Available online at www.sciencedirect.com
SCIENCE DIRECT®
Fungal Genetics and Biology 43 (2006) 111–123
FUNGAL GENETICS AND BIOLOGY
www.elsevier.com/locate/yfgbi

A cDNA-AFLP based strategy to identify transcripts associated with avirulence in *Phytophthora infestans* ☆

Jun Guo [a,b,c,1], Rays H.Y. Jiang [a,1], Lars G. Kamphuis [a,2], Francine Govers [a,*]

[a] *Plant Sciences Group, Laboratory of Phytopathology, Wageningen University, Binnenhaven 5, NL-6709 PD Wageningen and Graduate School Experimental Plant Sciences, The Netherlands*
[b] *Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China*
[c] *College of Plant Protection, Northwest A & F University, Yangling Shaanxi 712100, China*

Received 17 September 2005; accepted 23 November 2005

Abstract

Expression profiling using cDNA-AFLP is commonly used to display the transcriptome of a specific tissue or developmental stage. Here, cDNA-AFLP was used to identify transcripts in a segregating F1 population of *Phytophthora infestans*, the oomycete pathogen that causes late blight. To find transcripts derived from putative avirulence (*Avr*) genes germinated cyst cDNA from F1 progeny with defined avirulence phenotypes was pooled and used in a bulked segregant analysis (BSA). Over 30,000 transcript derived fragments (TDFs) were screened resulting in 99 *Avr*-associated TDFs as well as TDFs with opposite pattern. With 142 TDF sequences homology searches and database mining was carried out. cDNA-AFLP analysis on individual F1 progeny revealed 100% co-segregation of four TDFs with particular AVR phenotypes and this was confirmed by RT-PCR. Two match the same *P. infestans* EST with unknown sequence and this is a likely candidate for *Avr4*. The other two are associated with the *Avr3b-Avr10-Avr11* locus. This combined cDNA-AFLP/BSA strategy is an efficient approach to identify *Avr*-associated transcriptome markers that can complement positional cloning.

Keywords: cDNA-AFLP; Transcript profiling; *Phytophthora infestans*; Avirulence gene; BSA

1. Introduction

Many plant–pathogen interactions are governed by specific interactions between pathogen avirulence (*Avr*) genes and corresponding plant resistance (*R*) genes. An interaction where a corresponding pair of *R* gene and *Avr* gene is present and expressed, results in incompatibility and the plant is resistant. When one of the two is inactive or absent, the interaction is compatible and the plant susceptible. This cross talk between host and pathogen was assembled in the gene-for-gene model by Flor (1942), who extracted the concept from his work on the interactions between flax and flax rust. Since, the early nineties numerous *R* genes from model plant or crop species have been identified and cloned (Young, 2000; Dangl and Jones, 2001) and, in parallel, many *Avr* genes mainly from fungi and bacteria (White et al., 2000; Luderer and Joosten, 2001; van't Slot and Knogge, 2002). The availability of both a cloned *R* gene and its corresponding cloned *Avr* gene offers exciting opportunities to elucidate the gene-for-gene interaction at the molecular and cellular level. In recent years, the guard model has won ground particularly by studies on a few model pathosystems, such as the interactions between Arabidopsis or tomato and the bacterial speck pathogen *Pseudomonas syringae*, and tomato and the leaf mold fungus *Cladosporium fulvum* (Innes, 2004;

☆ The GenBank accession numbers for the TDFs described in this paper are DW010060 to DW010198 (to be provided upon acceptance).
* Corresponding author. Fax: +31 317 483 412.
E-mail address: Francine.Govers@wur.nl (F. Govers).
[1] These authors contributed equally to the work.
[2] Present address: Australian Centre for Necrotrophic Fungal Pathogens, SABC, Murdoch University, Perth, WA 6150, Australia.

1087-1845/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.fgb.2005.11.001

Rooney et al., 2005). In this model R proteins and pathogen effectors (i.e., AVR proteins) are part of a larger dynamic complex. The pathogen effectors target host cell proteins in order to suppress defense responses or elicit susceptible responses. R proteins evolved as a counter-defense and function to monitor the effector targets.

The subject of our studies is *Phytophthora infestans*, the notorious Irish potato famine pathogen and the causal agent of late blight (Govers and Latijnhouwers, 2004). *Phytophthora* species resemble fungi morphologically but in the tree of life they are classified as oomycetes, a unique group of eukaryotes that evolved independently from fungi. Oomycetes include significant pathogens of insects and animals and they are responsible for a wide variety of destructive plant diseases. All *Phytophthora* species (more than 65), and the majority of the *Pythium* species are plant pathogens, and also all downy mildew diseases and white rusts are caused by oomycetes (Agrios, 1997). Oomycetes not only look like fungi, they also behave like fungi and use the same weaponry to attack plants (Latijnhouwers et al., 2003). Similarly, the R proteins that plants use to defeat oomycetes have the same architecture as R proteins that stop fungal invasions (Ballvora et al., 2002; van der Vossen et al., 2003; Gao et al., 2005; Huang et al., 2005) and many oomycete–plant interactions follow the gene-for-gene model. Genetic analyses on host and pathogen have demonstrated that this model also suits the potato-*P. infestans* pathosystem (van der Lee et al., 2001).

Unlike R proteins, the pathogens' AVR proteins or effectors are highly divergent (Luderer and Joosten, 2001; van't Slot and Knogge, 2002). Many of the fungal *Avr* genes were cloned by reverse-genetics using purified elicitor preparations as starting material. For genetically more tractable fungi, like, for example, *Magnaporthe grisea*, positional cloning appeared to be a suitable approach, and for cloning bacterial *Avr* genes classical bacterial genetics, such as genetic complementation proved to be very efficient (van den Ackerveken and Bonas, 1997; Collmer, 1998). In the case of *Phytophthora*, however, *Avr* gene cloning has lagged behind (Tyler, 2001, 2002). Because of the (hemi-)biotrophic nature of many oomycete–plant interactions purifying elicitors is difficult and, in our hands attempts to identify race specific elicitors from *P. infestans* were unsuccessful (Alfonso and Govers, 1995). Therefore, reverse genetics is not an option. Moreover, low DNA transformation efficiencies and relatively large genome sizes hamper complementation or gene tagging approaches. A more suitable approach is positional cloning and recently three oomycete *Avr* genes have been identified starting off with this approach: *Avr1b*-1 from *Phytophthora sojae* (Shan et al., 2004), and *ATR13* and $ATR1^{NDWsB}$ from the Arabidopsis downy mildew pathogen *Hyaloperonospora parasitica* (Allen et al., 2004; Rehmany et al., 2005). These two species are homothallic and the number of inbred progeny that was generated was sufficient to obtain recombinants in the *Avr* regions and to identify closely linked markers.

For cloning *Avr* genes in *P. infestans* we also adopted a positional cloning approach and generated high-density maps of chromosomal regions carrying *Avr* genes (van der Lee et al., 2001). In addition, a BAC library of a strain carrying six dominant *Avr* genes and suitable for marker landing, is available (Whisson et al., 2001). However, *P. infestans* is heterothallic and the problem we face is the inability to generate large segregating mapping populations. Also the relatively large genome size (245 Mb) reduces the marker density and even with high-density linkage maps (van der Lee et al., 2004) we were not able to generate enough markers for efficient landing. To complement the positional cloning strategy we aimed at generating transcriptome markers. In this study, we combined a cDNA-AFLP based strategy with bulked segregant analysis (BSA) to identify *Avr*-associated transcripts. cDNA-AFLP is a relatively simple method to obtain a genome-wide display of differentially expressed genes and it has already been successfully used for gene discovery in *P. infestans* (Avrova et al., 2003; Dong et al., 2004). Many of the known *Avr* genes show a relatively high expression or a stage specific expression in pre-infection stages and therefore, we used germinating cysts as starting material for RNA isolation. cDNA-AFLP patterns obtained from pools of strains with identical AVR phenotypes revealed a high number of putative *Avr*-associated transcript derived fragments (TDFs) for each of the four *Avr* loci that were targeted. Subsequently, segregation of the *Avr*-associated TDFs in an F1 mapping population was analyzed resulting in transcriptome markers for two *Avr* loci.

