Theor Appl Genet (2002) 105:355–363
DOI 10.1007/s00122-002-0953-x

M. Quint · R. Mihaljevic · C. M. Dussle · M. L. Xu
A. E. Melchinger · T. Lübberstedt

Development of RGA-CAPS markers and genetic mapping of candidate genes for sugarcane mosaic virus resistance in maize

Received: 23 September 2001 / Accepted: 12 December 2001 / Published online: 19 June 2002

Abstract Three previously published resistance gene analogues (RGAs), pic13, pic21 and pic19, were mapped in relation to sugarcane mosaic virus (SCMV) resistance genes (*Scmv1*, *Scmv2*) in maize. We cloned these RGAs from six inbreds including three SCMV-resistant lines (D21, D32, FAP1360A) and three SCMV-susceptible lines (D145, D408, F7). Pairwise sequence alignments among the six inbreds revealed a frequency of one single nucleotide polymorphism (SNP) per 33 bp for the three RGAs, indicating a high degree of polymorphism and a high probability of success in converting RGAs into co-dominant cleaved amplified polymorphic sequence (CAPS) markers compared to other sequences. SNPs were used to develop CAPS markers for mapping of the three RGAs in relation to *Scmv1* (chromosome 6) and *Scmv2* (chromosome 3), and for pedigree analyses of resistant inbred lines. By genetic mapping pic21 was shown to be different from *Scmv2*, whereas pic19 and pic13 are still candidates for *Scmv1* and *Scmv2*, respectively, due to genetic mapping and consistent restriction patterns of ancestral lines.

Keywords RGA · SNP · CAPS · SCMV · Maize

Communicated by C. Möllers

M. Quint · R. Mihaljevic · C.M. Dussle · M.L. Xu
A.E. Melchinger · T. Lübberstedt (✉)
Institute of Plant Breeding, Seed Science,
and Population Genetics, University of Hohenheim,
Fruwirthstraße 21, 70593 Stuttgart, Germany
e-mail: Thomas.Luebberstedt@agrsci.dk
Fax: +45-58-11-33-01

Present addresses:
M.L. Xu, College of Bioscience and Biotechnology,
Yangzhou University, Yangzhou, Jiangsu, 225009, PR China

T. Lübberstedt, Danish Institute of Agricultural Sciences,
4200 Flakkebjerg, Denmark

Introduction

Sugarcane mosaic virus (SCMV) is one of the most important virus diseases of maize and causes serious yield losses in susceptible cultivars (Fuchs and Grüntzig 1995). Owing to the non-persistent transmission, control of aphid vectors by chemical means is not effective and, therefore, cultivation of resistant maize varieties is the most efficient method of virus control. In a study with 122 early maturing European maize inbreds, three lines (FAP1360A, D21 and D32) displayed complete resistance and four lines displayed partial resistance (FAP1396A, D06, D09 and R2306) against SCMV and maize dwarf mosaic virus (MDMV) (Kuntze et al. 1997). In field trials, resistance of all three European lines D21, D32 and FAP1360A seemed to be controlled by one to three genes (Melchinger et al. 1998). Linkage mapping and 'bulked segregant analysis' (BSA) mapped two major genes, *Scmv1* on the short arm of chromosome 6 and *Scmv2* near the centromere of chromosome 3 (Melchinger et al. 1998; Xia et al. 1999; Xu et al. 1999; Dußle et al. 2000). Minor quantitative trait loci (QTLs) affecting SCMV resistance were identified on chromosomes 1, 5, and 10 (Xia et al. 1999). For full resistance to SCMV, the presence of both *Scmv1* and *Scmv2* is essential. *Scmv1* suppresses symptom expression throughout all developmental growth stages at a high level, whereas *Scmv2* was expressed at later stages of infection (Xia et al. 1999; Dußle et al. 2000).

With the cloning of a number of disease resistance genes (R genes) from several plant species, it became obvious that these R genes share homologies in protein domains such as the nucleotide-binding site (NBS) and leucine-rich repeats (LRRs) (reviewed in Bent 1996; reviewed in Hammond-Kosack and Jones 1996). Degenerate primers based on the amino-acid sequence of these domains have meanwhile allowed successful PCR amplification of several RGAs from various plant species with significant homology to known plant disease R genes. Collins et al. (1998) identified 20 RGA loci in maize, which mapped preferentially to chromosomal regions

known to carry R genes (McMullen and Simcox 1995). These RGAs can be further analysed for their potential use in marker-assisted selection (MAS) or even the cloning of target genes. The latter approach has been successfully used in maize to clone the *Rp1-D* gene, a member of the complex locus composed of approximately nine gene homologues, conferring resistance to common rust, *Puccinia sorghi* (Collins et al. 1999).

Most sequence variation is attributable to single nucleotide polymorphisms (SNPs), with the rest attributable to insertions or deletions of one or more bases, repeat length polymorphisms and rearrangements (The International SNP Map Working Group 2001). In maize, one SNP between two randomly sampled sequences occurs approximately every 104 bp (Tenaillon et al. 2001). Therefore, SNPs are present at sufficient density for comprehensive haplotype analysis as applied in this study.

The objectives of the present study were to: (1) clone three RGAs previously mapped to chromosomal regions known to harbour SCMV R genes from six inbred lines resistant or susceptible to SCMV, (2) measure the frequency of SNP and CAPS occurring within RGAs and evaluate their usefulness as a source for marker development, and (3) map RGA-CAPS markers in relation to *Scmv1* and *Scmv2* to examine their role as candidates for SCMV R genes.

