* gene product. The three conserved motifs forming the NBS region are underlined. The two allele-specific substitutions (V239A and I247M) are double underlined. The C-terminal LRR is shown detached from the rest of the sequence.

complex has been associated with the generation of novel race specificities (RICHTER *et al.* 1995), and intralocus recombinants have also been identified which confer apparently race nonspecific partial resistance. The complete dominance of the race-specific *rp1* alleles is thought to be conferred by a highly efficient molecular recognition of the elicitor (HULBERT 1997). A change from partial to complete resistance may therefore be generated by quite a minor sequence change in the part of the *R* gene which affects the interaction with its corresponding *Avr* gene. Many of the important grass species carry sequences homologous to *rp1* (AYLIFFE *et al.* 2000). However, only a few of these orthologs have a demonstrated function. Although *Pi37* shares homol-

TABLE 2

Peptide sequence comparison between the *Pi37* gene products translated from a resistant and a susceptible rice cultivar

Rice cultivar	R type[a]	Allele code	Amino acid residue position								
			16	91	239	247	347	676	870	964	1040
St. No. 1	R	Pi37	I	C	A	M	D	M	L	Y	L
LTH	S	Pi37	T	R	V	I	D	V	L	Y	F
Q1063	S	Pi37	I	C	V	I	D	M	L	Y	L
Nipponbare	S	Pi37	I	C	V	I	D	M	L	Y	L
Nipponbare	S	Pi37-4	T	R	V	I	G	V	P	C	F

[a] R, resistant; S, susceptible.

ogy with *rp1* and confers race-specific resistance in rice to the blast fungus, it remains to be seen whether it has any effect in a heterologous situation (Elliott *et al.* 2002; Christensen *et al.* 2004; Huang *et al.* 2007).

The structure of *Pi37*: Comparative studies have shown that intron positions are highly conserved over long evolutionary periods (Roy *et al.* 2003; Roy and Penny 2006). This applies to the conserved NBS region in cereal NBS–LRR genes (Bai *et al.* 2002). The six characterized rice blast NBS–LRR *R* genes *Pita*, *Pib*, *Pi9*, *Pi2*, *Piz-t*, and *Pi36* all carry introns in their coding region. *Pita* has an intron in the beginning of its kinase-2 motif, one of the commonest positions for introns in cereal NBS–LRR genes (Bryan *et al.* 2000; Bai *et al.* 2002). *Pib* has a complex structure in which two tandemly repeated partial NBS regions carry two introns between the RNBS-B and the GLPL motif (Wang *et al.* 1999; Bai *et al.* 2002). *Pi9*, *Pi2*, and *Piz-t* all contain two introns, one upstream of the NBS domain and one downstream of the LRR region (Qu *et al.* 2006). The *Pi36* coding region is interrupted by four introns in the NBS and LRR domains (Liu *et al.* 2007). Of the four paralogs in the *Pi37* cluster, only *Pi37-1* has any introns (this article, data not shown). This copy is the most outlying of the four paralogs (supplemental Table S1 at http://www.genetics.org/supplemental/). *Pi37* thus appears be the first representative of a cereal NBS–LRR gene lacking an intron.

The genomic organization of the *Pi37* gene cluster: The *Pi37* gene cluster has been located in a recombination-suppressed 374-kb region, flanked by regions showing enhanced recombination (Chen *et al.* 2005). This pattern is reminiscent of the maize *rp1* gene complex. Suppression of recombination has also been noted in some other *R* gene regions, such as *Mi* (Van Daelen *et al.* 1993), *Mla* (Wei *et al.* 1999), *Pita*^{2} (Nakamura *et al.* 1997), *Pi-CO39(t)* (Chauhan *et al.* 2002), *Pi5* (Jeon *et al.* 2003), and *I* loci (Vallejos *et al.* 2006). Recombination frequency is severely reduced in the hemizygous state (Ozias-Akins *et al.* 1998; Goel *et al.* 2003). The *Pi37* region lies within a segment introgressed from *indica* into *japonica* rice (Yunoki *et al.* 1970), and the dominant mode of inheritance of the four markers (see FPSM4, FSTS4, FSTS1, and FSTS3 in Figure 1A) cosegregating with *Pi37* is suggestive that hybrids between *Pi37* carriers and noncarriers may well be effectively hemizygous for the introgression segment (Chen *et al.* 2005; Vallejos *et al.* 2006). Thus the absence of localized recombination between resistant and susceptible haplotypes may have driven the evolution of diversity at this *R* gene locus. In the "trench warfare" model for the evolution of the host–pathogen interaction (Stahl *et al.* 1999), natural selection maintains a dynamic equilibrium between susceptible and resistant alleles in a population but at a cost to fitness. When the relevant pathogens or

