### ORIGINAL PAPER

### The *Solanum pimpinellifolium Cf-ECP1* and *Cf-ECP4* genes for resistance to *Cladosporium fulvum* are located at the *Milky Way* locus on the short arm of chromosome 1

Eleni Soumpourou · Michael Iakovidis · Laetitia Chartrain · Verity Lyall · Colwyn M. Thomas

Received: 27 June 2007 / Accepted: 24 August 2007 / Published online: 15 September 2007 © Springer-Verlag 2007

Abstract The interaction between tomato and the leaf mould pathogen Cladosporium fulvum is an excellent model to study gene-for-gene interactions and plant disease resistance gene evolution. Most Cf genes were introgressed into cultivated tomato (Solanum lycopersicum) from wild relatives such as S. pimpinellifolium and novel Cf-ECP genes were recently identified in this species. Our objective is to isolate Cf-ECP1, Cf-ECP2, Cf-ECP4 and Cf-ECP5 to increase our understanding of Cf gene evolution, and the molecular basis for recognition specificity in Cf proteins. The map locations of Cf-ECP2 and Cf-ECP5 have been reported previously and we report here that Cf-ECP1 and Cf-ECP4 map to a different locus on the short arm of chromosome 1. The analysis of selected recombinants and allelism tests showed both genes are located at Milky Way together with Cf-9 and Cf-4. Our results emphasise the importance of this locus in generating novel Cf genes for resistance to C. fulvum. Candidate genes for Cf-ECP1 and Cf-ECP4 were also identified by DNA gel blot analysis of

Communicated by C. Gebhardt.

Eleni Soumpourou and Michael Iakovidis contributed equally to this work.

E. Soumpourou · M. Iakovidis · L. Chartrain · V. Lyall · C. M. Thomas (⊠) School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK e-mail: colwyn.thomas@uea.ac.uk

L. Chartrain Department of Disease and Stress Biology, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK bulked segregant pools. In addition, we generated functional cassettes for expression of the *C. fulvum* ECP1, ECP2, ECP4 and ECP5 proteins using recombinant *Potato Virus X*, and three *ECPs* were also expressed in stable transformed plants. Using marker-assisted selection we have also identified recombinants containing *Cf-ECP1*, *Cf-ECP2*, *Cf-ECP4* or *Cf-ECP5 in cis* with a linked T-DNA carrying the non-autonomous *Zea mays* transposon *Dissociation*. Using these resources it should now be possible to isolate all four *Cf-ECPs* using transposon tagging, or a candidate gene strategy.

### Introduction

The interaction between cultivated tomato (Solanum lycopersicum) and the leaf mould pathogen Cladosporium fulvum is an excellent model to study gene-for-gene interactions and plant disease resistance gene evolution (reviewed in Rivas and Thomas 2005). Resistance genes like Cf-9 and Cf-4 are members of a complex genetic locus on the short arm of chromosome 1 that were originally identified in the wild tomato species S. pimpinellifolium and S. hirsutum, and bred into cultivated tomato (Stevens and Rick 1988; Rivas and Thomas 2005). All Cf genes characterized to date encode type I trans-membrane extracellular leucine-rich repeat glycoproteins, which lack an obvious cytoplasmic signalling domain (Rivas and Thomas 2005). Cf proteins activate a hypersensitive response (HR) upon recognition of genetically defined C. fulvum avirulence (Avr) proteins, but the molecular mechanism of Avr protein perception, and activation of the HR, have yet to be determined (Rivas and Thomas 2005).

Several *Cf* genes from the *Milky Way (MW)* locus on the short arm of tomato chromosome 1, including *Cf-4*, *Hcr9-4E*,

*Cf-9* and *Hcr9-9B*, have been cloned and extensively characterised (Jones et al. 1994; Thomas et al. 1997; Takken et al. 1999; Panter et al. 2002) and these studies form the basis for our current understanding of *Cf* gene evolution. The *MW* locus is composed of variable numbers of tandemly duplicated homologues of *Cladosporium* resistance gene *Cf-9* (*Hcr9s*) and one mechanism that generates *Cf* gene variation is sequence exchange between *Hcr9* paralogues (Parniske et al. 1997; Kruijt et al. 2004). The isolation and characterisation of additional *Cf* genes at *MW* would significantly increase our understanding of the organisation and evolution of *Cf* genes at *MW*, together with the molecular basis for Cf protein recognition specificity.

C. fulvum extracellular proteins (ECPs) are secreted into the leaf apoplast during infection and they share several features with C. fulvum Avr proteins, including their low molecular weight. The mature processed forms of ECPs contain an even number of cysteine residues that appear to function in disulphide bridge formation (Van Den Ackerveken et al. 1993; Lauge et al. 2000) but some of these are dispensible, at least in terms of their necrosis-inducing activity (Luderer et al. 2002). Functional analysis of a limited number of S. pimpinellifolium accessions was sufficient to identify five novel Cf genes (Cf-ECP1, Cf-ECP2, Cf-ECP3, Cf-ECP4 and Cf-ECP5) that trigger an HR in response to the C. fulvum ECP1, ECP2, ECP3, ECP4 and ECP5 proteins (Lauge et al. 1998, 2000). Genetic mapping showed Cf-ECP2 and Cf-ECP3 defined a new complex locus for C. fulvum resistance at Orion (OR) on the short arm of chromosome 1 (Haanstra et al. 1999; Yuan et al. 2002) and the mapping of Cf-ECP5 also defined a new locus, located 3 cM proximal to MW, which was designated Aurora (AU; Haanstra et al. 2000). Solanum germplasm is therefore an important source of novel Cf genes (Rivas and Thomas 2005).

Mapping of the S. pimpinellifolium Cf-ECP1 and Cf-ECP4 genes has not been reported previously and we show here, by molecular analysis of selected recombinants and allelism tests, that both genes are located at MW along with Cf-9 and Cf-4. We used DNA gel blots to identify candidate Hcr9s for Cf-ECP1 and Cf-ECP4. As a preliminary step towards gene isolation by transposon tagging we generated lines containing Cf-ECP1, Cf-ECP2, Cf-ECP4 or Cf-ECP5 in cis with a linked T-DNA carrying the non-autonomous Zea mays transposon Dissociation (Ds) (Rommens et al. 1992). We also constructed cassettes for the in planta expression of ECP1, ECP2, ECP4 and ECP5 proteins using Agrobacterium-mediated delivery of recombinant Potato Virus X (PVX), or stable transformed plants. Together, these resources will facilitate the isolation and characterisation of all four Cf-ECP genes using a candidate gene approach, or transposon tagging with Ds.

#### Materials and methods

### Tomato stocks

Cf-ECP stocks were obtained from Dr. MHAJ Joosten and Dr. P Lindhout at the University of Wageningen, The Netherlands; S. pimpinellifolium containing Cf-ECP1 (LA1547) and Cf-ECP4 (LA1683); S. lycopersicum containing the introgressed S. pimpinellifolium genes Cf-ECP2 (Ontario 7518) and Cf-ECP5 (G1.1161) as reported by Haanstra et al. (1999, 2000). Other S. lycopersicum lines were supplied by Matthew Smoker at The Sainsbury Laboratory, UK; S. lycopersicum var. Moneymaker (Cf0), line FT33 (Rommens et al. 1992), and line M18, which contains a Ds-tagged allele of Cf-9 (Cf-9:Ds<sup>M18</sup>; Jones et al. 1994).

### ECP1, ECP2, ECP4 and ECP5 expression constructs

Oligonucleotides were obtained from MWG BIOTECH and are available on request. Sequences encoding *C. fulvum* ECP1, ECP2, ECP4 and ECP5 were constructed as described previously (Thomas et al. 1997). Each gene was amplified in two parts from *C. fulvum* race five genomic DNA. Complete *ECP* coding sequences were reconstituted using overlap PCR. Oligonucleotides to the 3' end of each reading frame included a *Bam*HI sequence for cloning into vector pSPNae (Thomas et al. 1997). All four sequences were ligated into *NaeI/Bam*HI digested pSP-Nae to generate in-frame fusions of each *ECP* with sequences encoding the *Nicotiana tabacum* PR1A signal peptide.

PR1A:ECP fusions were cloned as ClaI/BamHI fragments into pSLJ4K1 (Thomas et al. 1997) fusing the PR1A:ECP sequences to the cauliflower mosaic virus 35S promoter, and the Agrobacterium tumefaciens nopaline synthase (nos) transcription terminator. PR1A:ECP sequences were excised as EcoRI/HindIII fragments and cloned into pSLJ7291 (Jones et al. 1992) for transformation of S. lycopersicum. PR1A:ECP sequences were also cloned as ClaI/SalI fragments into pGR107, which contains a cDNA copy of PVX. Recombinant ECP clones were transferred into A. tumefaciens GV3101 by electroporation.

### DNA sequence analysis

Sequences of *PR1A:ECP* clones were confirmed using an ABI BigDye sequencing kit (Applied Biosystems). Reaction products were electrophoresed on an Applied Biosystems 377 automated sequencer and sequences were assembled using UNIX-based versions of the Staden Program Package.

#### Generation of tomato transformants

Binary vector plasmids were mobilized from *Escherichia coli* DH5 $\alpha$  into *A. tumefaciens* LBA4404 using tri-parental mating, as described by Jones et al. (1992). The transformation of tomato cotyledons (*S. lycopersicum* var. Money-maker), and plant regeneration procedures were performed essentially as described by Dixon et al. (1996).

### Delivery of recombinant PVX using Agrobacterium infiltration

A. tumefaciens GV3101 clones containing PR1A:ECPs were grown to saturation at 28°C in liquid culture containing 50 µg/ml kanamycin. Bacteria were pelleted at 5,000g for 10 min at 15°C and resuspended in two times the original volume of a solution containing  $1 \times$  Muraschige and Skoog salts (Duchefa), 2% w/v sucrose and 150 µM acetosyringone. Bacteria were incubated at room temperature for 2–3 h and then infiltrated into tomato cotyledons.