2. Materials and methods

2.1. P. infestans strains and mapping population

The *P. infestans* strains used in this study are two Dutch field isolates of opposite mating type (80029; A1 and 88133; A2) and 18 F1-progeny (designated as cross 71). The cross 71 mapping population was previously described and characterized by Drenth et al. (1995) and van der Lee et al. (1997). The nomenclature of genes, gene clusters, and phenotypes is according to van der Lee et al. (2001) with one exception; *Avr3* now has the suffix '*b*' to indicate that this avirulence gene elicits resistance on plants carrying resistance gene R3b and not R3a (Huang et al., 2004). Consequently, an avirulent and virulent phenotype on R3b plants is indicated by AVR3b and avr3b, respectively.

2.2. P. infestans culture conditions

Phytophthra infestans strains were routinely grown at 18 °C in the dark on rye agar medium supplemented with 2% sucrose (RSA) (Caten and Jinks, 1968). To obtain germinating cysts for RNA isolation, sporulating mycelium grown on RSA was flooded with ice-cold water and incubated at 4 °C. At this temperature, sporangia release the zoospores into the water. After 4 h incubation, the

zoospore suspension was filtered through a 10-μm nylon mesh to remove sporangia and mycelial fragments. Cysts were obtained by vigorous shaking of the zoospore suspension for 2 min. To allow germination the cyst suspension was incubated at 18 °C for at least 2 h. The germination rate and germ tube length were checked with regular time intervals. When more than half of the cysts were germinated and the length of their germ tubes was 4–6 times the diameter of the cysts the tissue was collected by centrifugation (5 min at 3000*g*), frozen in liquid nitrogen and stored at −80 °C.

2.3. cDNA-AFLP analysis

RNA isolation, cDNA synthesis, and cDNA-AFLP analysis were performed as described previously for *P. infestans* by Dong et al. (2004). Total RNA from germinated cysts was isolated using Trizol (Gibco-BRL) according to the manufacturer's instructions and subsequently purified using phenol–chloroform extraction. Poly A^+ RNA was isolated from 100 μg total RNA with the QIAGEN Oligotex mRNA kit. cDNA was synthesized using oligo dT (12–18) and superscript II reverse transcriptase (Gibco-BRL). The primary template for cDNA-AFLP was prepared in a one-step restriction–ligation reaction in which adapters were ligated to *Apo*I/*Taq*I digested cDNA fragments. The quality of each primary template was checked by performing a PCR on the diluted primary template using primers matching the adapters and by analyzing the PCR products on agarose gel. Based on the intensity on gel the quantity was estimated. Pre-amplification was performed in 25 cycles using primers corresponding to the *Apo*I and *Taq*I adapters without extension (A and T primers as in Dong et al., 2004). The diluted pre-amplification products were used as template for the selective amplification with two selective base extensions at the 3′-end of the primers (A + 2 and T + 2 primers). The A + 2 primers were either labeled by phosphorylating the 5′-end with [γ-^{32}P]ATP for detection of the cDNA-AFLP fragments by autoradiography, or with IRD700 or IRD800 for fluorescence detection using LI-COR Global IR^2 systems. For analysis of the cDNA-AFLP fragments by silver staining, the primers were not labeled. Separation of the cDNA-AFLP fragments was performed on 4–6% denaturating polyacrylamide gels as described by van der Lee et al. (1997).

2.4. Bulked segregant analysis

Bulked segregant analysis (BSA) was performed essentially following the procedure described by Michelmore et al. (1991). Ten F1 progeny of the cross 71 mapping population were selected and divided over four pools consisting of 2 or 3 F1 progeny with identical or nearly identical avirulence phenotypes (Table 1). Each phenotype is represented by 4-6 F1 progeny divided over two pools. From the six avirulence genes that segregate in cross 71 *Avr3* (renamed *Avr3b*), *Avr10* and *Avr* 11 are closely linked (van der Lee et al., 2001) and in this study we consider *Avr3b-Avr10-Avr11* as one locus. Primary templates of the 2 or 3 F1 progeny that made up one pool were mixed in equal amounts (based on the quantity and quality check described above) and served as template for the pre-amplification. In the selective amplification, all 256 *Apo*I + 2/*Taq*I + 2 primer combinations were used. In Figs. 1B–E the expected patterns for each of the pools are shown.

2.5. Isolation, cloning, and sequencing of TDFs

The cDNA-AFLP fragments (i.e., TDFs) of interest were excised from gels using a razor blade. The gel slices were rehydrated in 100 μl of water and incubated at 70 °C for 15 min. The eluted fragment was reamplified with the primers with the same two base pair extension as used in the cDNA-AFLP analysis. PCR products were purified using QIAquick PCR purification kit (Qiagen, Hilden) and cloned into pGEM-T Easy (Promega, Madison, WI,

Table 1
Composition of BSA pools for selecting *Avr*-associated TDFs

Pool	Strain	Phenotypes on differentials containing resistance gene[a]					
		R1	*R3b*	*R10*	*R11*	*R4*	*R2*
1	re11-16	AVR	avr	avr	avr	AVR	AVR
	T15-1	AVR	avr	avr	avr	AVR	AVR
	T30-2	AVR	avr	avr	avr	AVR	AVR
2	D12-2	avr	avr	avr	avr	avr	AVR
	D12-23	avr	avr	avr	avr	avr	AVR
	T35-3	avr	avr	avr	avr	avr	avr
3	D12-17	AVR	AVR	AVR	AVR	avr	avr
	T15-9	AVR	AVR	AVR	AVR	avr	avr
4	T20-2	avr	AVR	AVR	AVR	AVR	AVR
	E12-3	avr	AVR	AVR	AVR	AVR	avr

[a] AVR and avr indicate avirulence and virulence phenotype, respectively.

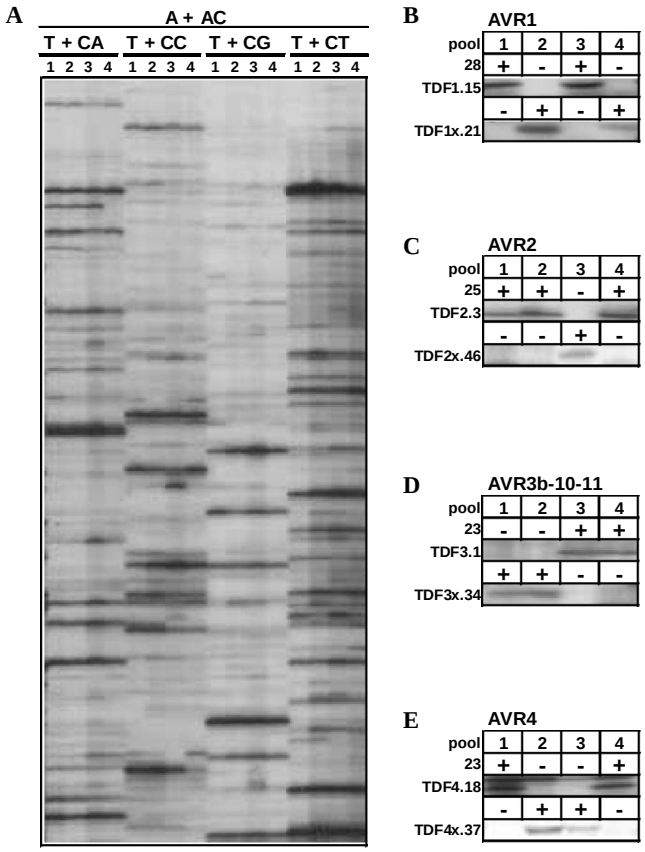

Fig. 1. cDNA-AFLP analysis. (A) Section of autoradiograph showing cDNA-AFLP fingerprints in four BSA pools generated with the indicated primer combinations. For the composition of pools 1, 2, 3, and 4 see Table 1. In (B–E) the second row shows the number of *Avr*-associated TDF candidates found in this study and the expected cDNA-AFLP patterns in the four pools for TDFs associated with AVR1, AVR2, AVR3b-AVR10-AVR11, and AVR4 phenotypes, respectively. The third row shows examples of candidate TDFs with the expected pattern. The fourth row shows the expected opposite pattern and the fifth row examples. In (B) TDF1.25 was obtained with primer combination A + TG/T + CG and TDF1 x .21 with A + GA/T + GT. In (C) TDF2.3 with A + AG/T + TT and TDF2 x .46 with A + TG/T + GG. In (D) TDF3.1 with A + AG/T + AC and TDF3 x .34 with A + GT/T + CC. In (E) TDF4.18 with A + TG/T + GG and TDF4 x .37 with A + GT/T + TA.

