The Plant Cell, Vol. 10, 1055–1068, June 1998, www.plantcell.org

Dissection of the Fusarium *I2* Gene Cluster in Tomato Reveals Six Homologs and One Active Gene Copy

Guus Simons,[a,1] John Groenendijk,[a] Jelle Wijbrandi,[a] Martin Reijans,[a] José Groenen,[a] Paul Diergaarde,[a] Theo Van der Lee,[a,2] Marjo Bleeker,[a] Joke Onstenk,[a] Michiel de Both,[a] Michel Haring,[b] Jurriaan Mes,[b] Ben Cornelissen,[b] Marc Zabeau,[a,3] and Pieter Vos[a]

[a]Keygene N.V., Agro Business Park 90, P.O. Box 216, 6700 AE Wageningen, The Netherlands
[b]University of Amsterdam, Institute for Molecular Cell Biology, Section Plant Pathology, Kruislaan 318, 1098 SM Amsterdam, The Netherlands

The *I2* locus in tomato confers resistance to race 2 of the soil-borne fungus *Fusarium oxysporum* f sp *lycopersici.* The selective restriction fragment amplification (AFLP) positional cloning strategy was used to identify *I2* in the tomato genome. A yeast artificial chromosome (YAC) clone covering ~750 kb encompassing the *I2* locus was isolated, and the AFLP technique was used to derive tightly linked AFLP markers from this YAC clone. Genetic complementation analysis in transgenic R_1 plants using a set of overlapping cosmids covering the *I2* locus revealed three cosmids giving full resistance to *F. o. lycopersici* race 2. These cosmids shared a 7-kb DNA fragment containing an open reading frame encoding a protein with similarity to the nucleotide binding site leucine-rich repeat family of resistance genes. At the *I2* locus, we identified six additional homologs that included the recently identified *I2C-1* and *I2C-2* genes. However, cosmids containing the *I2C-1* or *I2C-2* gene could not confer resistance to plants, indicating that these members are not the functional resistance genes. Alignments between the various members of the *I2* gene family revealed two significant variable regions within the leucine-rich repeat region. They consisted of deletions or duplications of one or more leucine-rich repeats. We propose that one or both of these leucine-rich repeats are involved in Fusarium wilt resistance with *I2* specificity.

INTRODUCTION

Fusarium oxysporum f sp *lycopersici* is a soil-borne fungus that causes wilting disease in tomato. The fungus infects plants through the roots via direct penetration or via wounds, after which the xylem vascular tissue of the plants is colonized. Entire plants or plant parts above the point of vascular invasion of the pathogen may die within a period of weeks after infection. Until now, three host-specific races (races 1, 2, and 3) of this pathogen have been identified (Stevens and Rick, 1986). The dominant *I2* gene in tomato confers resistance against *F. o. lycopersici* race 2. The gene has been introgressed from the wild tomato species *Lycopersicon pimpinellifolium* (Stall and Walter, 1965) and has been mapped genetically to chromosome 11 (Laterrot, 1976).

Fusarium wilt is an economically important disease of many agricultural and floricultural crops. Resistant cultivars are not always available, or the resistance is rapidly overcome by new races of the pathogen. We have been cloning the tomato *I2* resistance gene and identifying the determinants of specificity as a first step toward developing crops with durable resistance to *F. oxysporum.*

Several resistance genes have been cloned in recent years and have been shown to occur in gene clusters. The *Cf-9*, *Cf-2*, and *Pto*/*Prf* loci of tomato (Martin et al., 1993; Jones et al., 1994; Dixon et al., 1996; Salmeron et al., 1996), the *N* locus in tobacco (Whitham et al., 1994), the *RPP5* locus in Arabidopsis (Parker et al., 1997), the *Xa21* disease resistance gene family in rice (Song et al., 1995, 1997), and the unlinked *L* and *M* loci of flax (Lawrence et al., 1995; Anderson et al., 1997) are each composed of at least five related genes. Sequence analysis of an *Xa21* gene family (Song et al., 1997) indicated that evolution followed recombination, duplication, and transposition events and that not every member of an evolving gene family is functional in conferring resistance.

Recently, two members of a multigene family, designated complex *I2C*, were isolated from the *I2 F. o. lycopersici* race 2 resistance locus in tomato (Ori et al., 1997). The genes show similarity to the group of isolated plant resistance genes that encode cytoplasmic proteins containing a nucleotide binding site motif and leucine-rich repeats (LRRs). The members of the *I2C* family were mapped to five genomic positions. Two of these are clusters of several genes, both

[1]To whom correspondence should be addressed. E-mail keygene@euronet.nl; fax 31-317-424939.
[2]Current address: Agricultural University of Wageningen, Laboratory of Phytopathology, Binnenhaven 9, 6709 HA Wageningen, The Netherlands.
[3]Current address: PE/Genscope, Drie Eikenstraat 661, 2650 Edegem, Belgium.

located on chromosome 11. Only circumstantial evidence was provided for the involvement of the cloned *I2C-1* and *I2C-2* genes in Fusarium resistance.

Here, we report the cloning of the functional *I2* resistance gene using a map-based positional cloning approach. The selective restriction fragment amplification (AFLP) technique was used to identify tightly linked markers. The *I2* gene is a member of the tomato *I2C* multigene family and confers full resistance to *F. o. lycopersici* race 2. We also provide evidence that the previously characterized *I2C-1* and *I2C-2* members of the family (Ori et al., 1997) do not confer resistance to susceptible plants. The *I2C* gene family on chromosome 11 specifying *I2* resistance is composed of seven members. Alignment studies between members of the *I2C* gene family enabled us to determine two regions within the LRR region of the protein that might be involved in Fusarium resistance with *I2* specificity.

RESULTS

Yeast Artificial Chromosome 1/546 Contains the *I2* Gene

Detailed genetic studies (Segal et al., 1992) have oriented the *I2* resistance gene relative to the neighboring restriction fragment length polymorphism (RFLP) markers TG36, TG105, and TG26 on chromosome 11 of tomato. Genetic distances of 4.1 and 4.4 centimorgans (cM) between *I2* and the flanking RFLP markers TG26 and TG36 were determined (Figure 1A). To identify a DNA segment containing the *I2* gene, we screened a 2.2-genome equivalent tomato yeast artificial chromosome (YAC) library with two sets of polymerase chain reaction (PCR) primers derived from the RFLP markers TG36 and TG26. These primer sets generated PCR products of 2.0 and 0.7 kb, respectively, on tomato DNA. A total of seven YACs, four for TG36 and three for TG26, that varied in length from 510 to 975 kb were identified. A long-range physical map was constructed using indirect end-label mapping (Burke et al., 1987), resulting in a YAC contig of 2.5 Mb (Figure 1B). Hybridization experiments were performed with RFLP marker TG105 to localize this marker on the physical map. The *I2* gene has been mapped within 0.4 cM of TG105 between TG105 and TGL36 (Sarfatti et al., 1989). TG105 hybridized with YACs 1/546, 4/645, and 4/1045. Because the overlap between these three YACs was calculated to be ~100 kb, it was concluded that TG105 mapped closely to the left arm of YAC 1/546. The latter YAC (750 kb) also contains the flanking RFLP marker TG36 (Figure 1B).

Identification of Linked AFLP Markers

To identify AFLP markers closely linked to the *I2* gene, we performed AFLP fingerprinting by using the enzyme combination EcoRI and MseI with YAC 1/546. A total of 120 AFLP fragments were identified that were present in YAC strain 1/546 and absent in the yeast acceptor strain AB1380. These AFLP fragments were subsequently analyzed on the resistant and susceptible parental lines, which were used in crosses to identify recombinants. Eighteen polymorphic AFLP markers, designated EM01 to EM18, were identified. Subsequently, these 18 AFLP markers were analyzed on the overlapping YAC clone 1/538 (Figure 1B), which has an overlap with YAC 1/546 of ~250 kb. None of the 18 AFLP markers, however, was present on YAC 1/538, indicating that the *L. pimpinellifolium* introgression segment did not extend into this YAC.

Genetic Mapping of the Tomato *I2* Locus

Resistant tomato lines containing the *I2* gene were crossed with the susceptible tomato lines containing either of the recessive genes *a* (for anthocyaninless) or *sub* (for subtilis) (Figure 1A) to select for recombinants in the *I2* region. The resulting F_2 populations segregated for *I2* and for one of the morphological marker genes. Recombinant plants are either resistant and anthocyaninless or resistant and subtilis. Of 4500 F_2 seeds, a total of 169 resistant anthocyaninless and 18 resistant subtilis were obtained. An AFLP analysis with the 18 linked *I2* markers was performed with these recombinant plants. All recombinant plants, anthocyaninless as well as subtilis, contained all markers, with the exception of seven resistant anthocyaninless plants in which AFLP marker EM18 was lacking. Hence, AFLP marker EM18 flanks the *I2* gene on the proximal side in the direction of the *a* locus. The resistant subtilis plants appeared not to be informative because the recombination events in these plants had not taken place between the *I2* gene and the AFLP markers.