## Construction of Cf-ECP1 and Cf-ECP4 bulked segregant pools

Bulked segregant pools from Cf-ECP1 × Cf0 ( $F_21E$ ) and Cf-ECP4 × Cf0  $F_2$  populations ( $F_24E$ ) were constructed from c.80  $F_2$  seedlings. A single cotyledon was harvested from all  $F_2s$  in each population to constitute the *Cf-ECP/+* bulk. Each  $F_2$  population was then infected with PVX:*ECP1* or PVX:*ECP4* and scored for systemic necrosis at 14 days post-infection. Plant material from wildtype seedlings in each population was pooled and used to prepare DNA for the +/+ bulk.

### Nucleic acid preparations

DNA was prepared from Cf-ECP lines and for bulked segregant analysis as follows. Cotyledons or young leaves (2.5 g) were ground in liquid N<sub>2</sub> and transferred to a tube containing 5 ml phenol-chloroform-isoamyl alcohol (25:24:1) and 7.5 ml of NTES buffer (100 mM NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA and 1% w/v SDS). Samples were centrifuged at 8,000 rpm in a Sorvall SS34 rotor at 4°C for 10 min. The supernatant was adjusted to 150 mM NaAc pH 6 and nucleic acids were precipitated by adding an equal volume of isopropanol. Samples were centrifuged at 8,000 rpm in an SS34 rotor for 15 min at 4°C. Nucleic acids were resuspended in sterile water and LiCl was added to the supernatant to 2 M, incubated on ice for 1 h and centrifuged at 8,000 rpm in an SS34 rotor for 15 min at 4°C. DNA was ethanol precipitated from the supernatant.

### DNA gel blot analysis

Two to 5 µg of tomato DNA was digested at 37°C for 16 h, extracted with phenol–chloroform and ethanol precipitated. DNA was electrophoresed for 20 h at 2.5 v/cm in 0.8% w/v agarose gels in a vertical gel apparatus using 40 mM Tris–acetate, pH 7.9, 1 mM EDTA as running buffer. Nucleic acids were transferred to Hybond-N membrane (Amersham) and cross-linked by irradiation with UV light. A 1.5 kbp probe was PCR amplified from a *Cf-9* cDNA clone (Jones et al. 1994) using primers Cf95' (AAAGCAAACA TTTCTTGATTTCTT) and Cf9F2 (GCTTGGAATATGA CCTTCAAATCTG). Filters were hybridised with <sup>32</sup>P-labelled probes at 65°C for 16–20 h, and washed four times for 15 min in 2 × SSC (1× SSC is 0.15M sodium chloride, 0.015 M sodium citrate) and 1% w/v SDS at 65°C.

### Results

Construction of ECP expression cassettes and functional analysis using recombinant PVX

Genomic or cDNA clones encoding C. fulvum ECP1, ECP2, ECP4 and ECP5 were isolated previously based upon their partial amino acid sequences (Van Den Ackerveken et al. 1993; Lauge et al. 2000). Genomic sequences of ECP4 and ECP5 have not been reported and these were amplified from C. fulvum race 5 DNA using primers to the reported cDNA sequences (Lauge et al. 2000). Like ECP2, the coding regions of ECP4 and ECP5 are interrupted by a single intron (Fig. 1). ECP cassettes encoding the mature secreted forms of each ECP were generated by overlap PCR and fused in-frame to sequences encoding the N. tabacum PR1A signal peptide, which can target proteins to the leaf apoplast in planta. The four PR1A:ECP cassettes were also cloned into a binary vector containing a cDNA copy of PVX for Agrobacterium-mediated delivery of C. fulvum Avr and ECP proteins (Lauge et al. 2000; Luderer et al. 2002).

To determine if *PR1A:ECP* sequences can induce a *Cf*-dependent HR suspensions of *A. tumefaciens* were infiltrated into a range of *Cf-ECP* genotypes (Fig. 2). In most cases recombinant PVX only resulted in symptoms characteristic of wildtype PVX infection (Fig. 2a). However, each *PVX:ECP* induced necrotic lesions in systemically infected leaves of seedlings expressing its cognate *Cf-ECP* (Fig. 2). Leaf epinasty was followed by development of necrotic lesions that spread systemically, resulting in seedling death. These features are similar to other reports of PVX-mediated expression of *C. fulvum* ECP and AVR proteins in tomato (Hammond-Kosack et al. 1995; Thomas et al. 1997; Lauge

Fig. 1 Genomic sequences of C. fulvum ECP4 and ECP5. The coding regions of ECP4 (a) and ECP5 (b) were amplified using primers to the 5' and 3' ends of their corresponding cDNAs. Intron sequences are shown underlined and each exhibits the characteristic 5' donor and 3' acceptor sequences of eukaryotic introns. The amino acid sequence of each protein is shown above the nucleotide sequence. The predicted proteolytic cleavage sites of each signal peptide sequence (Lauge et al. 2000) is indicated with an arrowhead, and the mature secreted protein is shown *italicised* and highlighted in grey

