Title

Evolution of the rice Xa21 disease resistance gene family.

Journal

The Plant cell, 9(8)

ISSN

1040-4651

Authors

Song, W Y

Pi, L Y

Wang, G L

et al.

Publication Date

1997-08-01

Peer reviewed

The Plant Cell, Vol. 9, 1279–1287, August 1997 © 1997 American Society of Plant Physiologists

Evolution of the Rice *Xa21* Disease Resistance Gene Family

Wen-Yuan Song,[a] Li-Ya Pi,[a] Guo-Liang Wang,[a,1] John Gardner,[b] Tom Holsten,[a,2] and Pamela C. Ronald[a,b,3]

[a] Department of Plant Pathology, University of California, Davis, California 95616

[b] Center for Engineering Plants for Resistance against Pathogens, University of California, Davis, California 95616

The rice disease resistance gene *Xa21*, encoding a receptor-like kinase, is a member of a multigene family. Sequence analysis of seven family members revealed two distinct classes of genes. One member from each class encodes a receptor kinase–like open reading frame. The other five members encode truncated open reading frames of the predicted receptor kinase. A highly conserved 233-bp sequence (HC) was also identified among the seven family members. Recombination at the HC region between family members apparently resulted in the precise swapping of promoter regions. Large sequence duplications were generated by a presumed unequal crossover event in intergenic regions. Insertions of transposon-like sequences truncated two of the predicted open reading frames. A model for amplification and diversification of the *Xa21* gene family is presented.

INTRODUCTION

The survival of most organisms depends on the presence of specific genetic systems to maintain diversity in the face of a changing environment. Classic examples include antigenic variation in trypanosomes and immunoglobulin gene formation in mammals (Johnson et al., 1991; Plasterk, 1992). Similarly, most plant species contain many highly polymorphic disease resistance genes that are related in sequence (Pryor and Ellis, 1993). It has long been speculated that DNA alterations play a key role in the evolution of these genes, thus allowing plants to generate new resistances to match the changing pattern of pathogen virulence (Pryor and Ellis, 1993). In support of this hypothesis, recombination at disease resistance loci has been observed to be associated with the creation of novel resistance phenotypes (Richter et al., 1995). However, the molecular basis for the evolution of disease resistance genes remains largely unknown.

The recent cloning of several disease resistance genes and the characterization of their predicted products represent a breakthrough in our understanding of the molecular basis of plant disease resistance (Grant et al., 1995; Song et al., 1995; Staskawicz et al., 1995; Dixon et al., 1996). Based on the predicted protein structures, the cloned resistance genes can be grouped into five classes. The first class is represented by the maize *HM1* gene, which encodes a reductase that inactivates a fungal-produced toxin (Johal and Briggs, 1992). The second class includes the tomato gene *Pto*, which confers resistance to *Pseudomonas syringae* pv *tomato* containing the avirulence gene *avrPto* (Ronald et al., 1992a; Martin et al., 1993). *Pto* encodes an active serine/threonine kinase that interacts both with the *avrPto* gene product and with a serine/threonine kinase encoded by the tomato *Pti* gene (Loh and Martin et al., 1995; Zhou et al., 1995; Scofield et al., 1996; Tang et al., 1996). Resistance genes from Arabidopsis (*Rps2* and *Rpm1*), tobacco (*N*), flax (*L6*), and tomato (*Prf*) are members of the third class of plant resistance genes (Bent et al., 1994; Mindrinos et al., 1994; Whitham et al., 1994; Grant et al., 1995; Lawrence et al., 1995; Salmeron et al., 1996). Common features of this class are the presence of leucine-rich repeats (LRRs), a putative cytoplasmic signaling domain, and a nucleotide binding site. A fourth class of resistance genes is represented by the tomato *Cf* class (*Cf-9* and *Cf-2*). These genes mediate resistance to the fungal pathogen *Cladosporium fulvum* (Jones et al., 1994; Dixon et al., 1996). The *Cf* genes encode putative membrane-anchored proteins with the LRR motif in the presumed extracellular domain and a short C-terminal tail in the intracellular domain. The fifth class of disease resistance gene is represented by the rice gene *Xa21*, conferring resistance to *Xanthomonas oryzae* pv *oryzae* (*Xoo*) (Song et al., 1995; Wang et al., 1996). *Xa21* encodes a receptor-like kinase consisting of LRRs in the putative extracellular domain similar to the tomato CF-9 protein and a serine/threonine kinase in the putative intracellular domain similar to the PTO kinase. Thus, the structure of *Xa21* indicates an evolutionary link between different classes of plant disease resistance genes.

Genomic DNA gel blot analysis revealed that *Xa21* belongs to a multigene family containing at least eight members (Ronald et al., 1992b; Song et al., 1995). Most of these members can be mapped to a single locus on chromosome

[1] Current address: Institute of Molecular Agrobiology, National University of Singapore, 59A The Fleming, 1 Sci Park Drive, Singapore 118240.

[2] Current address: Calgene, 1920 Fifth Street, Davis, CA 95616.

[3] To whom correspondence should be addressed. E-mail pcronald@ucdavis.edu; fax 916-752-5674.

11 that is linked to at least nine major resistance genes and one quantitaive trait locus for resistance (Ronald et al., 1992b; Song et al., 1995; L.-H. Zhu, personal communication). Furthermore, pulsed-field gel electrophoresis analysis demonstrated that most members of the *Xa21* gene family are located in a 230-kb genomic region (Ronald et al., 1992b; Williams et al., 1996). We have cloned seven *Xa21* gene family members from the resistant rice line IRBB21 designated A1, A2, B (*Xa21*), C, D, E, and F (Song et al., 1995; Wang et al., 1995). In this study, we demonstrate that transposition, recombination, and duplication contribute to the genetic diversity of the rice *Xa21* disease resistance gene family.