USA). Recombinant clones were sequenced by BaseClear (Leiden, The Netherlands) or Shanghai Biotech (Shanghai, China).

2.6. DNA sequence analysis and bioinformatics

Sequences were analysed in Vector NTI 8. For BLAST searches, we used the NCBI BLAST program and the Standalone-BLAST Version 2.2.3 (Altschul et al., 1997). The *P. infestans* EST databases are accessible at http://www.pfgd.org and http://staff.vbi.vt.edu/estap (Kamoun et al., 1999; Randall et al., 2005). The genomic sequences and annotated protein sequences of *P. sojae* and *P. ramorum* were obtained from the website of the DOE Joint Genome Institute (http://www.jgi.doe.gov/genomes). TDF sequences were searched against GenBank and EST databases by BLASTX and BLASTN, respectively. A GenBank hit was considered to be a homologue if the BLASTX *E* value is less than 1E−3. A TDF was considered to be represented by an EST if the BLASTN identity is equal to or larger than 99%. RT-PCR primers were designed based on the cloned TDF sequence or the EST sequence if the TDF has a corresponding EST. Primer lengths were between 18 and 25 bp with melting temperatures higher than 55 °C in all cases. The primer sequences are available from the authors upon request.

2.7. RT-PCR analysis

To remove genomic DNA from RNA preparations, 10 μg total RNA was treated with 4 U RQ1 RNase-free DNase (Promega, Madison, WI) at 37 °C for 1 h. The removal of all DNA was verified in a PCR reaction under the same conditions as those used for the RT-PCR reaction, except that the cDNA synthesis step was omitted. The first-strand cDNA was synthesized using oligo dT(16) and Superscript II reverse transcriptase for 30 min at 40 °C (Gibco-BRL). Sequence-specific primers were used in the subsequent PCR with cDNA as template with 30 cycles (30 s at 94 °C, 30 s at 56–60 °C and 60 s at 72 °C).

2.8. Nomenclature of TDFs

The cDNA-AFLP fragments and the clones containing the fragments are named TDF followed by a number that refers to the *Avr* gene for which, according the BSA pattern, the TDF was a candidate. This *Avr*-associated number is then followed by a period and a random clone number. For *Avr3b-Avr10-Avr11*, the *Avr*-associated number is 3. In cases where an 'x' is added as suffix the TDF showed an opposite pattern in the BSA. Occasionally, an 's' is added at the very end to indicate that the TDF was selected in the BSA analysis on silver stained gels.

3. Results and discussion

3.1. BSA for selecting transcripts associated with avirulence

BSA was initially developed as a method for rapidly identifying polymorphic DNA markers linked to any specific gene or genomic region (Michelmore et al., 1991). Two bulked DNA samples are generated from a segregating population from a single cross. Each pool, or bulk, contains individuals that are identical for a particular trait or genomic region but arbitrary at all unlinked regions. The two bulks are therefore genetically dissimilar in the selected region but seemingly heterozygous at all other regions. Previously, the *P. infestans* cross 71 mapping population was succesfully used for BSA to identify AFLP markers linked to six *Avr* genes segregating in cross 71 (van der Lee et al., 2001). In the present study we used the same cross 71 mapping population, a similar pool design and the same pool sizes for a BSA approach aimed at selecting transcripts

derived from *Avr* genes. Instead of DNA, cDNA of different individuals from the cross was pooled. The phenotypes of the strains that constitute the four BSA pools are listed in Table 1. Anticipating that *Avr* genes are expressed just prior to infection we used RNA isolated from germinating cysts as starting material.

It is logical to combine a BSA approach with an efficient genome-wide transcriptional profiling method. Recently, Dong et al. (2004) described an optimized cDNA-AFLP protocol for *P. infestans* that was based on in silico cDNA-AFLP fingerprinting of a large set of *P. infestans* ESTs. The primer combination *Apo*I/*Taq*I and selective amplification using primers with two base extensions resulted in clear transcription profiles that were easy to score. Fig. 1A shows a section of a typical autoradiograph with cDNA-AFLP patterns obtained from the four BSA pools with four primer combinations following the protocol of Dong et al. (2004). All 256 *Apo*I + 2/*Taq*I + 2 primer combinations were used to generate radioactive TDFs that were visualized by autoradiography. A subset of the primer combinations was used to generate unlabeled TDFs and those were visualized by silver staining. Over 30,000 TDFs ranging in size from 40 to 600 bp were analyzed. Overall, the patterns obtained with autoradiography and silver staining were comparable but remarkably some TDFs that were detected with the radioactive primer were not visible as a band on silver stained gels and, vice versa, some clear bands on silver stained gels were absent on autoradiographs.

TDFs present in avirulent but not in virulent strains are expected to show up in two pools (or three in the case of *Avr2*) but not in the others. In fact, the pool design included internal controls, for example, an *Avr1* specific transcript should only be present in pool 1 and pool 3 whereas an *Avr4* specific transcript should be present in pool 1 and pool 4 but not in pool 2 nor pool 3. For each of the three *Avr* genes and the *Avr3b-Avr10-Avr11* locus 23 or more TDFs that behaved according to the predicted patterns were detected. In total 99 such *Avr*-associated TDFs were found, some of which were only visible by silver staining. In all cases, TDFs with opposite pattern were also found. Examples are shown in Figs. 1B–E. Although the observed BSA patterns suggest that the TDFs represent genes that are specifically expressed in either avirulent or virulent strains one should bear in mind that also polymorphisms in the *Apo*I or *Taq*I recognition site or in the two base pair extensions may result in differential cDNA-AFLP patterns.

3.2. Segregation of Avr-associated TDFs in cross 71

From previous studies in which the segregation of the avirulence phenotypes in cross 71 was analyzed, it was evident that the AVR1, AVR2 and AVR4 phenotypes behave as single dominant traits (Alfonso and Govers, 1995; van der Lee et al., 2001). AVR3b, AVR10, and AVR11 are also dominant but the genes are closely linked (van der Lee et al., 2001). The *Avr3b-Avr10-Avr11* locus might harbour three independent genes but it cannot be excluded that the locus contains a single gene that either controls other loci conferring avirulence on *R3b*, *R10*, and *R11* plants or that interacts with an uncharacterized R gene shared by *R3b*, *R10*, and *R11* plants. Many of the known avirulence factors from plant pathogens are effector proteins that are present in avirulent strains but absent, unstable or mutated in virulent strains (Westerink et al., 2004). Hence, the *Avr*-associated TDFs that were identified in the BSA may all represent candidate *Avr* genes. However, we hypothesize that from each set only TDFs derived from one transcript (or possibly three in the case of *Avr3b-Avr10-Avr11*) can represent the real *Avr* gene. To make a further selection, we performed fluorescent and silver stained cDNA-AFLP analyses on the two parental lines of cross 71 and 18

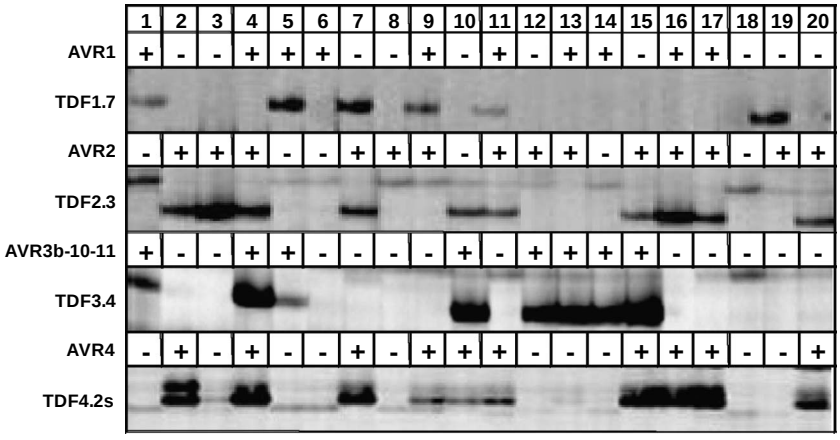

Fig. 2. cDNA-AFLP patterns showing the segregation of four *Avr*-associated TDFs in 18 F1 progeny of cross 71 (lanes 3–20). Lanes 1 and 2 show the cDNA-AFLP patterns obtained from the two parental isolates 80029 and 88133, respectively. The avirulence and virulence phenotypes of parents and progeny are indicated by + and −, respectively. TDF1.7 was obtained with primer combination A + AT/T + GA, TDF2.3 with A + AG/T + TT, TDF3.4 with A + AA/T + AG and TDF4.2s with A + TC/T + TC.