																	۲																				
м	Q	Ι	₽	т	s	L	L	A	L	т	A	т	L	т	F	A	V	A .	DP	s	F	R	F	s	E	Y	G	N	G	D	T	T	C	S I	P	DN	L
ATO	CA	AAT	TCC	AAC	CTC	CCTC	CTO	CGC	CCT	GAC	CGC	CAC	CCT	GAC	CTT	CGCC	GTC	GCC	GACCO	CCTO	CCTI	CCG	CTTO	CAG	CGAG	TAC	GGG	AAC	CGGG	CGAC	ACC	LACC.	TGO	TCC	CCG	GACA	ACCI
			10			20				30			40			50			00			10			00							100				·	
I	S	P	S	V	S	I	D	D	G	F	G	G	C	K	8	F	P .	A	N T	N	S	V	8	Y	A	L	L	E	G	S	R	A	G	C	I	VK	L
ATC	:AG	ccc	CIC	AGT	crc	CATC	GAG	CGA	cee	CIT	CGG	scee	CTG	CAA	ACA	STTC	CCA	IGCC.	AACA	CGA	ACTO	CGI	GCA	ATA	Cece	CIG	CIG	GAG	GGG	STCA	AGG	JGCG	GGI	TGC	ATT	GTCA	AGCI
			13	0		14	0			150			16	0		17	0		18	0		19	0		20	0		2	210			220			23	0	
F	T	N	N	а	с	N	E	s	а	D	S	E																		- i	н	v	P	A	v	GE	T
TTO	AC	GAA	CAA	CGC	TTG	CAAT	GAJ	ATC	GGC	TGA	TTC	CGGA	GTA	CGC	CAA	AGTG	CAA	GAT	GACG	ACCO	ссал	AAC	rcc/	AT	GACT	AAC	TT	GTO	GC	<u>rag</u> a	CAT	FGTT	CC7	GCA	GTC	GGAG	AGAC
			25	0		26	0			270			28	0		29	0		304	0		31	0		32	0		3	330			340			35	0	
K	с	8	₽	P	I	G	G	G	W	0	s	v	0	v	т	с	P	*																			
AAG	TG	CCA	GCC	TCC	GAT	TGGT	GG	rgg	ATG	GCA	GAG	TGT	GCA	GGT	CAC	TGC	CCG	TAA																			
			31	0		38	0			390			40	0		41	.0																				
в																																					
M	N	т	F	т	L	L	L	А	т	L	р	L	L	A	Y	G	R	G	DN	K	P	G	8	F	N	Y	I										
er.																																					
ATC	AA	CAC	TTT	TAC	TCT	ACTC	CTO	GGC	CAC	ACT	CCC	CACI	GCT	GGC	ATA	rggc	AGG	GGC	GACA	ATAJ	AGCC	CGG	CCA	FTT	CAAT	TAC	CAT	GT/	AG	TGGG	TC	TCG	ACC	CTG	GAT	CTGC	CCA
ATO	SAA	CAC	10	TAC	TCT	ACTC 20	CTO	GGC	CAC	ACT 30	ccc	CACI	GCT 40	GGC	ATA	reec 50	AGG	GGC	GACA 60	ATAJ	AGCC	CGG 70	CCAG	FTT	CAAT 80	TAC	TTA	T <u>GT</u>	AG 0	reee	TC	100	ACC	CTG	<u>GAT</u> 11	CTGC 0	CCAT
AT	SAA CCA	ATO	10	TAC TTT	TCT.	ACTC 20 ACAA	CT	GC	CAC	ACT 30	AGA	ACT	GCT 40	GGC	TGC	rggc 50	AGG	GGC	GACAI 60 CGCG	ATA)	AGCO	CGG 70 TAG	CAN	ACA	CAAT 80 TGT	GAG	:ATI	TC2	AGC	CAT	CG	TCG 100	AC	CTG	GAT 11	CTGC 0 GGTC	GTTO
ATO	GAA	ATO	TTT 10 TCC 13	TAC TTT 0	GTC	ACTC 20 ACAA 14		GC	CAC	ACT 30 GCT 150	AGA	ACI	GCT 40 <u>CAA</u> 16		ATA:	17	AGG	GGC	GACA 60 <u>CGCG</u> 18	ATAJ TGGC 0	AGCO	CGG 70 <u>TAG</u> 19	ссао <u>саа</u> л 0	ACA	CAAT 80 <u>3TGT</u> 20	GAG	GC2	TCJ	AGC	IGGG	CG	TCG 100 5TCC 220	AC	CTG	GAT 11 TTT 23	CTGC 0 GGTC 0	GTT
CAC	SAA ICA	ATO	TTT 10 TCC 13	TAC TTT 0	TCT.	ACTC 20 ACAA 14 CGCG	CTO CC	GGC(	CAC	ACT 30 GCT 150	AGA	CACI	GCT 40 (CAA 16 (ATC	GGC CGG 0	ATA:	rggc 50 <u>cerc</u> 17	AGG	GGC	GACAJ 60 CGCG 180 GGCG	ATAJ TGGC 0 AGTC	CATA	CGG 70 <u>TAG</u> 19	CCAC	ACA	CAAT 80 <u>STGT</u> 20	GAC 0	CAC	TC/	AGC 210	CAT	YCGO	TCG 100 TCC 220	AC	CTC	GAT 11 TTT 23 GAG	CTGC 0 GGTC 0 CCGG	GTT
CC1	SAA TCA	CAC	TTT 10 TCC 13 CGG 25	TAC TTT 0 CCC 0	TCT.	ACTC 20 ACAA 14 CGCG 26		GC	CAC CTC	ACT 30 <u>GCT</u> 150 <u>AGC</u> 270	AGA	AAGT	GCT 40 <u>CAA</u> 16 <u>CATC</u> 28	GGC 0 0 0 0 0	ATA:	rggc 50 <u>cerc</u> 17 <u>GATG</u> 29	AGG	iggc icga iaga	GACA 60 CGCG 180 GGCG 300	ATAJ TGG( 0 AGT( 0	CATA	CGG 70 <u>TAG</u> 19 <u>AAA</u> 31	ссао 0 с <u>саа</u> 0 с <u>саа</u>	ACA	CAAT 80 <u>FTGT</u> 20 AT <u>GT</u> 32	GAC 0 CAC	GC2	ACT	AGC 210 210 210	IGGG	CGC	TCG 100 5TCC 220 220 CTT 340	ACC	CTG	GAT 11 23 GAG 35	CTGC 0 GGTC 0 CCGG 0	GTTO CATI
CC1	SAA ICA	CAC	TTT 10 TCC 13 CGG 25	TAC TTT 0 CCC 0	TCT.	ACTC 20 ACAA 14 CGCG 26		FCT	CAC	ACT 30 CCT 150 AGC 270	AGA	AAGT	GCT 40 <u>CAA</u> 16 <u>CATC</u> 28	GGC 10 10 10	ATA:	regec 50 <u>cerc</u> 17 <u>GATG</u> 29	AGG	GGC	GACAJ 60 CGCG 18 GGCGJ 30	ATAJ TGGC 0 AGTC 0	AGCC	CGG 70 <u>TAG</u> 19 <u>AAA</u> 31	CCAC CAAJ 0 CGAJ 0	ACA	CAAT 80 <u>STGT</u> 20 AT <u>GT</u> 32	GAC 0 CAC 0	SGCI	SACT	AGC 210 210 210 330	ECAT	YCGC	TCG 100 3TCC 220 CCTT 340	ATS	CTG	GAT 11 23 GAG 35	CTGC 0 GGTC 0 CCGG 0	GTTO
CCI	SAA ICA	CAC ATC	TTT 10 <u>TCC</u> 13 <u>CGG</u> 25 <u>TGG</u> 37	TAC TTT 0 CCC 0 CGG 0	TCT.	ACTC 20 ACAA 14 CGCG 26 26 CATC 38		FCT	CAC CTC GGC	ACT 30 GCT 150 AGC 270 GCC 390	AGA	AAGT	GCT 40 CAA 16 CATC 28 CAG 40	GGC 0 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	ATA TGC GCG	regec 50 <u>cerc</u> 17 <u>GATG</u> 29 <u>cera</u> 41	AGG	GGC CGA AGA	GACAJ 60 <u>CGCG</u> 180 <u>GGCG</u> 300 <u>TCTA</u> 420	ATA/ TGGC 0 AGTC 0 CCG/ 0	AGCC CATA CGAG	CGG 70 <u>TAG</u> 19 <u>AAA</u> 31 <u>ACG</u> 43	ссао 0 с <u>саа</u> 0 с <u>саа</u> 0 т <u>тсс</u> 0	ACA AAC	CAAT 80 <u>FTGT</u> 20 ATGT 32 ATTC 44	GAC 0 CAC 0 TCI 0	2002	VTC2	AGC 210 210 210 210 210 210 210 210 210 210	IGGG SCAT	CGC TT AAZ	TCG 100 5TCC 220 CCTT 340 460	ACS ATS ATS	CTG	GAT 11 23 GAG 35 GAT 47	CTGC 0 GGTC 0 CCGG 0 TCTC 0	CCAT GTTC CATT
CAC	GAA CCA	CAC	TTT 10 TCC 13 CGG 25 TGG 37	TAC TTT 0 CCC 0 CGG 0	TCT.	ACTC 20 ACAA 14 CGCG 26 26 CATC 38	CCC 0 AAX 0 TCC	ETC	CAC CTC GGC	ACT 30 GCT 150 AGC 270 GCC 390	AGA	AAGT CGAC	GCT 40 CAA 16 CATC 28 CAG 40		ATA TGCC GCGC	regec 50 <u>certe</u> 17 <u>sare</u> 29 <u>certa</u> 41	AGG CGCG 0 CGTG 0 CGTG 0 CGTG 0	GGC	GACAI 60 CGCG 180 GGCG 300 TCTAC 420	ATA/ TGGC 0 AGTC 0 CCG/ 0	AGCC CATA CGAG AGAG	CGG 70 TAG 19 AAA 31 ACG 43	CCAC CAAZ 0 CGAZ 0 <u>PTCC</u> 0	ACA	CAAT 80 <u>FTGT</u> 20 ATGT 32 ATTC 44	GAC 0 CAC 0 TCT 0		GTZ SACT	AGC 210 210 210 210 210 210 210 210 210 210	CAT	TC CGC	TCG 100 3TCC 220 CTT 340 460	ACS ATS ATS	CTG	GAT 11 TTT 23 GAG 35 GAT 47	CTGC 0 GGTC 0 CCGG 0 TCTC 0	CAT
CAC	GAA	CAC ATG CAG AGA	TTT 10 <u>TCC</u> 13 <u>CGG</u> 25 <u>TGG</u> 37 E D	TAC TTT 0 CCC 0 CGG 0 I	TCT.	ACTC 20 ACAA 14 CGCG 26 26 CATC 38 C	CTO CCTO 0 AAM 0 TCO 0 D CCG		CAC CTC GGC TCA	ACT 30 GCT 150 AGC 270 GCC 390	AGA AGA	CACT	CAA 16 CATC 28 CAG 40 SACAG	GGC	ATA TGC GCG AAA	regec 50 <u>cerc</u> 17 <u>GATG</u> 29 <u>cera</u> 41 <i>L</i>	AGG CGCG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGCG CGCG CO CGCG CGCG CO CC CO CC CC CO CC CC CC CC CC CC CC	AGA	GACAI 60 CGCG 180 GGCG 300 TCTA0 420 C	ATAJ	AGCC CATA CGAG AGAG	CGG 70 TAG 19 AAA 31 ACG 43 TTG	CCAC CAAZ 0 CGAZ 0 TTCC 0 GAG	ACA ACA ACA COL	CAAT 80 <u>FTGT</u> 20 ATGT 32 ATTC 44 E	TAC GAC 0 CAC 0 TCI 0 L GCI	SATI	GTZ SACI	AGC 210 210 210 210 210 210 210 210 210 210	IGGG ICAI ICCG ICCG	TC CGG TT AAJ G	TCG 100 TCC 220 CCTT 340 ACAC 460	ACI ATI ATI TGI	CTG CTC CATC	GAT 11 23 GAG 35 GAT 47 T	CTGC 0 CCGG 0 CCGG 0 TCTC 0 TCTC 0	GTTO GTTO CATTO TATO
CAC	GAA GCT CG CG	CAC ATG CAG AGA GCG	TTT 10 <u>TCC</u> 13 <u>SCGG</u> 25 <u>SCGG</u> 37 5 <u>SAAG</u> 49	TAC TTT 0 CCC 0 CCC 0 CCC 0 I ATA 0	TCT.	ACTC 20 ACAA 14 CGCG 26 26 26 26 26 26 26 26 26 26 26 26 26	CTO CCC 0 AA1 0 TCC 0 D CCGJ 0	GGC( TCT) TCC GAG E ACG	CAC CTC GGC TCA E AAG	ACT 30 <u>GCT</u> 150 <u>AGC</u> 270 <u>GCC</u> 390 <u>GCC</u> 390 <u>GCC</u> 390 <u>GCC</u> 390	AGA AGA CATC	CGACI	YGCT 40 YCAA 16 28 CATC 28 40 40 50 SACA 52	CGGC CGG CGG CAGG CAGG CAGG CAGG CAGG C	ATA TGC GCG AAA T GGA	resc 50 <u>cerc</u> 17 <u>SATG</u> 29 <u>cera</u> 41 <u>L</u> TTCI 53	AGG CGCG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGTG CO CGCG CO CO CO CGCG CO CO CO CO CO CO CO CO CO CO CO CO CO	A GAC	GACAI 60 <u>CGCG</u> 180 <u>GGCG</u> 300 <u>TCTA0</u> 420 <i>C</i> 540	ATAJ	CATA CGAG AGAG G I GGAA	CGG 70 TAG 19 AAA 31 ACG 43 G TTG 55	CCAC CAA2 0 CGA2 0 PTCC 0 GAGC	ACA AACA AAC CCT T GGA	CAAT 80 STGT 20 ATGT 32 ATGT 32 ATTC 44 E CTGA 56	TAC GAC 0 CAC 0 TC 0 L GCI 0	CATI	A CCC	AGC 210 210 210 210 210 210 210 210 210 210	CCCG CCCG CCCG CCCG CCCG CCCG CCCG CCC	ACGA	TCG 100 5TCC 220 101 340 460 1 580	ACI ATI ATI TGJ X	CTG CTC CATC	GAT 11 TTT 23 GAG 35 GAT 47 T CAC 59	CTGC 0 CCGG 0 CCGG 0 TCTC 0 TCTC 0 TCTC 0	GTTO GTTO CATT TATO G 1 GGAO
CAC CAC	SAA ICA SCT ICG C G	CAC ATG CAG AGA GCG	TTTT 10 TTCC 13 SCGGG 25 SCGG 25 STGG 37 SCGG 37 SCGG 49	TAC TTT 0 CCCC 0 CCCC 0 CCCC 0 I ATA 0	TCT. GTC TAT TTC TCC	ACTC 20 ACAA 14 CGCG 26 26 26 26 26 38 C CTTG 50	CTV CCC 0 AAM 0 TCC 0 D CCC 0 0 D CCC 0	GGC	CAC CTC GGC TCA E AAG	ACT 30 2GCT 150 270 6GCC 390 510	AGA AGA	CGACT CGAC CCGAC CTCA	CAAA 16 CAAC 28 ACAG 40 SACAG 52	GGC CGG CGG CGG CAG CAG CAG CAG CAG CAG	ATAN TGCC GCGC AAAA TI GGAN	CCTC 50 CCTC 17 SATG 29 CCTA 41 L FTCT 53	AGG CGCG 0 CGTG 0 CGAT 0 Q CGCA 0	AGA	GACAI 60 CGCG 180 GGCG 300 TCTA0 420 C 540	ATAJ	AGCC CATA CGAG AGAG GGAA	CGGG 70 TAG 19 AAAA 31 ACG 43 CG TTG 55	CCAG CCAA 0 CCGAA 0 CCGAA 0 0 CCGAA 0 0 GCGAGC 0	ACA AACA AAC	CAAT 80 STGT 20 ATGT 32 ATGT 44 E CTGA 56	TAC GAC 0 CAC 0 TCJ 0 L GCJ 0	CATI	A CAGO	AGC 210 210 210 210 210 210 210 210 210 210	CCCG CCCG CCCG CCCCG CCCCG CCCCG CCCCG	CCGC TTT AAA G AGC	TCG 100 5TCC 220 CTT 340 460 I SGAT 580	ACX ATX ATX TGJ Y ATJ	CTG CTC CATC	GAT 11 777 23 GAG 35 GAT 47 7 CAC 59	CTGC 0 CCGG 0 CCGG 0 TCTC 0 TCTC 0 TCTC 0	CCA GTTV CAT TATV G 1 GGA