RESULTS

Physical Organization of the Cloned *Xa21* Gene Family Members

Sequence data and restriction enzyme analysis of cosmid and bacterial artificial chromosome clones indicated that the seven *Xa21* family members are contained on four contigs (Figure 1). The first contig, carrying *Xa21* and member C, spans a 40-kb region; the second contig includes members D, A1, and A2 and occupies a 150-kb region; and contigs of 130 and 40 kb contain members E and F, respectively. The observation that the member E–hybridizing DNA fragment is absent from the *Xa21* donor line (*Oryzae longistaminata*) but present in the recurrent parent IR24 and the introgression line IRBB21 suggests that member E is derived from the recurrent parent (Ronald et al., 1992b).

Transposable Element–like Sequences in the Coding Regions of the *Xa21* Gene Family

An interesting feature of *Xa21* family members is the presence of 15 transposable element–like sequences. Sequences were designated as transposons based on possession of terminal inverted repeats, duplication of target site sequence, or structural similarity and sequence identity with known transposons. The positions of these elements are shown in Figure 1. Thirteen elements occur in noncoding regions (W.-Y. Song, L.-Y. Pi, and P.C. Ronald, unpublished data), whereas two elements, named *Retrofit* and *Truncator*, occur in the coding regions of members D and E, respectively; this arrangement results in duplication of target sequences and disruption of the open reading frames (ORFs) of these two members (Figure 1, numbers 3 and 9). The only apparent cause for premature truncation of D and E is the insertion of *Retrofit* and *Truncator*, respectively. *Retrofit* belongs to the Drosophila copia class of retrotransposons and carries a large ORF showing greatest similarity to the ORF of maize *Hopscotch* (68.6% similarity; 54.6% identity) and tobacco *Tnt1* (51.4% similarity; 31.9% identity) (Grandbastien et al., 1989; White et al., 1994). The insertion site of this element is located between the 23rd (V) and 24th (P) amino acids of the 22nd LRR, creating a new stop codon. Translation of copy D would lead to a truncated molecule lacking the transmembrane and kinase domains (Figure 2A). Insertion of *Retrofit* into a presumed coding region contrasts with the observation in yeast and maize that integration of retrotransposons is biased toward noncoding regions (SanMiguel et al., 1996; Voytas et al., 1996).

Truncator (2913 bp) represents a novel transposon-like sequence carrying a 9-bp target site repeat. The sequence shows no substantial similarity to any sequence in the database and contains no obvious ORFs. Insertion of this element into the N terminus of the kinase domain of member E would presumably result in premature truncation of the receptor kinase, resulting in a receptor-like molecule structurally similar to the tomato fungal resistance gene products CF-9 and CF-2 (Figure 2B; Jones et al., 1994; Dixon et al., 1996).

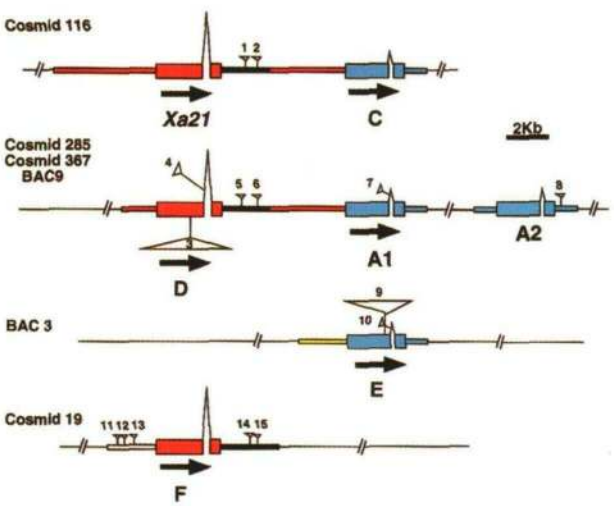

Figure 1. Genome Organization of the Seven *Xa21* Family Members and Location of 15 Transposon-like Elements.

Cosmid and bacterial artificial chromosome (BAC) clones carrying the family members are designated. Wide bars represent predicted coding regions; narrow bars represent noncoding regions. Introns are indicated by angled lines, and the nonsequenced regions are shown by thin lines. A gap in the sequence of BAC9 is indicated by the broken line. Sequences sharing high levels of identity (>95%) are labeled in the same colors. Letters refer to names of *Xa21* gene family members, and arrows indicate the direction of the ORFs. The 15 transposon-like elements are numbered and represented by open triangles. The scale of the coding and flanking regions is indicated by the scale bar. Introns and transposable elements are not drawn to scale.

Two Classes of Genes Encoding Receptor Kinase–like Proteins at the *Xa21* Locus

The entire coding region, the intron, and the 3′ flanking region of the seven family members can be grouped into two classes based on DNA similarity (Figures 1 and 3). One class (designated the *Xa21* class) contains *Xa21* as well as members D and F. The second class (designated the A2 class) contains members A1, A2, C, and E. Within each class, family members share striking nucleotide sequence identity (98.0% average identity for the members of the *Xa21* class; 95.2% average identity for the members of the A2 class). In contrast, a low level of DNA sequence identity was observed between members of the two classes (e.g., 63.5% identity between *Xa21* and A2) (Figure 3).

