The Product of the Tobacco Mosaic Virus Resistance Gene N: Similarity to Toll and the Interleukin-1 Receptor

Steve Whitham,* S. P. Dinesh-Kumar,† Dolil Choi,‡ Reinhard Hehl,§ Catherine Corr,† and Barbara Baker†§
*Department of Plant Pathology
†Department of Plant Biology
University of California
Berkeley, California 94720
§Plant Gene Expression Center
University of California
and United States Department of Agriculture
Agriculture Research Service
Albany, California 94710

Summary

The products of plant disease resistance genes are postulated to recognize invading pathogens and rapidly trigger host defense responses. Here we describe isolation of the resistance gene N of tobacco that mediates resistance to the viral pathogen tobacco mosaic virus (TMV). The N gene was isolated by transposon tagging using the maize Activator transposon. A genomic DNA fragment containing the N gene conferred TMV resistance to TMV susceptible tobacco. Sequence analysis of the N gene shows that it encodes a protein of 131.4 kDa with an amino-terminal domain similar to that of the cytoplasmic domain of the Drosophila Toll protein and the interleukin-1 receptor (IL-1R) in mammals, a nucleotide-binding site (NBS), and 4 imperfect leucine-rich repeats (LRR). The sequence similarity of N, Toll, and IL-1R suggests that N mediates rapid gene induction and TMV resistance through a Toll-IL-1-like pathway.

Introduction

Plants actively resist diseases caused by viruses, bacteria, fungi, and nematodes by elaborating effective local resistance responses that halt pathogen growth and spread. The hypersensitive response (HR) is the most commonly activated resistance response and is characterized by necrotic lesions at the site of pathogen ingress. Activation of HR is, however, specific and induced upon interaction of plants and pathogens genetically endowed with the capacity for mutual recognition (Keen, 1990). Recognition is postulated to result from the interaction of the product of a plant resistance gene with a corresponding pathogen avirulence gene product (Keen, 1990). Flor (1947) proposed this "corresponding genes" model of plant and pathogen interaction based on studies of the fungal rust pathogen of flax, Melampora lini. The model accurately accounts for the outcome of numerous plant–pathogen interactions. Resistance genes that mediate HR are often dominant traits and have been hypothesized to encode products that function as receptors for recognition of specific pathogen avirulence gene products and initiation of signal transduction pathways leading to expression of the resistance responses (Keen, 1992). If either the plant or pathogen partner lacks a functional allele of the corresponding gene pair, then resistance is not triggered and the plant becomes diseased. It follows that plants and pathogens can harbor a number of corresponding pairs of resistance and avirulence genes that interact to provide the signals for induction of disease resistance (Keen, 1990).

The corresponding genes hypothesis has gained support through the identification and isolation of avirulence genes from three major pathogen groups. The first pathogen avirulence gene was isolated from Pseudomonas syringae pv. glycinia, which causes bacterial blight on soybean (Saskawicz et al., 1984). Since that time, several avirulence genes have been isolated and characterized from bacteria, fungi, and viruses (Keen et al., 1993). In some cases, the role of the avirulence gene product as an elicitor has been established. In P. syringae pv. glycinia, the avrD gene mediates production of C-glycosyl lipid syringolide elicitors (Yuce et al., 1994) that induce HR in soybean containing the resistance gene Rps4. The fungal avirulence genes avr9 and avr4 of Cladosporium fulvum encode products that elicit HR in tomato carrying avr9 and avr4 resistance genes, respectively (Mathieu et al., 1994; van den Ackerveken et al., 1993). A viral avirulence gene, tobacco mosaic virus (TMV) coat protein, has also been identified that elicits the N gle-mediated HR in Nicotiana sylvestris (Culver and Dawson, 1991).

Elucidation of the nature of resistance genes has been facilitated by advances in gene isolation procedures such as positional cloning (Rommens et al., 1989; Bernatzky and Tanksley, 1986) and transposon tagging using maize transposons in maize as well as heterologous plant species (Baker et al., 1986; Chuck et al., 1993). The HML gene of maize, which confers resistance to the fungal pathogen Cochliobolus carbonum Nelson race 1, was cloned by transposon tagging (Mohal and Briggs, 1992). HML encodes an HC toxin reductase that inactivates the HC toxin produced by C. carbonum race 1. The Pto gene of tomato, which confers resistance to P. syringae strains carrying the avrPto gene, has been isolated by positional cloning. Sequence analysis of Pto indicates that it encodes a serine/threonine kinase (Martin et al., 1993). The Pto kinase may interact directly or indirectly with the avrPto-encoded elicitor molecule and then phosphorylate a subsequent modulator of the resistance response, thereby participating in a signal transduction cascade.

The N–TMV interaction has long served as a classical model system for the study of plant resistance responses to pathogens. The TMV-induced HR is characterized by the formation of necrotic lesions in tobacco bearing
Figure 1. TMV-Induced Resistance and Disease Responses in Tobacco

Tobacco leaves from plants inoculated with TMV. (A) HR of the TMV<sup>+</sup> tobacco cv. Samsun NN. (B) Mosaic symptoms of a systemic infection of the TMV<sup>-</sup> tobacco cv. SR1 nn. (C) The HR and mosaic symptoms of a TMV<sup>+</sup> plant bearing an Ao-induced unstable mutation of N. (D) HR of a TMV<sup>-</sup> transgenic (T0) SR1 nn transformed with the pTG38 T-DNA construct containing the N gene.
the N gene (NN tobacco). The dominant N gene was introgressed into TMV-sensitive (TMV\textsuperscript{s}) Nicotiana tabacum from the related TMV-resistant (TMV\textsuperscript{r}) species Nicotiana glutinosa (Holmes, 1938). TMV infection of NN tobacco induces HR within 48 hr postinfection, and TMV is restricted to the region immediately surrounding the necrotic lesions (Figure 1A). In contrast, TMV\textsuperscript{s} tobacco cultivars (nn tobacco) allow TMV to spread systemically and develop mosaic symptoms characterized by intermittent areas of light and dark green leaf tissue (Figure 1B). TMV is a mechanically transmitted positive sense RNA virus that encodes four proteins: two are required for viral replication (126 kDa and 163 kDa), one is required for viral cell-to-cell movement (30 kDa), and one is required for viral RNA encapsidation (17.5 kDa) (Dawson, 1992). Although the TMV avirulence gene corresponding to N has not been conclusively identified, one study suggests that the 126 kDa replicase protein is required for HR induction in NN tobacco (Padgett and Beachy, 1993). TMV-induced HR is accompanied by induction of defense responses such as production of antimicrobial compounds (phytoalexins), lignin deposition, and synthesis of hydrolytic enzymes such as chitinase (Lamb et al., 1989; Lamb, 1994). These events are rapidly induced in NN plants but not in nn plants. The precise causal relationships among TMV activation of HR, the induction of defense responses, and the prevention of viral spread has not been established.