To postulate a flanking AFLP marker on the other side, we screened a large number of randomly selected susceptible F_2 plants with the 18 AFLP markers. All markers were absent in these plants, except for one anthocyaninless plant that contained markers EM03, EM12, and EM16. The recombination event in this latter plant cannot be between *a* and *I2* because it is anthocyaninless and susceptible. Therefore, in this plant, the recombination event has to be at the other, distal side of the *I2* locus. From these data, we conclude that the *I2* gene is flanked by marker EM18 at one end of the DNA segment comprising the *I2* resistance gene and on the other end by the markers EM03, EM12, and EM16. All of the remaining markers completely cosegregated with the *I2* resistance gene based on the analysis of recombinants of crosses with the recessive morphological markers.

Physical Delineation of the *I2* Locus

As a first step in constructing an integrated genetic and physical map around the *I2* locus, the AFLP markers were used as hybridization probes with partial digests of YAC

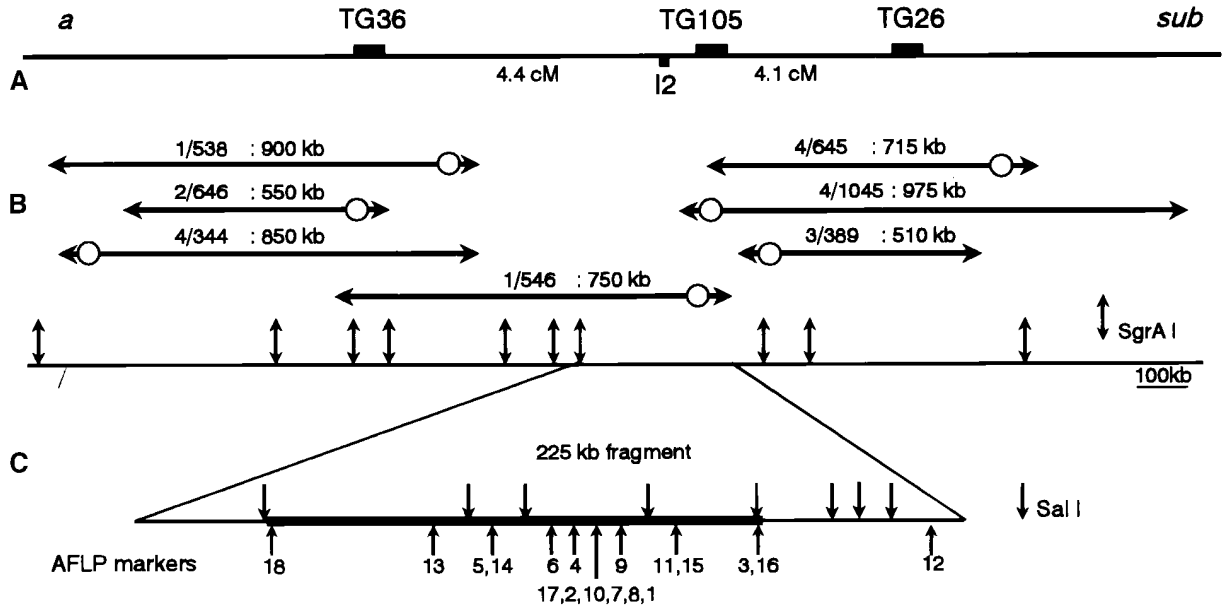


Figure 1. Genetic and Physical Map of the Tomato Genomic Region Containing the *I2* Locus.

(A) RFLP map around the *I2* locus on chromosome 11 of tomato (Segal et al., 1992). Map distance is indicated in centimorgans (cM). The *I2* locus and the recessive morphological markers *a* and *sub* are indicated.

(B) Schematic representation of the YAC contig around the *I2* locus. The seven YACs are indicated by horizontal lines with arrowheads on both ends. The circle and arrowhead combination represents the left arm of pYAC4. The lengths of YACs (in kilobases) and the high-resolution physical map of the region around the *I2* gene for the enzyme SgrAI are indicated.

(C) Physical fine map of the 225-kb fragment for the enzyme SalI and location of the 18 linked AFLP markers. No further resolution of AFLP markers 17, 2, 10, 7, 8, and 1 could be accomplished using SalI. The 130-kb DNA region between the flanking AFLP markers EM18 and EM03/EM16 is indicated as a wide bar. Vertical lines with arrowheads on both sides represent SgrAI sites; vertical lines with one arrowhead represent SalI sites.

1/546. It appeared that all of the AFLP markers were located on one large 225-kb fragment ranging from the left arm of the YAC to the first SgrAI site. To position the AFLP markers EM01 to EM18 on the physical map of this 225-kb fragment, we performed two types of hybridization analysis. First, the AFLP markers were used as probes on the SalI partial digests. Second, the AFLP markers were used as probes on DNAs of the cosmid contig (see next section). A high-resolution physical map for the enzyme SalI of the 225-kb fragment and the positions of the 18 AFLP markers EM01 to EM18 are shown in Figure 1C. The physical distance between the flanking AFLP marker EM18 on one side and markers EM03 and EM16 on the other side was calculated to be ~130 kb.

Contig Building Using AFLP

A cosmid library of YAC clone 1/546 was constructed in the binary vector pCLD04541 (Jones et al., 1992) to perform genetic complementation analysis in transgenic plants. In colony hybridization experiments, ~10,000 cosmids were screened with the purified 225-kb fragment as a probe. Approximately 150 hybridizing colonies were identified. A cosmid contig encompassing the previously identified 130-kb DNA segment was constructed with AFLP fingerprinting using the enzyme combination EcoRI and MseI and primers without selective nucleotides. The DNA fingerprints contained ~8 to 20 amplified fragments. Sets of DNA samples containing amplified fragments of identical size were selected and rerun on polyacrylamide gels until a contiguous array of all of the amplified fragments throughout the 130-kb fragment (flanked by the AFLP markers EM18 and EM03/EM16) was obtained. The final fingerprint of the cosmid contig between the flanking markers is shown in Figure 2.

Complementation Analysis in Transgenic Plants

Twenty genomic cosmid clones encompassing the 130-kb DNA segment were introduced into the susceptible tomato line 52201 by Agrobacterium-mediated transformation (Koornneef et al., 1987). For each cosmid clone, at least 10 independent transgenic R_0 plants were obtained. All R_0 plants were selfed

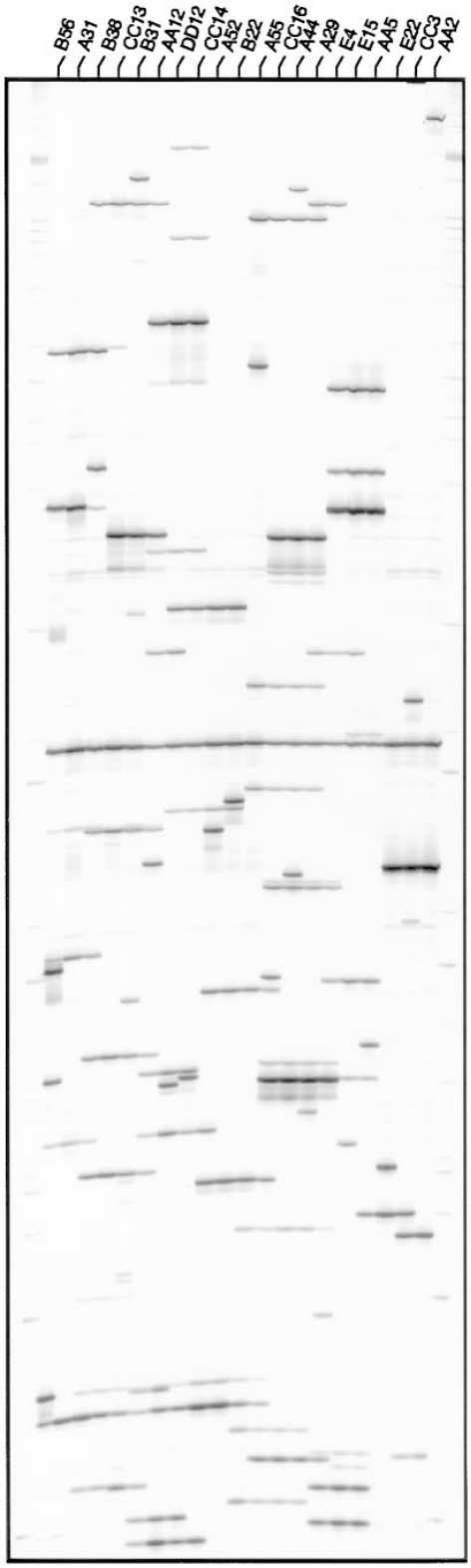


Figure 2. Cosmid Contig Building Using AFLP Fingerprinting.