Only the *Xa21* and A1 ORFs encode receptor kinase–like proteins. Sequences of the other family members contain alterations that cause a premature truncation of the predicted receptor kinase–like ORF (small deletions in F and C, base pair mutations in A2, or transposon insertions in D and E). A1 and XA21 share 68.6% amino acid identity overall (Figure 4). Domains I and II, carrying the presumed signal peptide and N terminus of the protein, are 100% identical, whereas the LRR domain (domain III) of XA21 and A1 share a low level of identity (59.5%) and differ in the number of LRRs (23 versus 22, respectively). Domain IV contains charged amino acid residues with 57.1% identity. The putative transmembrane domain (domain V) shows the lowest sequence identity (34.6%) between XA21 and A1. Compared with the relatively low level of identity (64%) for the juxtamembrane domain (domain VII), the catalytic domains (domain VIII) of XA21 and A1 are highly conserved (82% identity). Finally, there is low sequence identity between Xa21 and A1 in domain IX, which constitutes the C termini of the receptor kinase–like proteins (38.5% identity). The differences observed between members of the two classes suggest that they may differ in function. Indeed, we have found that transgenic plants containing the A1 sequence are susceptible to all *Xoo* isolates tested (G.-L. Wang and P.C. Ronald, unpublished data).

Identification of a Highly Conserved GC-Rich Region

Sequence comparisons revealed that a 233-bp region (in terms of the length in *Xa21*) located immediately downstream of the start codon is highly conserved among all of the sequenced members (Figure 3). In contrast to the 40.9% average GC content of the rice genome (Kwon and Son, 1985) and the 40.0% GC content of the 8318-bp sequence including and surrounding the *Xa21* gene, the highly conserved (HC) region has a high GC content (61.8% for *Xa21*) hallmarked by the typical GC-rich restriction enzyme recognition site NotI. The predicted amino acid sequence of the HC region spans domains I and II of XA21 and shares nearly 100% identity among most family members (Figure 4).

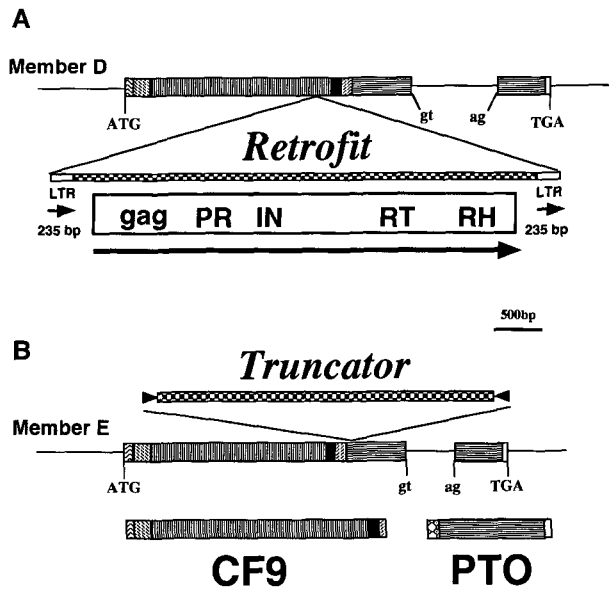

Figure 2. Position of Transposable-like Sequences in the Coding Regions of the *Xa21* Gene Family Members D and E.

(A) Family member D and insertion position of *Retrofit*. *Retrofit* carries long terminal repeats (LTRs; small arrows) and a single large ORF encoding a protein with the following domains: gag, protease (PR), integrase (IN), reverse transcriptase (RT), and RNase H (RH). The large arrow indicates direction of the ORF.

(B) Family member E and insertion position of *Truncator*. Arrowheads mark the orientation of the inverted repeats. The deduced protein structures of the tomato resistance genes *Cf-9* and *Pto* are shown below.

In both **(A)** and **(B),** the insertion elements are designated by hatched bars. The presumed deduced amino acid sequences of members D and E are shown by shaded rectangles. Domain representations are as described in the legend to Figure 4.

Evidence for HC Region–Mediated Recombination between Family Members within Intragenic Regions

In addition to the transposition events presented above, recombination at the HC region was also found to play an important role in the evolution of the *Xa21* gene family. Four classes of DNA sequences (~1.3 kb) can be delimited upstream of the HC region (Figure 3). For example, the 5′ flanking region of family member F is divergent from that of other family members (<40% identity). The precise break point (from sequence similarity to divergence) between *Xa21* and F is located within the HC region, 120 bp downstream from the start codon (Figure 3). This sudden change in sequence identity is not likely to result from random events, such as transposon insertion or deletion, because such events presumably would lead to an altered coding region. This is not the case; the deduced amino acid sequence of F

maintains the receptor kinase–like ORF (Figure 5). These results suggest that a recombination event occurred in the HC region, resulting in the formation of a chimeric sequence containing the 5′ flanking region of F and a downstream region (including coding region, intron, and 3′ flanking region) of the *Xa21* class.

In further support of the idea that the HC region mediates intragenic recombination, we also observed apparent recombination break points near or within the HC region for gene family members E, A1, and C (Figure 3). For E, the 5′ flanking region is divergent from all other members, whereas the 3′ downstream regions belong to the A2 class (Figure 3). The sudden change of DNA identity can be explained by a recombination event between a progenitor A2-type gene and an unknown family member. The likely recombination break point in E is located 105 bp upstream of the HC region because sequences upstream of this site are quite different when compared with the high level of DNA sequence identity downstream of this site.