Here we describe the isolation of the N gene from tobacco by transposon tagging with the maize transposon Activator (Ac). A positive selection scheme was employed to isolate mutants unable to respond to TMV with an HR (TMV\textsuperscript{r} mutants). One TMV\textsuperscript{r} mutant carried an unstable mutation that correlated with the presence of an Ac transposon. Genomic DNA sequences flanking this Ac were used to identify complementary DNA and genomic DNA clones containing N sequences. Confirmation that the N gene had been cloned came from complementation of the TMV\textsuperscript{s} phenotype with a genomic DNA fragment.
Sequence analysis of genomic and cDNA clones shows that the N gene encodes a protein of ~131.4 kDa with an amino-terminal domain with similarity to the cytoplasmic domains of the Drosophila Toll protein (Hashimoto et al., 1998) and the interleukin-1 receptor (IL-1R) (Sims et al., 1989) in mammals, a nucleotide-binding site, and a leucine-rich repeat (LRR) region. The presence of these domains in the predicted N protein is consistent with the prediction that it functions in a signal transduction pathway for induction of defense responses. The potential function of N in the TMV-dependent HR will be discussed with respect to features central to signal transduction in both the Toll and IL-1R systems.

Results

Isolation of an Unstable TMV-Susceptible Mutant

We utilized the Ac transposon of maize (McClintock, 1948) for insertional mutagenesis and isolation of the N gene. In tobacco, critical functional aspects of Ac transposition in maize are retained (Baker et al., 1986). These include conservation of Ac structural integrity following transposition, generation of an 8 bp duplication at the site of insertion, generation of footprints upon excision, and continued transposition over several generations (Heil and Baker, 1990). The Ac transposon was introduced into the TMV\(^{\circ}\) tobacco cv. Samsun NN, which is homozygous for the N gene (see Experimental Procedures).

The isolation of Ac-induced mutations of N required generation and screening of large populations of F1 seedlings heterozygous for N (Nn) to identify potential lossof-function mutations. These mutants would be identified as plants unable to mount a TMV-dependent HR. Therefore, we developed and utilized a positive mutant selection scheme that exploited the reversible temperature sensitivity of the N-mediated HR to TMV (see Experimental Procedures). At temperatures above 28°C, plants carrying the N gene do not generate HR in response to TMV infection, but allow TMV to spread systemically. When these systemically infected plants are shifted to a permissive temperature (21°C), they develop lethal systemic HR. Those plants bearing an inactive, mutant allele of N were anticipated to survive the temperature shift.

We used this positive selection scheme to isolate TMV\(^{\circ}\) mutants potentially bearing Ac-induced mutations of N. Nn seedlings (84,000) were produced from crosses between three NN::Ac parents (T1-9, T1-10, and T1-13; see Experimental Procedures) and an nn parent (SR1 nn). A control population of 29,000 Nn seedlings from one cross between Samsun NN and SR1 nn was also produced. Nn populations were subjected to the positive mutant selection described in Experimental Procedures. Seedlings that survived this treatment were termed TMV\(^{\circ}\) mutants.

We isolated a total of 47 TMV\(^{\circ}\) mutants from Nn populations. Of these, 11 mutants were isolated from F1 populations derived from a control cross between Samsun NN and SR1 nn, and 36 mutants were isolated from three F1 populations derived from crosses between NN::Ac and SR1 nn. The frequency of loss of resistance to TMV among Samsun NN × SR1 nn and NN::Ac × SR1 nn progeny was similar and was 3.8 × 10\(^{-4}\) and 5.6 × 10\(^{-4}\), respectively. The ability to obtain TMV\(^{\circ}\) mutants at similar frequencies from Nn populations, with or without Ac, indicated that the N gene has a high spontaneous mutation rate. For further molecular and genetic analyses, 15 mutant lines containing Ac were used.

Our first objective was to determine whether these mutants bore mutations of N. The only mutations we expected to isolate using the positive selection were N loss-of-function mutations or dominant suppressor mutations of N. To test this, each mutant line was crossed to both SR1 nn and Samsun NN, and the test cross progeny were analyzed for the TMV\(^{\circ}\) or TMV\(^{\circ}\) phenotypes. If a mutant line bore an unlinked dominant suppressor of N, then 25% of the progeny of the SR1 nn test cross would be expected to be TMV\(^{\circ}\). If a dominant suppressor mutation linked to N was present, then 50% of the offspring of the Samsun NN test cross were expected to be TMV\(^{\circ}\). All progeny of the mutant lines crossed to SR1 nn were TMV\(^{\circ}\), whereas all progeny of mutant lines crossed to Samsun NN were TMV\(^{\circ}\), demonstrating that they bore mutations that were recessive to N and not dominant suppressor mutations (data not shown).

Further studies required the generation of lines homozygous for their N loss-of-function mutations. For this purpose, we employed a molecular marker, Nt-1, that identified restriction fragment length polymorphisms (RFLPs) between nn tobacco (Nt-17) and NN tobacco (Nt-1G) (see Experimental Procedures; Figure 2A). Because the Nt-1G RFLP is linked to N by less than 0.25 cM, we assumed that plants homozygous for the Nt-1G RFLP were also homozygous for their mutations of N. DNA was isolated from 27–64 selfed progeny of 15 mutants and analyzed by Southern blot hybridization with the Nt-1 marker (data not shown). We found that eight of 15 lines carried deletions of the Nt-1G RFLP whereas seven lines retained the Nt-1G RFLP (unpublished data). The identification of deletion homozygotes supported the idea that at least eight mutant lines carried loss of N function mutations. The seven remaining mutant lines did not bear deletions of the Nt-1G RFLP and included the line that bore an unstable mutation of N (N\(^{\circ}\)) (see below).