CAC	SAA SCT CG CG CG CG VCG	CAC ATC CAC AGA AGA GCC	TTTT 10 STCCC 13 SCGGG 25 ATGG 37 SCGG 49 SAAG 49 SCGCCC 49 SCCCCCC 49	TAC TTT 0 CCC 0 CCC 0 CCC 0 I ATA 0 I CCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCC 1 CCCCC 1 CCCCCC 1 CCCCCCCC	TCT. GTC. TAT TTCC TCC	ACTC 20 ACAA 14 CGCG 26 26 26 26 26 26 26 26 26 26 26 26 26	CTV CCC 0 AAT 0 TCV 0 D CCGJ 0 CCCC 0 CCCCCC 0 CCCCCCCCCCCCCCCC	GGCC TCT TCC GAG E ACG L GCT	CAC CTC GGC TCA E AAG C CTCT	CACT 30 CGCT 150 CAGC 270 GGCC 390 CGC 390 CGC 390 CGC 390 CGCC 390 CGCC 390 CGCC 390 CGCC 270 CGCCC 270 CGCC CGCCC CGCC CGCC CGCC CGCCC CGCCC CGCC CGCC CGCCC CGCCC C	CCCC CAGA CATC CGCC CGCC CAA	AAGT CGAC CCGAC CCGAC CCGAC	CAAA 16 CAAA 16 CAAC 28 40 CAAG 40 CAAG 40 SACAG 52 52 52 52 52 52 52 52 52 52	CGGC CGGG CGGG CAGG IO CCAGG IO CCAGG IO CCAGG IO CCAGG IO CCAGG IO CCAGG IO CCAGG IO CCAGG IO CCG IO CCG IO C IO C	ATAN TGCC GCGC AAAA TI GGAN CGAN	regec 50 17 <u>GATC</u> 29 <u>CCTA</u> 41 <i>L</i> FTCT 53 <b>A</b>	AGG CGCG CO CGCG CO CGCG CO CGCA CO CGCA CO CGCA CO CGCA CO CGCG CO CGCG CGCG	GGGC GCGA BAGA CCAT A AGC D	GACAJ 60 CGCG 180 GGCGJ 300 FCTAC 420 C C ATGCJ 540 Y	ATAI TGGG 0 AGTC 0 CCGJ 0 I G ATCC 0 N :	AGCC CATA CGAG AGAG G I GGAM I S	CGGG 70 TAG 19 AAAA 31 ACG 43 : G TTG 55 : G GTG	CCAO CCAA 0 CCGA 0 CCGA 0 CCGA 0 CCGA 0 CCGA 0 CCCAO 0 CCCAO 0 CCAO 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCGAA CCGAA	ACA AACA AACA CCT T GGA	CAAT 80 <u>STGT</u> 20 ATGT 32 ATTC 44 E CTGA 56 P	TAC GAC 0 CAC 0 TCI 0 L GCI 0 P GCC	CATI SGC2 CCAC TTC TAC	SACI SACI SACI SACI SACI SACI SACI SACI	AGC 210 210 210 210 210 210 210 210 210 210	GCAT CCCG CCCG CCCG CCCG CCCG CCCG CCCG	TTT AAA G AGC	TCG 100 5TCC 220 220 CCTT 340 460 I 560 T	ACX ATX ATX TGI Y ATX	CTG CTC CATC	GAT 11 TTT 23 GAG 35 GAT 47 T CAC 59 V	CTGC 0 CCGG 0 CCGG 0 TCTC 0 TCTC 0 R TCGG	GTT GTT CAT GGA GGA
CAC CAC CAC L ACJ	SAA CCA SCT CG CG CG V CCG	CAC ATC CAC AGP GCC TTC	TTTT 10 5TCCC 13 5TCCC 25 37 5 5 5 5 5 5 5 5 5 5 5 5 5	TAC TTT 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 I ATA 0 I T CCC 0 I I I I I I I I I I I I I	TCT. GGTC. TTAT TTCC TCC A CGGG	ACTC 20 ACAA 14 CGCG 26 26 26 26 26 26 26 26 26 26 26 26 26	CTV CCC 0 0 TCC 0 0 TCC 0 0 CCC 0 0 CCC 0	FICE FICE FICE FICE FICE FICE FICE FICE	CAC CTC GGC TCA E AAG C TGT	ACT 30 GCT 150 270 GCC 270 GCC 390 CAGC 510 CAGC 510 CAGC 510 CAGC 510 CAGC 510 CAGC 510 CAGC 510 CAGC 50 CAGC CAGC CAGC 50 CAGC CAGC 50 CAGC 50 CAGC CAGC CAGC CAGC CAG	AGA AGA CGCC CGCC CGCC CGCC	AAGT CGAC CGAC CCGAC CCGAC CCGAC CCGAC CCGAC CCGAC CCGAC	CAAA 16 CAAC 28 40 CAAC 28 40 52 52 52 52 52 52 52 52 52 52	CGGC CGGG CAGG CAGG CAGG CAGG CAGG CAGG	ATAN TGCC GCGC AAAA GGAN K CGAA	regec 50 20 29 29 20 20 41 L TTCT 53 A AGGC 65	AGG CGCG 0 CGTG 0 CGAT 0 CGCA 0 CGCA 0 CGCA 0 CGCA	GGGC GCGA EAGA CAT A A A GC A C A C A C A C A C A C A C A	GACAJ 60 CGCG 180 GGCGJ 300 FCTAC 420 C : ATGC 540 Y I FTATZ 660	ATAJ TGG( 0 AGT( 0 CCGJ 0 CCGJ 0 CCGJ 0 N : AACJ 0	AGCC CATA CGAG AGAG G I GGAA I S ATCA	CGGG 70 TAG 19 AAAA 31 ACG 43 CG 55 CG GGTG 67	CCAO CCAA 0 CCGA 0 CCGA 0 CCGA 0 CCGA 0 CCGA 0 CCCAO 0 CCCAO 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCAA 0 CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAA CCGAC CCGAA CCGAC CCGAA CCCAA CCGAA CCCAA CCGAA CCGAA CCGAA CCCAA CCC	ACA AACA AAC T GGA	CAAT 80 3TGT 20 ATGT 32 ATTC 44 E CTGA 56 P ACCC 68	TAC GAC 0 CAC 0 TCI 0 L GCCI 0 P GCC 0	G CCAO TTO T CCAO T T CAGO	SACT SACT SACT SACT SACT SACT SACT SACT	AGC 210 210 210 210 210 210 210 330 200 150 H 200 570 A 200 200 200 200 200 200 200 200 200	GCAT CCCG GCCG GCTC G NTGG H SGCA	CGG TTT AAJ G AGC L	TCG 100 5TCC 220 220 CTT 340 460 1 580 T 580 T SGAC 700	ACX ATY ATY TGI Y ATY F	GCTG GCTC GCATU DACTU D CGAU D TGA!	GAT 11 23 GAG 35 GAG 47 T CAC 59 V IGT 71	CTGC 0 GGTC 0 CCGG 0 TCTC 0 TCTC 0 TCGG 0 R TCGG	GTTM GTTM CATT TATM G I GGGA GGGA GGGA
CAC	SAA SCT CG CG CG CG VCG	CAC ATG CAG AGA AGA F GCCG	TTTT 10 <u>STCC</u> 13 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u> 25 <u>STCC</u>	TAC TIT 0 CCC 0 CCC 0 CCC 0 I ATA 0 I GGA 0	TCT. GTC. TAT. TTTC. TTCC.	ACTC 20 ACAA 14 CGCG 26 26 26 26 38 CCTTG 50 CTTG 50 I CCGAT 62	CTV CC 0 AAT 0 TCC 0 D CC 0 CTV 0	E CT	CAC CTC GGC TCA E AAG C TGT	ACT 30 <u>CGCT</u> 150 <u>CAGC</u> 270 <u>GGCC</u> 390 <u>CGCC</u> 510 <u>CGCG</u> 630	CCCC CAGA	CGACI CGAC CGAC CCGAC CCGAC CCGAC CCGAC CCGAC CCGAC CCGAC CCGAC	CATC 28 CATC 28 CATC 28 CATC 28 CATC 28 CATC 20 S S CCATC 28 CATC 20 C CATC 20 CATC 20 CATC 20 CATC 20 C CATC 20 C CATC 20 C CATC 20 C	CGGC CGGG CAGG CAGG CAGG CAGG CAGG CAGG	ATAN TGCC GCGC AAAA T T GGAN CGAN	regec 50 20 29 29 29 20 20 20 20 20 20 20 20 20 20 20 20 20	AGG CCCC 0 CCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CGGC CGA CAT A AAGC. D TGA	GACAA 60 180 GGCGG 300 FCTAA 420 C : 540 S - 540 Y - 1 FTATZ 660	ATA) TGG( 0 AGT( 0 CCG) 0 CCG) 0 I ( 0 ATC) 0 N ( ) AAC) 0	AGCC CATA CGAG AGAG GGAA I S ATCA	CGGG 70 TAG 19 AAAA 31 AACG 43 : G TTG 55 : G GTG 67	CCAO CCAA 0 CGAA 0 CGAA 0 GGGA 0 CGGGA 0 CGGGA	ACA AACA AAC T SGA	CAAT 80 3TGT 20 ATGT 32 44 E CTGA 56 P ACCC 68	TAC GAC 0 CAC 0 TCI 0 L GCI 0 0 P GCC 0	CATI GGCJ CCAO TTY TAO GCAGO	SACI SACI SACI SACI SAGO SAGO	AGC 210 CTG 210 CTG 210 CTG 570 H CTG 590	GCAI CCCG CCCG CCCG CCCG CCCG CCCG ATGG H CCCCA	CCGC TTT AAJ G AGC L	TTCG 100 STCC 220 CCTT 340 ACAC 460 I SGAT 580 T SGAC 700	ACX ATY ATY TGJ Y ATJ F	CTG CTC CATC CATC CATC CATC CATC CATC CA	GAT 11 23 GAG 35 GAT 47 T CAC 59 V IGT 71	CTGC 0 GGTC 0 CCGG 0 CCGG 0 TCTC 0 TCTC 0 TCTC 0 R TCGG 0	GTT GTT CAT TAT GGA GGA GGGA
CAC CAC CAC	SAA ICA SCT ICG C Q Q T CG	CAC ATG CAG AGA GCG GCG	TTTT 10 TCC 13 SCGG 25 SCG 25 S	TAC TTT 0 CCCC 0 CCCC 0 CCCC 0 CCCC 0 I I GGG 0 I GGA 0	TTTT TAT TTTC TTC TTC TTC	ACTC 20 ACAA 14 CGCG 26 26 26 26 26 26 50 50 1 CCTTG 50 50 1 CCATC	CTV CC 0 AAM 0 TCC 0 D CC 0 0 CTV 0	SGC( ICT) ITC) SAG E ACG L SCT.	CAC CTC GGC TCA E AAG CTGT	ACT 30 <u>CGCT</u> 150 <u>CAGC</u> 270 <u>GGCC</u> 390 <u>CGCC</u> 390 <u>CGCC</u> 630	CCCC AGA CATC	AAGG	CAA 16 CAAC 28 CAAC 28 CAAC 40 0 S SACA 52 5 CAC 64	CGGC CGGG CGGG CGGG CGGG CAGG CGGG CAGG CGG CGGG CGG CGGG CG C	ATAN TGCC GCGC AAAA T GGAN GGAN	regec 50 20 20 20 20 20 20 20 20 20 20 20 20 20	AGG 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	A A A A A A A C A A A C A A A C A A A A	GACAJ 60 CGCCG 18 GGCCG 30 GGCCG 30 TCTAA 420 C : 54 54 54 54 54 54 54 54 54 54 54 54 54	ATAJ TGGG 0 AGT( 0 CCGJ 0 I ( AATCO 0 N 2 AACJ 0	AGCC CATA CGAG AGAG GGI I S ATCA	CGGG 70 TAG 19 AAAA 31 AACG 43 : G TTTG 55 : G GTG 67	CCAO CCAA 0 CCGAJ 0 CCGAJ 0 CCGAJ 0 CCGAJ 0 CCGAJ 0 CCGAJ	ACA AACA AAC CCT T GGA Y	CAAT 80 3TGT 20 ATGT 32 ATGT 44 E CTGA 56 P ACCC 68	TAC GAC 0 CAC 0 TCI 0 L GCI 0 P GCC 0 0	CATI SGC/ CCAO TTO TTO G	SACI SACI SACI SACI SACI SACI SACI SACI	AGC 00 AGC 210 CTG 330 CCTC 570 H CTC 570 A CTG 590	GCAT CCCG CCCG CCCG CCCG CCCG CCCG CCCG	CGC CCGC TTT AAJ G AGC L TTT	TTCG 100 STCC 220 CTT 340 ACAC 460 I SGAT 580 T NGAC 700	ACS ATS ATS TGJ Y ATJ F	CTG CTC CATU CATU D CGAU D TGA	GAT 11 TTT 23 GAG 35 GAT 47 T CAC 59 V IGT 71	CTGC 0 GGTC 0 CCGG 0 TCTC 0 TCTC 0 TCGG 0 R TCGG 0	GTTM GTTM CAT TATM GGA GGA A 1 GGGA
CAC CAC TCT L ACT	SAA SCI SCI CG CG CG V CG V CG W MCT	CAC ATC CAC AGA GCAC GCAC CAC AGA CAC ACA CAC	TTTT 10 TCC 13 SCGG 25 ATGG 25 ATGG 37 5 L SAAG 49 5 R GGCC 61 7 2 2 2 2 2 2 2 2 2 2 2 2 2	TAC TTT 0 CCCC 0 CCCC 0 CCCC 0 CCCC 0 I ATA	TTT TTT TTTC TTCC TTCC TTCC TCC TCC TCC	ACTC 20 ACAA 14 CGCG 26 26 26 26 26 26 26 26 50 27 50 2 1 CCTTG 50 2 CCTTG 62 CCTGG	CTV CCC 0 AAY 0 TCC 0 CCC 0 CCC 0 CCC 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 AAY 0 CCC 0 CCC 0 AAY 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCC 0 CCCC 0 CCCC 0 CCCC 0 CCCC 0 CCCC 0 CCCC 0 CCCC 0 CCCCC 0 CCCCC 0 CCCCCCC 0 CCCCCCCC	GGCI TTCI GAG E ACGI L GCT AGGG	CAC CTC GGC TCA EGA GAT	ACT 30 CCT 150 AGC 270 GCC 390 CAGC 510 CAGC 510 CAGC 510	CCCC AGA CATC CCCC CCCC CCCC CCCC CCCC C	AAGT CGAC TCA TCA FTCG R G	CAA 16 CAA 28 CAA 28 CAA 40 0 S SACA 52 54 52 54 52 54 52 54 52 54 52 54 52 54 54 52 54 54 54 54 54 54 54 54 54 54	CGGC CGGG CGGG	ATA <u>GCG</u> <u>AAAA</u> <u>I</u> <u>GGA</u> <u>K</u> <u>CGA</u>	regec 50 29 29 20 20 20 20 20 20 20 20 20 20 20 20 20	AGG CCCC 0 CCCCCCCCCCCCCCCCCCCCCCCCCCCCC	GGGC GCGA GAGA CCAT A AGC D TGA	GACAJ 60 <u>CGCCG</u> 30 <u>GGCCG</u> 30 <u>TCTA</u> 420 <i>C</i> 54 54 54 54 54 54 54 54 54	ATA) TGGG 0 AGT( 0 CCG) 0 I ( ATCC 0 N ( AAC) 0	AGCC CATA CGAG AGAG G 1 GGAA I S ATCA	CGGG 70 19 AAAA 31 AACG 43 55 55 67 67	CCAO CAAA 0 CGAA 0 PTCC 0 FTCC 0 GGAGC 0 0 CGAGC	ACA AACA AAC COT T GGA Y	CAAT 80 3TGT 20 ATGT 32 44 E CTGA 56 P ACCC 68	GAC 0 C CAC 0 C CAC 0 C CAC 0 C CAC 0 C CAC 0 C CAC 0 C CAC 0 C CAC 0 C CAC 0 C C C C	CATI SGC2 CCAC TTC TTC G CAGO	SACI SACI SACI SACI SAGU	AGC 210 210 210 210 210 210 210 210 210 210	GCAI CCCG GCTC G ATGG H CCCA	CGG TTT AAJ GG AGG L	TTCG 100 5TCC 220 CCTT 340 460 1 3GAT 580 T NGAC 700	ACX ATX TGI Y ATZ F	CTG CTC CAT CAT CACT CACT CACT CACT CACT	GAT 11 23 GAG 35 GAT 47 T CAC 59 V IGT 71	CTGC 0 GGTC 0 CCGG 0 TCTC 0 TCTC 0 TCGG 0 R TCGG 0	CCA GTT CAT TAT GGA GGA