The nearly identical DNA sequences of C and A1 provide the most striking example of a recombination event mediated by the HC region. For example, the 5′ flanking region of C shows nearly perfect identity (99.2%) to that of *Xa21*, whereas the downstream region of C belongs to the A2 class (Figure 3). The high level of identity between the 5′ flanking sequences of *Xa21* and C extends 3.8 kb upstream (data not shown). This upstream region includes the functional promoter for the *Xa21* gene (Song et al., 1995). These

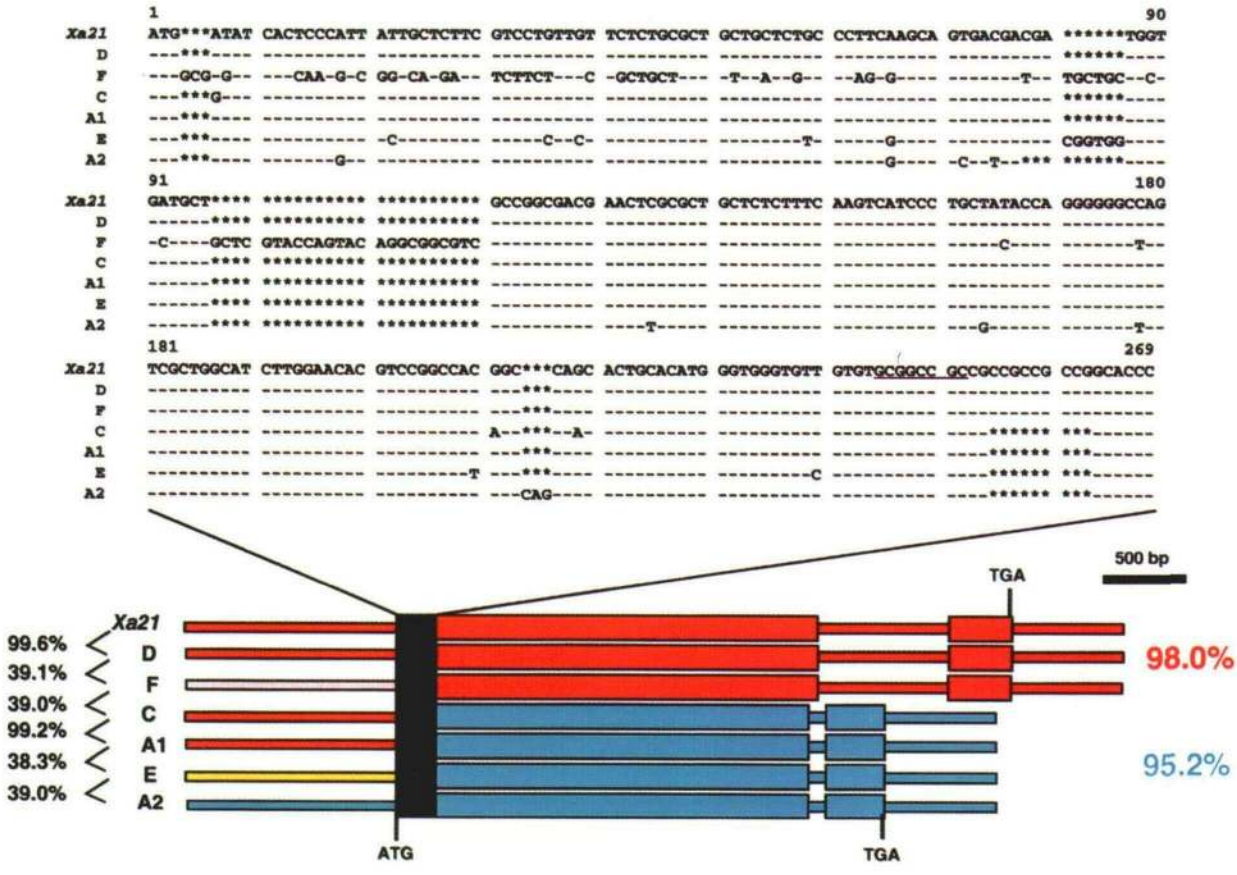

Figure 3. HC-Mediated Recombination Events between *Xa21* Gene Family Members.

Wide bars represent predicted coding regions; narrow bars represent noncoding regions. Start and stop codons are indicated. The 5′ flanking regions and downstream regions are grouped into four and two groups, respectively, and are shown in different colors based on sequence identity. The percentage of DNA sequence identity between promoter regions and between downstream classes is shown to the left and right, respectively. The HC region is indicated by a black bar. Sequence alignment of the HC region for different *Xa21* family members is shown above. Only nucleotides that differ from the *Xa21* sequence are indicated. Asterisks represent gaps for maximal alignment; dashes indicate identity. The NotI site is underlined. Transposon-like and microsatellite sequences (three stretches of microsatellite sequences with ATATAT repeats were identified at the 3′ regions of *Xa21* and members F and D; W.-Y. Song and P.C. Ronald, unpublished data) were removed for maximal sequence alignments (Figure 1).

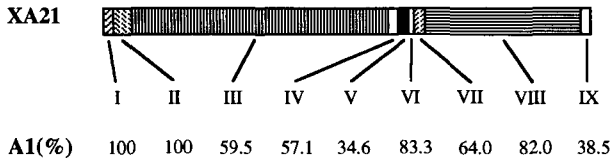

Figure 4. A Schematic Diagram Comparing the Predicted Amino Acid Sequences of XA21 and A1.