Materials and methods

Plant materials

Seventeen maize inbred lines were analysed for CAPS restriction pattern including: (1) three resistant European dent lines D21, D32, and FAP1360A, (2) three highly susceptible European lines D408 (dent), D145 (flint), and F7 (flint), (3) ancestral lines of the three resistant lines, A632, V3, WD, W401, Co158, Co125, FAP493B, and FAP954A, and (4) three partially resistant lines D06, D09, and FAP1396A.

The mapping populations consisted of: (1) 87 $F_{2:3}$ families derived from a cross between D32 and D145 (PopA) (Vuylsteke et al. 1999; Xia et al. 1999), (2) 27 resistant BC5 individuals from the cross (F7 × FAP1360A) × F7 (PopB) (Xu et al. 1999), and (3) 30 resistant BC7 individuals from the cross (D408 × D21) × D408 (PopC). PopB originally consisted of 40 individuals, but 13 were excluded after progeny testing in field trials for resistance to SCMV. In addition, the map of PopB was saturated with SSR markers *phi075*, *umc1002*, *umc1018*, *bmc1600*, *bmc1433*, *bngl107*, *bmc1538*, and *bngl426*. Primer sequences were obtained from the maize database (http://www.agron.missouri.edu/ssr.html). PCR amplification was performed as described by Xu et al. (1999). pic13 was also mapped in oat-maize chromosome addition lines (Ananiev et al. 1997).

Cloning and sequencing of RGAs

Genomic DNA was extracted from 0.1 g of freeze-dried leaf tissue following the CTAB procedure as described by Hoisington et al. (1994). RGAs were isolated and cloned from the parental lines using PCR with specific primers (Table 1) based on the original sequences of pic13 (pic13L2/pic13R2), pic21 (pic21L1/pic21R1), and pic19 (pic19L/pic19R) described by Collins et al. (1998). DNA amplifications were performed in a standard reaction mix containing 100 ng of genomic DNA, 10 mM of Tris–HCl (pH 9), 50 mM of KCl, 1.5 mM of $MgCl_2$, 0.3 mM of dNTPs, 0.5 pMol of each primer and 0.625 U of *Taq* DNA polymerase (Amersham Pharmacia, Freiburg). After an initial denaturation step at 94 °C for 2 min, template DNAs were amplified using 35 cycles with the following conditions: 1 min at 94 °C, 2 min at 55 °C, and 2 min at 72 °C. The final extension step was conducted at 72 °C for 2 min. Amplification products were excised from the agarose gel, extracted with the Nucleospin Kit (Macherey and Nagel, Düren), and blunt-end cloned into the pBluescript vector.

Table 1 Primer sequences used for specific amplifications of pic13, pic19 and pic21

RGA	Primer	Sequences 5′–3′
pic13	pic13L2	TTGAAGCCATTGCTGGTGAC
	pic13R2	GCCATGAGCTATCCATTGAG
pic19	pic19L	TAGATGATGTCTGGACGGCT
	pic19R	AGCCAATGGCAAACCATCAC
	pic19X	GCAGTTCCTCTCTGCAACGTG
	ippic19L	CCAGAGTTACATCAGTGTGG
	ippic19R	ACATCAGCCGTCCAGACATC
pic21	pic21L1	GGAAGACCACGCTGCTCAAC
	pic21R1	CTCATCAGGTGGTCGCCAAC

Elongation of pic19 was performed by inverse PCR. Genomic DNA (500 ng) of the resistant parent FAP1360A was digested with *Rsa*I for 1 h at 37 °C, self-ligated overnight at 16 °C, and amplified with inverse PCR primers ippic19L and ippic19R. The elongated pic19 can be recreated after amplification by using primers pic19L and pic19X (Table 1). DNA sequencing was performed using the ALFExpress automated DNA sequencer (Amersham Pharmacia, Freiburg). Sequencing reaction conditions were chosen as suggested by the manufacturer (Amersham Pharmacia, Freiburg). For verification, pic19 was again sequenced by SEQLAB Sequence Laboratories Göttingen GmbH. The DNA sequences and the deduced amino-acid sequences were analysed using the ALIGN Plus 2.0 software package (http://www.scied.com/ses_alim.htm).

Generation and mapping of CAPS markers

Sequence regions displaying single nucleotide changes within restriction enzyme recognition sites between parental lines of mapping populations were used to identify RGA-CAPS using the CLONE Manager 3.11 software package (http://www.scied.com/ses_cm6.htm). PCR products were generated using the primers (Table 1) and the amplification conditions listed above. PCR products were digested with restriction endonucleases *Nla*III (pic19), *Rsa*I (pic13), and *Mbo*II (pic21) at 37 °C for 4 h according to the manufacturers' suggestions. RGAs were separated on 3% MetaPhor agarose gels in 0.5 × TBE buffer except for pic21, which was separated on denaturing polyacrylamide gels (SequaGel6, Biozym, Hessisch Oldendorf). Mapping of pic13, pic21, and pic19 was performed with PopA. In addition, pic13 was mapped in oat-maize addition lines and pic19 in PopB (Xu et al. 1999).

Linkage and statistical analyses

RGAs were mapped using data generated in previous studies (Xia et al. 1999; Xu et al. 1999). Marker orders and genetic distances for PopA were calculated with MAPMAKER 3.0b (Lander et al. 1987) using a LOD threshold of 3.0 and the Kosambi mapping function. Marker orders and genetic distances for PopB were calculated with Cri-Map (http://biobase.dk/Embnetut/Crimap).