We determined whether any of the mutant lines bore mutations that affected their ability to elaborate an HR to other pathogens. Three progeny of each homozygous mutant line, Samsun NN, and SR1 nn were inoculated with two bacterial pathogens that elicit an HR on tobacco (see Experimental Procedures). The pathogenic bacteria, P. syringae pv. tomato (Pst) strain DC 3000 and P. syringae pv. phaseolicola (Pspl) strain MP 53121, elicited an HR on all plants whereas no HR was observed to the nonpathogenic Pst strain DC 3000 hps::Tn5 or the water control (data not shown). These results show that the TMV\(^{\circ}\) mutants were not affected in the ability to mount an HR to bacterial pathogens and suggested that the inability to elaborate an HR was probably specific to TMV.

In maize, mutations caused by Ac transposons can revert and lead to expression of an unstable phenotype (Fedorch, 1989). Therefore, we expected that those mutants bearing Ac insertions might give rise to germlinal and so-
matic revertants. Progeny bearing germinai reversion events were expected to display a fully TMV* phenotype upon TMV infection (see Figure 1A) while somatic revertants were expected to display a variegated phenotype, composed of both TMV* and TMV$^\text{ns}$ tissues (TMV$^\text{ns}$) (see Figure 1C). Selfed progeny (95–150) of the 15 homozygous mutant lines were inoculated with TMV and scored for their phenotypes. Among the 15 lines tested, only the D1-1 line gave rise to progeny with either the TMV$^\text{ns}$ or TMV$^*$ phenotype. Of the 145 D1-1 progeny scored, 20 were TMV$^*$, 68 were TMV$^\text{ns}$, and 57 plants displayed the TMV$^\text{ns}$ phenotype. In contrast, the 1888 progeny of the other 14 lines, D2-2 to D28-2, remained fully TMV$^*$ (data not shown), demonstrating that they possessed nonreverting mutations. The identification of TMV$^\text{ns}$ and TMV$^\text{ns}$ individuals in the D11-1 line indicated that it possessed an unstable TMV$^\text{ns}$ (N') mutation. Therefore, we postulated that the D11-1 mutant line bore an Ac insertion in the N gene.

The Unstable TMV$^\text{ns}$ Mutation and Two Ac Transposons Cosegregate with N

If Ac was inserted in the N gene in the D11-1 line, then the TMV$^\text{ns}$ phenotype and the Ac transposon causing the mutation were expected to cosegregate with the Nt-1G RFLP. We tested the linkage of the TMV$^\text{ns}$ phenotype with the Nt-1G RFLP in the test cross progeny of the progenitor of the D111-1 mutant line, C2-2(N'), crossed to SR1 nn. The test cross progeny, termed the D111 population, were inoculated with TMV and scored for their phenotypes. Of 80 D111 plants scored, 28 displayed the TMV$^\text{ns}$ phenotype and 52 were TMV$^\text{*}$ (Table 1). To determine the genotype of these 80 plants, DNA from each was digested with EcoRl and subjected to Southern blot hybridization with the Nt-1 marker. In D111 plants, 39 were heterozygous for the Nt-1G RFLP (Nn genotype) and 41 were homozygous for the Nt-17 RFLP (nn genotype) (Table 1). The 26 plants displaying the TMV$^\text{ns}$ phenotype were also heterozygous for the Nt-1G RFLP (Table 1). All plants of the nn genotype were TMV$^\text{*}$ (Table 1). An example of the Nt-1 hybridization data summarized in Table 1 is shown in Figure 2B. Ten plants were heterozygous for the Nt-1G RFLP (Nn genotype) as demonstrated by the presence of both the 14.3 kb Nt-1G and 15.5 kb Nt-17 RFLPs (Figure 2B, lanes 2, 4–11, and 14). Of these ten plants, six displayed the TMV$^\text{ns}$ phenotype (Figure 2B, lanes 2, 4, 7, 9, 11, and 14). Four TMV$^\text{*}$ plants were homozygous for the Nt-1T

---

| Table 1. The TMV$^\text{ns}$ Phenotype Cosegregates with the Nt-1G RFLP |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| TMV Phenotype$^*$ | Nt-1 RFLPs | TMV$^\text{ns}$ | TMV$^\text{*}$ | Total |
| Nt-1/Gnt-1T | 0 | 26 | 13 | 39 |
| Nt-1/Tn-1T | 0 | 0 | 41 | 41 |

$^*$ From the cross of the TMV$^\text{ns}$ susceptible mutant C2-2 to SR1 nn (the D111 population), 80 plants were inoculated with TMV and scored for their phenotypes.

---

| Table 2. Two Ac Transposons Cosegregate with the Nt-1G RFLP |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Nt-1 RFLPs | Ac10 | Ac6 | Ac8 | Ac8 | Total |
| Nt-1/Gnt-1T | 30 | 3 | 5 | 1 | 39 |
| Nt-1/Tn-1T | 0 | 0 | 41 | 41 |

$^*$ Following hybridization with the Nt-1 probe, Southern blots containing D111 DNAs digested with EcoRl were stripped and hybridized with the 5' Ac probe.

$^*$ Two Ac bands were identified that cosegregate with Nt-1G.

$^*$ Plants containing neither Ac10 nor Ac8.

---

| Table 3. Excision of Ac10 Is Correlated with Reversion to TMV Resistance |
|-----------------|-----------------|-----------------|-----------------|
| N-Linked Ac | TMV Phenotype$^*$ | TMV$^\text{ns}$ | TMV$^\text{*}$ | TMV$^\text{ns}$ |
| Ac10 | 0 | 1 | 1 |
| Ac10/Ac6 | 0 | 18 | 1 |
| Ac8 | 52 | 1 | 18 |
| Ac10/Ac8 | 2 | 1 | 0 |

$^*$ From the cross of the TMV$^\text{*}$ germinai revertant D112-15 to SR1 nn (the ES01 population), 95 plants were inoculated with TMV and scored for their phenotypes.