Domains are numbered as follows: I, presumed signal peptide; II, presumed N terminus; III, LRR; IV, charged amino acid; V, presumed transmembrane; VI, charged amino acid; VII, juxtamembrane; VIII, serine/threonine kinase; IX, C-terminal tail. The numbers below each domain indicate amino acid identity between XA21 and A1.

results strongly suggest that C was created by a recombination event in the HC region between progenitors of the *Xa21* and A2 classes. The likely recombination break point in member C is delimited by two characteristic deletions: one is located at position −37 and is present only in *Xa21* class members (*Xa21*, D, C, and A1) (data not shown); another deletion is located at position 255 and occurs in all A2 class members (Figure 3).

Evidence for Recombination in Intergenic Regions of the *Xa21* Gene Family Members

Evidence for recombination in the intergenic regions of the *Xa21* family members was also observed. First, sequences in the 5′ flanking region of members C and *Xa21* are identical for 3.8 kb and then abruptly diverge. The same site of divergence can be observed in the 3′ flanking regions of *Xa21* and member F (Figure 6). The presence of a conserved site of divergence suggests not only that this is a recombination break point but also that the *Xa21*/C cluster and member F are generated from the same progenitor (Figure 7). Second, the sequence of a 14,742-bp region spanning the *Xa21*/C cluster shows 97.7% identity to the corresponding sequence (14,871 bp) of the D/A1 cluster (Figure 1), suggesting that these regions evolved through sequence duplication. This duplication process can be explained by a presumed unequal crossover event in the flanking region of these two clusters.

DISCUSSION

We have sequenced the coding region and flanking sequences of seven *Xa21* gene family members. Based on sequence comparisons, recombination, duplication, and transposition were identified as the major events contributing to the evolution of this complex resistance locus.

Diversification of the *Xa21* Gene Family Members by Transposable-like Elements

We have identified 15 transposon-like elements at the *Xa21* locus. Thirteen of these elements insert into noncoding regions, whereas two elements insert into presumed coding regions. Integration of *Truncator* and *Retrofit* into the E and D coding sequences, respectively, creates ORFs that encode truncated proteins. For example, insertion of *Truncator* leads to the generation of a membrane-bound receptor similar to CF-9 (Jones et al., 1994). Thus, it is possible to envision the evolution of *Cf-9*– and *Pto*-type genes through separation of the receptor and kinase domains by insertion of a transposon into a tomato receptor kinase (Figure 2B), although in this case, the next methione is too far downstream to serve as a start codon for a functional kinase. Similarly, insertion of *Retrofit* could generate a presumably soluble receptor similar in structure to PGIP, an LRR-containing glycoprotein that specifically inhibits polygalacturonase activity in vitro (Stotz et al., 1994). Thus, our nucleotide sequence analysis suggests that the insertion of transposable-like elements can provide an effective tool for diversification of receptor kinases.

The HC Region Is Highly Conserved among All Sequenced *Xa21* Family Members

The HC region is present in all sequenced family members, including six members (A1, A2, *Xa21*, C, D, and F) inherited from the wild rice species *O. longistaminata* and one family member (E) presumably inherited from the cultivated rice species *O. sativa* IR24 (Ronald et al., 1992b). Moreover, the sequence of RG103 (from *O. sativa* IR36), which spans the HC region, also shows this conservation (McCouch et al., 1988; Ronald et al., 1992b; Song et al., 1995). The HC region is located immediately downstream of the start codon and spans domains I and II of XA21. The predicted amino acid sequence of domain I suggests that it functions as a signal peptide, whereas domain II has no known function (Song et al., 1995). These two domains have also been identified in several plant receptor–like kinase proteins containing

	I	II
F	MARSPTSVMISSLLLLLLIGPASS	DDDAAAAAARTSTGGV
XA21	MISLP*LLLFVLLFSALLLCPSSS	DDDGDA

Figure 5. The Deduced Amino Acid Sequence of F in the HC Region Maintains the Receptor Kinase–like ORF.

The predicted amino acid sequences encoded by the first 120 bp of the HC region of homolog F and the corresponding region of XA21 preceding the presumed recombination break point are shown. The asterisk in XA21 represents a gap for maximal alignment. Identical amino acids are indicated in boldface type.

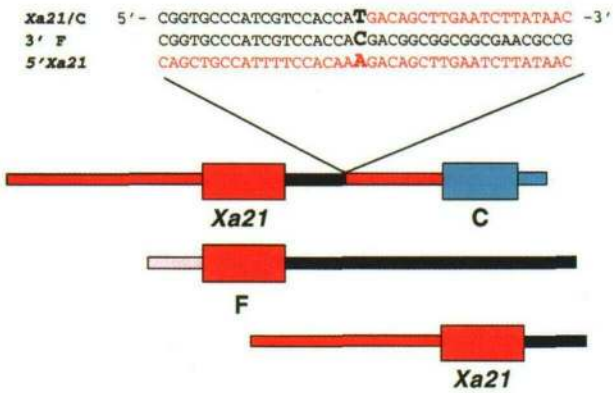

Figure 6. Presumed Recombination Break Point in the Intergenic Region of *Xa21* Family Members *Xa21*, C, and F.