Significant differences among the six inbred lines (resistant: D21, D32, FAP1360A; susceptible: D408, D145, F7) for SNP and CAPS frequencies per 100 bp of RGA sequence were tested by ANOVA using the software package PLABSTAT (http://www.uni-

Table 2 Genetic similarities (GS) of RGAs and frequency of SNPs per 100 bp between six maize inbred lines

Inbred lines	F7	FAP1360A	D21	D408	D32	D145	
pic19							
F7		6.38	5.10	6.54	5.58	6.34	GS
FAP1360A	0.34		4.78	6.22	5.90	2.55	
D21	0.50	0.53		2.55	1.91	5.10	
D408	0.34	0.34	0.74		1.59	5.58	
D32	0.45	0.41	0.84	0.86		4.15	
D145	0.38	0.81	0.55	0.47	0.57		
	SNP/100 bp					$\Sigma^a = 58$	
pic13							
F7		0.92	4.94	1.23	3.94	0.56	GS
FAP1360A	0.79		4.80	1.83	4.23	0.89	
D21	0.50	0.43		4.62	3.34	3.30	
D408	0.93	0.86	0.57		3.90	0.84	
D32	0.29	0.36	0.50	0.36		3.62	
D145	0.79	0.86	0.43	0.86	0.36		
	SNP/100 bp					$\Sigma^a = 14$	
pic21							
F7		1.17	0.70	0.70	0.70	3.29	GS
FAP1360A	1.00		0.47	0.94	0.70	3.99	
D21	1.00	1.00		0.47	0.23	3.29	
D408	0.93	0.93	0.93		0.47	3.52	
D32	0.93	0.93	0.93	0.87		3.52	
D145	0.13	0.13	0.13	0.07	0.07		
	SNP/100 bp					$\Sigma^a = 15$	

[a] Total number of SNPs between all six investigated inbred lines

hohenheim.de/~ipspwww/soft.html). Analyses of variance for randomised complete block designs were used to obtain a mean value for each RGA sequence. Least significant differences were calculated to test for differences between individual entry means. To determine SNP frequencies, pairwise comparisons between each two of the six inbred lines were performed. CAPS frequencies were calculated in pairwise comparisons of the three parental pairs of the mapping populations PopA, PopB, and PopC.

Genetic similarity of the RGA sequences between the six investigated inbred lines was measured on the basis of SNPs (Table 2) to test the six inbred lines for clustering into the heterotic flint and dent groups or resistant and susceptible lines using the software package NTSYSpc Version 2.02i (Rohlf 1989). The data set consisted of 58 detected SNPs between all six inbred lines for pic19, 14 for pic13, and 15 for pic21. Graphical genotypes of pic19 were displayed by software package GGT:GraphicalGenoTyping (http://www.spg.wau.nl/pv/pub/ggt).

Results

Isolation of pic13, pic19, and pic21 sequences from six inbred lines

All six RGA alleles of pic19 and pic21 were identical in size and similar to the sequences published by Collins et al. (1998). All pic19 sequences contained continuous open reading frames (ORFs), except those for D145 and D408. Likewise, pic21 sequences of all six inbreds showed continuous ORFs, except F7. The sequence of pic13 was incomplete, i.e. it was sequenced from both ends without identifying overlaps. pic13 showed a continuous ORF for both partial sequences of FAP1360A, D32, and F7. Lines D21, D408, and D145 displayed at least one stop codon.

Frequency of SNPs and probability of conversion into RGA-CAPS

The overall mean for the six genotypes over the three RGAs was 3.05 SNPs per 100 bp (= 1 SNP per 33 bp) of the RGA sequence. SNP frequencies of pic19 (4.70 SNPs per 100 bp) were significantly higher ($P < 0.01$) than those of pic13 (2.86 SNPs per 100 bp) and pic21 (1.61 SNPs per 100 bp). The number of CAPS sites per 100 bp of the RGA sequence was measured for the three parental pairs of the mapping populations FAP1360A/F7, D32/D145, and D21/D408. The number of CAPS varied between 0.00 and 4.15 per 100 bp, with an overall mean of 2.20 CAPS sites per 100 bp. CAPS frequencies of pic19 (3.72) were significantly higher ($P < 0.01$) than those of pic13 (1.69) and pic21 (1.18). PopA revealed the highest degree of polymorphism (3.13), followed by PopC (2.18), and PopB (1.82).

Genetic similarity analyses revealed no clustering into the heterotic flint and dent groups or resistant and susceptible lines (Fig. 1). The graphical haplotype for pic 19 was constructed on the basis of the same SNP data for all six inbred lines, revealing conserved sequence blocks.

Development of CAPS markers

CAPS polymorphisms were identified for all investigated RGAs in at least one of the three mapping populations. Sequence alignment between the cloned pic13 fragments revealed a SNP within a *Rsa*I recognition site resulting in polymorphism between D32 and D145