Table 2
GenBank accession numbers and AFLP codes showing the primer extensions and fragment size of 142 cloned TDFs

TDF	AFLP code	Accession number
1.1	A + AA/T + AAs169	DW010060
1.2	A + AA/T + CTs232	DW010070
1.3	A + AA/T + GGs345	DW010173
1.5	A + AT/T + ACs234	DW010078
1.6	A + AT/T + CTs71	DW010079
1.7	A + AT/T + GAs233	DW010080
1.8a	A + AT/T + TTs72	DW010081
1.8b	A + AT/T + TTs72	DW010081
1.9	A + CA/T + AGs362	DW010174
1.10	A + CA/T + GGs360	DW010172
1.11	A + CC/T + CGs152	DW010061
1.12	A + CC/T + TAs65	DW010062
1.13	A + TC/T + ACs219	DW010063
1.14	A + TC/T + ACs212	DW010064
1.15	A + TC/T + GTs179	DW010065
1.16	A + CT/T + ATs222	DW010066
1.17	A + CT/T + ATs222	DW010066
1.18	A + CT/T + CTs240	DW010067
1.19	A + GC/T + AGs225	DW010068
1.20	A + GG/T + TAs124	DW010071
1.21	A + GG/T + TTs161	DW010072
1.22	A + GT/T + CAs268	DW010073
1.23	A + TA/T + TGs132	DW010074
1.24	A + TA/T + GTs110	DW010075
1.25	A + TG/T + CGs109	DW010076
1.1s	A + TA/T + GAs150	DW010069
1.2s	A + TC/T + CAs200	DW010077
1 x .15	A + GC/T + CTs169	DW010082
1 x .21	A + GA/T + GTs137	DW010083
2.1	A + AG/T + GCs155	DW010084
2.3	A + AG/T + TTs137	DW010096
2.4	A + AA/T + AAs251	DW010097
2.5	A + AC/T + CCs117	DW010098
2.6	A + AC/T + GTs152	DW010099
2.7	A + AC/T + TGs183	DW010100
2.8	A + AT/T + CTs319	DW010180
2.9	A + CA/T + CCs117	DW010101
2.10	A + CC/T + GCs101	DW010085
2.11	A + TC/T + GCs97	DW010086
2.12	A + TC/T + GGs72	DW010087
2.13	A + TC/T + ACs97	DW010088
2.14	A + CT/T + ATs144	DW010089
2.15	A + CT/T + CCs113	DW010090
2.16	A + CG/T + GAs412	DW010175
2.17	A + CG/T + GGs401	DW010176
2.18	A + GA/T + GAs238	DW010091
2.19	A + GT/T + ATs274	DW010092
2.20	A + TA/T + CGs168	DW010094
2.21	A + TG/T + GAs399	DW010177
2.22	A + TT/T + CGs314	DW010178
2.23	A + TT/T + CGs312	DW010179
2.1s	A + CC/T + TAs210	DW010093
2.2s	A + GG/T + GCs220	DW010095
2 x .8	A + CC/T + CCs93	DW010103
2 x .39	A + TA/T + ACs346	DW010181
2 x .46	A + TG/T + GGs254	DW010102
3.1	A + AG/T + ACs153	DW010104
3.2	A + AG/T + ACs104	DW010114
3.3	A + AG/T + TTs242	DW010116
3.4	A + AA/T + AGs156	DW010117
3.6	A + AT/T + GAs160	DW010118
3.7	A + AT/T + TCs315	DW010184
3.8	A + AT/T + TCs252	DW010119
3.9	A + AT/T + TCs108	DW010120

Table 2 (*continued*)

TDF	AFLP code	Accession number
3.10	A + GC/T + CAs147	DW010105
3.11	A + GC/T + GAs250	DW010106
3.12	A + GC/T + GTs265	DW010107
3.13	A + GC/T + TAs156	DW010108
3.14	A + GA/T + CAs281	DW010109
3.15	A + GG/T + CCs164	DW010110
3.16	A + GG/T + CCs84	DW010111
3.17	A + GG/T + CCs84	DW010111
3.19	A + GT/T + TAs168	DW010112
3.20	A + TT/T + AAs347	DW010182
3.21	A + TT/T + GCs293	DW010183
3.22	A + TT/T + TTs110	DW010115
3.1s	A + TA/T + AGs90	DW010113
3 x .2	A + AG/T + ATs93	DW010126
3 x .4	A + AC/T + ACs257	DW010136
3 x .5	A + AT/T + GAs158	DW010141
3 x .6	A + CA/T + ACs163	DW010143
3 x .7	A + CA/T + CCs130	DW010144
3 x .9	A + TC/T + GGs164	DW010145
3 x .10	A + CG/T + AAs169	DW010121
3 x .11	A + CG/T + GAs215	DW010122
3 x .13	A + GC/T + ATs172	DW010123
3 x .14	A + GC/T + CCs333	DW010185
3 x .16	A + GC/T + GCs216	DW010124
3 x .17	A + GC/T + GTs270	DW010125
3 x .19	A + GA/T + CAs400	DW010186
3 x .22	A + GA/T + TCs215	DW010127
3 x .23	A + GG/T + ATs151	DW010128
3 x .24	A + GG/T + CAs167	DW010129
3 x .25	A + GG/T + GGs352	DW010187
3 x .27	A + GG/T + GGs236	DW010130
3 x .28	A + GG/T + CAs218	DW010131
3 x .29	A + GG/T + GGs193	DW010132
3 x .30	A + GG/T + GGs120	DW010133
3 x .33	A + GT/T + AGs178	DW010134
3 x .34	A + GT/T + CCs137	DW010135
3 x .35	A + GT/T + GTs381	DW010188
3 x .42	A + TA/T + GTs303	DW010189
3 x .43	A + TA/T + TTs77	DW010137
3 x .45	A + TA/T + GTs112	DW010138
3 x .47	A + TG/T + TGs257	DW010139
3 x .48	A + TG/T + AGs227	DW010140
3 x .51	A + TT/T + GTs291	DW010190
3 x .52	A + TT/T + CCs218	DW010142
4.1	A + AA/T + AGs233	DW010146
4.2	A + AA/T + GGs105	DW010155
4.3	A + AC/T + AGs139	DW010158
4.4	A + AC/T + CGs183	DW010159
4.5	A + AC/T + GTs69	DW010160
4.7	A + AT/T + AAs232	DW010161
4.8	A + TC/T + AAs123	DW010162
4.9	A + CT/T + CGs96	DW010163
4.10	A + GC/T + CCs114	DW010147
4.11	A + GA/T + AGs503	DW010191
4.12	A + GA/T + CGs238	DW010148
4.13	A + GG/T + AGs85	DW010149
4.14	A + GT/T + AGs296	DW010192
4.15	A + TA/T + GTs354	DW010193
4.16	A + TG/T + AAs240	DW010150
4.17	A + TG/T + AGs117	DW010151
4.18	A + TG/T + GGs116	DW010152
4.19	A + TT/T + CAs95	DW010153
4.20	A + TT/T + GGs117	DW010156
4.1s	A + TC/T + GAs125	DW010154
4.2s	A + TC/T + TCs144	DW010157
4.3s	A + GT/T + TTs180	DW010198

Table 2 (*continued*)

TDF	AFLP code	Accession number
4 x .3	A + AC/T + TTs74	DW010167
4 x .10	A + TC/T + TCs133	DW010164
4 x .12	A + CG/T + TAs196	DW010165
4 x .18	A + GC/T + TTs483	DW010194
4 x .20	A + GA/T + GCs187	DW010166
4 x .31	A + GT/T + ACs364	DW010195
4 x .32	A + GT/T + ACs375	DW010196
4 x .37	A + GT/T + TAs361	DW010197
4 x .38	A + GT/T + TGs216	DW010168
4 x .40	A + TA/T + CTs286	DW010169
4 x .49	A + TT/T + TAs258	DW010170
4 x .50	A + TT/T + TGs206	DW010171

F1 progeny, and screened for presence or absence of TDFs. For 25 *Avr*-associated TDFs, there was segregation in the F1 progeny, 8 of which were associated with *Avr1*, 8 with *Avr2*, 4 with *Avr3b-Avr10-Avr11* and 5 with *Avr4*. Representative patterns are shown in Fig. 2. Two of the 25 were not polymorphic in the parental lines and are thus unlikely candidates to represent an *Avr* gene. However, for four of the 25 *Avr*-associated TDFs the presence/absence pattern matched exactly with the avirulence phenotypes of the two parental lines and the 18 F1 progeny making them ideal transcriptome markers representing an *Avr* gene. Two are associated with the *Avr3b-Avr10-Avr11* locus (TDF3.1 and TDF3.4), and two with *Avr4* (TDF4.1s and TDF4.2s) (Fig. 2).