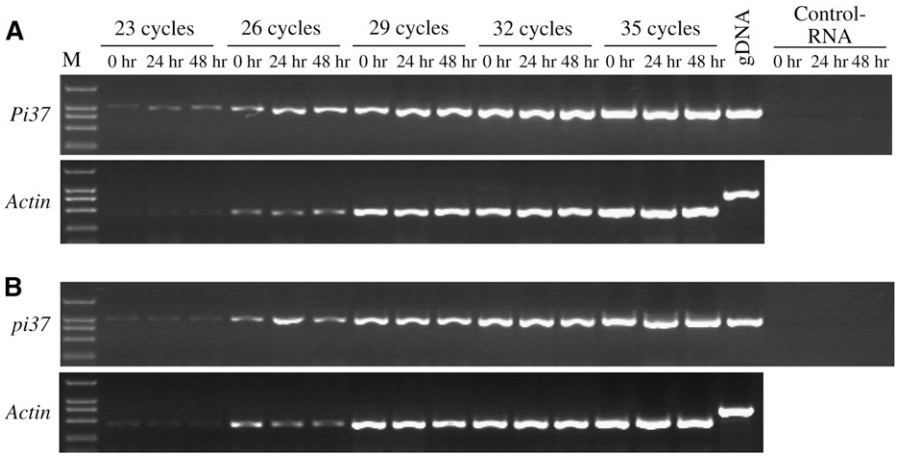

Figure 4.—Semiquantitative RT–PCR analysis of *Pi37* expression. Two-week-old (A) resistant (St. No. 1) and (B) susceptible (LTH) seedlings were inoculated with blast isolate CHL1159. The expression of *Pi37* was examined 0, 24, and 48 hr after inoculation. The rice *Actin1* gene was used as a positive control, and total RNA as a negative control. The amplicon was sampled every three PCR cycles starting at the 23rd cycle.