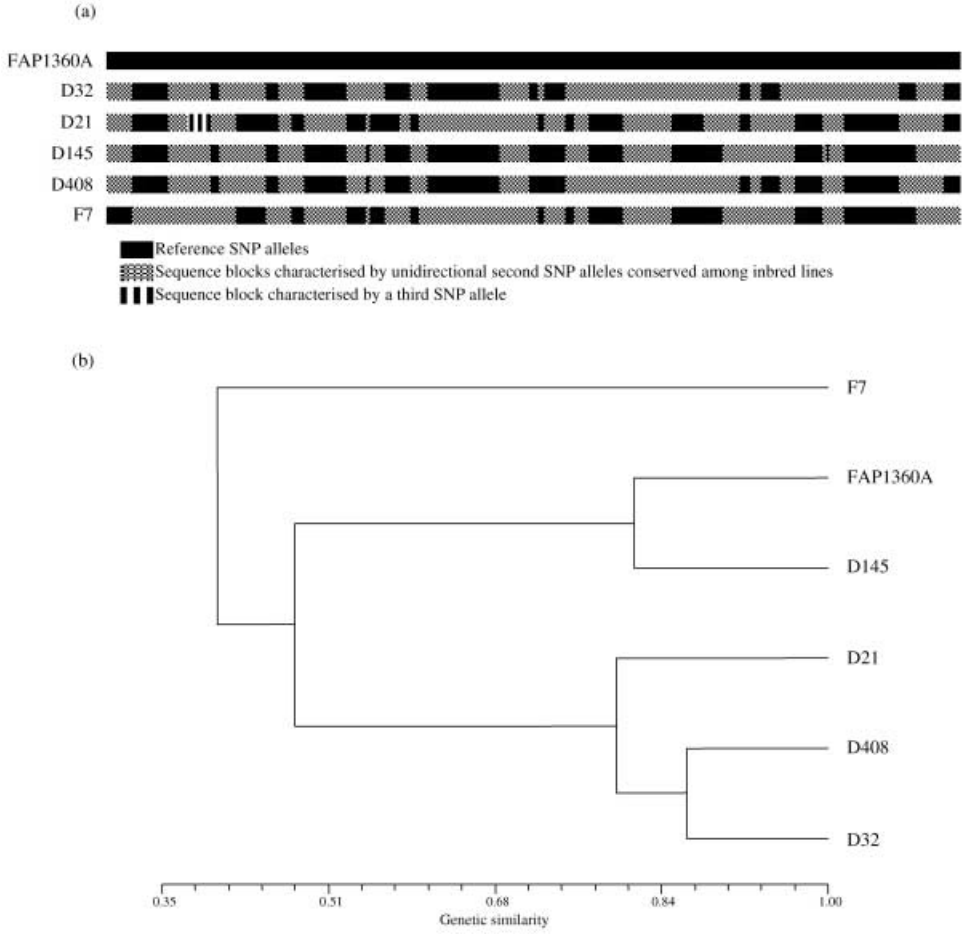

Fig. 1 (**a**) Graphical haplotyping for the example of pic19 (624 bp) in six European inbred lines based on 58 SNPs between the six maize inbred lines FAP1360A, D32, D21, D145, D408 and F7. (**b**) Genetic similarities among the six inbred lines based on pic19 SNPs

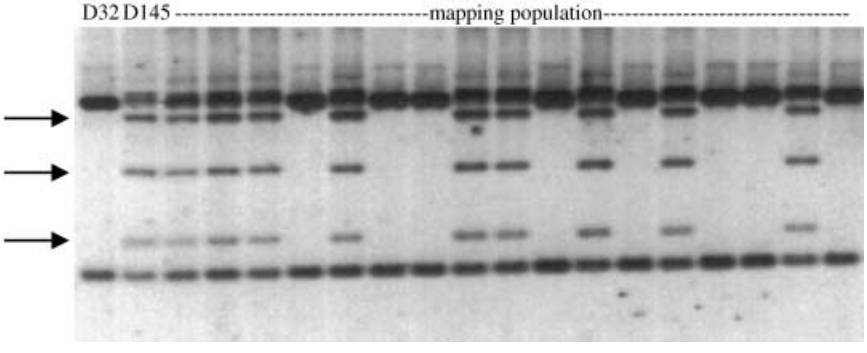

Fig. 2 *Rsa*I restriction pattern of pic13 PCR products from parental inbred lines D32 (SCMV resistant) and D145 (SCMV susceptible) and part of PopA

(PopA) (Fig. 2) and between D21 and D408 (PopC) (data not shown). The sequences of pic21 were polymorphic between D32 and D145 at a *Mbo*II recognition site. Digestion of elongated pic19 amplification products with *Nla*III generated polymorphisms in all three parental pairs (data not shown).

Genetic mapping of pic13, pic21, and pic19 relative to *Scmv1* and *Scmv2*

Amplification of pic13 from oat-maize addition lines showed a PCR product of the expected size only in the oat line carrying an additional maize chromosome 3. Amplification of pic19 and pic21 did not result in differential amplification of the oat line carrying the respective additional maize chromosome. For PopA, pic13 was mapped into the interval of *umc102/csu285b* on chromosome 3L (Fig. 3), which completely overlaps with the *Scmv2*-QTL region previously identified by Xia et al. (1999). However, pic13 could not be mapped in PopB because of identical banding patterns between heterozygotes and the recurrent parent F7 for all tested restriction enzymes.

Using PopA, pic21 was located on chromosome 3L but, contrary to pic13, it was located outside the *Scmv2*