None of the *Avr1* and *Avr2* candidates cosegregated with avirulence and it is therefore unlikely that these TDFs are derived from *Avr1* or *Avr2*. Nevertheless, based on the segregation patterns of the remaining 19 TDFs we anticipate that some of them are linked to the *Avr* locus (data not shown). If the polymorphism represents a DNA polymorphism they could be used as markers for fine mapping the *Avr* regions. Alternatively, they could be used for the construction of a transcriptome map (Brugmans et al., 2002).

3.3. TDF cloning and sequencing

To enable further analysis of the TDFs we cloned the majority of the 99 *Avr*-associated TDFs and a number of TDFs with opposite pattern. TDFs were excised from gel, re-amplified, cloned, and sequenced. Based on the size of the clone insert and the presence or absence of the expected two-base primer extension in the sequence, it was concluded that 142 TDFs were successfully cloned. Overall, the success rate of cloning was over 94%. GenBank accession numbers of the cloned TDFs and AFLP codes showing the primer extensions and fragment size, are listed in Table 2. The TDF nucleotide sequences were used to design primers for RT-PCR analysis (see Section 3.4) and for similarity searches in various databases (see Section 3.5).

3.4. RT-PCR expression analysis of Avr-associated TDFs

In parallel to the segregation analysis of the TDFs in cross 71, expression of cloned TDFs was analyzed by RT-PCR. For 42 TDFs suitable primers were designed and for 38, RT-PCR products were obtained. The primer design was based on the TDF sequence itself or on the sequence of a matching *P. infestans* EST with a sequence similarity higher than 99%. The majority of the RT-PCR products could be visualized on agarose gels but for several the small size of the RT-PCR product or the occurrence of multiple bands with size differences of only a few base pairs required an electrophoresis system with a higher resolution (i.e., polyacrylamide gels). Table 3 lists the 38 TDFs including the amplicon sizes. The RT-PCR analysis included the two parental strains of cross 71 and 9 F1 progeny. Of the 38 TDFs four showed an RT-PCR expression pattern that perfectly matched the avirulence phenotypes in parents and F1 progeny, and these are the same four TDFs that matched in the segregation analysis based on cDNA-AFLP patterns: TDF3.1, TDF3.4, TDF4.1s, and TDF4.2s (Table 3). In the avirulent parent and progeny, the RT-PCR product was present and in the virulent parent and progeny it was absent. Since both RT-PCR and cDNA-AFLP give this black and white pattern it is very likely that the difference is caused by presence versus absence of mRNA and not by polymorphisms in the sequences. Hence, the genes corresponding to these TDFs seem to be regulated at the transcriptional level.

Several of the other 34 TDFs showed differential RT-PCR patterns but there was no association with the avirulence phenotypes. A substantial number, however, showed no differential expression at all. Again none of the *Avr*-associated TDFs tested by RT-PCR appeared to be a candidate for *Avr1* or *Avr2*.

3.5. Sequence similarity of TDFs with Phytophthora sequences and known sequences

Sequence similarity to known sequences may help in assigning a function to the genes from which the TDFs are derived. All TDF sequences were compared by BLAST algorithm to the NCBI GenBank and *P. infestans* EST databases with an *E* value cutoff of 1E−03. Of the 142 TDFs 56% had no match at all. A small percentage (16%) had a match in GenBank and 39% had high sequence similarity to *P. infestans* ESTs (Fig. 3A). The *P. infestans* EST database comprises over 75,000 ESTs obtained from cDNA libraries representing a broad range of growth conditions, stress responses, and developmental stages (Randall et al., 2005). It is likely that more TDFs have matching cDNA clones in the EST libraries but because many of the ESTs are only partially sequenced, matching cDNA clones may not always be recognizable. On the other hand, cDNA-AFLP is a very sensitive method and is able to detect very low abundance mRNA that may not be present in the EST database.

Table 3
RT-PCR analysis and sequence similarity of 38 *P. infestans* TDFs

TDF	*P. infestans* EST hit[a]	Amplicon size (bp)[b]	RT-PCR[c]	*P. sojae* homologue[d]	*P. ramorum* homologue[d]	SwissProt hit of *P. sojae* homologue	BLAST identity (%)	*E* value
1.1		117	—	pro135357	pro75828	RB38_HUMAN (P57729) Ras-related protein Rab-38	40	3.00E−33
1.2		190	—	pro135623	pro71960	RDPO_SCHPO (Q05654) Retrotransposable element Tf2 155 kDa protein	29	1.00E−99
1.3	CON_001_13933	304*	—	pro135300	pro75790	PTPJ_HUMAN (Q12913) Protein-tyrosine phosphatase	36	1.00E−36
1.5		202	—					
1.6		38	—					
1.7	CON_003_04202	390*	—	pro140341	pro83108			
1.14		90	—	pro143752	pro87069	RDPO_SCHPO (Q05654) Retrotransposable element Tf2 155 kDa protein	27	2.00E−66
1.22	CON_016_07340	400*	—	pro125097	pro83808	SYM_ARATH (Q9SVN5) Probable methionyl-tRNA synthetase	37	9.00E−21
2.3	CON_001_16821	191*	—	pro129917	pro73127	VTL2_MOUSE (O89116) Vesicle transport v-SNARE protein Vti1-like 2	29	4.00E−21
2.7	CON_001_30638	472*	—	pro109725	pro87143	AQP3_HUMAN (Q92482) Aquaporin 3	43	1.00E−34
2.11		59	—					
2.13		59	—					
2.15	CON_001_14541	380*	—	pro131502	pro84862	GTT2_HUMAN (P30712) Glutathione *S*-transferase theta 2	33	1.00E−20
3.1		54	Yes					
3.3		148	—	pro108156	pro39196	ENGA_RICPR (Q9ZCP6) Probable GTP-binding protein engA	24	1.00E−03
3.4		115	Yes	pro133266	pro80794	MYH3_CHICK (P02565) Myosin heavy chain, fast skeletal muscle	22	2.00E−07
3.7		282	—	pro131930	pro74150	TRHY_SHEEP (P22793) Trichohyalin	18	1.00E−11
3.8		172	—					
3.9		62	—					
3.16		52	—					
3.19	CON_002_01106	377*	—	pro133266	pro80794	MYH3_CHICK (P02565) Myosin heavy chain, fast skeletal muscle	22	2.00E−07
3.20	CON_010_06936	490*	—	pro137091	pro85962	NSB1_HUMAN (P82970) Nucleosomal binding protein 1	23	3.00E−04
3 x .7		94	—		pro80914			
3 x .11		182	—	pro143645	pro80057	MYSJ_DICDI (P54697) Myosin IJ heavy chain	33	1.00E−114
3 x .22		182	—	pro131604	pro80644	DSPP_HUMAN (Q9NZW4) Dentin sialophosphoprotein precursor	17	7.00E−09
3 x .33	CON_001_10962	236*	—	pro131005	pro85669			
4.1		191	—	pro131094	pro86402	BFR1_SCHPO (P41820) Brefeldin A resistance protein	20	6.00E−17
4.2		67	—					
4.10		80	—	pro138207	pro81288			
4.13		42	—					
4.14	CON_001_33999	271*	—	pro140951	pro72858	CATL_DROME (Q95029) Cathepsin L precursor	40	5.00E−58
4.18		71	—					
4.19	CON_001_29569	452*	—	pro131364	pro73340			
4.20		75	—	pro134550	pro74902	CSK_CHICK (P41239) Tyrosine-protein kinase	28	4.00E−18
4.1s	CON_001_33634	186*	Yes	pro109418	pro83335			
4.2s	CON_001_33634	186*	Yes	pro109418	pro83335			
4 x .3		41	—					
4 x .50	CON_014_07231	473*	—	pro138318	pro82098			

[a] *Phytophthora infestans* EST hits with *E* value < 1E−50 and identity >99% are listed.
[b] RT-PCR amplicon size was calculated based on TDF or EST sequence information; *The primers were designed on the EST sequence.
[c] 'Yes' indicates that the RT-PCR polymorphism correlates with the AVR phenotypes of the parents and 9 F1 progeny; '—' indicates no polymorphism or no correlation with the AVR phenotypes.
[d] *Phytophthora sojae* and *P. ramorum* whole genome sequences and the gene annotation at the JGI website (http://www.jgi.doe.gov/genomes) were used for analysis. Genes with BLAST *E* value less than 1E−3 were considered homologues.

