Theor Appl Genet (2003) 106:1171–1177
DOI 10.1007/s00122-002-1105-z

M. Quint · C. M. Dußle · A. E. Melchinger
T. Lübberstedt

Identification of genetically linked RGAs by BAC screening in maize and implications for gene cloning, mapping and MAS

Received: 4 March 2002 / Accepted: 7 May 2002 / Published online: 25 January 2003
© Springer-Verlag 2003

Abstract The resistance gene analogue (RGA) pic19 in maize, a candidate for sugarcane mosaic virus (SCMV) resistance gene (R gene) *Scmv1*, was used to screen a maize BAC library to identify homologous sequences in the maize genome and to investigate their genomic organisation. Fifteen positive BAC clones were identified and could be classified into five physically independent contigs consisting of overlapping clones. Genetic mapping clustered three contigs into the same genomic region as *Scmv1* on chromosome 6S. The two remaining contigs mapped to the same region as a QTL for SCMV resistance on chromosome 1. Thus, RGAs mapping to a target region can be successfully used to identify further-linked candidate sequences. The pic19 homologous sequences of these clones revealed a sequence similarity of 94–98% on the nucleotide level. The high sequence similarity reveals potential problems for the use of RGAs as molecular markers. Their application in marker-assisted selection (MAS) and the construction of high-density genetic maps is complicated by the existence of closely linked homologues resulting in 'ghost' marker loci analogous to 'ghost' QTLs. Therefore, implementation of genomic library screening, including genetic mapping of potential homologues, seems necessary for the safe application of RGA markers in MAS and gene isolation.

Keywords RGA · SCMV · Maize · pic19 · Ghost marker · BAC

Communicated by G. Wenzel

M. Quint · C. M. Dußle · A. E. Melchinger · T. Lübberstedt
Institute of Plant Breeding, Seed Science,
and Population Genetics, University of Hohenheim,
Fruwirthstraße 21, 70593 Stuttgart, Germany,

T. Lübberstedt (✉)
Danish Institute of Agricultural Sciences,
Research Centre Flakkebjerg, DK-4200 Slagelse, Denmark,
e-mail: Thomas.Luebberstedt@agrsci.dk
Fax: ++45-58-11-33 01

Introduction

Plant pathogens exhibit a high mutation rate from avirulence to virulence, and natural selection favours the spread of new virulent races. In response to this, there is a co-evolution in plants that generate novel resistance protein variants, which are capable of recognising either the modified avirulence determinant or other pathogen components (Hammond-Kosack and Jones 1997). One mechanism for this is the clustering of R genes within large complex loci. Recombination between clustered loci leads to new pathotype-specific resistances. In a classical study with maize, Saxena and Hooker (1968) identified genes conferring resistance to *Puccinia sorghi*. They found 16 linked, but separate, loci conferring resistance to 16 different pathogen isolates (*Rp1A* to *Rp1N, Rp5, Rp6*). Another mode of clustering of R genes is displayed by linkage groups containing genes for recognition of different pathogens. In lettuce, linkage group I contains eight genes for resistance to downy mildew (*Bremia lactucae*) and a gene for aphid resistance (*Pemphigus bursarius*), whereas in linkage group II two additional downy mildew R genes are clustered with a gene for turnip mosaic virus resistance and a gene for resistance to the root pathogen *Plasmopara lactucae-radicis* (Landry et al. 1987; Kesseli et al. 1993; Witsenboer et al. 1995; Meyers et al. 1998a).

The recent cloning of a number of R genes by transposon tagging and positional cloning uncovered sequence homology in conserved amino-acid domains. This has been exploited to isolate hundreds of sequences with homology to R genes, RGAs, based on degenerate PCR primers derived from the conserved amino-acid domains. RGAs belonging to the same gene cluster displayed an even higher degree of sequence homology than unlinked RGAs (Ellis et al. 1995; Parniske et al. 1997; Meyers et al. 1998a), which can be explained by the evolution of many R genes by duplication. Circumstancial evidence would suggest that all organisms have experienced at least one round of genome duplication in their phylogenetic past. Thus all eukaryotes probably are

ancient polyploids (Leipoldt and Schmidtke 1982). Therefore, in the plant kingdom a high degree of homology between unlinked RGAs is also very likely. This has been shown for Arabidopsis (Meyers et al. 1999) and may also be valid for maize with its potential allotetraploid evolution (Gaut et al. 1997). Therefore, the maize genome contains duplicated chromosome segments with colinear gene arrangements (Helentjaris et al. 1995).