Boxes represent the ORFs of the designated family members; bars represent the flanking regions. The same colors indicate a high level of sequence homology. The nucleotides of the presumed recombination break points are indicated in large, boldface type. Sequences surrounding the recombination break point are also shown.

the LRR motif (e.g., RLK5, TMK1, TMKL1, and ERECTA) and the transmembrane receptor–like proteins CF-9 and CF-2 (Jones et al., 1994; Walker, 1994; Dixon et al., 1996; Torii et al., 1996); however, the DNA sequences of the regions encoding domains I and II of these proteins vary in length and nucleotide composition.

Duplication, Recombination, and Diversification of *Xa21* Gene Family Members

Like many plant genes, most cloned resistance genes are members of multigene families, indicating that gene duplication and subsequent diversification are common processes in plant gene evolution (Ronald et al., 1992b; Martin et al., 1993; Jones et al., 1994; Whitham et al., 1994; Lawrence et al., 1995; Song et al., 1995; Dixon et al., 1996). For example, the presumed duplication and diversification of the *Pto* gene family lead to the generation of alternative recognition capabilities of the encoded proteins (Martin et al., 1994; Zhou et al., 1995). Family member FEN carries 87% amino acid identity to PTO but confers sensitivity to the insecticide fenthion rather than resistance to *P. s. tomato*.

Studies of the plant-specific polyketide synthase supergene family provide another example of duplication and functional divergence. In this model, the stilbene synthase (SS) and GCHS2 genes evolved from chalcone synthase (CHS) through gene duplication (Tropf et al., 1994; Helariutta et al., 1996). The DNA sequence and protein catalytic properties of the SS and GCHS2 genes and encoded products differ from those of CHS (Fliegmann et al., 1992; Helariutta et al., 1995). These results indicate that various related functions can arise from an ancestral gene during evolution. Similarly, our data suggest that duplication and subsequent divergence generated two *Xa21* subfamilies, although in this case, a functional role for the A2 family remains to be determined.

In addition to duplication events, recombination contributes to the observed diversity of plant gene families. For example, recombination at the maize disease resistance locus *Rp1* has been observed to be associated with the creation of novel resistance phenotypes (Richter et al., 1995). However, genes at the *Rp1* locus have not yet been cloned, which prevents a detailed analysis of the recombinational events that lead to the novel recognition. In contrast, recombinants at the maize alcohol dehydrogenase1 (*Adh1*) locus have been characterized extensively. These studies have demonstrated that intraallelic but not interallelic recombination occurs among *Adh1* alleles (Freeling, 1978; Osterman and Dennis, 1989; Gaut and Clegg, 1991, 1993; Akada and Dube, 1995). Molecular characterization of the *Adh1* alleles revealed substantial sequence divergence in the 3′ flanking regions of these alleles, suggesting that interallelic recombination is blocked because of disruption of pairing in this region (Osterman and Dennis, 1989).

The absence of interallelic recombination among *Adh1* alleles indicates that sequence diversification tends to reduce the frequency of subsequent recombination events. However, in the case of the *Xa21* gene family, recombination events between diverged family members (such as *Xa21* and A2) do occur and presumably are maintained by the presence of the HC region. The high GC content of the HC

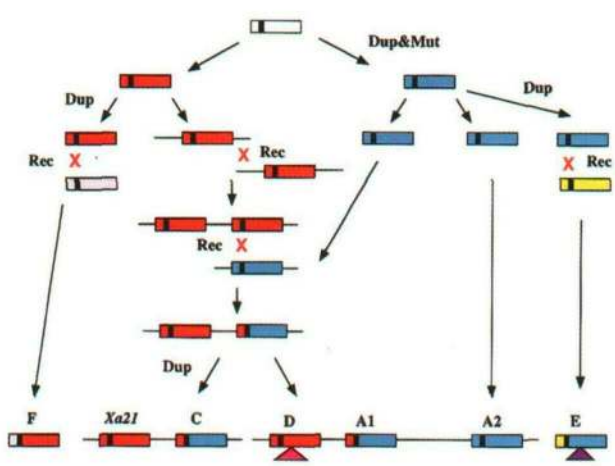

Figure 7. Model for the Evolution of *Xa21* Family Members.

Wide bars represent *Xa21* family members. Colors indicate relationships among these members: white, presumed receptor kinase progenitor; red, *Xa21*-type family members; blue, A2-type family members. The HC region is shown in black. Pink and yellow boxes indicate unknown family members. The presumed recombination sites are indicated by X's. Triangles represent integration of *Retrofit* (pink) and *Truncator* (purple). Evolutionary events are indicated as follows: Dup, duplication; Mut, mutation; Rec, recombination.

region would serve to stabilize pairing in this region, allowing for establishment of a recombination initiation site. Thus, the HC region may function to promote exchange of sequence information between divergent family members. Precedents exist in animal systems in which recombination between duplicated, diverged gene family members is associated with conserved DNA stretches (Plasterk, 1992). However, to our knowledge, recombination mediated by a conserved region has not been previously observed in plants.

Although the mechanism for HC region–mediated recombination is not known, at least two models can be envisioned. First, this region may mediate programmed recombination similar to that observed in African trypanosomes (Plasterk, 1992). In trypanosomes, antigenic variation is controlled by a variant surface glycoprotein (VSG), which is encoded by a member of a multigene family containing >1000 members. Recombination at stretches of highly conserved nucleotides between silent and expressed members of the VSG gene family leads to expression of new antigens. Alternatively, HC-mediated recombination may be an example of an ectopic recombination event wherein the HC region serves as a recombination initiation site (Petes and Hill, 1988). Frequent recombination in this region would maintain the conservation of the HC region but allow flanking sequences to diverge. Over time, mismatch repair would lead to homogenization of the HC region and result in an overall increased GC content, as has been observed in yeast (Brown and Jiricny, 1988). In both of these models, the outcome of the recombination is precise swapping of promoter regions of the *Xa21* gene family members.