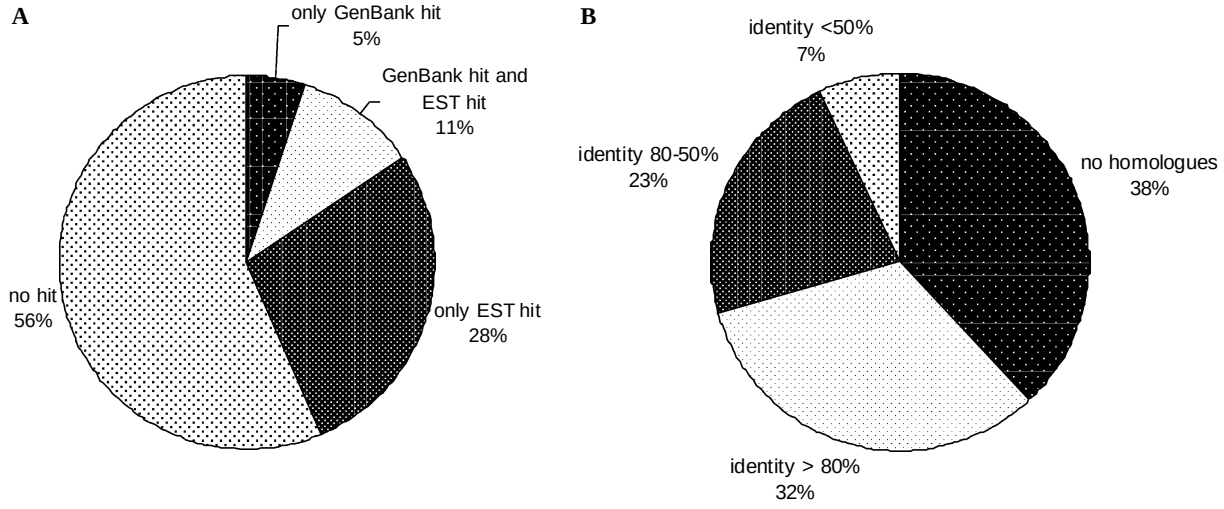

Fig. 3. Percentages of 142 cloned *P. infestans* TDFs with (A) sequence homology in GenBank and *P. infestans* EST databases and (B) homologues in *P. sojae*. Homologues were counted if the BLASTX *E* value was less than 1E−3. In the GenBank sequences, the *P. infestans* ESTs deposited in GenBank were not included.

The 142 TDFs were also BLASTed against the fully sequenced genomes of *P. sojae* and *P. ramorum*. Over one-third had no homologues in *P. sojae* and one-third had homologues with a similarity higher than 80% (Fig. 3B). The homologues in the *P. sojae* proteome that were assigned to represent the TDFs were subsequently BLASTed against the SwissProt database. A wide range of hits was found, such as proteins that function as phosphatase or kinase but also an ABC transporter, a water channel protein, and molecular motor proteins. As expected, there are also TDFs that do have a match in the *P. sojae* proteome but no hit in SwissProt. Table 3 shows the results for 38 of the 142 cloned TDFs.

TDFs represented by *P. infestans* ESTs and TDFs with *P. sojae* homologues with a variety of putative functions were taken for further data mining and bioinformatics analysis, such as in silico expression analysis, signal peptide prediction, and gene copy number of the homologues in *P. sojae* and *P. ramorum* (Table 4). In the *P. infestans* EST database (Randall et al., 2005), we analyzed the distribution of ESTs representing the TDFs over the various libraries and, based on these numbers, we predicted stage specific expression patterns and expression levels. For example, for TDF3 x .34 many ESTs are found in the germinating cysts library and zoospore library but none in a mycelium library. This indicates that the TDF3 x .34 gene is specifically expressed at a relatively high level in zoospores and germinating cysts. In contrast, for TDF2.7 there is only one EST in the database, which indicates that this aquaporin-like gene is transcribed at a low level in wall-less zoospores. Of the 55 TDFs for which we found a matching *P. infestans* EST, only 16 have ESTs in germinating cyst stages. Our screening strategy did not include a stage specific selection but since we used germinating cysts as starting material one would expect to find matching ESTs in that stage. This is true for only one-third of the TDFs confirming that cDNA-AFLP is a very sensitive method that can reveal very low abundance transcripts.

Many of the fungal and oomycete elicitors identified to date are small secreted proteins with an even number of cysteine residues that usually form disulfide bridges to stabilize the protein (van't Slot and Knogge, 2002). Another feature typical for oomycete elicitors is the RXLR motif, a motif shared by four oomycete avirulence factors that lack cysteines (Allen et al., 2004; Shan et al., 2004; Armstrong et al., 2005; Rehmany et al., 2005). For the *Avr*-associated TDFs, the presence of a signal peptide combined with a particular cysteine signature or an RXLR motif can be indicative for elicitor function. Two proteins representing TDF3 x .34 and TDF4 x .49 are predicted to be secreted by the program SignalPv2.0 (Nielsen et al., 1997; Nielsen and Krogh, 1998) and are also rich in cysteine residues. The protein represented by TDF4.1s has a homologue in *P. sojae* that is a secreted protein so we anticipate that the full length TDF4.1s protein also has a signal peptide (Table 4). These three proteins have no homology with any known protein but, interestingly, the *P. sojae* homologue of the TDF4.1s protein has an RXLR motif making TDF4.1s a promising candidate for an *Avr* gene.

With the exception of one *Phytophthora* elicitor, i.e., NIP1 (Fellbrich et al., 2002; Qutob et al., 2002), all oomycete elicitors and avirulence factors identified to date are unique for oomycetes: there are no homologous in organisms other than oomycetes. This is true for elicitins (Jiang et al., 2006) and the glycoprotein elicitor (gpe) containing pep13 (Sacks et al., 1995; Brunner et al., 2002), two protein families which are ubiquitous in the *Phytophthora* genus and have elicitor activity on a large variety of plant species. This is also true for the four ecotype- or cultivar-specific oomycete avirulence factors with the RXLR motif (Allen et al., 2004; Shan et al., 2004; Armstrong et al., 2005;

Table 4
Analysis of *Avr*-associated TDFs using data mining and bioinformatics

TDF	*P. infestans* EST hit[a]	Protein size[b]	Signal peptide[c]	Cys spacing pattern[d]	Transcripts in *P. infestans* EST database[e]	*P. sojae* homologue[f]	SwissProt BLAST hit of *P. sojae* homologue	Putative function	Phylogenetic distribution[g]	Genes in *P. sojae*[h]	Genes in *P. ramorum*[h]
1.7	CON_003_04202	nd	—		ZO(1) SP(1) MY(1)	pro140341		Uknown	Only in *Phytophthora*	1	1
2.3	CON_001_16821	nd	—		MY(1)	pro129917	VTL2_MOUSE (O89116)	Vesicle transporter	Other species	1	1
2.7	CON_001_30638	nd	—		ZO(1)	pro109725	AQP3_HUMAN (Q92482	Water channel	Other species	>10	>10
3.4		nd	—		—	pro133266	MYH3_CHICK (P02565)	Cyto-skeleton related	Other species	1	1
3 x .34	CON_020_07430	159	SP	C-n20-C-n9-C-n8-C	ZO(13) CY(7)SP(1)	pro138143		Unknown secreted protein	Only in *Phytophthora*	8	3
4.1s	CON_001_33634	>150		None	MY(1)	pro109418		Unknown secreted protein	Only in *Phytophthora*	2	1
4.1		nd	—		—	pro131094	BFR1_SCHPO (P41820)	ABC transporter	Other species	>10	>10
4.20		nd	—		—	pro134550	CSK_CHICK (P41239)	Kinase	Other species	1	1
4 x .49	CON_011_07076	300	SP	C-n28-C-n3-C-n10-C-n17-C-n23-C-n103-C	MY(8)	pro144423		Unknown secreted protein	Only in *Phytophthora*	>10	>10

[a] *Phytophthora infestans* EST hits with *E* value < 1E−50 and identity > 99% are listed.
[b] 'nd' indicates that the protein sequence is incomplete.
[c] SP indicates that a signal peptide is predicted at the N-terminus by the program SignalP (Nielsen et al., 1997; Nielsen and Krogh, 1998).
[d] The proteins with signal peptide were used for cysteine spacing analysis.
[e] The tissue types from which the EST libraries are derived are zoospores (ZO), germinated cysts (CY), sporangia (SP), and mycelia (MY). The numbers in brackets indicate the number of ESTs present in the various libraries (Randall et al., 2005).
[f] *Phytophthora sojae* whole genome sequences and the gene annotation at the JGI website (http://www.jgi.doe.gov/genomes) were used for analysis. Genes with BLAST *E* value less than 1E−3 were considered homologues.
[g] Homologues in species other than *Phytophthora* were considered as homologues when the BLAST *E* value was less than 1E−3 and the similarity >30%.
[h] *Phytophthora sojae* and *P. ramorum* whole genome sequences and gene annotation at the JGI website (http://www.jgi.doe.gov/genomes) were used for analysis. Genes with BLAST similarity higher than 50% were considered to be members of the same gene family. Numbers indicate the size of the family.