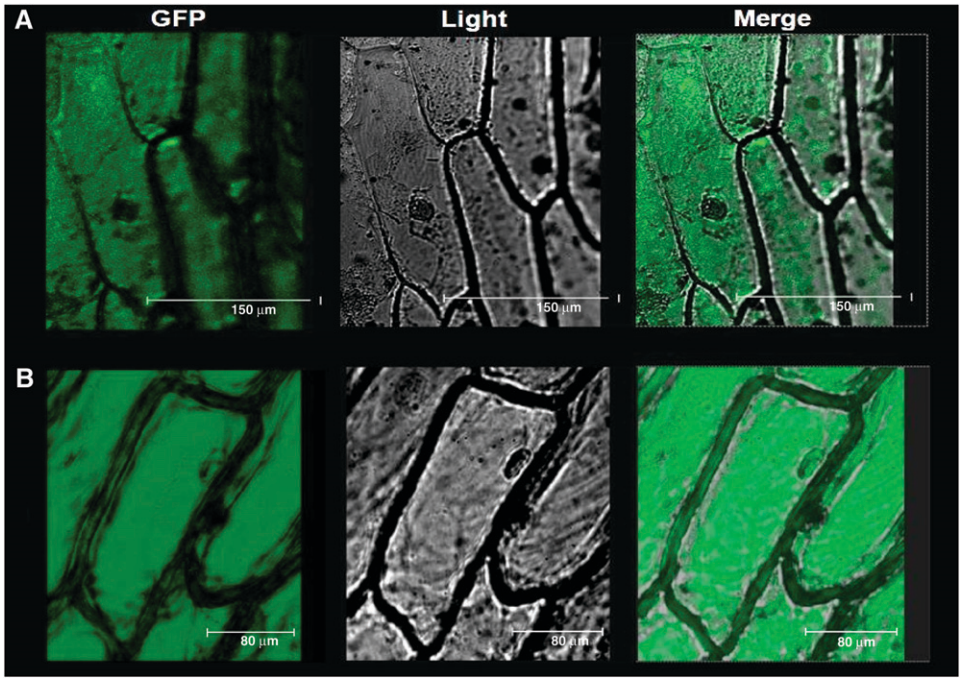

FIGURE 5.—Subcelluar localization of the *Pi37* gene product. (A) eGFP::Pi37 fusion protein under the control of CaMV 35S promoter transiently expressed in onion epidermal cells following ballistic transformation. (B) eGFP under the control of CaMV 35S promoter. Green fluorescent protein (GFP) images are captured in a dark field, cellular structures are visualized under a light field (Light), and the two images are superimposed (Merge).

even pathogen races are absent, the maintenance of specific *R* genes can therefore confer an evolutionary cost (TIAN *et al.* 2002, 2003; ZHOU *et al.* 2004). While the existence of the *Pi37* gene cluster may allow for a greater rate of novel *R* gene generation via intragenic recombination, the maintenance of the *R* gene might therefore become a liability to the host. As a result, suppression of recombination at the *Pi37* locus may represent an adaptive mechanism supporting balancing selection between resistant and susceptible alleles.

The *Pi37* gene cluster contains four NBS–LRR genes, arranged in tandem in the same orientation. The complex shares substantial homology to the maize *rp1* complex (PENG *et al.* 1999; AYLIFFE *et al.* 2000). *rp1* haplotypes vary in the number of homologs present, ranging from as few as 1 to as many as 50 (MONOSI *et al.* 2004). On the basis of sequence comparisons, the *rp1* homologs of maize, sorghum, and barley are thought to all have arisen independently from a single *rp1*-like gene present in the common grass ancestor (RAMAKRISHNA *et al.* 2002). *Pi37* may well have evolved from the same common *rp1*-like gene ancestor. The simplest predicted series of events leading the present status of the *Pi37* locus starts with an ancient triplication of a single copy *Pi37-1* gene to give rise to *Pi37-2* and *-3*, and this was followed later by a second duplication event in which the *Pi37-3* sequence gave rise to the present day *-3* and *-4*. A more detailed analysis of the intergenic sequences separating the *Pi37* members in various haplotypes should provide evidence to test this model and will allow for further elucidation of the evolution of this *R* gene cluster.

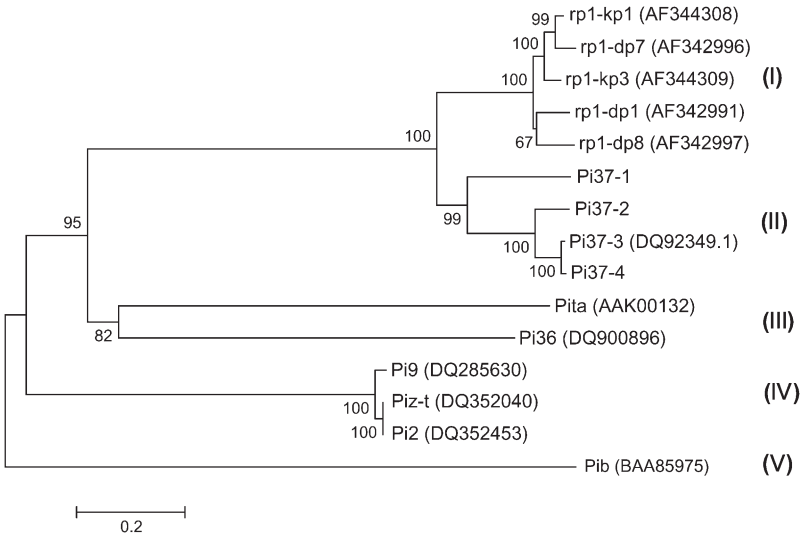

FIGURE 6.—Phylogenetic analysis of the *Pi37* cluster and the other 11 *R* genes. Numbers on the branches indicate bootstrap percentages. The unit branch length is equivalent to 0.2 nucleotide substitutions per site, as indicated by the bar at the bottom left corner.