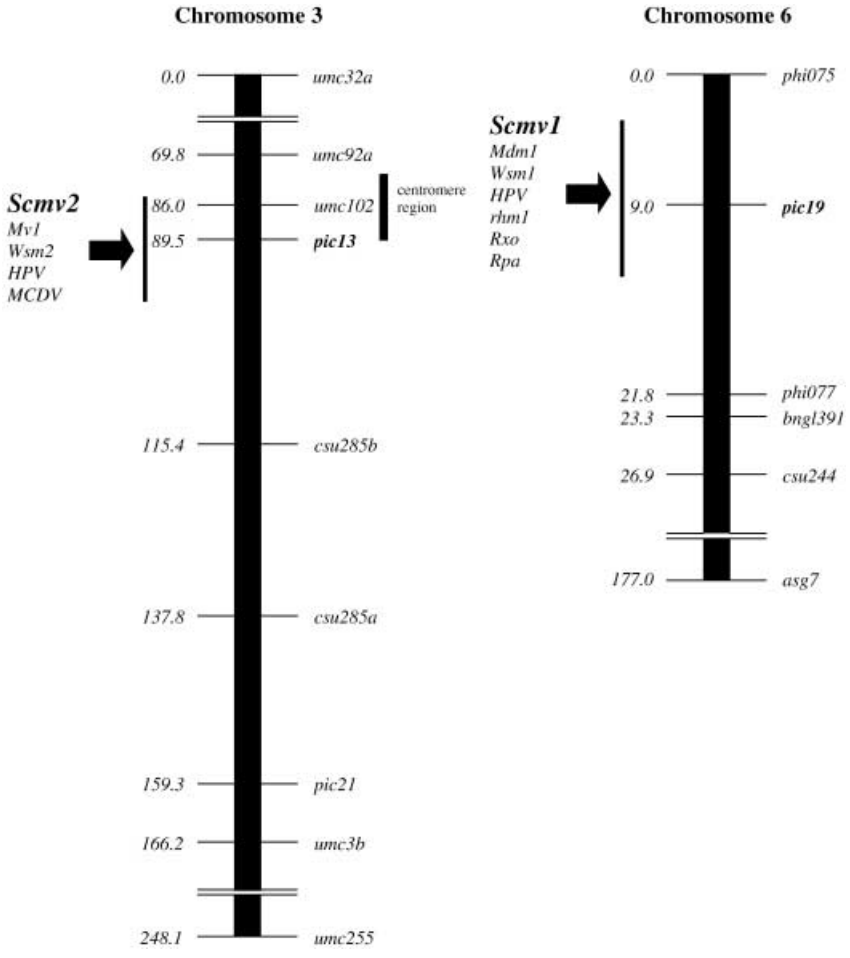

Fig. 3 Genetic map of genome regions conferring resistance to SCMV on maize chromosomes 3 and 6 derived from D32 × D145 $F_{2:3}$ mapping populations (PopA). Chromosome 3 containing RGAs pic13 and pic21, and genes for resistance to sugarcane mosaic virus (*Scmv2*), maize mosaic virus (*Mv1*), wheat streak mosaic virus (*Wsm2*), high plains virus (*HPV*) and maize chlorotic dwarf virus (*MCDV*). Chromosome 6 containing pic19 and genes for resistance to sugarcane mosaic virsu (*Scmv1*), wheat streak mosaic virus (*Wsm1*), high plains virus (HPV), southern corn leaf blight (*rhm1*), rice bacterial streak (*Rxo*) and sorghum bacterial stripe (*Rpa*)

Table 3 Evaluation of restriction pattern of pic19 and pic13 in resistant, susceptible, partially resistant, and ancestral inbred lines

Inbred lines		pic13		pic19		
		Genotype A	Genotype B	Genotype A	Genotype B	Genotype C
Resistant	FAP1360A		X	X		
	D21	X		X		
	D32	X		X		
Susceptible	F7		X		X	
	D408	X				X
	D145	X			X	
Partially resistant	D06	X		X		
	D09	X		X		
	FAP1396A	X		X		
Ancestors	A632	X		X		
	V3	X			X	
	WD	X			X	
	W401	X				X
	Co158	X			X	
	Co125		X		X	
	FAP493B	X			X	
	FAP954A	X		X		

region flanked by markers *csu285a* and *umc3b* (Fig. 3). In PopA, pic19 mapped to maize chromosome 6S between markers *phi075* and *phi077b* (Fig. 3). In PopB, pic19 mapped into the marker interval *umc1023/E2M7-1* with a map distance of 3.5 cM to *Scmv1*.

RGA analyses in ancestral lines and partially resistant inbreds

Two genotypes were observed when pic13 amplification products were digested with *Rsa*I. F7, FAP1360A, and Co125, the direct ancestor and potential donor of *Scmv2*,

showed the same restriction banding pattern (genotype B). The remaining inbred lines displayed a different restriction pattern (genotype A) (Table 3).

Three genotypes were observed for pic19 when digested with *Nla*III. Besides the three resistant lines FAP1360A, D21, and D32, the potential donors of *Scmv1*, A632 and FAP954A, and the three partial resistant inbreds D06, D09, and FAP1396A, as well as FAP1360A, showed the same restriction pattern (genotype A). The second restriction banding pattern (genotype B) was observed for the following six lines: V3, WD, Co158, Co125, FAP493B, and F7. The remaining two lines D408 and W401 lack one *Nla*III restriction site and thus revealed a third restriction pattern (genotype C) (Table 3). In contrast to pic13, the resistant lines revealed the same restriction pattern.

Discussion

One of the long-term aims of molecular marker technology in plant breeding is the selection of superior individuals directly at the level of DNA. RGAs provide an excellent source for the development of molecular markers, especially for resistance traits, because of their high level of polymorphism and their putatively functional character. Thereby, RGAs can be converted to single-copy PCR markers like CAPS (Konieczny and Ausubel 1993), while RGA-RFLPs frequently result in complex banding patterns because of sequence homology to related RGA sequences (Collins et al. 1998). Diagnostic markers like allele-specific RGA-CAPS will have wider applications in MAS strategies in the future.

SNP and CAPS frequency in three maize RGAs

Rafalski et al. (2001) analysed random cDNA clones in a collection of over 30 public and private maize lines representative for the North American corn germplasm. This analysis was restricted to coding regions. Sequence alignment revealed one single nucleotide change per 70 bp among the 30 lines. However, the authors pointed out that pairwise comparison between any two lines – such as in the results presented in this article – reveals a lower degree of polymorphism. Useche et al. (2001) performed similar experiments in maize and reported a SNP frequency of only one SNP per 138 bp in coding regions. Another study revealed that maize has an average of one SNP every 104 bp between two randomly sampled sequences (Tenaillon et al. 2001). Hence, the frequency of one SNP per 33 bp originating from single nucleotide changes in pairwise comparisons is significantly higher in the present study. This result confirms the highly polymorphic character of plant RGAs and possibly R genes, which is an important feature in evolutionary processes at R gene loci and especially R gene clusters (Parniske and Jones 1999).