A Model for the Evolution of the *Xa21* Gene Family

From the sequence analysis, a model for the evolution of the *Xa21* family members can be envisioned (Figure 7). First, duplication and subsequent divergence of a progenitor receptor kinase gene create the *Xa21* and A2 families. Second, unequal recombination between the family members generates additional members. Third, recombination at the HC region allows formation of novel promoter/coding region combinations, as exemplified by the chimeric structures of members F, E, and C. Fourth, a large duplication of the *Xa21*/C cluster further amplifies the resistance gene family. Finally, recent integration of transposable elements into *Xa21* family members results in genes with the capacity to encode novel proteins.

METHODS

DNA Analysis

Plasmid DNA isolation, restriction digests, subcloning, DNA blotting, and gel blot hybridization were achieved according to standard protocols (Sambrook et al., 1989).

DNA Sequencing and Analysis

DNA sequencing was accomplished by use of an automated sequencer (model 4000L; Li-Cor, Lincoln, NE) and the Sequitherm long read cycle sequencing kit (Epicentre Technologies, Madison, WI). To fill in gaps, a primer walking strategy was performed with synthesized primers (Operon, Alameda, CA) and a DNA sequencer (model 373; Applied Biosystems, Foster City, CA). The Genetics Computer Group (Madison, WI) sequence analysis programs GAP and PILEUP were used to calculate the percentage of identity and to perform multiple alignments of DNA and protein sequences, respectively. GenBank accession numbers are as follows: A1, U72725; A2, U72727; Xa21/C, U72723; D, U72726; E, U72724; F, U72728; 3′ flanking region of F, U72729; and RG103, U82168.

ACKNOWLEDGMENTS

We thank John Luke and Dean Lavelle for technical assistance and Scot Hulbert, Charles Langley, Brandon Gaut, and Chao-Qiang Lai for valuable discussions. We are grateful to Veronique Szabo, Maureen Whalen, and Steve Scofield for critical reading of the manuscript. This work was supported by the National Institutes of Health, the U.S. Department of Agriculture, the Rockefeller Foundation, and the National Science Foundation.

Received April 3, 1997; accepted May 21, 1997.

REFERENCES

Akada, S., and Dube, S. (1995). Organization of soybean chalcone synthase gene cluster and characterization of a new member of the family. Plant Mol. Biol. **29,** 189–199.

Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J., and Staskawicz, B.J. (1994). *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. Science **265,** 1856–1859.

Brown, T., and Jiricny, J. (1988). Different base/base mispairs are corrected with different efficiencies and specificities in monkey kidney cells. Cell **54,** 705–711.

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K., and Jones, J.D.G. (1996). The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell **84,** 451–459.

Fliegmann, J., Schroder, G., Schanz, S., Britsch, L., and Schroder, J. (1992). Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. Plant Mol. Biol. **18,** 489–503.

Freeling, M. (1978). Allelic variation at the level of intragenic recombination. Genetics **98,** 211–224.

Gaut, B.S., and Clegg, M.T. (1991). Molecular evolution of alcohol dehydrogenase 1 in members of the grass family. Proc. Natl. Acad. Sci. USA **88,** 2060–2064.

Gaut, B.S., and Clegg, M.T. (1993). Molecular evolution of the *Adh1* locus in the genus *Zea*. Proc. Natl. Acad. Sci. USA **90,** 5095–5099.

Grandbastien, M.A., Spielmann, A., and Caboche, M. (1989). *Tnt1*, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature **337,** 376–380.

Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W., and Dangl, J.L. (1995). Structure of the *Arabidopsis Rpm1* gene enabling dual specificity disease resistance. Science **269,** 843–846.

Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R.J., Schroder, J., and Teeri, T.H. (1995). Chalcone synthase–like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (Asteraceae). Plant Mol. Biol. **28,** 47–60.

Helariutta,Y., Kotilainen, M., Elomaa, P., Kalkkinen, N., Bremer, K., Teeri, T.H., and Albert, V.A. (1996). Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: Evolution with substrate change and catalytic simplification. Proc. Natl. Acad. Sci. USA **93,** 9033–9038.

Johal, G.S., and Briggs, S.P. (1992). Reductase activity encoded by the HM1 disease resistance gene in maize. Science **258,** 985–987.

Johnson, D.E., Lu, J., Chen, H., Werner, S., and Williams, L.T. (1991). The human fibroblast growth factor genes: A common structured arrangement underlies the mechanisms for generating receptor forms that differ in their third immunoglobulin domain. Mol. Cell. Biol. **11,** 4627–4634.

Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J., and Jones, J.D.G. (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science **266,** 789–793.

Kwon, Y.-W., and Son, C.-H. (1985). Characteristics of nuclear DNA in rice roots of *japonica* and *indica* × *japonica* varieties. Seoul Natl. Univ. Coll. Agric. Bull. **10,** 63–68.

Lawrence, G.J., Finnegan, E.J., Ayliffe, M.A., and Ellis, J.G. (1995). The *L6* gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. Plant Cell **7,** 1195–1206.

Loh, Y.T., and Martin, G.B. (1995). The *Pto* bacterial resistance gene and the *Fen* insecticide sensitivity gene encode functional protein kinase with serine/threonine specificity. Plant Physiol. **108,** 1735–1739.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D., and Tanksley, S.D. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science **262,** 1432–1436.