$^*$ DNA isolated from the ES01 plants was digested with EcoRl and hybridized with the 5' Ac probe.

$^*$ Plants containing neither Ac10 nor Ac8.
TMV\textsuperscript{n}, germinally revertant progeny of the unstable mutant line. We identified a fully TMV\textsuperscript{n} plant, D112-15, from the selfed progeny of the TMV\textsuperscript{n} mutant C2-2. Genomic DNA from D112-15 was isolated and subjected to Southern blot hybridization with the Nt-1 and 5' Ac probes. These analyses showed that D112-15 was homozygous for the Nt-1G RFLP and carried both Ac10 and Ac8 (data not shown). We predicted that Ac10 or Ac8 had germinally excised from one of the N\textsuperscript{*} alleles of D112-15 leading to the fully resistant phenotype and that this excision event was masked in the Southern analysis by the presence of these elements in the other N\textsuperscript{*} allele. To test this possibility, D112-15 was test crossed to SR1 nn, giving rise to the E501 population. The E501 population was expected to segregate 1:1 for the TMV\textsuperscript{n} to the TMV\textsuperscript{v} and TMV\textsuperscript{v}\textsuperscript{n} phenotypes. We inoculated 95 E501 plants with TMV and scored their phenotypes, and DNA from these plants was digested with EcoRI for Southern blot hybridization with the 5' Ac probe. In E501 plants, 54 were TMV\textsuperscript{n}, 21 were TMV\textsuperscript{v}, and 20 were TMV\textsuperscript{v}\textsuperscript{n} (Table 3). Southern blot hybridization with the 5' Ac probe showed that none of the 54 TMV\textsuperscript{n} individuals carried Ac10 whereas Ac8 was present in 52 TMV\textsuperscript{v} plants (Table 3). The absence of Ac10 and the presence of Ac8 among the TMV\textsuperscript{n} E501 progeny strongly suggested that Ac10 had inserted into the N gene.

To establish further the insertion of Ac10 in the N gene, we confirmed the correlation between the insertion and excision of Ac10 and the TMV\textsuperscript{n} and TMV\textsuperscript{v} phenotypes, respectively. A TMV\textsuperscript{n} mutant, E501-70, with only the Ac10 transposon, was identified (Figure 3B, lane 4). We analyzed
500 selfed progeny of E501-70 (F501 population) for their phenotypes following TMV inoculation. Seven TMV® progeny plants were identified as germinal revertants. DNA from the seven revertants was isolated and digested with EcoRI for Southern blot hybridization analysis with the Nt-1 and 5' Ac probes. These analyses showed that three germinal revertants were heterozygous for the Nt-1G RFLP (data not shown) and did not harbor Ac10 (two of which are shown in Figure 3B, lanes 6 and 7). The remaining four germinal revertants were homozygous for the Nt-1G RFLP (data not shown) and contained Ac10 (two of which are shown in Figure 3B, lanes 5 and 8). As with D112-15, we predicted that germlinal excision of Ac10 was masked by Ac10 residing in the second N* allele in the Nt-1G homozygotes. These data further supported our conclusion that Ac10 was inserted in the N gene.

Molecular confirmation of the Ac10 insertion into the N gene was obtained by isolation of genomic sequences flanking Ac10 by the inverse polymerase chain reaction (IPCR) (see Experimental Procedures). Using genomic fragments as probes, we expected to detect Ac10-disrupted restriction enzyme fragments in TMV® or TMV® plants and a wild-type genomic fragment in TMV® plants (Figure 3A). The IPCR products from the 5' and the 3' ends of Ac10 were isolated and were termed N-1 (419 bp) and N-2 (116 bp), respectively (see Experimental Procedures). Sequence analysis of N-1 and N-2 confirmed that we had isolated the 5' and 3' Ac10 genomic junction fragments and that Ac10 was flanked by an 8 bp target site duplication (data not shown). The N-1 and N-2 probes were suboptimal for Southern blot hybridization analysis owing to either their short length or repetitive nature (see Figure 6). Therefore, N-1 was utilized to isolate a cDNA clone from which a large and unique probe, N-5 (Figure 3A), was derived for Southern blot hybridization analyses of putative Ac mutants and revertants.

The N-5 probe was used to test whether genomic sequences were disrupted by Ac10 insertion and restored to wild-type in germinal revertants. Based on Ac hybridization, we constructed an EcoRI restriction map of the mutant and wild-type N gene that predicted a 7.9 kb wild-type EcoRI fragment (Figure 3A). Hybridization of the N-5 probe to N. glutinosa and Samsun NN DNA digested with EcoRI detected the expected 7.9 kb wild-type fragment whereas, in E501-70, N-5 detected the 10.2 kb insertion fragment (Figure 3C). Significantly, all 54 TMV® E501 siblings of E501-70 contained only the 7.9 kb excision fragment (data not shown). The N-5 probe detected only the 7.9 kb wild-type fragment in three TMV® revertant progeny of E501-70 that were Nt-1G heterozygotes (two of which are shown in Figure 3C), demonstrating that upon Ac10 excision the wild-type genomic fragment was restored. The four TMV® progeny of E501-70, homozygous for Nt-1G, possessed both the 7.9 kb excision fragment and the 10.2 kb insertion fragment (two of which are shown in Figure 3C). These results demonstrated that Ac10 insertion led to inactivation of the N gene and excision of Ac10 led to reversion to the TMV® phenotype.

Southern blot hybridization of the N-5 probe to EcoRI-digested DNA from a TMV® F501 plant is shown in Figure 3C (lane 9). This plant displays both the 10.2 kb insertion band and the 7.9 kb excision band. Interestingly, 19 of 21 TMV® plants from the E501 generation have both the 10.2 kb and 7.9 kb N-5 hybridizing bands. The presence of both bands indicated that the collected tissue was comprised of a mixture of cells whose genomes contained either the Ac10 insertion or excision of Ac10 from the N gene.