Rehmany et al., 2005). In contrast to elicitin genes and *gpe* genes, however, none of these four avirulence genes belongs to a conserved gene family. Apart from the conserved RXLR motif they all show high sequence divergence with their homologues in other species and this may be a hallmark for host- or cultivar-specific avirulence genes. To evaluate the likelihood that the TDFs are derived from *Avr* genes we analyzed the phylogenetic distribution and we investigated whether the cloned TDFs belong to a gene family. Of the 88 TDFs that have homologues in *P. sojae* and *P. ramorum*, 25 seem to be unique for *Phytophthora*. Noticeably, the two secreted proteins listed in Table 4 only occur in *Phytophthora*. The homologues of TDF 4 x .49 form a large gene family with over 10 members in both, *P. sojae* and *P. ramorum*, and those of TDF3 x .34 appear to form a larger family in *P. sojae* than in *P. ramorum*. TDF4.1s has only two weak homologues (BLAST identity $< 40\%$) in *P. sojae* and one weak homologue in *P. ramorum*, which suggests that this gene is of high sequence divergence among *Phytophthora* species.

4. Conclusions

In this study, we demonstrate that combining a bulked segregant analysis strategy with a highly efficient transcriptional profiling method can be very effective in selecting *Avr*-associated transcripts. We focused on four *Avr* genes and for two of these we found TDFs that fulfill all criteria that make the TDF a likely *Avr* candidate. First of all, the TDFs occurred in germinating cysts, a preinfection stage in which an *Avr* gene is most likely to be expressed. Secondly, the TDFs were present in pools consisting of strains having an AVR phenotype but were absent in pools consisting of virulent strains (avr phenotype). Thirdly, segregation of the TDFs in F1 progeny correlated entirely with segregation of the AVR/avr phenotypes and, fourthly, RT-PCR confirmed the *Avr* associated segregation in the F1 progeny.

The two TDFs that were assigned as candidates for *Avr4*, TDF4.1s and TDF4.2s, appear to match to the same *P. infestans* EST contig but the deduced protein is an unknown protein. Since the homologues in *P. sojae* and *P. ramorum* are very divergent the protein seems to be unique for *P. infestans*. The *P. sojae* homologue though, has all the hallmarks of the family of RXLR proteins: a signal peptide, an RXRL motif and high sequence divergence with the other family members. All four oomycete *Avr* genes identified so far, *P. sojae Avr1b*-1 (Shan et al., 2004), *P. infestans Avr3a* (Armstrong et al., 2005) and the two *Hyaloperonospora parasitica* ecotype-specific *Avr* genes, *ATR13* and *ATR1* (Allen et al., 2004; Rehmany et al., 2005), belong to this RXLR super family and sequencing of the full-length gene represented by TDF4.1s and TDF4.2s showed that this gene is also an RXLR family member (Pieter van Poppel, J.G and F.G, unpublished). Hence, the TDF4.1/TDF4.2 gene is a likely candidate for *Avr4*. Functional characterization is in progress.

The two TDFs that associate with the *Avr3b-Avr10-Avr11* locus, TDF3.1 and TDF3.4 are more mysterious. They fulfill all selection criteria but there are no matching *P. infestans* ESTs, and only TDF3.4 has an obvious homologue in the *P. sojae* proteome. These TDFs have recently been used as probes and markers to zoom in on the *Avr3b-Avr10-Avr11* locus and physical mapping showed that *Avr3b-Avr10-Avr11* linked AFLP and the TDFs are located on the same BAC contig (R.H.Y.J., Rob Weide and F.G., unpublished).

For *Avr1* and *Avr2* no candidates were recovered. Previously, we used the same mapping population and a similar pooling strategy to identify AFLP markers (van der Lee et al., 2001). Also in that study the selection for *Avr4* and *Avr3b-Avr10-Avr11* linked markers was much more successful. For *Avr2* this could be explained by the fact that it was not included in the BSA, only random markers were selected. In the present study, the pooling for *Avr2* was not optimal which may have caused a lower efficiency. For *Avr1*, however, it is unclear why the screening was unsuccessful. In both studies the BSA screening resulted in the highest number of candidates for *Avr1* but just one AFLP marker (van der Lee et al., 2001) and none of the TDFs passed the next, more stringent selection steps.

Previous studies in *P. infestans* have demonstrated that cDNA-AFLP is a powerful technique that complements other expression profiling approaches, such as EST sequencing (Avrova et al., 2003; Dong et al., 2004). Here, we showed that cDNA-AFLP can be combined with BSA to find transcripts associated with particular phenotypes. Since the *Avr*-linked genetic markers and the *Avr*-associated TDFs were generated from the same mapping population we can now integrate the various gene discovery approaches to identify *P. infestans Avr* genes.

Acknowledgments

We are grateful to Jacqueline Gerdsen for help with cloning, Drs. Qin Ling and Xiaowu Wang for useful suggestions and discussions, Prof. Dongyu Qu for support and encouragement, and Rob Weide for critically reading the manuscript. This work was financially supported by the Wageningen University Interdisciplinary Research and Education Fund (INREF) in the framework of the joint WU-CAAS PhD training programme and by NWO-Aspasia Grant 015.000.057. The authors acknowledge Syngenta for access to the Syngenta *Phytophthora* Consortium EST Database prior to public release and the DOE Joint Genome Institute for making *P. sojae* and *P. ramorum* sequence data publicly available.

References

Agrios, G.N., 1997. Plant Pathology, fourth ed. Academic Press, USA.

Alfonso, C., Govers, F., 1995. A search for determinants of race-specificity in the *Phytophthora infestans*-potato pathosystem. In: Dowley, L.J.,

Bannon, E., Cooke, L., Keane, T., O'Sullivan, E. (Eds.), *Phytophthora infestans*. Boole Press Ltd., Dublin, pp. 107–115.

Allen, R.L., Bittner-Eddy, P.D., Grenville-Briggs, L.J., Meitz, J.C., Rehmany, A.P., Rose, L.E., Beynon, J.L., 2004. Host–parasite coevolutionary conflict between Arabidopsis and downy mildew. Science 306, 1957–1960.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Armstrong, M.R., Whisson, S.C., Pritchard, L., Bos, J.I., Venter, E., Avrova, A.O., Rehmany, A.P., Bohme, U., Brooks, K., Cherevach, I., Hamlin, N., White, B., Fraser, A., Lord, A., Quail, M.A., Churcher, C., Hall, N., Berriman, M., Huang, S., Kamoun, S., Beynon, J.L., Birch, P.R., 2005. An ancestral oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is recognized in the host cytoplasm. Proc. Natl. Acad. Sci. USA 102, 7766–7771.

Avrova, A.O., Venter, E., Birch, P.R.J., Whisson, S.C., 2003. Profiling and quantifying differential gene transcription in *Phytophthora infestans* prior to and during the early stages of potato infection. Fungal Genet. Biol. 40, 4–14.

Ballvora, A., Ercolano, M.R., Weiss, J., Meksem, K., Bormann, C.A., Oberhagemann, P., Salamini, F., Gebhardt, C., 2002. The *R*1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. Plant J. 30, 361–371.

Brugmans, B., Fernandez del Carmen, A., Bachem, C.W., van Os, H., van Eck, H.J., Visser, R.G., 2002. A novel method for the construction of genome wide transcriptome maps. Plant J. 31, 211–222.