Among the RGAs, pic19 displayed a significantly higher degree of polymorphism than pic13 and pic21 (Table 2); pic19 might be phylogenetically older than pic13 and pic21 and, therefore, displays a significantly higher degree of polymorphism. A second possibility would be the coding character of pic13 and pic21 in contrast to pic19. Since a SNP in the first two positions within a triplet generates more amino-acid changes, the mutations in coding regions should be mainly in the third position of the triplet. Hence, the significantly higher SNP frequency at the third position of triplets for pic19 suggests expression of the respective sequence, but was not observed in the current study.

The generally high degree of polymorphism of the three RGAs might also be influenced by their map position. Two of the three RGAs map to potential R gene clusters on chromosomes 3L close to the centromere and 6S close to the nucleolus organiser region (nor). In gene clusters the recombination frequency is expected to be very high because of the high density of coding sequences. In many plant species, recombination rates can vary up to an order of magnitude over relatively small intervals (reviewed by Schnable et al. 1998). Detailed physical and genetic mapping of grass genomes revealed the clustering of genes, and supports the hypothesis that much of the meiotic recombination occurs in genes and most recombination events are restricted to few chromosome regions containing gene clusters (Gill et al. 1996; Künzel et al. 2000). The maize genome exhibits a very striking gene distribution with almost all genes present in 10–20% of the genome (Carels et al. 1995). In fact, genes exhibit recombination rates 10 to 100-fold higher than the genome average (reviewed by Lichten and Goldman 1995). In maize, one-fifth of all recombination events in a 140-kb interval between the *anthocyaninless1* (*a1*) and the *shrunken2* (*sh2*) genes were resolved within a 377-bp region of the *a1* gene (Xu et al. 1995). It is not known whether the *Scmv* regions on chromosomes 6 and 3 contain R gene clusters or single pleiotropically acting genes. Zhao et al. (2001) mapped two maize R genes, *Rxo* and *Rpa*, to the same position on the short arm of maize chromosome 6. The authors identified 5–6 RGAs in the direct neighbourhood of *Rxo* and *Rpa*. These findings and the high degree of polymorphism for pic19 support the presence of a R gene cluster on chromosome 6S. However, one member of the putative R gene cluster on chromosome 6S, *Mdm1*, shows 100% linkage with nor (Simcox et al. 1995), whereas *Scmv1* and nor were not absolutely linked (Xu et al. 1999). Very low recombination rates are typically exhibited by regions surrounding the centromeres and the nor. Therefore, the putative R gene cluster on chromosome 6S is located in the direct neighbourhood of regions with suppressed recombination. Interestingly, the situation seems to be similar on chromosome 3 with *Scmv2* in the direct neighbourhood of the centromere.

The structure of plant materials (flint-dent, SCMV resistant-susceptible; relationship by descent of the three resistant lines) was not reflected by the pic19-based genetic similarity analysis (Fig. 1). Explanations for the lack of grouping of the six inbred lines are either a miss-

ing correlation of the RGA allele sequences with these factors or the above mentioned high degree of polymorphism within RGAs. Furthermore, Xu et al. (2000) identified a 7.2 cM interval containing eight molecular markers as well as pic19 flanking *Scmv1* conserved across the three resistant lines D21, D32, and FAP1360A, whereas the susceptible lines displayed a different segregation pattern. Therefore, this 7.2 cM chromosomal segment seemed to be identical in the three resistant lines. However, disagreement of RGA-based genetic similarity analysis with phenotypic, pedigree, and marker data question the utility of SNPs for association or disequilibrium mapping (reviewed in Lazzeroni 2001), at least for this genome region.

Single nucleotide changes in all three RGAs between all six inbred lines generated only two of the possible four SNP alleles except one. The formation of the second SNP alleles was not randomly distributed. Several consecutive SNPs were ordered in sequence blocks (haplotypes) with unidirectional mutations. Most of the sequence blocks contain more than one single SNP (Fig. 1a). Presuming a low number of RGA alleles in European founder materials, the driving force for the development of the high degree of polymorphism and variability within these RGAs seems to be recombination.

In conclusion, RGA-derived markers are especially interesting because of their genomic localisation within R gene clusters which putatively contain numerous genes relevant for plant breeding purposes. The frequencies of SNPs and CAPS are promising prerequisites for conversion of even short RGA sequences into molecular markers. This should further enhance the mapping of RGAs as potential candidates for genes conditioning resistance to pathogens in plant species.

Candidate gene evaluation

The identification of CAPS restriction sites between the parents of three mapping populations facilitated mapping of the three RGAs, pic13, pic19, and pic21, to genomic regions known to be involved in the inheritance of SCMV resistance on chromosomes 3 and 6. To address the question of whether any of the RGAs is a likely candidate for one of the target genes *Scmv1* or *Scmv2*, respectively, we evaluated the following criteria: (1) *Linkage with the target gene*. A potential role as a candidate gene can be ruled out in case of a large map distance to the target gene. Cosegregation and even tight linkage to the R gene indicates a functional role of the RGA because of: (i) incomplete penetrance of SCMV resistance and escapes resulting in mis-scoring, as well as (ii) the potential presence of more than one SCMV R gene within both regions. (2) *Comparison of the R alleles of the six inbred lines with those of the ancestor and partially resistant lines*. Sharing the same allele among resistant, partially resistant inbred lines and putative R gene donors on the one hand and different alleles in the other lines on the other hand, indicates a functional role in the inheritance of SCMV resistance. (3) *Continuous ORFs as a prerequisite for coding regions*. A continuous ORF spanning the complete RGA sequence strongly suggests a coding character on the basis of the statistical probability of the occurrence of stop codons.