In conclusion, we have applied the *in silico* map-based cloning method to successfully isolate the functional gene *Pi37* and have characterized the *R* gene cluster in which it lies. This general approach is proving to be an effective means for the genetic dissection of gene complexes in recombination-suppressed regions.

This research was supported by grants from the National 973 project (2006CB1002006), the National 863 projects (2006AA10A103; 2006AA100101), the National Natural Science Foundation (30570994), the Innovation Research Team Project from the Ministry of Education of China (IRT0448), and the Guangdong Provincial Natural Science Foundation (039254).

LITERATURE CITED

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang *et al.*, 1997 Gapped BLAST and PSLBLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.

Ayliffe, M. A., N. C. Collins, J. G. Ellis and A. Pryor, 2000 The maize *rp1* rust resistance gene identifies homologues in barley that have been subjected to diversifying selection. Theor. Appl. Genet. **100:** 1144–1154.

Bai, J., L. A. Pennill, J. Ning, S. W. Lee, J. Ramalingam *et al.*, 2002 Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. Genome Res. **12:** 1871–1884.

Bryan, G. T., K. S. Wu, L. Farrall, Y. L. Jia, H. P. Hershey *et al.*, 2000 A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. Plant Cell **12:** 2033–2046.

Chauhan, S., M. L. Farman, H. B. Zhang and A. Leong, 2002 Genetic and physical mapping of a rice blast resistance locus, *Pi-CO39*(t), that corresponds to the avirulence gene *AVR1–CO39* of *Magnaporthe grisea*. Mol. Genet. Genomics **267:** 603–612.

Chen, S, L. Wang, Z. Q. Que, R. Q. Pan and Q. H. Pan, 2005 Genetic and physical mapping of *Pi37*(t), a new gene conferring resitance to rice blast in the famous cultivar St. No. 1. Theor. Appl. Genet. **111:** 1563–1570.

Chen, X. W., J. J. Shang, D. X. Chen, C. Lei, Y. Zou *et al.*, 2006 A β-lectin receptor kinase gene conferring rice blast resistance. Plant J. **46:** 794–804.

Christensen, A. B., H. Thordal-Christensen, G. Zimmermann, T. Gjetting, F. Michael *et al.*, 2004 The germinlike protein GLP4 exhibits superoxide dismutase activity and is an important component of quantitative resistance in wheat and barley. Mol. Plant-Microbe Interact. **17:** 109–117.

Collins, N., J. Drake, M. Ayliffe, Q. Sun, J. Ellis *et al.*, 1999 Molecular characterization of the maize *Rp1-D* rust resistance haplotype and its mutants. Plant Cell **11:** 1365–1376.

Dangl, J. L., and J. D. G. Jones, 2001 Plant pathogens and integrated defence responses to infection. Nature **411:** 826–833.

Elliott, C., F. S. Zhou, W. Spielmeye, R. Panstruga and P. Schulze-Lefert, 2002 Functional conservation of wheat and rice *Mlo* orthologs in defense modulation to the Powdery mildew fungus. Mol. Plant-Microbe Interact. **15:** 1069–1077.

Ellis, J. G., P. N. Dodds and T. Pryor, 2000 The generation of plant disease resistance gene specificities. Trends Plant Sci. **5:** 373–379.

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:** 495–507.

Ellis, J. G., G. J. Lawrence and P. N. Dodds, 2007 Further analysis of gene-for-gene disease resistance specificity in flax. Mol. Plant Path. **8:** 103–109.

Ezuka, A., T. Yunoki, Y. Sakurai, H. Shinoda and K. Toriyama, 1969a Studies on the varietal resistance of rice to blast. I. Tests for genotype of "true resistance" (in Japanese with English summary). Bull. Chugoku Natl. Agric. Exp. Stn. **E4:** 1–31.