**Fig. 2** PVX-mediated delivery of *PR1A:ECP* sequences and stable expression in *S. lycopersicum* induces a HR on tomato genotypes expressing the cognate *Cf-ECP* gene. **a** Symptoms of wild type PVX infection (vein clearing and chlorosis) on leaves lacking *Cf-ECP* genes. **b** Leaf epinasty and necrotic lesions on leaves of Cf-ECP1 plants (LA1547) induced by PVX:*ECP1*. **c** Necrotic lesions on leaves of Cf-ECP2 plants (Ontario 7518) induced by PVX:*ECP2*. **d** PVX:*ECP4*-induced leaf epinasty and necrosis in Cf-ECP4 plants

et al. 2000; Luderer et al. 2002) and demonstrates all four *PVX:ECP* sequences are functional *in planta*.

### Analysis of PR1A:ECP expression in transformed plants

Stable expression of *C. fulvum Avr* genes in tomato can be achieved in genotypes lacking the cognate Cf gene (Hammond-Kosack et al. 1994; Thomas et al. 1997). The

(LA1683) **e** Developing necrotic lesions on Cf-ECP5 plants (G1.1161) induced by PVX:*ECP5*. **f** Seedling lethality in progeny from a cross between an ECP2-expressing line (UEA721) and *Cf-ECP2* at 18 dpg. **g** Lethality in progeny from a cross between an ECP4-expressing line (UEA 678) and *Cf-ECP4* at 18 dpg. **h** Progeny expressing *ECP5* and *Cf-ECP5* show wilting of the cotyledons and lack leaf primordia compared to wildtype segregants (**i**)

 $F_1$  progeny from crosses between *Avr*-expressing plants and lines containing the corresponding *Cf* gene exhibit seedling lethality due to activation of a systemic HR at about 12 days post-germination (Hammond-Kosack et al. 1994; Thomas et al. 1997).