On maize chromosomes 6 and 3, R gene clusters include genes conferring resistance to several virus and bacterial diseases. Chromosome 6S contains R genes for SCMV, maize dwarf mosaic virus, wheat streak mosaic virus, high plains virus, rice bacterial streak, sorghum bacterial stripe and Southern corn leaf blight, whereas the cluster on chromosome 3L contains genes conferring resistance to SCMV, wheat streak mosaic virus, high plains virus, maize mosaic virus and maize chlorotic dwarf virus (for a review see Quint et al. 2000). There is strong evidence for the clustering of R genes on chromosomes 6 and 3, but the issue of pleiotropy versus clustering still needs to be solved for both regions.

Collins et al. (1998) isolated and mapped several RGAs to the maize genome. One of them, pic19, was mapped as a single-copy RGA to the *Scmv1* region on chromosome 6, and more intense investigations in SCMV-specific materials suggested pic19 to be a candidate gene for *Scmv1* (Quint et al. 2002). Hence, RGAs cosegregating or linked to a target R gene may provide a tool for the isolation of further linked RGAs based on homology to members of the same gene cluster. R gene and RGA clusters have the potential to be highly duplicated and span large physical distances (Meyers et al. 1998a) which might even be interrupted by other sequences. Therefore, the isolation of overlapping BAC/YAC clones spanning flanking markers by chromosome walking may be time-consuming and ineffective. In contrast, the approach outlined in this study fosters isolation of further candidate genes.

The objectives of this study were to: (1) identify homologues of pic19, (2) investigate their genomic organisation, (3) determine their map position to clarify whether they map to (a) the direct neighbourhood within the same cluster, (b) independent loci or (c) duplicated genome regions in the maize genome, and (4) evaluate the use of RGAs for genetic mapping, MAS, and gene cloning on the basis of these results.

Materials and methods

BAC screening

Scmv1 candidate pic19 was previously elongated to 624 bp (Quint et al. 2002). It was used as a probe to screen a B73 BAC library, purchased from Texas A & M BAC Center (http://hbz.tamu.edu/bac.html), covering the genome approximately 4 x. Labelling was performed by α-32P-dATP/dCTP random priming. Hybridisation and washing procedures of BAC membranes were conducted at 65 °C as proposed by the supplier (http://hbz.tamu.edu/bac.html). Autoradiography was carried out for 1 to 2 days at –80 °C with an intensifying screen.

Table 1 Primer sequences used for genetic mapping of BAC contigs 1 to 4

BAC contig	Primer	5′–3′
1	71L1	GGA AGC ATA TTG TCG TTG T
	71R1	GCA TGC TCC GTC GTA TG
2	pic19L	TAG ATG ATG TCT GGA CGG CT
	230R1	GGC ACA ATA CAG GGA A
3	239L2	AGC CCT TGT GCC AAT AA
	pic19X	GCA GTT CCT CTC TGC AAC GTG
4	UBC860	$(TG)_8$ RA

Restriction analysis

For identification of overlapping BAC clones and classification of independent contigs, BAC DNA was isolated using the Macherey & Nagel (Düren) Nucleobond BAC 100 Kit. 2µg of BAC DNA was digested with 10 U of either *Bam*HI or *Hin*dIII in a 30 µl reaction for 4 h at 37 °C. Digested DNA was separated on a 0.8% agarose gel. The presence or absence of bands of the same size was converted into a 1/0 matrix to identify overlapping BAC clones by scoring for fragments of common size. This process was performed for both restriction enzymes independently. To identify false positives to be excluded from further experiments, gels were blotted and hybridised with pic19 following the same protocols as mentioned above.