Ezuka, A., T. Yunoki, Y. Sakurai, H. Shinoda and K. Toriyama, 1969b Studies on the varietal resistance of rice to blast. II. Tests for field resistance in paddy fields and upland nursery beds (in Japanese with English summary). Bull. Chugoku Natl. Agric. Exp. Stn. **E4:** 32–53.

Flor, H. H., 1971 Current status of the gene for gene concept. Annu. Rev. Phytopathol. **9:** 275–296.

Fujii, K., Y. Hayano-Saito, K. Saito, N. Sugiura, N. Hayashi *et al.*, 2000 Identification of a RFLP marker tightly linked to the panicle blast resistance gene, *Pb1*, in rice. Breed Sci. **50:** 183–188.

Fukuoka, S., and K. Okuno, 2001 QTL analysis and mapping of *pi21*, a recessive gene for field resistance to rice blast in Japanese upland rice. Theor. Appl. Genet. **103:** 185–190.

Goel, S., Z. B. Chen, J. A. Conner, Y. Akiyama, W. W. Hanna *et al.*, 2003 Delineation by fluorescence in situ hybridization of a single hemizygous chromosomal region associated with aposporous embryo sac formation in *Pennisetum squamulatum* and *Cenchrus ciliaris*. Genetics **163:** 1069–1082.

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.

Hammond-Kosack, K. E., and J. D. G. Jones, 1997 Plant disease resistance genes. Annu. Rev. Plant Physiol. Plant Mol. Biol. **48:** 575–607.

Hiei, Y., S. Ohta, T. Komari and T. Kumashiro, 1994 Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J. **6:** 271–282.

Huang, H. E., C. A. Liu, M. J. Lee, C. G. Kuo, H. M. Chen *et al.*, 2007 Resistance enhancement of transgenic tomato to bacterial pathogens by the heterologous expression of sweet pepper ferredoxin-I protein. Phytopathology **97:** 900–906.

Hulbert, S. H., 1997 Structure and evolution of rust resistance in maize. Annu. Rev. Phytopathol. **35:** 293–310.

Islam, M. R., K. W. Shepherd and G. M. E. Mayo, 1989 Recombination among genes at the *L* group in flax conferring resistance to rust. Theor. Appl. Genet. **77:** 540–546.

Jia, Y., S. A. McAdams, G. T. Bryan, H. P. Hershey and B. Valent, 2000 Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO J. **19:** 4004–4014.

Jeon, J.S., D. Chen, G. H. Yi, G. L. Wang and P. C. Ronald, 2003 Genetic and physical mapping of *Pi5*(t), a locus associated with broad-spectrum resistance to rice blast. Mol. Genet. Genomics **269:** 280–289.

Jones, D. A., and J. D. G. Jones, 1997 The roles of leucine-rich repeat protein defences. Adv. Bot. Res. **24:** 89–167.

Liu, X. Q., F. Lin, L. Wang and Q. H. Pan, 2007 The *in silico* map-based cloning of *Pi36*, a rice coiled-coil–nucleotide-binding site–leucine-rich repeat gene that confers race-specific resistance to blast fungus. Genetics **176:** 2541–2549.

Liu, X. Q., L. Wang, S. Chen, F. Lin and Q. H. Pan, 2005 Genetic and physical mapping of *Pi36*(t), a novel rice blast resistance gene located on rice chromosome 8. Mol. Genet. Genomics **274:** 394–401.

Mestre, P., and D. C. Baulcombe, 2006 Elicitor-mediated oligomerization of the tobacco N disease resistance protein. Plant Cell **11:** 1833–1846.

Meyers, B. C., A. Kozik, A. Griego, H. Kuang and R. W. Michelmore, 2003 Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. Plant Cell **15:** 809–834.

Meyers, B. C., K. A. Shen, P. Rohani, B. S. Gaut and R. W. Michelmore, 1998 Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell **11:** 1833–1846.