We generated ten primary transformants for each *PR1A:ECP* construct in the *S. lycopersicum* cultivar "Moneymaker" (Cf0). Approximately 25% of transformants

were sterile polyploids and c. 30% lacked the *ECP* transgene (results not shown). The remaining transformants were used as female parents in test crosses to Cf-ECP lines to identify transformants expressing functional ECPs (Table 1). Genetic analysis of self progeny (Table 1) from each transformant enabled us to identify active lines containing single copy T-DNAs for *ECP2* (UEA721), *ECP4* (UEA678 and UEA682), and *ECP5* (UEA726, UEA729 and UEA730). No transformants expressing *ECP1* were obtained in this study.

Development of the seedling lethal phenotype induced by *ECP2* and *ECP4* was similar to that reported for the *C. fulvum Avr9* and *Avr4* genes in Cf-9 and Cf-4 plants (Hammond-Kosack et al. 1994; Thomas et al. 1997). Wilting of the cotyledons was observed at 14 days post-germination (dpg) followed by necrosis in the cotyledons and true leaves at 18 dpg (Fig. 2f, g). The seedling lethal phenotype induced in all three *ECP5*-expressing plants was distinct from *ECP2* and *ECP4* (Fig. 2h). In a segregating population seedlings expressing *Cf-ECP5* and *ECP5* were stunted at 14 dpg compared to wildtype segregants and showed wilting of the cotyledons. The seedlings failed to produce true leaves (Fig. 2h) and seedling death was observed at approximately 21 dpg (not shown).

The Cf-ECP1 and Cf-ECP4 genes are located on the short arm of chromosome 1

Our preliminary analysis showed *Cf-ECP1* and *Cf-ECP4* are located on chromosome 1 (results not shown). To map the genes more accurately we prepared bulked segregant pools from Cf-ECP1 × Cf0 and Cf-ECP4 × Cf0  $F_2$  populations ( $F_21E$  and  $F_24E$ ) for marker analysis. We used cleaved amplified polymorphic sequence (CAPS) markers located at or close to *MW* (CT116, TG236, LoxR and CP46), which have been characterised previously (Balint-Kurti et al. 1994; Haanstra et al. 1999; Wulff et al. 2004).

LoxR is derived from sequences at the 3' end of the lipoxygenase *LoxR* gene at the distal end of *MW* (Wulff et al. 2004) and the Cf-ECP1 allele (LoxR<sup>Cf-ECP1</sup>) lacks a site for *Hinc*II (Fig. 3a). CAPS analysis of *Cf-ECP1/+* and +/+ bulks showed LoxR<sup>Cf-ECP1</sup> cosegregates with *Cf-ECP1* (Fig. 3a). The *S. lycopersicum* allele of TG236 (TG236<sup>lyc</sup>), which is located proximal to *MW* (Balint-Kurti et al. 1994; Haanstra et al. 1999, 2000) contains a site for *Ssp*I, which is absent from the Cf-ECP4 allele (TG236<sup>Cf-ECP4</sup>; see Fig. 3b). CAPS analysis of *Cf-ECP4/+* and +/+ bulks showed TG236<sup>Cf-ECP4</sup> cosegregates with *Cf-ECP4*. These data show *Cf-ECP1* and *Cf-ECP4* are both located on the short arm of chromosome 1.

Kan<sup>R</sup>; Kan<sup>S a</sup> Stock Genotype of TC parent Ratio of necrotic: wildtype TC progeny Transgene **UEA721** ECP2 11:5(1) Cf-ECP2 16:13 **UEA678** ECP4 20:5(1) Cf-ECP4 15:11 **UEA682** ECP4 14:4(1) Cf-ECP4 11:14 **UEA726** ECP5 38:11(1) Cf-ECP5 10:14 **UEA729** ECP5 16:5(1) Cf-ECP5 15:12 **UEA730** ECP5 14:3(1)Cf-ECP5 9:11

**Table 1** Genetic analysis of ECP transgenics

<sup>a</sup> The ratio of kanamycin resistant (Kan<sup>R</sup>) to susceptible (Kan<sup>S</sup>) progeny is shown and the number of predicted T-DNA loci based on a  $\chi^2$  goodness-of-fit test to a 3:1 ratio is shown in parentheses



**Fig. 3** CAPS analysis of  $F_21E$  and  $F_24E$  bulked segregant pools. **a** PCR amplification with LoxR-F1 and LoxR-R1 primers generates a 496 bp product from Cf0 and Cf-ECP1. Only the Cf0 allele (*lane 1*) contains a target for *Hinc*II, which generates cleavage products of 262 and 234 bp. The *Hinc*II site is absent from the Cf-ECP1 allele (*lane 2*). The Cf-ECP1/+ bulk (*lane 3*) contains both alleles but the Cf-ECP1 allele is absent from the +/+ bulk (*lane 4*). **b** The TG236 marker was

amplified with the primers TG236F and TG236R to generate a product of c.1,000 bp. The Cf0 allele contains a target sequence for *SspI (lane 1)* that generates fragments of 830 and 170 bp, which is absent in the Cf-ECP4 allele (*lane 2*). The Cf-ECP4/+ bulk contains both alleles (*lane 3*) and the +/+ bulk contains only the Cf0 allele (*lane 4*). Size markers are in lane '*M*'

### Cf-ECP1 and Cf-ECP4 are located distal to T-DNA<sup>FT33</sup> at MW

To map *Cf-ECP1* and *Cf-ECP4* more precisely we determined their location relative to the T-DNA in line FT33 (T-DNA<sup>*FT33*</sup>), which is located c.3 cM proximal to *MW* (Rommens et al. 1992; Jones et al. 1994). The T-DNA contains a copy of the non-autonomous *Z. mays* transposon *Ds*, which carries the *E. coli uidA* (*GUS*) gene. FT33 was crossed to four Cf-ECP lines (Cf-ECP1, Cf-ECP2, Cf-ECP4 and Cf-ECP5) as the female parent and F<sub>1</sub> progeny were used to pollinate Cf0 plants. Test cross (TC) progeny were assayed histochemically for GUS activity and then infected with the appropriate PVX:*ECP* construct to determine their *Cf-ECP* genotype.

A  $\chi^2$  analysis confirmed that none of the four *Cf-ECP* genes assort independently of the *GUS* reporter in T-DNA<sup>*FT33*</sup> demonstrating they are linked in repulsion (results not shown). Most progeny were either GUS-positive and exhibited wildtype symptoms of PVX infection (GUS/WT), or lacked the *GUS* reporter and exhibited PVX:*ECP*-induced necrosis (+/N). However, recombinant chromosomes were also detected (Table 2). The recombinant fraction between T-DNA<sup>*FT33*</sup> and *Cf-ECP2* (located proximal to T-DNA<sup>*FT33*</sup>) and T-DNA<sup>*FT33*</sup> with *Cf-ECP5* appear consistent with previous reports (Haanstra et al. 1999, 2000).

Recombinant progeny from both crosses (GUS/N and +/WT seedlings) were used in a molecular analysis to determine their location relative to T-DNA<sup>*FT33*</sup> (Fig. 4). FT33 × Cf-ECP4 +/WT recombinants lacked the LoxR<sup>*Cf-ECP4*</sup> allele, which is located distal to T-DNA<sup>*FT33*</sup>, but retained TG236<sup>*Cf-ECP4*</sup>, which is located proximal to it. In contrast, GUS/N recombinants lacked TG236<sup>*Cf-ECP4*</sup> but retained LoxR<sup>*Cf-ECP4*</sup> (Fig. 4). Similarly, all +/WT recombinants from the FT33 × Cf-ECP1 cross lacked the LoxR<sup>*Cf-ECP1*</sup> allele, but this was retained in GUS/N recombinants (Fig. 4). Analysis of the same recombinants with TG236 and CT116 was not possible as the Cf-ECP1 allele

Table 2 Linkage analysis of Cf-ECP genes

is not polymorphic with the *S. lycopersicum* allele, at least with the enzymes tested (results not shown). These data show *Cf-ECP1* and *Cf-ECP4* are located distal to T-DNA<sup>*FT33*</sup> on the short arm of chromosome 1.

The genetic distance between Cf-ECP1 and Cf-ECP4 with T-DNA<sup>FT33</sup> (Table 2) is similar to that reported for Cf-9 and T-DNAFT33 (Jones et al. 1994) suggesting they may also be located at MW. To confirm this, allelism tests with Cf-9 were performed by crossing Cf-ECP1 and Cf-ECP4 lines to line M18, which contains a Ds-tagged allele of Cf-9 (Cf-9:: $Ds^{M18}$ ) that originated from T-DNA<sup>FT33</sup> (Jones et al. 1994). The F<sub>1</sub>s were used to pollinate Cf0 plants and TC progeny from both crosses were assayed for GUS (as a marker for Cf-9) and then infected with recombinant PVX as described above. No recombinant chromosomes were detected in 491 TC progeny from the M18  $\times$  Cf-EC4 cross, or in 572 progeny from the  $M18 \times Cf$ -ECP1 cross (Table 2). These data show that Cf-ECP1 and Cf-ECP4 are allelic to Cf-9 at MW, or very tightly linked.

# Identifying T-DNA<sup>FT33</sup>/ Cf-ECP recombinants for transposon tagging

Our objective is to isolate all four *Cf-ECP* genes by transposon tagging so we used marker-assisted selection to identify recombinants containing T-DNA<sup>*FT33*</sup> in cis with each *Cf-ECP*. Approximately 50 GUS<sup>+</sup> TC seedlings from crosses between the Cf-ECP lines with FT33 were screened by CAPS analysis. Progeny containing *Cf-ECP1* or *Cf-ECP4* were identified using the LoxR marker (see Fig. 4); progeny containing *Cf-ECP2* or *Cf-ECP5* were identified using the markers CT116 and Cf4F3R3, respectively, which were previously shown to cosegregate with these genes (Haanstra et al. 1999, 2000). The presence of each *Cf-ECP* was confirmed by infecting self progeny with the appropriate PVX:*ECP* construct. In each case the progeny segregated at an approximate ratio of 3:1 for necrotic:wildtype individuals (results not shown).