Chromosome 6

Mapping in PopA located pic19 within the QTL peak of *Scmv1*. Mapping in PopB showed two recombinant individuals. However, field experiments, BSA (Xu et al. 1999), and QTL analyses (Xia et al. 1999; Dußle et al. 2000) did not exclude the presence of more than one SCMV R gene in the *Scmv1* region. Furthermore, incomplete penetrance of virus resistance and escapes might result in mis-scorings during the phenotypic evaluation of the mapping populations. Therefore, a functional role of pic19 in the expression of SCMV resistance cannot be ruled out by genetic mapping. Moreover, the three resistant, the partially resistant lines, as well as the putative donors of *Scmv1*, A632, and FAP954A, displayed the same restriction patterns for pic19 (Table 3). Xu et al. (2000) investigated the chromosomal regions harbouring *Scmv1* and *Scmv2* by AFLP analyses in the same three resistant European inbreds, their ancestral lines and partially resistant inbred lines. The banding patterns indicated the identity of a 7.2 cM window harbouring the *Scmv1* locus in D21, D32, and FAP1360A. Therefore, it is most likely that these three European inbred lines share the same *Scmv1* allele at this Mega-locus. Absence of stop codons in all three resistant lines strengthens the possible functional involvement in the expression of SCMV resistance. However, a continuous ORF is not restricted to the resistant lines because the susceptible inbred F7 also showed a continuous ORF. In conclusion, pic19 remains a candidate for *Scmv1*, but our findings cannot distinguish between close linkage and identity of pic19 with *Scmv1*.

Chromosome 3

In PopA, pic21 mapped to chromosome 3, and was located approximately 50 cM outside the marker interval flanking *Scmv2*. Therefore, it was excluded as a candidate gene for *Scmv2*.

In contrast, pic13 mapped in PopA within the QTL peak of *Scmv2*. According to Xu et al. (2000), the *Scmv2* region in FAP1360A most likely originated from Co125, while all other ancestral lines showed AFLP patterns in this region different from FAP1360A. The authors proposed the following explanations for the origin of the *Scmv2* region: (1) *Scmv2* in FAP1360A originates from Co125, or (2) Co125 is also the *Scmv2* donor of D32 displaying a very short donor region not detectable by the employed flanking markers or another inbred line. The fact that restriction patterns of the potential donor of *Scmv2*, Co125, and FAP1360A, are identical and show a

different allele at this locus than D21 and D32, confirms the previously published hypothesis of different *Scmv2* genes in FAP1360A and D32 (Xu et al. 2000). The presence of *Scmv1* seems to be sufficient for conferring partial resistance. Hence, *Scmv2* is not necessarily expected to be present in the partially resistant lines, explaining a different genotype in all three partially resistant inbred lines. No continuous ORF was found for pic13 in D21, in agreement with the lack of evidence for *Scmv2* in this inbred line.

The mapping results and analyses of ancestor lines cannot distinguish between identity and close linkage of pic19 and *Scmv1*, and pic13 and *Scmv2*, respectively. The previously mentioned results of Zhao et al. (2001), as well as studies on the *Pto*, *Cf*, and *Dm* clusters in tomato and lettuce (Martin et al. 1994; Thomas et al. 1997; Meyers et al. 1998), revealed the existence of numerous RGAs in the direct physical neighbourhood of the R genes. At least seven apple *Vf* gene analogues have been identified from a BAC contig of 290 kb, encompassing the *Vf* locus (Xu 2001, personal communication). These RGAs showed very similar amino-acid domains except for the LRRs, where deletion of several LRR units and point mutations occurred frequently. Therefore, even in the case of non-identity of pic19 and pic13 with *Scmv1* and *Scmv2*, respectively, they should provide excellent starting points for a map-based approach for cloning of the target genes themselves as well as other members of these clusters, such as MDMV, wheat streak mosaic virus (WSMV), maize mosaic virus (MMV), High Plains virus (HPV), and maize chlorotic dwarf virus (MCDV) R genes. Large and often continuous ORFs of the RGAs investigated in this study indicate that they are part of coding sequences. This is in agreement with large-scale sequencing of the above mentioned tomato and lettuce R gene loci. The majority of RGAs seemed to be functional and expressed, with only a few of them clearly being pseudogenes. Screening of DNA libraries will extend the number of RGAs in these regions and, therefore, broaden the probability of identifying the target genes. In contrast to the simple procedures to clone homologues, establishing potential functions of the RGAs remains challenging. Since the proof of function of a given candidate for a certain resistance is highly dependent on pathogen-specific plant material, exchange of R gene candidates between research groups working on different resistances is essential and can greatly accelerate the cloning of R genes by the RGA approach.

Acknowledgements The present study was supported by a grant from the Deutsche Forschungsgemeinschaft, Grant No. LU601/3. The authors gratefully acknowledge the mapping of SSRs to PopB by Susanne Dreissigacker and the skilled technical assistance of Elisabeth Kokai-Kota.