**N Is Contained within a 10.6 kb Genomic Fragment That Confers Resistance to TMV**

To confirm the identification of the N gene and to determine whether this gene was sufficient to confer the TMV® phenotype, we introduced a genomic DNA fragment containing the full-length N gene into TMV® SR1 in tobacco. The DNA fragment used for transformation was isolated from an N. glutinosa λ DNA library screened with probes N-5 and N-9 (Figure 4A) (see Experimental Procedures). Three clones (G25, G34, and G38) were isolated that contained identical internal structures as determined by re-
Figure 5. The Amino Acid Sequence of the N Protein

(A) The region of similarity to the cytoplasmic domains of Toll and IL-1R is underlined. Amino acids that comprise the predicted ATP/GTP-binding site are underlined twice. The LRR region is indicated in italics. Potential N-glycosylation sites are indicated in bold. Abbreviations for the single-letter amino acid codes are standard.

(B) Primary structure of the LRR region of N (amino acids 590–928) and comparison of its consensus sequence with the LRR consensus sequences of yeast adenyl cyclase (AdCycl; Kataoka et al., 1985), Drosophila Toll (Hashimoto et al., 1988), human platelet membrane glycoprotein Ib α chain (GpIbα; Titan et al., 1987), Drosophila Chaoptin (Reineke et al., 1988), and Arabidopsis receptor-like transmembrane kinase TMK1 (Chang et al., 1992), TMKL1 (Yalon et al., 1992), and RLK5 (Walker, 1993).

(C) Alignment of the amino-terminal domain of N (amino acids 8–150) with the cytoplasmic domains of Drosophila Toll (amino acids 804–996; Hashimoto et al., 1988) and the human IL-1R (H IL-1R; amino acids 317–524; Sims et al., 1989). Amino acids that are identical or similar are indicated by the boxed region. Conservative substitutions used are: hydrophobic amino acid, L/I/V/M/F/A; ionic amino acid, K/R/D/E/Q/N/H; aromatic amino acid, F/Y.
striction enzyme mapping (Figure 4A). We predicted that G38 contained all cis sequences necessary for proper expression of the transgene based on comparative restriction mapping of the genomic and cDNA clones. The 10.6 kb G38 XhoI fragment (Figure 4A) was subcloned into the T-DNA vector pOCA28 to generate the pTG38 construct. pTG38 was transformed into SR1 nn by Agrobacterium-mediated transformation (see Experimental Procedures). Untransformed controls SR1 nn and Samsun NN, as well as SR1 nn transformed with pOCA28 alone, were regenerated simultaneously.

Transformed T0 plants containing pTG38 and control plants were tested for resistance to TMV at 1 week following transfer to soil. Of 12 T0 plants containing pTG38, 11 responded to TMV infection with an HR. The HR on a leaf from a T0 plant transformed with pTG38 is shown in Figure 1D. The onset and appearance of HR was identical to that of Samsun NN plants (see Figure 1A). The SR1 control plants did not develop HR and later developed mosaic symptoms. These results demonstrated that this single gene is necessary and sufficient to confer HR to N1 tobacco infected with TMV.

Structure and Sequence Analysis of the N Gene Product
We deduced the structure and organization of the N gene from sequence analysis of the C7, C16, and C18 cDNA clones and partial sequencing and restriction enzyme digestion of the G38 genomic clone (Figure 4A). Taken together, C7 and C18 predict that five exons are spliced to form an open reading frame of 3432 bp that encodes a polypeptide of 1144 amino acids (N in Figure 4B). C16 encodes a truncated polypeptide of 652 amino acids (N’ in Figure 4B). N’ is apparently the result of alternative splicing of a 70 bp exon (AE1) to form an open reading frame of 1956 bp (Figure 4A). The 3’-untranslated regions of the cDNAs vary in length, indicating that different polyadenylation signals are used (data not shown). Partial sequencing of the G38 genomic clone shows that all sequences necessary to encode the C7, C16, and C18 cDNAs are present in the N gene. Thus, these cDNAs are encoded by a single gene.

The predicted amino acid sequences of the N and N’ proteins were analyzed with a variety of sequence analysis programs. The predicted molecular masses of N and N’ are 131.4 kDa and 75.3 kDa, respectively. With the exception of the carboxyl-terminal 22 amino acids, N’ is identical to the amino terminus of N (Figure 4B). Sequence analysis of N’ does not suggest a possible function for these 22 amino acids. Eight putative N-linked glycosylation sites are predicted in the amino acid sequence of N and have the consensus amino acid sequence NX(S/T) (Figure 5A).

Analysis of the N protein sequence with the program UALOM (Jones et al., 1994) and signalase (N. Mantei, Swiss Federal Institute of Technology) indicated that no transmembrane domain nor signal sequence is present in the N protein. Thus, based on sequence analysis, N appears to be cytoplasmic.

The amino acid sequence of the full-length N protein was compared with the GenBank data base (release 82.0) using the BLAST program (Altschul et al., 1990). In the amino-terminal region, N encodes three motifs found in various proteins with ATP/GTP binding activity (Figure 5A) (Traut, 1994). The first motif, the P loop, is functional in binding the phosphates of ATP/GTP and has the consensus sequence (AG)XXGK(S/T). The sequence GMGGVGGKT at amino acids 216–223 of N fits the P loop consensus (Figure 5A). The families of proteins containing the P loop include the adenylyl kinases, Ras family of proteins, elongation factors, ATP synthase β subunits, thymidine kinases, and phosphoglycerate kinases (Saraste et al., 1990). A second motif, kinase 2, is defined as four consecutive hydrophobic amino acids followed by an invariant aspartate (D) spaced, on average, at 61 amino acids from the P loop. Inspection of the N sequence shows that the amino acids LIVLD, 74 amino acids from the P loop, fit the kinase 2 motif. The third motif, kinase 3a, is involved in purine or ribose binding. Kinase 3a is spaced an average of 126 amino acids from the P loop and commonly contains a tyrosine (Y) or arginine (R). In N, the sequence FNGNSR, 97 amino acids from the P loop, may fit the kinase 3a motif.