Cross	Non-recombinant	classes	Recombinant	classes	Recombinant fraction		
	GUS/WT	+/N	GUS/N	+/WT			
$Cf0 \times (FT33 \times Cf-ECP4)$	139	149	4	5	2.4 %		
$Cf0 \times (FT33 \times Cf-ECP1)$	116	128	3	5	3.2 %		
$Cf0 \times (FT33 \times Cf-ECP5)$	112	100	5	3	4.1 %		
$Cf0 \times (FT33 \times Cf-ECP2)$	89	99	7	5	6.0 %		
$Cf0 \times (M18 \times Cf-ECP4)$	223	268	0	0	0 (+ 0.61%) <sup>a</sup>		
$Cf0 \times (M18 \times Cf-ECP1)$	274	298	0	0	0 (+ 0.53 %) <sup>a</sup>		

<sup>a</sup> For recombination values of zero the upper limit recombination value at P = 0.05 for *n* gametes was calculated according to the formula  $1-P^{1/n}$  according to Fisher and Yates (1963)



Fig. 4 Molecular analysis of recombinant TC progeny from FT33 × Cf-ECP1 and FT33 × Cf-ECP4 crosses. a Schematic showing molecular markers in S. lycopersicum (open box) and S. pimpinellifo*lium (grev box)*, pairing between non-sister chromatids, and the predicted haplotypes of recombinant chromosomes. b Molecular analysis of three GUS/N and three +/WT recombinants identified in progeny from FT33  $\times$  Cf-ECP1 and FT33  $\times$  Cf-ECP4 testcrosses using LoxR and TG236 CAPS markers. Panel I shows HincII digestion products of the LoxR marker on FT33 × Cf-ECP1 recombinants; panels II and III HincII digestion of LoxR and SspI digestion of TG236 from FT33 × Cf-ECP4 recombinants, respectively. Cf0 alleles are indicated with open arrowheads and Cf-ECP alleles with filled arrowheads. HincII digestion of the Cf0 LoxR allele generates two fragments of 262 and 234 bp, whereas digestion of the Cf-ECP4 allele generates a single product of 234 bp due to a short deletion within LoxR. Heterozygotes can be distinguished from Cf0 homozygotes based on the relative intensities of the 262 and 234 bp fragments (see lanes 3 and 4 in panel II). Lane 1, Cf0, lane 2 Cf-ECP parent, lane 3 Cf-ECP/+ bulk, Lane 4 +/+ bulk, and lanes corresponding to GUS/N and +/WT recombinants as indicated

DNA gel blot analysis of bulked segregant pools identifies candidate Hcr9s for Cf-ECP1 and Cf-ECP4

Having shown *Cf-ECP1* and *Cf-ECP4* are located at *MW* we identified cosegregating *Hcr9s* as candidate *Cf-ECPs*. DNA gel blots of Cf0 and Cf-ECP lines, together with the bulked segregant pools described above (Fig. 5) were hybridised with a *Cf-9* 5' probe PCR-amplified from pCD-NAL9 (Jones et al. 1994). Analysis of *Bgl*II-digested DNA identified several polymorphic *Hcr9s* that cosegregate with *Cf-ECP1*, and with *Cf-ECP4* (Fig. 5).



**Fig. 5** DNA gel blot analysis of  $F_21E$  and  $F_24E$  bulked segregant pools using a *Cf-9* probe identifies candidate *Cf-ECP1* and *Cf-ECP4* genes. DNA from Cf0, Cf-ECP1, Cf-ECP4 and bulked segregant pools were digested with *Bg*/II, electrophoresed on 0.8% w/v agarose gels, and blotted as described in methods. **a** The gel blot includes *Bg*/II-digested DNA from Cf0 (*lane 1*), Cf-ECP1 (*lane 2*), the Cf-ECP1/+ bulk (*lane 3*) and the +/+ bulk (*lane 4*). **b** *Bg*/II-digested DNA from Cf0 (*lane 1*), Cf-ECP4 (*lane 2*), the Cf-ECP4/+ bulk (*lane 3*) and the +/+ bulk (*lane 4*). In both panels polymorphic *Hcr9s* that cosegregate with the corresponding *Cf-ECP* gene are indicated with *open arrowheads* and their approximate sizes are indicated in kbp

#### Discussion

Durable resistance to the leaf mould pathogen C. fulvum has been a major objective for tomato breeders (Stevens and Rick 1988; Rivas and Thomas 2005). Successive introgressions of Cf genes into S. lycopersicum provided a wealth of genetic resources that formed the basis for informative studies on Cf gene organisation and evolution (Jones et al. 1994; Parniske et al. 1997; Thomas et al. 1997; Kruijt et al. 2004), and Cf protein recognition specificity (Van Der Hoorn et al. 2001a; Wulff et al. 2001). Functional Cf genes from several Solanum species have been mapped to chromosome 6 (Jones et al. 1993; Dixon et al. 1996) and chromosome 11 (EA Anderson and CM Thomas, in preparation), but most Cf genes and Hcr9 loci have been mapped to loci on the short arm of chromosome 1 including MW, Southern Cross (SC), Northern Lights (NL), OR and AU (Jones et al. 1993; Parniske and Jones 1999; Parniske et al. 1999; Haanstra et al. 1999, 2000; Yuan et al. 2002). Several studies suggest that Hcr9 amplification and the duplication of these loci occurred prior to speciation (Parniske and Jones 1999; Parniske et al. 1999; De Kock et al. 2005).

To learn more about functional *Cf* gene loci genetic strategies have been used to identify novel genes on the short arm of tomato chromosome 1 (Haanstra et al. 2000), and genes that induce an HR upon recognition of specific *C. fulvum* ECPs, like the *S. pimpinellifolium Cf-ECP1*, *Cf-ECP2*, *Cf-ECP3*, *Cf-ECP4* and *Cf-ECP5* genes (Lauge

et al. 1998, 2000). The chromosomal locations of Cf-ECP1 and Cf-ECP4 were previously unknown and we demonstrate both genes are located distal to T-DNAFT33 at MW (Fig. 4; Table 2). The isolation and characterisation of both genes will increase our understanding of the molecular basis for Cf gene evolution at MW. Cf-9 and Cf-4 were originally introgressed from S. pimpinellifolium and S. hirsutum but they apparently originated prior to speciation and are present in many wild Solanum accessions (Kruijt et al. 2005). These two haplotypes are the most extensively characterized at the molecular level and these analyses form the basis for our current understanding of Cf gene evolution (Jones et al. 1994; Parniske et al. 1997 Thomas et al. 1997; Kruijt et al. 2004). Each haplotype is comprised of tandemly duplicated Hcr9s and Cf gene variation is generated by sequence exchange between Hcr9 paralogues, and non-synonymous nucleotide substitutions in sequences encoding the solvent-exposed residues of a conserved LRR structural motif, which determine recognition specificity (Parniske et al. 1997; Thomas et al. 1997; Van Der Hoorn et al. 2001a; Wulff et al. 2001).

Previous mapping studies showed Cf-ECP2 and Cf-ECP3 are located at OR c.10 cM proximal to MW (Fig. 6) and together they define another complex locus for C. fulvum resistance (Haanstra et al. 1999; Yuan et al. 2002) while Cf-ECP5 was mapped 3 cM proximal to MW at AU (Haanstra et al. 2000), and these reports are consistent with our test cross analyses (Table 2). The demonstration that Cf-ECP1 and Cf-ECP4 are located at MW increases the number of distinct Cf recognition specificities at this locus to seven (Fig. 6) and emphasises its importance in generating Cf gene diversity. This is particularly true in S. pimpinellifolium where Cf-9, 9DC1, 9DC2, and 9DC3 (which all induce an Avr9-dependent HR) and Hcr9-9B originated (Stevens and Rick 1988; Jones et al. 1994; Parniske et al. 1997; Kruijt et al. 2004).



Fig. 6 A genetic map of the short arm of chromosome 1 showing the locations of ten distinct recognition specificities at the MW, AU and OR loci

It is assumed Cf-ECP genes also encode Hcr9 proteins and molecular analysis of populations where Cf-ECP2 or Cf-ECP5 segregated did identify candidate Hcr9s (Haanstra et al. 2000; De Kock et al. 2005). However, this strategy failed to identify an Hcr9 that could induce an ECP2-dependent HR (De Kock et al. 2005). Significantly, the Cf-ECP2 locus has not been delimited with flanking markers and it is possible that not all candidate Hcr9s have been tested (De Kock et al. 2005). Our analysis of bulked segregant pools identified at least four polymorphic Hcr9s that cosegregate with Cf-ECP1 (Fig. 5a) and six that cosegregate with Cf-ECP4 (Fig. 5b) so it appears the MW haplotypes in S. pimpinellifolium LA1547 (Cf-ECP1) and LA1683 (Cf-ECP4) are also comprised of duplicated Hcr9s. Candidate Hcr9s could now be tested by transient expression in N. tabacum or S. lycopersicum together with the cognate ECP.

An alternative strategy to cloning Cf-ECPs is transposon tagging using the genetically linked Ds element in line FT33 (Rommens et al. 1992; Fig. 6), as in the isolation of Cf-9 (Jones et al. 1994). We identified recombinants containing each Cf-ECP in cis with T-DNAFT33 and these will be crossed to a line expressing the Activator transposase to induce Ds transposition (Jones et al. 1994). Identification of Ds-tagged mutants of Cf-ECP2, Cf-ECP4 and Cf-ECP5 will be achieved by testcrossing tagging parents to the appropriate ECP line to identify suppressors of the seedling lethal phenotype (Fig. 2). The frequency of Ds-tagged mutants of Cf-ECP1, Cf-ECP4 and Cf-ECP5 should be similar to that reported for Cf-9 i.e., c.1 in  $2.5 \times 10^3$  (Jones et al. 1994), but less frequent for Cf-ECP2, which is located further from FT33 (Haanstra et al. 1999). This assumes each line contains a single copy of each Cf-ECP but in some haplotypes, such as the S. pimpinellifolium Cf-2 locus on chromosome 6 and the MW locus in LA1301, multiple copies of the functional genes have arisen through sequence duplications (Dixon et al. 1996; Van Der Hoorn et al. 2001b; Kruijt et al. 2004).