Cloning of pic19 homologues

Specific primers for pic19 (pic19L + pic19X, Table 1) were used to amplify pic19 homologous sequences out of each BAC clone. DNA amplification was performed with a standard reaction mix containing 10 ng of BAC 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 DNA was amplified using 35 cycles under the following conditions: 1 min at 94 °C, 2 min at 55 °C and 2 min at 72 °C. Final extension was conducted at 72 °C for 2 min. Amplification products were excised from the agarose gel, extracted with the Nucleospin Kit (Macherey & Nagel, Düren) and ligated into pGEM-T (Promega, Mannheim) plasmids. 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). Sequence alignment was performed using the ALIGN Plus 2.0 software package (http://www.scied.com/ses_alim.htm).

Mapping populations

Since the BAC library was derived from maize inbred line B73, all BAC contigs except one were mapped using the commercially available B73 × Mo17 RI (IBM) mapping population (PopA) consisting of a subset of 94 individuals. PopA contains more than 570 RFLP and SSR markers (Davis et al. 2000; http://www.cafnr.missouri.edu/mmp/ibmmaps.htm). The mapping population is derived from a cross of B73 × Mo17, which underwent random mating for four generations.

The last contig was mapped to a subset of 84 individuals of a D32 × D145 $F_{2:3}$ mapping population (PopB) containing 100 RFLP and SSR markers (Xia et al. 1999).

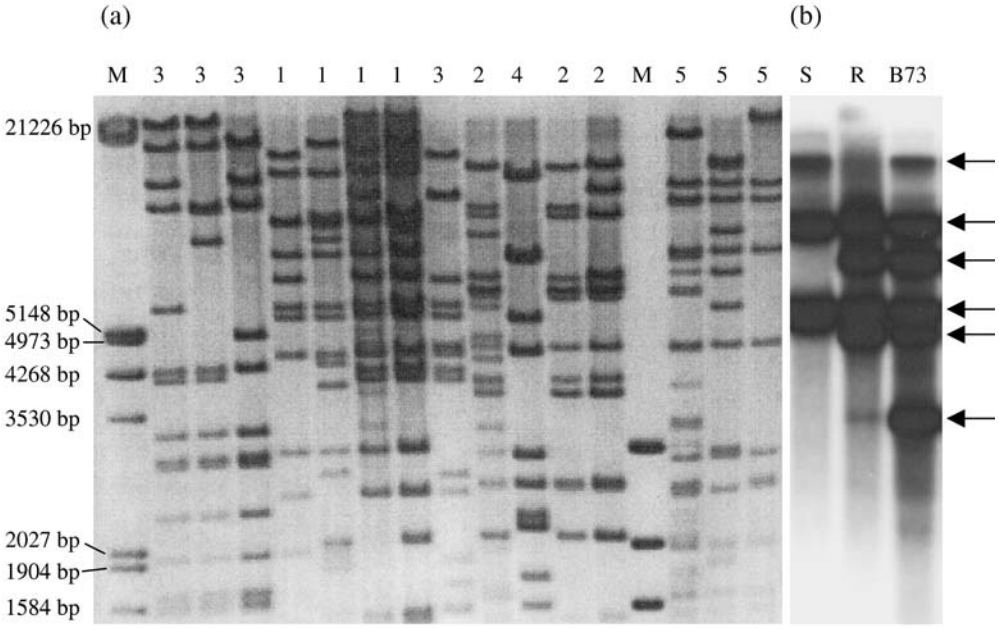

Fig. 1 (a) Identification of overlapping BAC clones and classification by *Hin*dIII restriction analysis into contigs 1–5; (b) Autoradiogram of *Hin*dIII-digested DNA of SCMV susceptible maize inbred line F7 (*S*), SCMV resistant maize inbred line FAP1360A (*R*) and B73 using pic19 as a probe

Mapping strategy and BAC mapping

The BAC mapping strategy included two steps: (1) the design of BAC contig-specific primers (Table 1) based on the DNA sequence of the pic19 homologues; (2) amplification of contig-specific PCR products discriminating both parents of the PopA, B73 and Mo17 (PCR conditions a, b). In the case of identical size of PCR fragments, pic19 homologues were converted to CAPS markers (c). Independent from the pic19 homologues, contig-specific amplification of inter-simple sequence repeat (ISSR) fragments in PopA (d) or contig-specific single-copy RFLPs of BAC subclones in PopB (e) were employed as alternative strategies for BAC mapping.