The final AFLP fingerprint of the 20 cosmids located between the flanking AFLP markers EM18 and EM03/EM16 is shown. The enzyme combination EcoRI and MseI without selective nucleotides was used. The amplified fragments with identical length between two cosmids were aligned until a contiguous array of bands was obtained. Cosmid numbers are indicated above the lanes. Length markers (10-bp ladder) flank the cosmid lanes.

to obtain R_1 seeds. Twenty to 25 seedlings of each R_1 line were inoculated with *F. o. lycopersici* race 2. Only three overlapping cosmids, designated A52, B22, and A55, were able to complement the susceptible phenotype in transgenic plants (Table 1). In most cases, a 3:1 segregation of the resistant phenotype was observed. The other 17 cosmids did not give rise to transformants that exhibited *F. o. lycopersici* race 2 resistance.

An illustration of the resistance assay is shown in Figure 3. After 3 weeks, as shown in Figure 3A, seedlings of the susceptible line 52201 had died, whereas seedlings of a plant transformed with cosmid B22 and seedlings of a plant transformed with A52 segregated 3:1 or 1:0 (resistant to susceptible), respectively (Figures 3B and 3C). Seedlings of a plant transformed with cosmid A29 (Figure 3D) all showed disease symptoms characteristic for Fusarium wilt and were classified as susceptible.

Additional genetic evidence that the *I2* gene is present on cosmids A52, B22, and A55 was obtained in succeeding generations after the selfing of the resistant R_1 plants or backcrossing with the susceptible tomato line 52201. Most R_2 lines segregated 3:1 or were completely resistant, whereas most backcross lines segregated 1:1 or were completely resistant (Table 1). The results obtained indicate that R_1 plants were either heterozygous (segregation ratio of 3:1 or 1:1) or homozygous for *I2* (completely resistant).

Structure of the *I2* Gene

A high-resolution physical map of the three overlapping complementing cosmids A52, B22, and A55 was constructed (data not shown). We concluded that the three cosmids share a DNA segment of ~7 kb. This region from cosmid B22 was subcloned into pBluescript II SK+, and the DNA sequence was determined. One large open reading frame consisting of 3801 bp encoding a polypeptide of 1266 amino acid residues with a relative molecular mass of 144.826 kD was deduced. A PCR-based approach was used to determine whether the *I2* sequence contains any introns. A cDNA library of poly(A)$^+$ RNA isolated from a resistant tomato line was used as a template for PCR amplification. The primers were designed based on the genomic *I2* sequence and covered the complete *I2* region. In addition, 5′ and 3′ rapid amplification of cDNA end (RACE) products

Table 1. Complementation Data of the Cosmids DD2 through A29 Representing the Most Relevant Part of the Cosmid Contig

Cosmid	No. of R_1 Lines[a]	Segregation Ratio[b] 1:0	3:1	1:1	0:1	No. of R_2 Lines[a]	Segregation Ratio[b] 1:0	3:1	1:1	0:1	No. of R_1BC Lines[a]	Segregation Ratio[b] 1:0	3:1	1:1	0:1
DD2	11				11										
CC14	10				10										
A52	10	1	4	1	4	13	2	3	8	0	8	1	3	3	1
B22	11	0	8	0	3	14	5	7	2	0	9	2	1	6	0
A55	13	0	3	2	8	9	1	5	3	0	5	0	1	4	0
CC16	15				15										
A44	10				10										
A29	11				11										

[a] R_1, R_2, and R_1BC progenies of independent transgenic R_0 plants, transformed with one of the cosmid clones, have been tested for resistance against *Fusarium oxysporum* f sp *lycopersici* race 2. R_1 is the selfed progeny of a primary transformant. R_2 is the selfed progeny of a resistant R_1 plant. R_1BC is the progeny of a resistant R_1 plant backcrossed with the susceptible line 52201.

[b] Of each progeny line, 20 to 25 seedlings were assayed. The segregation ratio for resistant to susceptible plants (R:S) was determined for each individual line.

was obtained using the Marathon cDNA amplification kit (Clontech, Palo Alto, CA), cloned into the TA cloning vector pCRII (Invitrogen), and sequenced. These nucleotide sequences were aligned with the 7-kb genomic sequence, and we deduced three intron sequences. An intron of 86 nucleotides was located just upstream of the ATG initiation codon, whereas two other intron sequences of 399 and 82 nucleotides were located in the 3′ untranslated region. No introns were found in the *I2* coding region. A schematic representation of the *I2* gene is depicted in Figure 4A.

Analysis of the *I2* Protein

Computer analysis of the I2 protein with proteins in the GenBank database showed homology to the nucleotide binding site (NBS)-LRR–containing type of *R* genes (reviewed in Baker et al., 1997). The I2 protein can be subdivided into six structural domains (a through f) and are outlined in Figure 4B. Domain a is a hydrophobic region without a cleavable signal sequence. It contains two overlapping potential myristoylation sites (GLAVGG and GAFLSS). Domain b harbors a stretch of basic residues embedding a potential nuclear localization signal (RKHK). The remaining part is rich in acidic residues (20% aspartic acid and glutamic acid). Domain c contains a putative NBS. Of the motifs comprising this predicted NBS, the P loop (GMGGQGKTT) occurs at residues 203 to 209, followed by the kinase 2 (KFLIVLDDV) and 3a (GSKIIVTTRKDSV) regions and the two conserved regions 2 (CKGLPLALKTLAG) and 3 (LKRCFSFC). Domain d starts at residue 499 and is a leucine-rich region with a leucine zipper (LZ) motif LPDSICGLYNLETLLLSSCADL (underlined residues indicate the motif). This part of the protein also contains two potential tyrosine phosphorylation sites (KKVEISGY and RVVTEEFY). Domain e is the LRR region consisting of 17 repeats that fit the proposed consensus for cytoplasmic LRR regions (Jones and Jones, 1997). The proposed central hydrophilic β-sheet is delimited by vertical lines. In 13 out of 17 repeats, the residue following this region is a cystein. Within the final four repeats, the most perfect repeat structure was present. Domain f is the C-terminal region without any homology to known protein regions.

Homology of the I2 Protein Sequence to Members of the *I2C* Gene Family

Recently, four members of the multigene *I2C* family, designated *I2C-1*, *I2C-2*, *I2C-3*, and *I2C-4*, were characterized (Ori et al., 1997). It was shown that the cosmid containing the *I2C-1* gene partially complements *F. o. lycopersici* race 2 susceptibility in only two transgenic plants tested (Ori et al., 1997). Because we identified overlapping cosmids that fully complement the susceptible phenotype, we compared the predicted I2 protein sequence encoded by the overlapping DNA fragment with the other members of the *I2C* gene family as a first step to unravel the *I2* specificity. The I2 protein sequence was 82 and 88% identical to the I2C-1 and I2C-2 protein sequences, respectively. A comparative analysis among these three members of the *I2C* gene family showed three major differences (deletions/insertions; Figure 5, boxes 1, 2, and 3). The first difference (box 1) represents a 12– to 14–amino acid deletion in I2 and I2C-2 compared with I2C-1 and is located in the NBS domain. A second deletion of 28 amino acids (box 2, residue positions 887 to 914) was observed in the I2C-1 protein sequence but not in I2C-2. At the C-terminal end at residues 1191 to 1213 within the I2 protein, a 23–amino acid deletion (box 3) in the I2C-1 and I2C-2 protein sequences was found. Deletion of the 28 or 23 amino acids results in the removal of one complete