No lines expressing ECP1 were obtained in this study (Table 1), which might reflect the limited number of transformants analysed, or toxicty associated with constitutive expression of ECP1 *in planta*. Despite this, *Ds*-tagged mutants of *Cf-ECP1* could be identified in a high-throughput screen based upon suppression of PVX:*ECP1*-induced necrosis (Fig. 2). The availability of stable transgenic lines expressing ECP2, ECP4 or ECP5 and PVX-mediated transient expression of all four ECPs will prove useful for the functional analysis of cloned *Cf-ECPs* (Fig. 2). Expression of *C. fulvum* AVRs and ECPs in tomato and tobacco obviates the requirement for *C. fulvum* infections and has proved invaluable for structure-function studies of Cf-9 and Cf-4 (Van Der Hoorn et al. 2001a; Wulff et al. 2001), functional analysis of AVRs and ECPs (Kooman-Gersmann et al. 1997; Luderer et al. 2002), and identifying novel *Cf* genes (Lauge et al. 1998, 2000).

The actual number of Cf gene loci in Solanum species has yet to be determined. Duplicated Hcr9s are present at SC and NL (Parniske et al. 1999; Parniske and Jones 1999) but some are pseudogenes and functional genes have yet to be identified at these loci. Only a limited analysis of S. pimpinellifolium accessions were required to identify Cf-ECP genes (Lauge et al. 1998, 2000), and more extensive analyses have shown that functional variants of Cf-9 and Cf-4 are widely distributed in Solanum species (Kruijt et al. 2005). This demonstrates that wild Solanum germplasm is an important source of Cf gene variation, which could be exploited for breeding, and for evolutionary analyses. Our results further highlight the importance of the MW locus in generating Cf gene novelty. Effective use of the genetic resources described here will make the isolation of all four *Cf-ECPs* by sequence homology or transposon tagging a realistic possibility. The molecular characterisation of all four Cf-ECPs should increase our understanding of the molecular mechanisms underlying Cf gene evolution at MW, and the molecular basis for Cf protein recognition specificity.

**Acknowledgments** Part of this work was funded by a BBSRC grant (P15199) awarded to C.M.T. Both E.S. and M.I. gratefully acknowledge receipt of Greek State Scholarships Foundation (IKY) awards for postgraduate study. Research with recombinant PVX is authorized under DEFRA licence PHL 170/5467.

### References

- Balint-Kurti PJ, Dixon MS, Jones DA, Norcott KA, Jones JDG (1994) RFLP linkage analysis of the Cf-4 and Cf-9 genes for resistance to Cladosporium fulvum in tomato. Theor Appl Genet 88:691–700
- Dixon MS, Jones DA, Keddie JS, Thomas CM, Harrison K, Jones JDG (1996) The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 84:451–459
- De Kock MJD, Brandwagt BF, Bonnema G, De Wit PJGM, Lindhout P (2005) The tomato *Orion* locus comprises a unique class of *Hcr9* genes. Mol Breed 15:409–422
- Fisher RA, Yates F (1963) Statistical tables for biological medical and agricultural research. Oliver and Boyd, Edinburgh
- Haanstra JPW, Lauge R, Meijer-Dekens F, Bonnema G, De Wit PJGM, Lindhout P (1999) The *Cf-ECP2* gene is linked to, but not part of the *Cf-4/Cf-9* cluster on the short arm of chromosome 1 of tomato. Mol Gen Genet 262:839–845
- Haanstra JPW, Meijer-Dekens F, Lauge R, Seetanah DC, Joosten MHAJ, De Wit PJGM, Lindhout P (2000) Mapping strategy for resistance genes against *Cladopsorium fulvum* on the short arm of chromosome 1 of tomato: *Cf-ECP5* near the *Hcr9* Milky Way cluster. Theor Appl Genet 101:661–668
- Hammond-Kosack KE, Harrison K, Jones JDG (1994) Developmentally regulated cell death on expression of the fungal avirulence gene Avr9 in tomato seedlings carrying the disease resistance gene Cf-9. Proc Natl Acad Sci USA 91:10445–49

- Hammond-Kosack KE, Staskawicz BJ, Jones JDG, Baulcombe DC (1995) Functional expression of a fungal avirulence gene from a modified potato virus X genome. Mol Plant Microbe Interact 8:181–185
- Jones DA, Dickinson MJ, Balint-Kurti PJ, Dixon MS, Jones JDG (1993) Two complex resistance loci revealed in tomato by classical and RFLP mapping of the Cf-2, Cf-4, Cf-5 and Cf-9 genes for resistance to Cladosporium fulvum. Mol Plant Microbe Interact 6:348–357
- Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JDG (1994) Isolation of the tomato Cf-9 gene for resistance to Cladosporium fulvum by transposon tagging. Science 266:789– 793
- Kooman-Gersmann M, Vogelsang R, Hoogendijk ECM, De Wit PJGM (1997) Assignment of amino acid residues of the Avr9 peptide of *Cladosporium fulvum* that determine elicitor activity. Mol Plant Microbe Interact 10:821–829
- Kruijt M, Brandwagt BF, De Wit PJGM (2004) Rearrangements in the *Cf-9* disease resistance gene cluster of wild tomato have resulted in three genes that mediate Avr9 responsiveness. Genetics 168:1655–1663
- Kruijt MDJ, Kip DJ, Joosten MHAJ, Brandwagt BF, de Wit PJGM (2005) The Cf-4 and Cf-9 resistance genes against Cladosporium fulvum are conserved in wild tomato species. Mol Plant Microbe Interact 18:1011–1021
- Lauge R, Joosten MHAJ, Haanstra JPW, Goodwin PH, Lindhout P, De Wit PJGM (1998) Successful search for a resistance gene in tomato targeted against a virulence factor of a fungal pathogen. Proc Natl Acad Sci USA 95:9014–9018
- Lauge R, Goodwin PH, De Wit PJGM, Joosten MHAJ (2000) Specific HR-associated recognition of secreted proteins from *Cladosporium fulvum* occurs in both host and non-host plants. Plant J 23:735–745
- Luderer R, De Kock MJD, Dees RHL, De Wit PJGM, Joosten MHAJ (2002) Functional analysis of cysteine residues of ECP elicitor proteins of the fungal tomato pathogen *Cladosporium fulvum*. Mol Plant Pathol 3:91–95
- Panter SN, Hammond-Kosack KE, Harrison K, Jones JDG, Jones DA (2002) Developmental control of promoter activity is not responsible for mature onset of Cf-9B-mediated resistance to leaf mold in tomato. Mol Plant Microbe Interact 15:1099–1107
- Parniske M, Jones JDG (1999) Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene family. Proc Natl Acad Sci USA 96:5850–5855
- Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BBH, Jones JDG (1997) Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. Cell 91:821–832
- Parniske M, Wulff B, Bonnema G, Thomas CM, Jones DA, Jones JDG (1999) Homologues of the Cf-9 disease resistance gene (Hcr9s) are present at multiple loci on the short arm of tomato chromosome 1. Mol Plant Microbe Interact 12:93–102
- Rivas S, Thomas CM (2005) Molecular interactions between tomato and the leaf mold pathogen *Cladosporium fulvum*. Ann Rev Phytopathol 43:395–436
- Rommens CMT, Rudenko GN, Dijkwell PP, Van Haaren MJJ, Ouwerberk PBF, Blok KM, Nijkamp HJJ, Hille J (1992) Characterization of the *Ac/Ds* behaviour in transgenic tomato plants using plasmid rescue. Plant Mol Biol 20:61–70
- Stevens MA, Rick CM (1988) Genetics and breeding. In: Atherton JG, Rudich J (eds) The tomato crop. Chapman Hall, London, pp 35– 109
- Takken FLW, Thomas CM, Joosten MHAJ, Golstein C, Westerink N, Hille J, Nijkamp HJJ, De Wit PJGM, Jones JDG (1999) A second gene at the tomato *Cf-4* locus confers resistance to *Cladosporium*

*fulvum* through recognition of a novel avirulence determinant. Plant J 20:279–28

- Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K, Jones JDG (1997) Characterization of the tomato *Cf-4* gene for resistance to *Cladopsorium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. Plant Cell 9:2209–2224
- Van Den Ackerveken GFJM, Van Kan JA, Joosten MHAJ, Mnisers JM, Verbakel HM, De Wit PJGM (1993) Characterization of two putative pathogenicity genes of the fungal tomato pathogen *Clad*osporium fulvum. Mol Plant Microbe Interact 6:210–215
- Van Der Hoorn RAL, Roth R, De Wit PJGM (2001a) Identification of distinct specificity determinants in resistance protein Cf-4 allows construction of a Cf-9 mutant that confers recognition of avirulence protein Avr4. Plant Cell 13:273–285
- Van der Hoorn RAL, Kruijt M, Roth R, Brandwagt BF, Joosten MHAJ, De Wit PJGM (2001b) Intragenic recombination generated two distinct *Cf* genes that mediate Avr9 recognition in the natural population of *Lycopersicon pimpinellifolium*. Proc Natl Acad Sci USA 98:10493–10498
- Wulff BHW, Thomas CM, Smoker M, Grant M, Jones JDG (2001) Domain swapping and gene shuffling identify sequences required for induction of an Avr-dependent hypersensitive response by the tomato Cf-4 and Cf-9 proteins. Plant Cell 13:255–272
- Wulff BBH, Thomas CM, Parniske M, Jones JDG (2004) Genetic variation at the tomato *Cf-4/Cf-9* locus induced by EMS mutagenesis and intralocus recombination. Genetics 167:459–470
- Yuan YN, Haanstra J, Lindhout P, Bonnema G (2002) The *Cladopso*rium fulvum resistance gene Cf-ECP3 is part of the Orion cluster on the short arm of chromosome 1. Mol Breeding 10:45–50