Monosi, B., R. J. Wisser, L. Pennill and S. H. Hulbert, 2004 Full-genome analysis of resistance gene homologues in rice. Theor. Appl. Genet. **109:** 1434–1447.

Nakamura, S., S. Asakawa, N. Ohmido, K. Fukui, N. Shimizu *et al.*, 1997 Construction of an 800-kb contig in the near-centromeric region of the rice blast resistance gene *Pi-ta2* using a highly representative rice BAC library. Mol. Genet. Genomics **254:** 611–620.

Ou, S. H., 1985 *Rice Disease*, Ed. 2, pp. 109–201. Commonwealth Mycological Institute, Kew Surrey, UK.

Ozias-Akins, P., D. Roche and W. W. Hanna, 1998 Tight clustering and hemizygosity of apomixes-linked molecular markers in *Penniselim squamulalum* implies genetic control of apospory by divergent locus that may have no allelic form in sexual genotypes. Proc. Natl. Acad. Sci. USA **95:** 5127–5132.

Pan, Q.H., Z. D. Hu, T. Tanisaka and L. Wang, 2003 Fine mapping of the blast resistance gene *Pi15*, linked to *Pii*, on rice chromosome 9. Acta. Bot. Sin. **45:** 871–877.

Parker, J. E., M. J. Coleman, V. Szabo, L. N. Frost, R. Schmidt *et al.*, 1997 The *Arabidopsis* downy mildew resistance gene *RPP5* shares similarity to the toll and interleukin-1 receptors with *N* and *L6*. Plant Cell **9:** 879–894.

Parniske, M., K. E. Hammond-Kosack, C. Golstein, C. M. Thomas, D. A. Jones *et al.*, 1997 Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. Cell **91:** 821–832.

Peng, Y., K. F. Schertz, S. Cartinhour and G. E. Hart, 1999 Comparative genome mapping of Sorghum bicolor (*L.*) Moench using an RFLP map constructed in a population of recombinant inbred lines. Plant Breed. **118:** 225–235.

Qu, S., G. Coaker, D. Francis, B. Zhou and G. L. Wang, 2003 Development of a new transformation-competent artificial chromosome (TAC) vector and construction of tomato and rice TAC libraries. Mol. Breed. **12:** 297–308.

Qu, S., G. Liu, B. Zhou, M. Bellizzi, L. Zeng *et al.*, 2006 The broad-spectrum blast resistance gene *Pi9* encodes a nucleotide-binding site-leucine-rich repeat protein and is a member of a multigene family in rice. Genetics **172:** 1901–1914.

Ramakrishna, W., J. Dobcovsky, Y. J. Park, C. Busso, J. Emberton *et al.*, 2002 Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. Genetics **162:** 1389–1400.

Roy, S. W., and D. Penny, 2006 Large-scale intron conservation and order-of-magnitude variation in intron loss/gain rates in apicomplexan evlotion. Genome Res. **16:** 1270–1275.

Richter, T., T. Pryor, J. Bennetzen and S. Hulbert, 1995 New rust resistance specificities associated with recombination in the *Rp1* complex in maize. Genetics **141:** 373–381.

Roy, S. W., A. Fedorrov and W. Gilbert, 2003 Large-scale comparison of intron in mammalian genes shows intron loss but no gain. Proc. Natl. Acad. Sci. USA **100:** 7158–7162.

Saxena, K. M. S., and A. L. Hooker, 1968 On the structure of a gene for disease resistance in maize. Proc. Natl. Acad. Sci. USA **61:** 1300.

Shepherd, K. W., and G. M. E. Mayo, 1972 Genes conferring specific plant disease resistance. Science **175:** 375–380.

Simmonds, N. W., 1991 Genetics of horizontal resistance to disease of crop. Biol. Rev. **66:** 189–241.

Smith, S. M., and S. H. Hulbert, 2005 Recombination events generating a novel *Rp1* race specificity. Mol. Plant-Microbe Interact. **18:** 220–228.