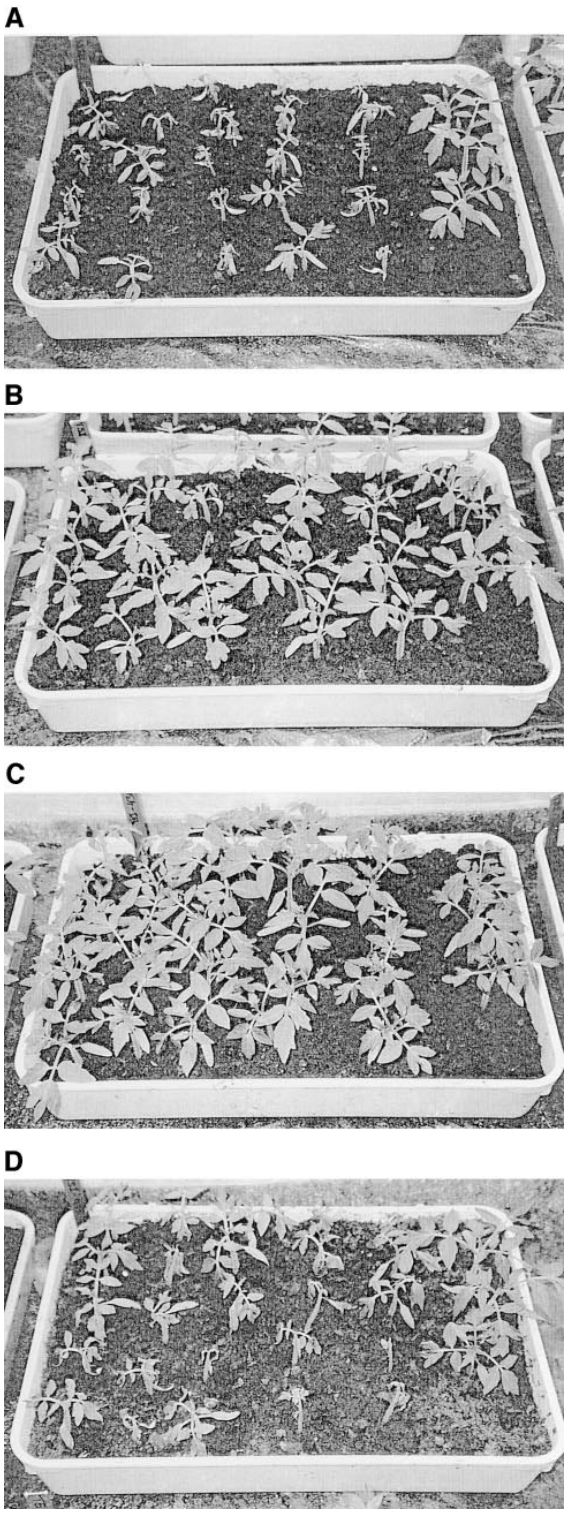


Figure 3. Resistance Assay.

Each picture shows a tray with 25 seedlings 3 weeks after inoculation with *F. o. lycopersici* race 2. At the right of each panel are three uninoculated seedlings that serve as controls. Each tray contains one tomato line.

(A) Susceptible line 52201.

(B) Transgenic R_1 line transformed with cosmid B22 and segregating 3:1.

(C) Transgenic R_1 line transformed with cosmid A52 and segregating 1:0.

(D) Transgenic R_1 line transformed with cosmid A29 and segregating 0:1.

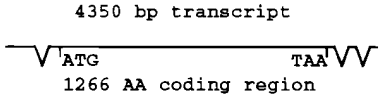
LRR from the protein sequence. Hence, the LRR regions of I2C-1 and I2C-2 consist of 15 or 16 repeats, respectively, compared with the 17 LRRs of the I2 protein sequence.

In addition, we compared the LRRs of the I2 protein sequence with the LRR region of two partial cDNA clones from the *I2C* family, designated *I2C-3* and *I2C-4* (Ori et al., 1997). Major differences (deletions and insertions) were detected in the same regions as where the previous 23– and 28–amino acid deletions were found in the I2C-1/I2C-2 protein sequences. The 28–amino acid deletion at nucleotide positions 887 to 914 was also found in the I2C-3 cDNA protein sequence but not in I2C-4. More striking is the situation at the C-terminal end of the LRR around the 23–amino acid deletion in the I2C-1 and I2C-2 protein sequence compared with the I2 protein sequence, as shown in Figure 6. The conserved 23–amino acid sequence NLQSLAESALPSSLSHLEIDDCP is present in three tandemly oriented copies within the I2 protein sequence and in two copies within the I2C-1 and I2C-2 sequence. In the I2C-3 and I2C-4 protein sequence, five or six copies, respectively, of this 23–amino acid sequence are present, which is indicative of the variability of this region among the various members of the I2-C gene family.

The *I2C* Gene Family at Chromosome 11 Is Composed of Seven Members

PCR using primers adjacent to the two variable regions in domain e and primers in domain d were analyzed on the complete cosmid contig to determine the number of *I2C* copies at the *I2* locus at chromosome 11 (Figure 7). No PCR products could be detected with cosmids E15 through AA2 (Figure 7, lanes 16 to 20), indicating that cosmid E4 (lane 15) represents the right border of the *I2C* gene family. Only small differences in the size of the PCR products could be detected with the primer set located in domain d of the I2 protein sequence (Figure 7, top). Two different PCR products spanning the 28–amino acid variable region were detected with primer set B (Figure 7, middle), whereas three different PCR products were obtained with primer set C spanning the 23–amino acid variable region (Figure 7, bottom). The sizes of the various PCR products using primer set B or C were in agreement with a deletion or insertion of a

A



B

a MEIGLAVGGAFLSSALNVLF

b DRLAPNGDLLNMFRKHKDHVKLLKKLKMTLR
GIQIVLSDAENKQASNPSVRDWLNELRDAVD
SAENLIEEVNYEALRLKVEGQHQNFSETSNQ
QVSDDFFLNIKDKLEDTIETLKDLQEQIGLL
GLKEYFDSTKLETRRPSTSVDDESDIFGRQS
EIEDLIDRLLSEGASGKKLT

c VVPIVGMGGQGKTTLAKAVYNDERVKNHFDL
KAWYCVSEGFDALRITKELLQEIGKFDSKDV
HNNLNQLQVKLKESLKGKKFLIVLDDVWNEN
YNEWNDLRNIFAQGDIGSKIIVTTRKDSVAL
MMGNEQIRMGNLSTEASWSLFQRHAFENMDP
MGHPELEEVGRQIAAKCKGLPLALKTLAGML
RSKSEVEEWKRILRSEIWELPHNDILPALML
SYNDLPAHLKRCFSFCAIFPKDYPFRKEQVI
HLWIANGLVPVKDEINQDLGNQYFLELRSRS
LFEKVPNPSKRNIEELFLMHDLVN

d DLAQLASSKLCIRLEESQGSH
MLEQCRHLSYSIGFNGEFK
KLTPLYKLEQLRTLLPIRIEF
RLHNLSKRVLHNILPTLR
SLRALSFSQYKIKELPNDLFT
KLKLLRFLDISRTWITK
LPDSICGLYNLETLLLSSCADL
EELPLQMEKLINLR
HLDVSNTRRLKMPLHLSRLK
SLQVLVGPKFFVDGWRMEDLGEAQNLHG
SLSVVKLENVVDRREAVKAKMREKNHVE
QLSLEWSESSIADNSQTESDILDELCPHK
NIKKVEISGYRGTNFPNWVADPLFL
KLVNLSLRNCKDCYSLPALGQLP
CLKFLSVKGMHGIRVVTEEFYGRLSSKKPFN
SLEKLEFEDMTEWKQWHALG

e

IGEFPTL	ENLSIKN	CPELSLEIP	1
IQFSSL	KRLEVSD	CPVVFDDAQLFRSQ	2
LEAMKQI	EEIDICD	CNSVTSFPFS	3
ILPTTL	KRIQISR	CPKLKLEAP	4
VGEMFV	EYLRVND	CGCVDDISP	5
EFLPTA	RQLSIEN	CQNVTRFL	6
IPTAT	ETLRISN	CENVEKLSVA	7
CGGAAQM	TSLNIWG	CKKLKCLPE	8
LLPSL	KELRLSD	CPEIEGELP	9
FNL	EILRIIY	CKKLVNGRK	10
EWHLQRL	TELWIDH	DGSDEDIEHWE	11
LPCSI	QRLTIKN	LKTLSSQH	12
LKSLTSL	QYLCIEG	YLSQIQSQGQLSS	13
FSHLTSL	QTLQIWN	FLNLQSLAE	14
SALPSSL	SHLEIDD	CPNLQSLFE	15
SALPSSL	SQLFIQD	CPNLQSLPF	16
KGMPSSL	SKLSIFN	CPLLTPLLE	17

f FDKGEYWPQIAHIPIINIDWKYI

Figure 4. Structure of the *I2* Gene and the I2 Protein.

(A) Schematic representation of the physical structure of the *I2* gene. The diagram represents the *I2* transcript, with three introns indicated by lines angled into vees. The initiation (ATG) and termination (TAA) codons are indicated.

(B) The amino acid sequence predicted from the DNA sequence of the *I2* gene is shown divided into six domains (a to f), as described in the text. Relevant sites are underlined. The two overlapping potential myristoylation sites are in domain a; the potential nuclear localization signal is in domain b; the P loop, kinases 2 and 3a, and the two conserved regions 2 and 3, respectively, are in domain c; the LZ motif (LPDSICGLYNLETLLLSSCADL) and the two potential tyrosine phosphorylation sites are in domain d. The LRR alignment is represented in domain e. Domain f is the C-terminal region of the protein. The numbers to the right of domain e refer to the specific LRR number.