The amino acid sequence of N from 590 to 928 contains an LRR region comprised of 4 imperfect repeats of -26 amino acids in length (Figure 5A). LRRs are found in a wide variety of proteins involved in signal transduction, cell adhesion, and various other functions. The consensus sequence derived for the aligned LRRs of N is similar to the consensus found in yeast adenylyl cyclases (Katoaka et al., 1985), Drosophila Toll (Hashimoto et al., 1988), human platelet membrane glycoprotein Ibα chain (Titi et al., 1987), Drosophila Chaoptin (Reinke et al., 1988), and Arabidopsis receptor-like transmembrane kinases (Chang et al., 1992; Valon et al., 1993; Walker, 1993) (Figure 5B). The amino-terminal amino acids (8–150) of N are similar to the cytoplasmic domains of the Drosophila Toll protein and the human IL-1R. The amino terminus of N is 65% similar to amino acids 864–998 of Toll and 49% similar to amino acids 393–526 of IL-1R (Figure 5C).

Discussion
We have isolated the N gene from tobacco by insertional mutagenesis with the maize transposon Ac. The N gene is a novel plant disease resistance gene from tobacco that confers resistance to TMV. Complementation tests in SR1 nn tobacco demonstrated that this single gene is necessary and sufficient to confer the HR to TMV. This result supports the basic assumption of the corresponding genes hypothesis that single dominant genes are involved in pathogen perception and subsequent induction of plant defense responses (Flor, 1947). The most likely role of genes governing such critical control points in the resistance pathway is that of a receptor for a ligand produced by the pathogen, in this case TMV. The product of the N gene contains sequence motifs that suggest that it could be a receptor molecule or another important component of a signal transduction pathway.
The N-t Probe Hybridizes to Repetitive Sequences Clusters near N

The N-t probe was hybridized to Southern blots of EcoRI-digested DNA of SR1 (lane 1), Samosun NN (lane 2), N. glutinosa (lane 3), N. sylvestris (lane 4), and N. tomentosiformis (lane 5). RFLPs introgressed into Samosun NN from N. glutinosa are indicated with an arrow. The N. tomentosiformis RFLPs replaced in Samosun NN by N. glutinosa RFLPs are indicated by the bars.

Complex Nature of N Locus in Nicotiana

In our mutant selection, we observed a high mutation rate (≈1 in 2000) of the N locus in Nn populations both with and without Ac. A similar high spontaneous mutation rate of ≈1 in 500 was observed for alleles of the Rl locus of maize, which confer resistance to specific races of the rust pathogen Puccinia sorghi (Bennetzen et al., 1988). The high natural mutation rate of Rl might be due to unequal crossing over, causing deletions between repetitive sequences (Hulbert and Bennetzen, 1991). At the N locus, we have identified a cluster of related repetitive sequences. The genome of N. tabacum is amphidiploid and is derived from the genomes of N. sylvestris and N. tomentosiformis. The N gene was introgressed into tobacco cultivars, such as Samosun NN, by crossing TMV 9 tobacco to N. glutinosa and by repeated backcrosses to a recurrent TMV 9 tobacco parent. We have shown with the N-t marker that introgression of the N gene coincided with introgression of DNA-bearing restriction enzyme fragments of N. glutinosa origin, such as the N-t 1G RFLP, and simultaneous loss of restriction fragments of the N. tomentosiformis origin, such as the N-t 1F RFLP (Figure 2A). Unlike the unique N-t marker, the N-1 marker (Figure 4A) detected at least ten genomic restriction enzyme fragments when hybridized to DNA from the pertinent Nicotiana species (Figure 6). At least four of these restriction enzyme fragments were present exclusively in Samosun NN and N. glutinosa (see Figure 6, lane 2 and 3), suggesting that they were introgressed with the N gene from N. glutinosa into Samosun NN. Conversely, at least five restriction fragments of N. tomentosiformis origin were present in SR1 nn, but not Samosun NN, suggesting that they were replaced by their N. glutinosa homologs upon introgression of N. The exchange of N-1-hybridizing repeats of N. tomentosiformis origin for their homologs from N. glutinosa demonstrated that these repeats are physically linked to each other and clustered near the N locus. These data indicate that a family of genes related to N may be present in the genus Nicotiana. In tobacco, these clustered repeats may contribute to the instability at the N locus by providing substrates for unequal crossing-over events. The presence of a family of genes with similarity to N suggests that these could be resistance genes, but their function remains to be investigated. Comparison of their sequences with that of N will be informative.

Possible Structural and Functional Properties of N

Analysis of the cDNA clones suggested that the N gene might encode truncated (N") and full-length (N) forms of proteins. The N" form may arise from alternative splicing of the 70 bp AE1 exon that introduces a stop codon such that a protein without the LRR region is synthesized. Multiple forms of proteins encoded by single genes can be generated by alternative splicing mechanisms (McKeown, 1992). In the human fibronectin and erythropoietin genes, differential splicing results in multiple forms of proteins that are expressed in different developmental stages and cell types (Mardon et al., 1987; Nakamura et al., 1992). In the case of fibronectin, two cis elements (GAAGAGA and CAAGG) within the 270 bp EPA exon modulates the exclusion or inclusion of this exon (Caputi et al., 1994). The presence of similar elements in the AE1 exon (data not shown) suggests that generation of the N" form might involve a similar type of splicing event. The role of this proposed splicing event in N-mediated resistance remains to be established.

Analysis of the predicted N protein sequence suggests that it is a cytoplasmic protein. The deduced amino acid sequence of N is comprised of three domains that may be of functional significance: a nucleotide-binding domain, an LRR region, and an amino-terminal domain with homology to the cytoplasmic domains of Toll and IL-1R. The nucleotide-binding domain is found in kinases, ATPases, and Ras proteins. The N protein contains three appropriately spaced structural motifs that are commonly found in proteins known to bind ATP/GTP (Traut, 1994). This observation raises the possibility that binding of ATP or GTP might be required for N protein function. In the ras-encoded protein p21, the GTP-bound state is involved in signaling effector proteins during cell growth and differentiation (Bourse et al., 1990). Purification of the N protein...
and subsequent biochemical assays are necessary to establish the role of nucleotide binding in N function. The LRR region of the N protein contains at least 4 imperfect tandem repeats. These repeats are present in many different classes of proteins and may be extracellular or cytoplasmic. In general, LRRs are thought to mediate specific protein–protein interactions. The LRR region of yeast adenylyl cyclase is required for activation by Ras proteins (Suzuki et al., 1990), while the LRR of human placental ribonuclease inhibitor binds pancreatic ribonuclease (Lee and Vallee, 1990). Binding of the von Willebrand factor to the LRR region of human glycoprotein Ibα heterodimer induces adhesion of human platelets to the subendothelium (Taitt et al., 1987). In the lutropin–choriogonadotropin receptor, LRRs form amphipathic peptide surfaces that mediate hormone binding (McFarland et al., 1989). In addition to protein–protein interactions, LRRs of yeast adenylyl cyclase and chaoptin have also been implicated in cell adhesion and membrane association functions (Kataoka et al., 1985; Van Vactor et al., 1998). The presence of similar repeats in the N protein suggests that it might act through a protein–protein contact.