References

Ananiev EV, Riera-Lizarazu O, Rines HW, Phillips RL (1997) Oat-maize chromosome addition lines: a new system for mapping the maize genome. Proc Natl Acad Sci USA 94:3524–3529

Bent AF (1996) Plant disease resistance genes: function meets structure. Plant Cell 8:1757–1771

Carels N, Barakat A, Bernardi G (1995) The gene distribution of the maize genome. Proc Natl Acad Sci USA 92:11,057–11,060

Collins NC, Webb CA, Seah S, Ellis JG, Hulbert SH, Pryor A (1998) The isolation and mapping of disease resistance gene analogues in maize. Mol Plant–Microbe Interact 11:968–78

Collins NC, Drake J, Aycliffe M, Sun Q, Ellis J (1999) Molecular characterisation of the maize *Rp1-D* rust resistance haplotype and its mutants. Plant Cell 11:1365–1376

Dußle CM, Melchinger AE, Kuntze L, Stork A, Lübberstedt T (2000) Molecular mapping and gene action of *Scm1* and *Scm2*, two major QTLs contributing to SCMV resistance in maize. Plant Breed 119:299–303

Fuchs E, Grüntzig M (1995) Influence of sugarcane mosaic virus (SCMV) and maize dwarf mosaic virus (MDMV) on the growth and yield of two maize varieties. J Plant Dis Protection 102:44–50

Gill KS, Gill BS, Endo TR, Taylor T (1996) Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. Genetics 144:1883–1891

Hammond-Kosack KE, Jones JDG (1996) Resistance gene-dependent plant defense responses. Plant Cell 8:1773–1791

Hoisington DA, Khairallah MM, Gonzales-de-Leon D (1994) Laboratory protocols. CIMMYT Applied Molecular Genetics Laboratory, Mexico, DF

Konieczny A, Ausubel FA (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

Kuntze L, Fuchs E, Grüntzig M, Schulz B, Klein D, Melchinger AE (1997) Resistance of early maturing European maize germplasm to sugarcane mosaic virus (SCMV) and maize dwarf mosaic virus (MDMV). Plant Breed 116:499–501

Künzel G, Korzun L, Meister A (2000) Cytogenetically integrated physical restriction fragment length polymorphism maps for barley genome based on translocation breakpoints. Genetics 154:397–412

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an Interactive computer package for constructing primary genetic linkage maps for experimental and natural populations. Genomics 1:174–181

Lazzeroni (2001) A chronology of fine-scale gene mapping by linkage disequilibrium. Stat Methods Med Res 10:57–76

Lichten M, Goldman AS (1995) Meiotic recombination hotspots. Annu Rev Genet 29:423–444

Martin GB, Frary A, Wu T, Brommonschenkel S, Chunwongse J, Earle ED, Tanksley SD (1994) A member of the tomato *Pto* gene family confers sensitivity to fenthion resulting in rapid cell death. Plant Cell 6:1543–1552

McMullen MD, Simcox KD (1995) Genomic organisation of disease and insect resistance genes in maize. Mol Plant–Microbe Interact 8:811–815

Melchinger AE, Kuntze L, Gumber RK, Lübberstedt T, Fuchs E (1998) Genetic basis of resistance to sugarcane mosaic virus in European maize germplasm. Theor Appl Genet 96:1151–1161

Meyers BC, Chin, DB, Shen, KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, Michelmore RW (1998) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell 10:1817–1831

Parniske M, Jones JD (1999) Recombination between diverged clusters of the tomato *Cf-9* plant disease resistance gene family. Proc Natl Acad Sci USA 96:5850–5

Rafalski JA, Ching A, Bhattramakki D, Morgante M, Dolan M, Register JC, Smith OS, Tingey S (2001) SNP markers in maize: discovery and applications. Plant and Animal Genome IXth Conference, p 44

Rohlf FJ (1989) NTSYS-pc numerical taxonomy and multivariate analysis system. Setauket, Exeter Publishing Company Limited

Schnable PS, Hsia AP, Nikolau BJ (1998) Genetic recombination in plants. Curr Opin Plant Biol 1:123–129

Simcox KD, McMullen MD, Louie R (1995) Co-segregation of the maize dwarf mosaic virus resistance gene, *Mdm1*, with the nucleolus organiser in maize. Theor Appl Genet 90:341–346

Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF, Gaut BS (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). Proc Natl Acad Sci USA 98:9161–9166

The International SNP Map Working Group (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933

Thomas CM, Jones DA, Parniske M, Harrison K, Balint-Kurti PJ, Hatzixanthis K, Jones JDG (1997) Characterisation of the tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *Cf-4* and *Cf-9*. Plant Cell 9:2209–2224

Useche F, Morgante M, Hanafey M, Tingey S, Gao GR, Rafalski JA (2001) Computer detection of single nucleotide polymorphisms (SNPs) in maize ESTs. Plant and Animal Genome IXth Conference, p 144

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lübberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Xia XC, Melchinger AE, Kuntze L, Lübberstedt T (1999) Quantitative trait loci mapping of resistance to sugarcane mosaic virus in maize. Phytopathology 89:660–667

Xu ML, Melchinger AE, Xia XC, Lübberstedt T (1999) High-resolution mapping of loci conferring resistance to sugarcane mosaic virus in maize using RFLP, SSR, and AFLP markers. Mol Gen Genet 261:574–581

Xu ML, Melchinger AE, Lübberstedt T (2000) Origin of *Scm1* and *Scm2* – two loci conferring resistance to sugarcane mosaic virus (SCMV) in maize. Theor Appl Genet 100:934–941

Xu X, Hsia AP, Zhang L, Nikolau BJ, Schnable PS (1995) Meiotic recombination breakpoints resolve at high rates at the 5′end of a maize coding sequence. Plant Cell 12:2151–2161

Zhao B, Lin X, Leach JE, Hulbert S (2001) Molecular analysis of the maize gene conferring a non-host hypersensitive reaction to rice bacterial streak. Plant and Animal Genome IXth Conference, p 90