DNA fragment encoding the 28– or 23–amino acid protein sequence. Using primer set C, we identified no PCR products encompassing five or six copies of the 23–amino acid sequence that correspond to the C-terminal region of I2C-3 or I2C-4. Amplification on genomic DNA did produce fragments of this size, indicating that these members are located elsewhere in the genome.

Subsequently, the nucleotide sequences of all of the PCR products generated with primer set C (Figure 7, bottom) on cosmid B56 through E4 were determined. Seven different sequences were obtained and assigned to the various cosmids. For instance, the DNA sequences of the 240-bp PCR product present in cosmids B56, A31, and B38 (Figure 7, lanes 1 to 3) were identical, and therefore, it represents one member of the gene family. Moreover, several additional selected primer sets originating from domains a to f of the I2 protein (Figure 4) were analyzed on the cosmid contig. Even using low annealing temperatures (45°C) and allowing mismatching, no additional PCR products appeared in the region between cosmids B56 and E4, and no PCR products could be detected in the region upstream of cosmid B56 and downstream of E4. Because no PCR products were generated with these additional primer sets, we conclude that the *I2* locus is composed of seven members.

Location of *I2*, *I2C-1*, and *I2C-2* on the Cosmid Contig

The 310-bp PCR product generated with primer set C (Figure 7, bottom) and present in cosmids A52, B22, A55, CC16, and A44 was identical to the C-terminal part of the *I2* gene. This sequence encodes a polypeptide composed of three copies of the NLQSLAESALPSSLSHLEIDDCP repeat, and such a structure is present only in the *I2* gene sequence and is absent in *I2C-1*, *I2C-2*, *I2C-3*, and *I2C-4* (Figure 6). This PCR product is unique throughout the cosmid contig and present in the three complementing cosmids A52, B22, and A55 and the two downstream-located cosmids CC16 and A44 (Figure 7, bottom). The latter two cosmids, however, contain an incomplete *I2* gene (the 5′ end is missing) and are unable to complement the susceptible phenotype. The sequence of the PCR product with a length of ~240 bp

Figure 5. Comparison of the I2, I2C-1, and I2C-2 Protein Sequences.

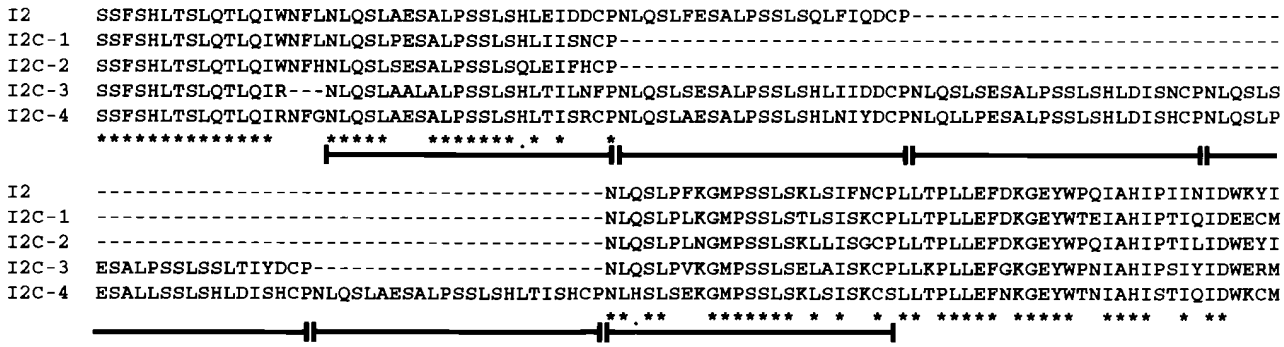


Figure 6. The C-Terminal Parts of the I2, I2C-1, I2C-2, I2C-3, and I2C-4 Homologs Contain Distinct Numbers of an LRR Unit.

Alignment of the C-terminal parts was as described in the legend to Figure 4. The 23–amino acid repeat sequence unit is indicated with a bar. Asterisks indicate identical residues, dots represent similar amino acids, and dashes indicate deletions.

and present in cosmids CC16, A44, A29, and E4 (Figure 7, bottom, open circles) completely matches the C-terminal part of the *I2C-1* sequence. Hence, we conclude that the *I2C-1* gene is located on cosmids CC16, A44, and A29 and partially on cosmid E4 because no PCR product was generated with primer set A on this cosmid (Figure 7, top).

The nucleotide sequence of the 240-bp PCR product generated by primer set C and present in cosmids DD2, CC14, A52, and B22 (Figure 7, bottom, asterisks) was identical to the C-terminal part of the *I2C-2* sequence. Based on the physical map of these cosmids, the *I2C-2* gene could be assigned to a DNA segment ~8 kb upstream of the *I2* gene shared by cosmids A52, B22, and A55.

I2 Specificity

In our complementation analysis, cosmids DD2 and CC14, containing the complete *I2C-2* gene, and cosmids CC16, A44, and A29, containing the complete *I2C-1* gene, did not reveal any increased resistance level compared with control 52201 plants in progeny tests after inoculation with *F. o. lycopersici* race 2. Only the *I2* gene located on the 7-kb DNA segment shared by cosmids A52, B22, and A55 was able to complement the susceptible phenotype. This member of the *I2C* gene family (designated *I2* gene) contains three in-tandem copies of the NLQSLAESALPSSLSHLEIDDCP repeat. Because the remaining six members located on the cosmid contig and the two cDNA members located on another locus do not possess the three in-tandem copies of the NLQSLAESALPSSLSHLEIDDCP repeat but instead harbor either two or more than three copies of this repeat, we postulate that the capability of disease resistance with the I2 specificity is likely to be determined by the number of these repeats present in the C-terminal region of the protein.

DISCUSSION

Isolation and Structure of the *I2* Gene

Here, we report the cloning of the *I2* gene of tomato conferring resistance to race 2 of the fungus *F. o. lycopersici.* Eighteen AFLP markers delineating the *I2* locus were derived from a 750-kb YAC clone. These markers span a DNA region of ~130 kb, corresponding to one marker every 10 kb. A set of 20 overlapping cosmids originating from this region was selected for complementation studies. Cosmid contig building was performed using AFLP fingerprinting without selective nucleotides.

Stable transformants conferring full resistance to race 2 of the fungus were obtained with three overlapping cosmids. None of the 20 cosmids, however, conferred resistance to race 1 of the fungus (G. Simons and J. Wijbrandi, unpublished data), thereby confirming the race-specific nature of the *I2* resistance gene. The *I2* gene encodes a polypeptide

Figure 5. (continued).

The I2, I2C-1, and I2C-2 amino acid sequences were aligned using the sequence comparison program CLUSTAL W (Higgins and Sharp, 1989). Numbers at right indicate the positions of the residues in the protein sequence. Box 1 represents a deletion in the NBS domain; boxes 2 and 3 represent deletions in the LRR region of the protein. Asterisks indicate identical residues, dots represent similar amino acids, and dashes indicate deletions.