Soderlund, C., K. Haller, V. Pampanwar, D. Ebbole, M. Farman *et al.*, 2006 MGOS: a resource for studing *Magnaporthe grisea* and *Oryza sativa* interactions. Mol. Plant-Microbe Interact. **19:** 1055–1061.

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

Sun, Q., N. C. Collins, M. Ayliffe, S. M. Smith, J. Drake *et al.*, 2001 Recombination between paralogues at the *rp1* rust resistance locus in maize. Genetics **158:** 423–438.

Sun, X. L., Y. L. Cao, Z. F. Yang, C. G. Xu, X. H. Li *et al.*, 2004 *Xa26*, a gene conferring resistance to *Xanthomonas oryzae* pv. *Oryzae* in rice, encodes an LRR receptor kinase-like protein. Plant J. **37:** 517–527.

Tian, D., H. Akashi, E. Stahl, J. Bergelson and M. Kreitman, 2002 Signature of balancing selection in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **99:** 11525–11530.

Tian, D., B. Shaw, J. Chen, M. Kreitman and J. Bergelson, 2003 Fitness cost of *R*-gene mediated resistance in *Arabidopsis thaliana*. Nature **424:** 74–77.

Traut, T. W., 1994 The function and consensus motifs of nine types of peptide segments that form different types of nucleotide binding sites. Eur. J. Biochem. **229:** 9–19.

Vallejos, C. E., G. Astua-Monge, V. Jones, T. R. Plyler, N. S. Sakiyama *et al.*, 2006 Genetic and molecular characterization of the *I* locus of *Phaseolus vulgaris*. Genetics **172:** 1229–1242.

Van Daelen, R. A., J. J. F. Gerbens, F. Van Rusissen, J. Aarts, J. Honteleez *et al.*, 1993 Long-range physical maps of two loci (*Aps-1* and *GP79*) flanking the root-knot nematode resistance gene (*Mi*) near the centromere of tomato chromosome 6. Plant Mol. Biol. **23:** 185–192.

Van Der Biezen, E. A, and J. D. G. Jones, 1998 Plant disease-resistance protein and the gene-for-gene concept. Trends Biochem. Sci. **12:** 454–456.

Velent, B., 1990 Rice blast as a model system for plant pathology. Phytopathology **80:** 33–36.

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 repeats class of plant disease resistance genes. Plant J. **19:** 55–64.

Wei, F., K. Gobelman-Werner, S. M. Morroll, J. Kurth, L. Mao *et al.*, 1999 The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. Genetics **153:** 1929–1948.

Wei, F. S., R. A. Wing and R. P. Wise, 2002 Genome dynamics and evolution of the *Mla* (Powdery mildew) resistance locus in barley. Plant Cell **14:** 1903–1917.

Yahiaoui, N., P. Srichumpa, R. Dudler and B. Keller, 2004 Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. Plant J. **37:** 528–538.

Yunoki, T., A. Ezuka, T. Morinaka, Y. Sakurai, H. Shinoda *et al.*, 1970 Studies on the varietal resistance of rice to blast. IV. Varietion of field resistance due to fungus strains (in Japanese with English summary). Bull. Chugoku Natl. Agric. Exp. Stn. **E6:** 21–41.

Zenbayashi-Sawata, K., S. Fukuoka, S. Katagiri, M. Fujisawa, T. Matsumoto *et al.*, 2006 Genetic and physical mapping of the partial resistance gene, *Pi34*, to blast in rice. Phytopathology **97:** 598–602.

Zhou, B., S. Qu, G. Liu, M. Dolan, H. Sakai *et al.*, 2006 The eight amino-acid differences witin three leucine-rich repeats between *Pi2* and *Piz-t* resistance proteins determine the resistance specificity to *Magnaporthe grisea*. Mol. Plant-Microbe Interact. **11:** 1216–1228.

Zhou, T., Y. Wang, J. Q. Chen, H. Araki, Z. Jing *et al.*, 2004 Genome-wide identification of NBS genes in *japonica* rice reveals significant expansion of divergent non-TIR NBS-LRR genes. Mol. Genet. Genomics **271:** 402–415.

Communicating editor: A. Paterson