DNA amplification was performed in a standard reaction mix containing 50 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 3 min, template DNA was amplified using 30 cycles with one of the three conditions: 45 s at 94 °C, 45 s at 62 °C (a), 48 °C (b) or 45 °C (c), and 1 min at 72 °C. The final extension step was conducted at 72 °C for 10 min. For CAPS analysis a *Mse*I recognition site was used to construct polymorphism between the parental DNAs. Therefore, the PCR product was digested with *Mse*I (10 µl of the PCR product with 5 U of *Mse*I in a 20-µl volume for 4 h at 37 °C) and separated on 1.5% agarose gels.

The ISSR PCR reaction mix contained 50 ng of genomic DNA, 10 mM of Tris–HCl (pH 9), 50 mM of KCl, 2.5 mM of $MgCl_2$, 0.5 mM of dNTPs, 1.0 pMol of the primer and 1.0 U of *Taq* DNA polymerase (Amersham Pharmacia, Freiburg). After initial denaturation at 94 °C for 1 min, template DNA was amplified using 35 cycles under the following conditions: 30 s at 94 °C, 30 s at 45 °C and 2 min at 72 °C. Final extension was conducted at 72 °C for 10 min.

Subcloning was performed by double-digesting BAC clones with 1 U of both *Bam*HI and *Hin*dIII in a 30-µl reaction using 200 ng of BAC DNA (Amersham Pharmacia Biotech, Freiburg). The restriction reaction was conducted at 37 °C for 4 h. After purification, 50 ng of double-digested BAC DNA was shotgun-cloned into pBluescript. For identification of single-copy clones, plasmid DNA was Southern blotted and hybridised using 100 ng of genomic maize DNA as a probe. Plasmids revealing no signals were tested for polymorphism between parents of PopB, because PopA was not available for Southern analysis. Genomic DNA (10 µg) of 84 individuals of PopB was *Hin*dIII-digested and Southern blotted to the Hybond N+ nylon membrane (Amersham Pharmacia Biotech, Freiburg). Hybridisation procedures were identical to those from BAC screening.

Linkage and statistical analyses

For PopA, linkage groups were constructed using MapMaker for UNIX version 3.0 (Whitehead Institute, Cambridge, Mass.). Initially, linkage groups were defined at a LOD of 6.0. The maps were constructed using the RI self option and the Haldane map function. All remaining markers were assigned at LOD 4.0. These computations were conducted at the University of Columbia, Missouri. The IBM population is an intermated recombinant inbred population. The random mating process increased the average number of recombination events per individual by approximately 3-fold compared with F_2 or RI-derived mapping lines. This is equivalent with a 3-fold expansion of map distances (Liu et al. 1996). Therefore, all map distances in the IBM map were divided by three.

For PopB, marker orders and genetic distances were calculated with MapMaker 3.0b (Lander et al. 1987) using a LOD threshold of 3.0 and the Kosambi mapping function.

Results

BAC analysis

Southern hybridisation of *Hin*dIII-digested B73 genomic DNA using pic19 as a probe identified six prominent bands (Fig. 1b). Screening of the BAC library identified 19 positive clones. Four clones were excluded that showed no positive signals when hybridised to pic19. Evaluation of the *Bam*HI or *Hin*dIII restriction gels of the remaining 15 BACs resulted in five groups of overlapping BAC clones (Table 2).

Table 2 Grouping of overlapping BAC clones based on *Bam*HI and *Hin*dIII restriction analysis

Contig 1 – pic19-1	Contig 2 – pic19-2	Contig 3 – pic19-3	Contig 4 – pic19-4	Contig 5 – pic19-5
71E12	44E24	191F2	181B13	62H1
103E15	230A6	198O20		192O11
1,75O11	267B13	212D24		241P7
2,05O18		239B24		