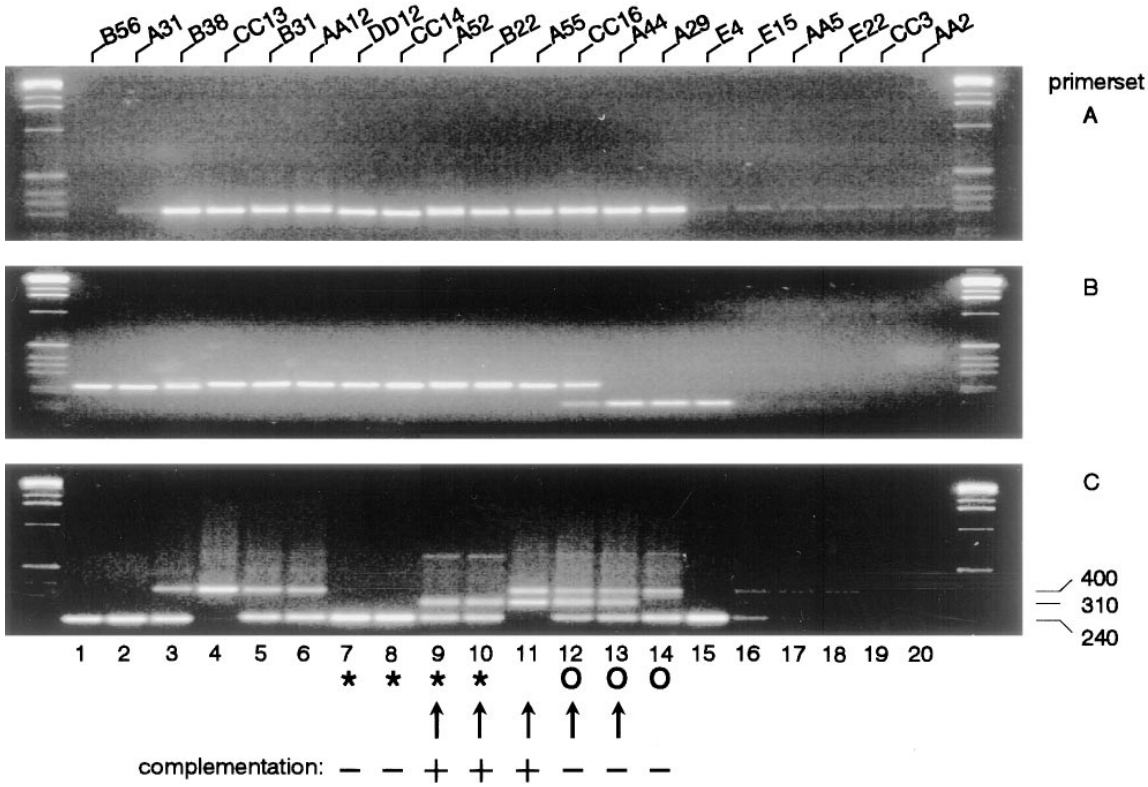


Figure 7. A Gene Cluster of *I2* Homologs.

Ethidium bromide–stained agarose gel of PCR products generated using primer set A, primer set B, and primer set C. Primer set A originates from domain d, whereas primer sets B and C are derived from domain e and encompass repeats 3 and 16 of the *I2* LRR region. The 20 cosmids spanning the region between the flanking AFLP markers are represented by lanes 1 to 20. The sizes of the PCR markers are indicated on the right. The cosmids containing the complete or part of the *I2* gene are indicated with arrows, the cosmids containing the *I2C-2* gene with asterisks, and the cosmids containing the *I2C-1* gene with open circles. The complementation data of these cosmids are given at the bottom of the figure (+, resistant; −, susceptible).

of 1266 amino acids, and comparison with the GenBank database sequences showed that it belongs to the NBS-LRR–containing type of *R* genes (Baker et al., 1997). In addition, the *I2* gene contains an LZ. This LZ lies between the NBS and the LRR domain. In the *RPS2* type of LZ-NBS-LRR–containing type of *R* genes, the LZ is located in the NH_2-terminal region of the protein. The aberrant location of the LZ in the *I2* gene might suggest another subclass of *R* genes. It is predicted that the *I2* gene encodes a cytoplasmic protein. The C-terminal LRR domain spans 17 repeats after the consensus LRR sequence. These 17 repeats are in contrast to the 27 LRRs postulated by Ori et al. (1997). However, they applied a less strict LRR rule than we did. Ori et al. allowed longer LRRs (up to 29 amino acids) and introduced gaps for better alignment (eight in the first 10 LRRs). Moreover, the cystein consensus is lost in the first 10 LRRs of the 27-LRR alignment.

Only the *I2* Gene Confers Full Resistance

The *I2* gene product described in this study is highly homologous to members of the *I2C* family that have been characterized recently by Ori et al. (1997). Two members of this family, *I2C-1* and *I2C-2*, map at the same locus and were found to be 82 and 88% identical, respectively, to *I2.* The I2 gene product fully complements the susceptible phenotype, whereas Ori et al. found only circumstantial evidence for the involvement of *I2C-1* and *I2C-2* in race 2 resistance. Expression of partial gene fragments of *I2C-1* and *I2C-3* in a sense or antisense orientation abrogated immunity to *F. o. lycopersici* race 2, and transgenic plants containing the *I2C-1* gene did not confer complete resistance that would be expected for a dominant gene. Only partial resistance was accomplished by the *I2C-1* gene. Making use of the sequence differences between *I2* and *I2C-1*/*I2C-2*, we were able to

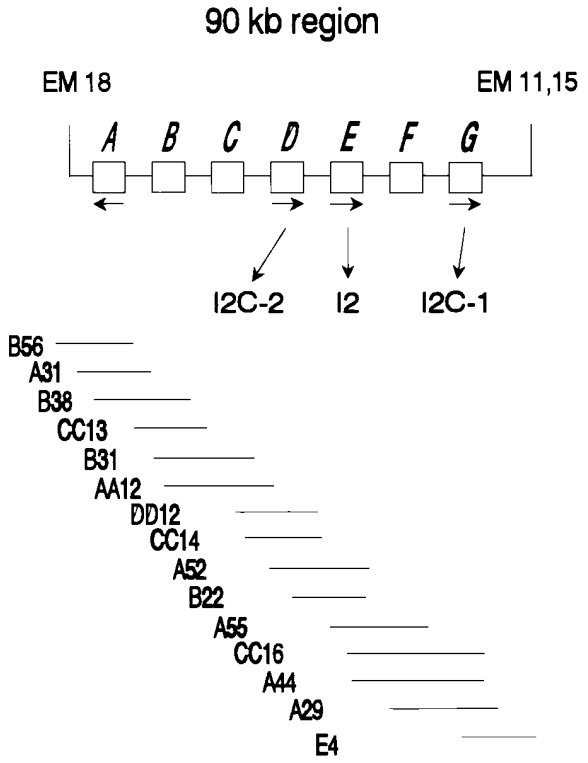


Figure 8. *I2* Locus on Chromosome 11.

Open squares represent the I2 homologs, designated A through G. Horizontal lines represent noncoding regions. The size of the DNA segment containing the homologs and the position of the flanking AFLP markers are indicated. Arrows indicate the direction of the open reading frames. In the lower part, the cosmid contig comprising the seven homologs is given.

position the latter genes on the cosmid contig. These genes map proximal and distal of *I2*, and cosmids containing these genes scored susceptible in our disease assay. Hence, we conclude that the *I2C-1* and *I2C-2* genes are not essential for resistance to *F. o. lycopersici* race 2. These data are in contrast to the data obtained by Ori et al., who claim that the *I2C-1* gene confers partial resistance to race 2 of the fungus.

The *I2C-1* gene is located on cosmid A29, and as shown in Figure 3D, plants transformed with this cosmid all score susceptible. To account for this discrepancy, we compared the two disease assays with regard to the fungal isolate used, the method applied, and the genetic background of the susceptible host. The major difference consists of the genetic background of the susceptible host: Ori et al. (1997) used tomato line VF36, which is resistant to *F. o. lycopersici* race 1, whereas tomato line 52201 is susceptible to both races 1 and 2. Tomato lines carrying race 1 resistance are known to show less severe disease symptoms when infected with race 2 (G. Simons and J. Wijbrandi, unpublished data). Moreover, we can exclude the possibility that resistance necessitates the concerted action of the *I2* gene and an additional member of the gene family, because sequence analysis showed that cosmid B22 contains the full-length *I2* gene but only the C-terminal part of the *I2C-2* gene.

I2 Gene Family

The *I2* locus on chromosome 11 is composed of seven homologs. These seven homologs span a region of ~90 kb. The intergenic region between homologs of the gene family was calculated to be ~8 to 10 kb. The genetic organization of the seven homologs, designated A through G, is depicted schematically in Figure 8. All homologs could be assigned to specific cosmids on the contig. The functional *I2* gene, homolog E, is flanked on one side by *I2C-2* (homolog D) and on the other side by *I2C-1* (homolog G) separated by an additional gene copy (homolog F). The orientation of homolog A is opposite to that of homologs D, E, and G, whereas the orientation of the other homologs is not known. PCR and sequence analysis have clearly demonstrated differences in the LRR region at approximately repeat 16 among the various homologs of the gene family. Sequencing of the complete *I2* locus and characterization of all homologs of the gene family are now in progress.

I2 Specificity Determinants

Examination of the LRR region of three members at the *I2* locus and two (cDNA) members not located at the *I2* locus revealed, besides amino acid substitutions, three sites with a large deletion or insertion. The insertion highlighted in box 1 (Figure 5) is in the NBS domain. Potentially, this insertion could influence the functionality of the *I2C-1* gene; however, there are two deletions (Figure 5, boxes 2 and 3) in the LRR domain to be taken into account. These deletions involve a 28– or 23–complete amino acid repeat unit giving rise to a deletion of one single LRR unit. The *I2* gene product possesses three tandemly oriented copies of a 23–amino acid repeat unit, whereas *I2C-1* and *I2C-2* contain only two and *I2C-3* and *I2C-4* contain five or six copies, respectively, of this repeat. The number of this LRR unit at the C terminus of the protein could be crucial for the expression of *I2* resistance.