**Proposed Role of N Protein in TMV Perception and Signal Transduction Leading to HR**

A model depicting the possible role of the N gene product in signal transduction is presented in Figure 7. The model is based on the sequence similarity of the amino terminus of the N protein to the cytoplasmic domains of the Drosophila Toll protein and the human IL-1R (Figure 7). The similarity in amino acid sequence among these three proteins raises the possibility that the N protein might trigger an intracellular signal transduction cascade similar to those triggered by Toll and IL-1R. In the IL-1 pathway, interaction of the cytokine IL-1 with IL-1R results in the activation and translocation of the Rel-related transcription factor NF-κB. NF-κB induces the synthesis of a variety of defense and signaling proteins involved in inflammatory, immune, and acute phase responses through binding to the decameric κB-binding motif (Baueuerle, 1991). In Drosophila, perception of an unknown signal by the extra cellular domain of Toll activates the Rel-related transcription factor, dorsal.

The dorsal gene product regulates the transcription of a set of zygotic genes involved in dorsoventral polarity (St Johnston and Nusslein-Volhard, 1992). In the immune response of Drosophila, another Rel-related transcription factor, Dif (Ip et al., 1993), induces various antimicrobial genes upon interaction of bacterial pathogens with an unknown receptor. Mutations in the cytoplasmic domains of the IL-1R and Toll proteins indicate that this domain is dispensable for activation of downstream events in the IL-1 and Toll signal transduction pathways (Heguy et al., 1992; Schneider et al., 1991). Interestingly, some of these amino acids are conserved in the amino-terminal domain of the N protein and might have an important role in signal transduction.

Analogous to the above-mentioned immune and developmental responses, the N protein may function as a receptor that interacts directly with the gene product of TMV that elicits the N-mediated HR. The introduction of the N gene alone was necessary and sufficient to confer the TMV phenotype to SR1 nn, which suggests that the N gene product may be the determinant for TMV perception and induction of HR. The elicitor of the N-mediated HR has not been strictly defined; however, the 126 kDa TMV replicase gene product may be involved (Padgett and Beachy, 1993). TMV is a rod-shaped virus whose virion is comprised solely of coat protein and the RNA genome. Therefore, for TMV to produce the replicase molecule, the RNA genome of TMV must be uncoated and expressed. TMV gains access to the cytoplasm of plant cells by mechanical damage. In the cytoplasm, virus particles are uncoated by cytoplasmic ribosomes and the genome is translated (Shaw et al., 1986). The first protein produced is the replicase, which in turn leads to multiplication of the viral genome within the cytoplasm. These events might be necessary to produce the elicitor of the N-mediated HR because neither the virion nor coat protein alone can induce HR. Indeed, TMV replication and gene expression occur before the onset of HR. The life cycle of TMV is carried out in the cytoplasm of the cells of the host plant, and therefore it would seem appropriate that N would encode a cytoplasmic protein that could interact with the TMV-encoded elicitor. Based on the homology of N with Toll and IL-1R, it is tempting to speculate that upon interaction with its corresponding elicitor, N might activate a Rel-like transcription factor that induces the expression of genes responsible for the N-mediated HR. Such transcription factors might be activated directly by N or through some effector molecule such as a kinase, phosphatase, or protease. Since the TMV–N interaction is a classical model system for many plant–pathogen studies in virus–host interaction, the cloned N gene will facilitate systematic dissection of the biochemical and molecular basis of this signal transduction pathway.

Cell death due to HR in plant defense has several similarities to programmed cell death or apoptosis seen in mammals and insects (Greenberg et al., 1994). The HR is preceded by rapid induction of a number of known biochemical, physiological, and molecular processes, such as production of reactive oxygen intermediates (O₂⁻, HO₂, and OH⁻), increase in lipoxygenase activity, production
no evidence that HR, per se, directly inhibits these processes. The molecular and biochemical events that control HR in plants are largely unknown. Recently, a number of mutant loci that lead to spontaneous HR have been described in Arabidopsis (Dietrich et al., 1994; Greenberg et al., 1994). The availability of these genes along with the cloned N gene might help us to understand the basis of HR in the TMV–N interaction.

Comparison of Viral Resistance Gene N and a Bacterial Resistance Gene, RPS2

In an accompanying paper (Mindrinos et al., 1994 [this issue of Cell]), the isolation of another plant disease resistance gene, RPS2, is described. Surprisingly, the gene products of N and RPS2 share common domains and overall sequence similarity (Figure 8). N and RPS2 are 49% similar and 24% identical. Like N, RPS2 encodes a P loop near the amino terminus, and the carboxyl terminus is comprised largely of an LRR region. There is another region of the sequence, GLPLAL, between the P loop and the LRR that is conserved.

It is interesting that a viral resistance gene from tobacco and a bacterial resistance gene from Arabidopsis thaliana have a similar structure. These similarities suggest that N and RPS2 might be members of a class of genes in plants that are necessary for specific plant–pathogen interactions that result in HR. Perhaps the mechanisms of pathogen perception and induction of HR have been reiterated throughout the plant kingdom. The isolation of N and RPS2 provides the means to identify other members of the class for comparative functional and sequence analysis. Such studies would aid in defining the domains involved in pathogen specificity and signal transduction. In addition, we hope that the isolation of resistance genes will provide the means to introduce pathogen resistance across species barriers that were not possible to breach by conventional breeding methods.