Brunner, F., Rosahl, S., Lee, J., Rudd, J.J., Geiler, C., Kauppinen, S., Rasmussen, G., Scheel, D., Nurnberger, T., 2002. Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora* transglutaminases. EMBO J. 21, 6681–6688.

Caten, C.E., Jinks, J.L., 1968. Spontaneous variability of single isolates of *Phytophthora infestans*. Can. J. Bot. 46, 329–348.

Collmer, A., 1998. Determinants of pathogenicity and avirulence in plant pathogenic bacteria. Curr. Opin. Plant Biol. 1, 329–335.

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

Dong, W.B., Latijnhouwers, M., Jiang, R.H.Y., Meijer, H.J.G., Govers, F., 2004. Downstream targets of the *Phytophthora infestans* G alpha subunit PiGPA1 revealed by cDNA-AFLP. Mol. Plant Pathol. 5, 483–494.

Drenth, A., Janssen, E.M., Govers, F., 1995. Formation and survival of oospores of *Phytophthora infestans* under natural conditions. Plant Pathol. 44, 86–94.

Fellbrich, G., Romanski, A., Varet, A., Blume, B., Brunner, F., Engelhardt, S., Felix, G., Kemmerling, B., Krzymowska, M., Nurnberger, T., 2002. NPP1, a *Phytophthora* associated trigger of plant defense in parsley and *Arabidopsis*. Plant J. 32, 375–390.

Flor, H.H., 1942. Inheritance of pathogenicity of *Melampsora lini*. Phytopathology 32, 653–669.

Gao, H., Narayanan, N.N., Ellison, L., Bhattacharyya, M.K., 2005. Two classes of highly similar coiled coil-nucleotide binding-leucine rich repeat genes isolated from the Rps1-k locus encode *Phytophthora* resistance in soybean. Mol. Plant Microbe Interact. 18, 1035–1045.

Govers, F., Latijnhouwers, M., 2004. Late blight. In: Goodman, R.M. (Ed.), Encyclopedia of Plant and Crop Science. Marcel Dekker, New York, pp. 1–5. doi:10.1081/E-EPCS-120019918.

Huang, S., van der Vossen, E.A., Kuang, H., Vleeshouwers, V.G., Zhang, N., Borm, T.J., van Eck, H.J., Baker, B., Jacobsen, E., Visser, R.G., 2005. Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. Plant J. 42, 251–261.

Huang, S.W., Vleeshouwers, V., Werij, J.S., Hutten, R.C.B., van Eck, H.J., Visser, R.G.F., Jacobsen, E., 2004. The *R3* resistance to *Phytophthora infestans* in potato is conferred by two closely linked *R* genes with distinct specificities. Mol. Plant Microbe Interact. 17, 428–435.

Innes, R.W., 2004. Guarding the goods. New insights into the central alarm system of plants. Plant Physiol. 135, 695–701.

Jiang, R.H.Y., Tyler, B.M., Whisson, S.C., Hardham, A.R., Govers, F., 2006. Ancient origin of elicitin gene clusters in *Phytophthora* genomes. Mol. Biol. Evol. in press, doi:10.1093/molbev/msj039.

Kamoun, S., Hraber, P., Sobral, B., Nuss, D., Govers, F., 1999. Initial assessment of gene diversity for the oomycete pathogen *Phytophthora infestans* based on expressed sequences. Fungal Genet. Biol. 28, 94–106.

Latijnhouwers, M., de Wit, P.J., Govers, F., 2003. Oomycetes and fungi: similar weaponry to attack plants. Trends Microbiol. 11, 462–469.

Luderer, R., Joosten, M.H.A.J., 2001. Avirulence proteins of plant pathogens: determinants of victory and defeat. Mol. Plant Pathol. 2, 355–364.

Michelmore, R.W., Paran, I., Kesseli, R.V., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88, 9828–9832.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1–6.

Nielsen, H., Krogh, A., 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 122–130.

Qutob, D., Kamoun, S., Gijzen, M., 2002. Expression of a *Phytophthora sojae* necrosis-inducing protein occurs during transition from biotrophy to necrotrophy. Plant J. 32, 361–373.

Randall, T.A., Dwyer, R.A., Huitema, E., Beyer, K., Cvitanich, C., Kelkar, H., Fong, A.M., Gates, K., Roberts, S., Yatzkan, E., Gaffney, T., Law, M., Testa, A., Torto-Alalibo, T., Zhang, M., Zheng, L., Mueller, E., Windass, J., Binder, A., Birch, P.R., Gisi, U., Govers, F., Gow, N.A., Mauch, F., van West, P., Waugh, M.E., Yu, J., Boller, T., Kamoun, S., Lam, S.T., Judelson, H.S., 2005. Large-scale gene discovery in the oomycete *Phytophthora infestans* reveals likely components of phytopathogenicity shared with true fungi. Mol. Plant Microbe Interact. 18, 229–243.

Rehmany, A.P., Gordon, A., Rose, L.E., Allen, R.L., Armstrong, M.R., Whisson, S.C., Kamoun, S., Tyler, B.M., Birch, P.R., Beynon, J.L., 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. Plant Cell 17, 1839–1850.

Rooney, H.C., van't Klooster, J.W., van der Hoorn, R.A., Joosten, M.H., Jones, J.D., de Wit, P.J., 2005. *Cladosporium* Avr2 inhibits tomato Rcr3 protease required for *Cf-2*-dependent disease resistance. Science 308, 1783–1786.

Sacks, W., Nurnberger, T., Hahlbrock, K., Scheel, D., 1995. Molecular characterization of nucleotide sequences encoding the extracellular glycoprotein elicitor from *Phytophthora megasperma*. Mol. Gen. Genet. 246, 45–55.

Shan, W., Cao, M., Leung, D., Tyler, B.M., 2004. The *Avr1b* locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene *Rps1b*. Mol. Plant Microbe Interact. 17, 394–403.

Tyler, B.M., 2001. Genetics and genomics of the oomycete host interface. Trends Genet. 17, 611–614.

Tyler, B.M., 2002. Molecular basis of recognition between *Phytophthora* pathogens and their hosts. Ann. Rev. Phytopathol. 40, 137–167.

van den Ackerveken, G., Bonas, U., 1997. Bacterial avirulence proteins as triggers of plant disease resistance. Trends Microbiol. 5, 394–398.

van der Lee, T., De Witte, I., Drenth, A., Alfonso, C., Govers, F., 1997. AFLP linkage map of the oomycete *Phytophthora infestans*. Fungal Genet. Biol. 21, 278–291.

van der Lee, T., Robold, A., Testa, A., van't Klooster, J.W., Govers, F., 2001. Mapping of avirulence genes in *Phytophthora infestans* with amplified fragment length polymorphism markers selected by bulked segregant analysis. Genetics 157, 949–956.

van der Lee, T., Testa, A., Robold, A., van 't Klooster, J., Govers, F., 2004. High-density genetic linkage maps of *Phytophthora infestans* reveal trisomic progeny and chromosomal rearrangements. Genetics 167, 1643–1661.

van der Vossen, E., Sikkema, A., Hekkert, B.L., Gros, J., Stevens, P., Muskens, M., Wouters, D., Pereira, A., Stiekema, W., Allefs, S., 2003. An ancient *R* gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. Plant J. 36, 867–882.

van't Slot, K.A.E., Knogge, W., 2002. A dual role for microbial pathogen-derived effector proteins in plant disease and resistance. Crit. Rev. Plant Sci. 21, 229–271.

Westerink, N., Joosten, M.H.A.J., de Wit, P.J.G.M., 2004. Fungal (A) virulence factors at the crossroads of disease susceptibility and resistance. In: Punja, Z. (Ed.), Fungal Disease Resistance in Plants. Biochemistry, Molecular Biology and Genetic Engineering. Haworth Press, pp. 93–127.

Whisson, S.C., van derLee, T., Bryan, G.J., Waugh, R., Govers, F., Birch, P.R.J., 2001. Physical mapping across an avirulence locus of *Phytophthora infestans* using a highly representative, large-insert bacterial artificial chromosome library. Mol. Genet. Genomics 266, 289–295.

White, F.F., Yang, B., Johnson, L.B., 2000. Prospects for understanding avirulence gene function. Curr. Opin. Plant Biol. 3, 291–298.

Young, N.D., 2000. The genetic architecture of resistance. Curr. Opin. Plant Biol. 3, 285–290.