These deletions and duplications might arise through DNA rearrangements, such as intragenic recombinations or slipped alignment during replication, as postulated for the flax rust resistance gene *M* (Anderson et al., 1997). Three mutants involving two direct repeats in the 3′ end of the coding region were identified for this gene. The mutants encode an internally deleted M protein containing a single repeat unit. The mutants lost M resistance, indicating that the double repeat structure is essential for resistance gene

function or specificity. Other examples of DNA rearrangements within *R* genes themselves are given by the *Cf-2* and *RPP5* genes. The alternating pattern of A- and B-type LRRs of Cf-2 (Dixon et al., 1996) is consistent with an increase in LRR numbers by intragenic unequal crossing over. A mutant allele for *RPP5* was identified with an in-frame intragenic duplication of 270 bp encoding 90 amino acids, giving rise to four additional LRRs in the open reading frame (Parker et al., 1997).

These examples clearly show the capacity of the LRR domain to evolve new configurations and potentially novel specificities. In the case of the *I2* gene, the new configuration could be pinpointed to a 23–amino acid repeat unit corresponding to one LRR at the C terminus of the protein. We propose that within the *I2C* gene family, evolution of specificities is ruled by the number of the 23–amino acid repeat unit, although we cannot exclude that the 28–amino acid repeat at the N terminus of the LRR is involved as well. Sequence data of all of the homologs of this family and additional domain swaps should give us insight into the structure and evolution of this locus to resolve the underlying mechanisms by which new specificities arise.

METHODS

Plant Material and Genetic Crosses

Tomato (*Lycopersicon esculentum*) lines RZ5, DR9, and E22, homozygous for the *I2* gene, and tomato line 52201, which does not contain the *I2* gene, were obtained from the following seed companies: Rijk Zwaan Zaadteelt en Zaadhandel (De Lier, The Netherlands), De Ruiter Seeds (Bergschenhoek, The Netherlands), and Enza Zaden (Enkhuizen, The Netherlands). The susceptible tomato line GCR210, which is homozygous for the recessive morphological marker gene *a* (for anthocyaninless), and the susceptible line GCR508, which is homozygous for the recessive morphological marker gene *sub* (for subtilis) (Stevens and Rick, 1986), were used in crosses with the resistant tomato lines RZ5 and DR9. Both susceptible lines were obtained from the Institute of Horticultural Research (Littlehampton, UK). The dominant *I2* gene maps at position 85 of chromosome 11, whereas the recessive genes *a* and *sub* map at position 68 (17 centimorgans [cM] from *I2*) and at position 89 (4 cM from *I2*). F_2 seedlings were scored for purple-colored hypocotyls (wild type) or green hypocotyls (anthocyaninless; homozygous *a*). Four-week-old F_2 plants were scored for growth habit: normal (wild type) or small with short internodes (subtilis; homozygous *sub*). The same F_2 plants were tested for resistance to *Fusarium oxysporum* f sp *lycopersici* race 2.

Disease Assay

F. o. lycopersici race 2 was maintained on Czapek Dox Agar (Difco Laboratories, Detroit, MI). Conidial suspensions were obtained by culturing the fungus in Czapek Dox Broth (Difco Laboratories) on a reciprocal shaker for 4 to 7 days at 25°C. The conidia were separated from mycelium fragments by filtration through a stainless steel filter with a pore size of 50 μm. The suspensions were adjusted to a concentration of 2×10^6 conidia per mL by diluting with water.

Tomato seeds were germinated in soil in the greenhouse at 25°C. Ten- to 14-day-old seedlings were used for inoculation with the fungus. The seedlings were carefully pulled out of the soil, and the roots were dipped in water to remove most of the adhering soil. Subsequently, the roots were immersed in the conidial suspension for 2 min, and then the plants were repotted in soil. The plantlets were grown in the greenhouse at 25°C during the day (16 hr) and 22°C at night (8 hr). After 3 to 4 weeks, the plants were scored for disease symptoms. Susceptible plants were small with yellow, wilting leaves and severe browning of stem tissue, and they were often dead. Resistant plants had no such symptoms.

Yeast Artificial Chromosome Library Construction and Screening

The tomato line *L. esculentum* E22 (Enza Zaden), homozygous for the *I2* locus, was used as source material to construct a yeast artificial chromosome (YAC) library. Protoplasts were isolated from the leaves of in vitro–grown shoots, which were 2 to 3 weeks old, as described by Van Daelen et al. (1989). Viable protoplasts (concentration of 50 million protoplasts per mL) were collected and mixed with an equal volume of agarose (1%, Seaplaque; FMC Bioproducts, Rockland, ME) to form a plug. The protoplasts embedded into the plugs were lysed with lysis mix (0.5 M EDTA, 1% *N*-laurylsarcosinate, and 1 mg/mL proteinase K, pH 8.0). After lysis, the plugs were stored at 4°C in storage buffer (fresh lysis mix) until used. Plasmid pYAC4 containing a unique EcoRI cloning site was used as cloning vector, and the yeast strain AB1380 was used as a host (Burke et al., 1987). High molecular weight DNA isolation, partial digestion with EcoRI in the presence of EcoRI methylase, ligation of vector arms to genomic DNA, size selection by pulsed-field gel electrophoresis, and transformation of the yeast host were performed as described by Burke et al. (1987) and Larin et al. (1991). We finally obtained 3840 clones with an average insert size of 520 kb, which corresponds to 2.2 genome equivalents, and the individual clones were stored in 40 96-well microtiter plates containing 75 μL of YPD solution (1% yeast extract, 2% peptone, and 2% dextrose).

To retrieve YAC clones that are positive for the restriction fragment length polymorphism (RFLP) markers TG26 or TG36 (Segal et al., 1992), we arranged the clones in a three-dimensional matrix (cube) and pooled them according to the rows, columns, and plates (X, Y, and Z coordinates, respectively). The library was divided into 4×3-dimensional cubes. Each cube consisted of 960 clones ($Z = 10$). Every matrix contained $8(X) + 12(Y) + 10(Z) = 30$ DNAs, and a total of 120 DNAs were screened.

Selective Restriction Fragment Amplification Screening for Linked Markers

The selective restriction fragment amplification (AFLP) technique (AFLP is a trademark filed by Keygene N.V.) is based on the amplification of subsets of genomic fragments by using polymerase chain reaction (PCR) and has recently been described in detail by Vos et al. (1995). To identify polymorphic markers linked to the *I2* gene, amplification reactions were performed using DNA templates for the enzyme combination EcoRI/MseI. Total YAC clone DNA was analyzed by AFLP fingerprinting, using all 64 +1/+2 primer combinations

(Büschges et al., 1997; Simons et al., 1997). Additional selective nucleotides at the EcoRI and MseI site were determined in a stepwise approach using the four different nucleotides. AFLP fingerprinting with tomato lines was performed with three selective nucleotides at the rare as well as frequent cutter sites. After amplification reactions, the samples were analyzed on a 5% denaturating polyacrylamide gel (Vos et al., 1995). Fingerprint patterns were visualized using a Fuji BAS-2000 PhosphorImager analysis system (Fuji Photo Film Company Ltd., Tokyo, Japan).

Cosmid Library

Total DNA of the *Saccharomyces cerevisae* AB1380 containing YAC 1/546 was isolated using zymolyase to make protoplasts, according to Green and Olsen (1990). DNA was partially digested with Sau3A-generating molecules with an average size of 15 to 25 kb. After centrifugation through a 10 to 35% sucrose gradient at 22,000 rpm for 22 hr at 20°C in a rotor (model SW41; Beckman Instruments, Palo Alto, CA), the fractions containing DNA molecules with a size of ~20 kb were pooled and concentrated by ethanol precipitation. The cohesive ends were partially filled in with dATP and dGTP, using the strategy of partial filling of 5′ extensions of DNA produced by type II restriction endonuclease, as described by Korch (1987) and Loftus et al. (1992). The binary T-DNA cosmid vector pCLD04541 (derivative of pJJ1881; Jones et al., 1992) was digested completely with XhoI, and the linear fragment was partially filled in with dTTP and dCTP, as described by Korch (1987). The 20-kb fragments were ligated to the cosmid vector and transduced to *Escherichia coli* strain XL1-Blue MR (Stratagene, La Jolla, CA) by using phage λ Gigapack II XL packaging extracts (Stratagene), as recommended by the manufacturer. Selection was performed on Luria-Bertani (1% bacto-tryptone, 0.5% bacto-yeast extract, and 1% NaCl, pH 7.5) agar plates containing 10 mg/L of tetracyclin.