Experimental Procedures

Plant Germplasm and Pathogen Inoculations

The N. tabacum (tobacco) cv. Petite Havana SR1, termed SR1 nn, and Samsun NN (provided by H. Loerz, Universität Hamburg) and N. sylvestris, N. tomentosiformis, and N. glutinosa (provided by V. Sisson, United States Department of Agriculture) were used for these studies. The U1 strain of TMV (gift of M. Zaitlin, Cornell University), bacterial pathogen Pst strain DC 3000 and Psp strain NP 53121, and the nonpathogenic Pst strain DC 3000 hspS: Tn5 (gifts of B. Staskawicz, University of California, Berkeley) were used for pathogen inoculations as described (Takahashi, 1986; Klement, 1990).

Transgenic Tobacco Bearing Active Ac Transposons

The TMV+ N. tabacum (tobacco) cv. Samsun NN was transformed with pGV3850 HPT:pKU8 (Baker et al., 1987) according to the method of Horsch et al. (1985). Transgenic calli were regenerated on 100 mg/l kanamycin to select for transgenic tissue harboring transposing Ac elements (Baker et al., 1987). Primary transgenics are referred to as the T0 generation. One plant, T0-3, containing transposed Ac elements was crossed to Samsun NN to generate T1 plants. Three T1 plants, T1-5, T1-10, and T1-13, shown to have transposing Ac elements, were crossed to SR1 nn to generate the three F1 Nm:Ac populations used in the mutant selection. A control F1 Nm population without Ac was also generated for the mutant selection by crossing Samsun NN and SR1 nn.
Isolation of TMV Susceptible Mutants

We subjected 64,000 Nn:Ac seedlings and 29,000 Nn seedlings to a positive mutant selection scheme. Seed from each cross was sown at a density of approximately three seeds per square centimeter. When seedlings were 8 weeks old, they were placed at 30°C and inoculated with a suspension of the U strain of TMV and Cellope (Fisher) using an artist's airbrush (Paasche VL) (Fulton, 1979). Seedlings were shifted to 21°C 3 days postinoculation and 2 days later were scored for lethal systemic necrosis or survival. The procedure was repeated twice to minimize isolation of escapes.

Isolation of the Nt 1 RFLP Marker

A DNA marker, Nt 1, flanking an Ac insertion (Hehl and Baker, 1990) was isolated that detected an RFLP between SRII n and Samsum NN (Figure 2A). The Nt 1 marker detected RFLPs specific for each of the diploid Nicotiana species. The 13.1 kb DNA fragment is present in N. sylvestris, SRII n, and Samsum NN (Figure 2A, lanes 2, 3, and 4, respectively). The 15.5 kb DNA fragment present in N. sylvestris and SRII n (Figure 2A, lanes 5 and 3, respectively) and the 14.3 kb DNA fragment is present in N. glutinosa and Samsum NN (Figure 2A, lanes 3 and 5, respectively). Samsum NN carries an RFLP identical in size with the 14.3 kb RFLP (Nt 1 G) in N. glutinosa, but lacks the 15.5 kb N. sylvestris-specific RFLP (Nt 1 T).

Linkage between the Nt 1 RFLP and N was tested in F2 progeny of a cross between SRII n and Samsum NN, segregating 3:1 for TMVn to TMVb and 1:2:1 for the Nt 1/G:Nt 1 T, Nt 1/G:Nt 1 G, and Nt 1 T:Nt 1 T genotypes, respectively. DNA from 420 TMV b F2 plants was digested with EcoRI and hybridized with the Nt 1 probe. Only one TMV b plant had an Nt 1 G RFLP, demonstrating that this RFLP is linked to N by less than 0.25 cm (data not shown).

Isolation of IPCR Clones

Genomic tobacco sequences flanking Ac10 were isolated by IPCR. Template DNA from plant D111-95 that carried only Ac10 and Ac8 (Figure 2C, lane 9) was digested with Hpal, ligated, linearized with ClaI, and amplified according to the following parameters: 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min for 35 cycles. A 419 bp product (N-1) was amplified using the Ac-specific primers CCG9 (5'-CCGCAGATCCACAGTACGTCGTCGAGGGA-3') and CCG2 (5'-CCGGACGAGAAGCAGGTATGAGGAG-3') to obtain the Ac10 3' flanking sequence (N-2). D111-95 DNA was digested with EcoRI, ligated, and linearized with Accl. A 122 bp product was amplified using primers CCG1 (5'-CAGCGTCCACGAGATCCATCTGCTCTCGACGAAAGTTACG-3') and CCG3 (5'-CCGGACGAGAAGCAGGTATGAGGAG-3') to obtain the Ac10 3' flanking sequence (N-3). The amplified DNA was separated on a 2% agarose gel and purified using the QIAquick gel extraction kit (Qiagen). A total of 2 × 10^6 pfu were screened with the N-1 and N-5 markers (Figure 4A).

The genomic library was constructed using N. glutinosa DNA that was partially digested with MboI and size fractionated by gel electrophoresis. DNA fragments larger than 12 kb were ligated to BamHI-digested, dephosphorylated arms of the bacteriophage λ vector Gem-11 (Promega) and packaged. The N-9 and N-5 probes were hybridized to the library and screened as described (Baker et al., 1987).

DNA Sequencing and Analysis

DNA was sequenced by the dideoxy chain termination method (Sanger et al., 1977) using the Sequenase version 2.0 system (U.S. Biochemical Corp.). For sequencing of the C7 cDNA, nested deletions were prepared by the exonuclease III method (Henikoff, 1987). cDNAs C7 and C7 were sequenced using primers derived from the sequence of C7.

GenBank data base (release 82.0) comparisons were done with the program BLAST (Altschul et al., 1990). Alignments of the N protein to Toll, the IL-1R, and RPS2 were performed using the Genetics Computer Group sequence analysis programs GAP, PILE-UP, and PRETTY (University of Wisconsin).

Acknowledgments

We are grateful to Choong-I Cheon for sequencing the C16 and C17 cDNAs. We thank Brian Osborne for stimulating discussions, advice, and critical reading of the manuscript. We thank David Hantz for expert advice and help on propagation and pathogen treatment of plants. We thank Jeff Schell for thoughtful advice. B. B. thanks D. C. for helpful discussions. This work was supported in part by grants from the McKnight Foundation (M108) and the United States Department of Agriculture (5355-22000-504-003).

Received August 9, 1994; revised August 25, 1994.

References