Martin, G.B., Frary, A., Wu, T., Brommonschenkel, S.H., Chunwongse, J., Earle, E.D., and Tanksley, S.D. (1994). A member of the tomato *Pto* gene family confers sensitivity to fenthion resulting in rapid cell death. Plant Cell **6,** 1543–1552.

McCouch, S., Kochert, G., Yu, Z., Wang, Z., Khush, G., Coffman, W., and Tanksley, S.D. (1988). Molecular mapping of the rice nuclear genome. Theor. Appl. Genet. **76,** 815–829.

Mindrinos, M., Katagiri, F., Yu, G.-L., and Ausubel, F.M. (1994). The A. thaliana disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. Cell **78,** 1089–1099.

Nicolas, A., Treco, D., Schultes, N.P., and Szostak, J.W. (1989). An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*. Nature **338,** 35–39.

Osterman, J.C., and Dennis, E.S. (1989). Molecular analysis of the *ADH1-C*^{m} allele of maize. Plant Mol. Biol. **13,** 203–212.

Petes, T.D., and Hill, C.W. (1988). Recombination between repeated genes in microorganisms. Annu. Rev. Genet. **22,** 147–168.

Plasterk, R.H. (1992). Genetic switches: Mechanism and function. Trends Genet. **8,** 403–406.

Pryor, T., and Ellis, J. (1993). The genetic complexity of fungal resistance genes in plants. Adv. Plant Pathol. **10,** 281–305.

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

Ronald, P.C., Salmeron, J.M., Carland, F.M., and Staskawicz, B.J. (1992a). The cloned avirulence gene *avrPto* induces disease resistance in tomato cultivars containing the *Pto* resistance gene. J. Bacteriol. **174,** 1604–1611.

Ronald, P.C., Albano, B., Tabien, R., Abenes, L., Wu, K.S., McCouch, S., and Tanksley, S.D. (1992b). Genetic and physical analysis of the rice bacterial blight disease resistance locus, *Xa21*. Mol. Gen. Genet. **236,** 113–120.

Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H.S., Lavelle, D.T., Dahlbeck, D., and Staskawicz, B.J. (1996). Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. Cell **86,** 123–133.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

SanMiguel, P., Tikhonov, A., Jin, Y.-K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z., and Bennetzen, J.L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science **274,** 765–768.

Scofield, S.R., Tobias, C.M., Rathjen, J.P., Chang, J.H., Lavelle, D.T., Michelmore, R.W., and Staskawicz, B.J. (1996). Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. Science **274,** 2063–2065.

Song, W.-Y., Wang, G.-L., Chen, L.-L., Kim, H.S., Pi, L.-Y., Gardner, J., Wang, B., Holsten, T., Zhai, W.-X., Zhu, L.-H., Fauquet, C., and Ronald, P.C. (1995). A receptor kinase–like protein encoded by the rice disease resistance gene *Xa21*. Science **270,** 1804–1806.

Staskawicz, B.J., Ausubel, F.M., Baker, B., Ellis, J.G., and Jones, J.D.G. (1995). Molecular genetics of plant disease resistance. Science **268,** 661–667.

Stotz, H.H., Contos, J.J., Powell, A.L., Bennett, A.B., and Labavitch, J.M. (1994). Structure and expression of an inhibitor of fungal polygalacturonases from tomato. Plant Mol. Biol. **25,** 607–617.

Tang, X., Frederick, R.D., Zhou, J., Halterman, D.A., Jia, Y., and Martin, G.B. (1996). Initiation of plant disease resistance by physical interaction of AvrPto and Pto kinase. Science **274,** 2060–2063.

Torii, K.-U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R.F., and Komeda, Y. (1996). The Arabidopsis *ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. Plant Cell **8,** 735–746.

Tropf, S., Lanz, T., Rensing, S.A., Schroder, J., and Schroder, G. (1994). Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. J. Mol. Evol. **38,** 610–618.

Voytas, D.F. (1996). Retroelements in genome organization. Science **274,** 737–738.

Walker, J.C. (1994). Structure and function of the receptor-like protein kinases of higher plants. Plant Mol. Biol. **26,** 1599–1609.

Wang, G.-L., Holsten, T.E., Song, W.-Y., Wang, H.-P., and Ronald, P.C. (1995). Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa21* disease resistance locus. Plant J. **7,** 525–533.

Wang, G.-L., Song, W.-Y., Wu, R.-L., Sideris, S., and Ronald, P.C. (1996). The cloned gene, *Xa21*, confers resistance to multiple *Xanthomonas oryzae* pv. *oryzae* isolates in transgenic plants. Mol. Plant-Microbe Interact. **9,** 850–855.

White, S.E., Habera, L.F., and Wessler, S.R. (1994). Retrotransposons in the flanking regions of normal plant genes: A role for copia-like elements in the evolution of gene structure and expression. Proc. Natl. Acad. Sci. USA **91,** 11792–11796.

Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of the tobacco mosaic virus resistance gene *N*: Similarity to Toll and the interleukin-1 receptor. Cell **78,** 1101–1115.

Williams, C.E., Wang, B., Holsten, T.E., Scambray, J., da Silva, F., and Ronald, P.C. (1996). High resolution linkage map of the region containing the rice *Xa21* disease resistance gene. Theor. Appl. Genet. **93,** 1119–1122.

Zhou, J., Loh, Y.T., Bressan, R.A., and Martin, G.B. (1995). The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by *Pto* and is involved in the hypersensitive response. Cell **83,** 925–935.