Construction of a Cosmid Contig of the 255-kb Fragment

Restriction fragment amplification, as described by Vos et al. (1995), was used to construct a cosmid contig of the 225-kb fragment. Approximately 500 ng of cosmid DNA was used for template preparation, and the primers used in the amplification of restriction fragments were the EcoRI primer 5′-GACTGCGTACCAATTC-3′ and the MseI primer 5′-GATGAGTCCTGAGTAA-3′ having no selective nucleotides. The EcoRI primer was labeled at the 5′ end, and each of the DNAs was amplified using the described primer set.

Complementation

The cosmid clones were introduced into *Agrobacterium tumefaciens* through conjugative transfer in a triparental mating with helper strain HB101 (pRK2013), essentially according to Deblaere et al. (1987). The susceptible tomato line 52201 was used for transformation with Agrobacterium. Transformation of tomato was performed according to Koornneef et al. (1987) by using 8-day-old cotyledon explants. The individual transformants of every cosmid clone were transferred to the greenhouse for seed production. R_1 lines of ~9 to 15 R_0 plants of each cosmid were tested for disease symptoms. Plants containing the binary cosmid vector without insert were included as a control.

PCR Analysis, DNA Sequencing, and Computer Analysis

For PCR analysis of the cosmid contig, we used the following primers: primer set A (sense, 5′-GGTGTTGCATAACATACTGCC-3′; antisense, 5′-TCAACCTGCTCAGATGTAGTGG-3′), primer set B (sense, 5′-TGGAGAGTTCCCTACACTTGAG-3′; antisense, 5′-TTCTCTTCAAGGTAGTTGGCAG-3′), and primer set C (sense, 5′-CCTCCTTTTCTCACCTCACTTCGC-3′; antisense, 5′-ATTTGTGGCCAGTATTCCCC-3′). To determine the sequence of the DNA segment containing the *I2* gene, we subcloned fragments of cosmid B22 in pBluescript II SK+ (Stratagene). Double-stranded DNA sequencing was performed using the Autoread Sequencing Kit and the A.L.F. DNA Sequencer device (Pharmacia LKB, Uppsala, Sweden). All sequencing was conducted on both strands. The data were assembled using the Staden sequence and analysis program (Dear and Staden, 1991). Multiple sequence alignments were performed using Clustal W (1.60) software (Higgins and Sharp, 1989).

ACKNOWLEDGMENTS

We thank Jerina Pot and Maarten Van Riel for artwork. We acknowledge the help of René Brinkman in sequencing the LRR-PCR products of the *I2* locus. This work was financed by Keygene N.V. and five Dutch breeding companies: Cebeco Zaden, De Ruiter Seeds, Enza Zaden, Rijk Zwaan, and Van der Have. The work was subsidized by the Dutch Ministry of Economic Affairs through a Programmatische Bedrijfsgerichte Technologie Stimulering grant.

Received December 29, 1997; accepted April 20, 1998.

REFERENCES

Anderson, P.A., Lawrence, G.J., Morrish, B.C., Ayliffe, M.A., Finnegan, E.J., and Ellis, J.G. (1997). Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. Plant Cell **9,** 641–651.

Baker, B., Zambryski, P., Staskawicz, B., and Dinesh-Kumar, S.P. (1997). Signaling in plant–microbe interactions. Science **276,** 726–733.

Burke, D.T., Carle, G.F., and Olson, M.V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science **236,** 806–812.

Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., Van Daelen, R., Van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F., and Schulze-Lefert, P. (1997). The barley *Mlo* gene: A novel control element of plant pathogen resistance. Cell **88,** 695–705.

Dear, S., and Staden, R. (1991). A sequence assembly and editing program for efficient management of large projects. Nucleic Acids Res. **14,** 3907–3911.

Deblaere, R., Reynaerts, A., Höfte, H., Hernalsteens, J.P., Leemans, J., and Van Montagu, M. (1987). Vectors for cloning in plant cells. Methods Enzymol. **153,** 277–292.

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.T., 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.

Green, E.D., and Olsen, M.V. (1990). Systematic screening of yeast artificial chromosome libraries using the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **87,** 1213–1217.

Higgins, D.G., and Sharp, P.M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. Comput. Appl. Biosci. **5,** 151–153.

Jones, D.A., and Jones, J.D.G. (1997). The role of leucine-rich repeat proteins in plant defences. Adv. Bot. Res. **24,** 89–167.

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

Jones, J.D.G., Shlumukov, L., Carland, F., English, J., Scofield, S.R., Bishop, G.J., and Harrison, K. (1992). Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. Transgen. Res. **1,** 285–297.

Koornneef, M., Jongsma, M., Weide, R., Zabel, P., and Hille, J. (1987). Transformation of tomato. In Tomato Biotechnology, D.J. Nevins and R.A. Jones, eds (New York: Alan R. Liss Inc.), pp. 169–178.

Korch, C. (1987). Cross index for improving cloning selectivity by partial filling of 5′-extensions of DNA produced by type II restriction endonucleases. Nucleic Acids Res. **15,** 3199–3220.

Larin, Z., Monaco, A.P., and Lehrach, H. (1991). Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc. Natl. Acad. Sci. USA **88,** 4123–4127.

Laterrot, H. (1976). Localisation chromosomique de *I2* chez la tomate controlant la resistance au pathotype 2 de *Fusarium oxysporum* f.sp. *lycopersici.* Ann. Amelior. Plant. **26,** 485–491.

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.

Loftus, M.G., Foster, L.M., and Ross, I.A. (1992). A rapid method for cosmid cloning. BioTechniques **12,** 172–176.

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.

Ori, N., Eshed, Y., Paran, I., Presting, G., Aviv, D., Tanksley, S., Zamir, D., and Fluhr, R. (1997). The *I2C* family from the wilt disease resistance locus *I2* belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. Plant Cell **9,** 521–532.

Parker, J.E., Coleman, M.J., Szabò, V., Frost, L.N., Schmidt, R., Van der Biezen, E.A., Moores, T., Dean, C., Daniels, M.J., and Jones, J.D.G. (1997). The Arabidopsis downy mildew resistance gene *RPP5* shares similarity to the Toll and interleukin-1 receptors with *N* and *L6.* Plant Cell **9,** 879–894.

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.

Sarfatti, M., Katan, J., Fluhr, R., and Zamir, D. (1989). An RFLP marker in tomato linked to the *Fusarium oxysporum* resistance gene *I2.* Theor. Appl. Genet. **78,** 755–759.

Segal, G., Sarfatti, M., Schaffer, M.A., Ori, N., Zamir, D., and Fluhr, R. (1992). Correlation of genetic and physical structure in the region surrounding the *I2 Fusarium oxysporum* locus in tomato. Mol. Gen. Genet. **231,** 179–185.

Simons, G., Van der Lee, T., Diergaarde, P., Van Daelen, R., Groenendijk, J., Frijters, A., Büschges, R., Hollricher, K., Töpsch, S., Schulze-Lefert, P., Salamini, F., Zabeau, M., and Vos, P. (1997). AFLP-based fine mapping of the *Mlo* gene to a 30-kb DNA segment of the barley genome. Genomics **44,** 61–70.

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.

Song, W.-Y., Pi, L.-Y., Wang, G.-L., Gardner, J., Holsten, T., and Ronald, P.C. (1997). Evolution of the rice *Xa21* disease resistance gene family. Plant Cell **9,** 1279–1287.

Stall, R.E., and Walter, J.M. (1965). Selection and inheritance of resistance in tomato to isolates of races 1 and 2 of the *Fusarium* wilt organism. Phytopathology **55,** 1213–1215.

Stevens, M.A., and Rick, C.M. (1986). Genetics and breeding. In The Tomato Crop, J.G. Atherton and J. Rudich, eds (London: Chapman and Hall Ltd.), pp. 35–109.

Van Daelen, R.A.J., Jonkers, J.J., and Zabel, P. (1989). Preparation of megabase-sized tomato DNA and separation of large restriction fragments by field inversion gel electrophoresis (FIGE). Plant Mol. Biol. **12,** 341–352.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van der Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. **23,** 4407–4414.

Whitham, S., Dinesh-Kumar, S.P., Choll, 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.

Dissection of the Fusarium *I2* Gene Cluster in Tomato Reveals Six Homologs and One Active Gene Copy

Guus Simons, John Groenendijk, Jelle Wijbrandi, Martin Reijans, José Groenen, Paul Diergaarde, Theo Van der Lee, Marjo Bleeker, Joke Onstenk, Michiel de Both, Michel Haring, Jurriaan Mes, Ben Cornelissen, Marc Zabeau and Pieter Vos

Plant Cell 1998;10;1055-1068
DOI 10.1105/tpc.10.6.1055