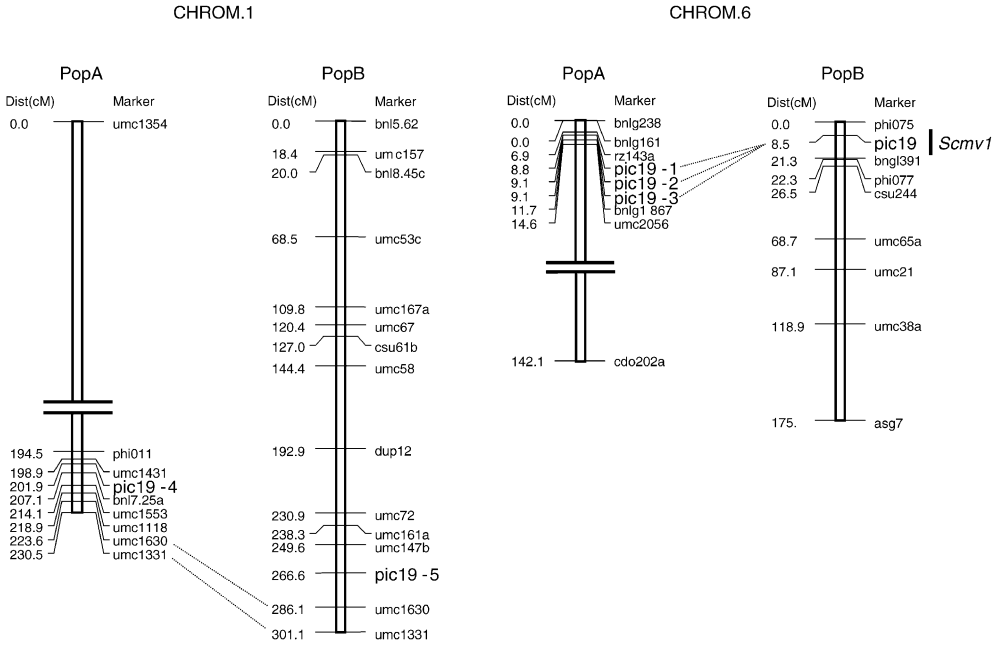

Fig. 2 Genetic mapping of pic19 homologues to maize chromosomes 1 and 6

Sequence alignments

RGA sequences homologous to pic19 could be amplified in four out of the five BAC contigs. Sequence comparisons revealed an identity of 94 to 98% among the pic19 RGA homologues of contigs 1 to 4. All differences were due to single nucleotide changes. Only contig 4 showed a continuous open reading frame (ORF). Contig 3 showed one stop codon. Contigs 1 and 2 revealed several frame shifts due to InDels.

Genetic mapping

PCR amplification of the pic19 homologues was not possible for contig 5. For contigs 1 and 2 [corresponding to mapping conditions (a) and (b) from the 'Materials and methods' section], SNP primers were designed resulting in differential amplification on the BAC as well as the genomic DNA level (Table 2). Genetic mapping placed both contig 1 and contig 2 into the chromosomal bins 6.00–6.01 within the marker interval *rz143a–bnlg1867* (Fig. 2) in PopA. Contig 1 and 2 displayed close linkage with a genetic distance of 0.37 cM.

For contig 3 (c), differential amplification was possible on the BAC level. Using the same primer pair, the parents of the mapping population did not show polymorphism. Cloning and sequencing of the regarding parental fragments revealed a SNP in a *Mse*I recognition site, which was used as a CAPS marker to place contig 3 in the same marker interval as contigs 1 and 2, cosegregating with contig 2. The evaluation of the marker scores revealed only two recombinants between contigs 1 and 2. Both individuals did not amplify in contig 3. For the remaining individuals, contig 3 scores were identical to both contig 1 and contig 2 scores.

For contig 4 (d), none of the previous approaches was successful. Screening of the five contigs with ISSR primers identified a contig 4-specific fragment using ISSR primer UBC860. The polymorphism was conserved between the two parental lines B73 and Mo17 on the genomic DNA level. Contig 4 was mapped to chromosomal bin 1.10 within the marker interval *umc1431–bnl7.25* (Fig. 2).

It was not possible to amplify pic19 homologues by PCR from contig 5 (e). Therefore, one member of the contig, BAC 192O11, was shotgun-subcloned. Subclone III-17 was used as a single-copy RFLP probe in PopB and this mapped contig 5 to the end of chromosome 1 into the marker interval *umc161a–umc147b* (Fig. 2). Since contigs 4 and 5 were mapped in different populations, their map distance was not determined.

Discussion

Genomic organisation of the pic19 RGA family

Southern hybridisation of pic19 to B73 genomic DNA revealed six prominent bands (Fig. 1b). Taking into account the size of the bands on the Southern blot and the size of the random primed pic19 probe, these six bands represent between three and six loci in the B73 genome. The five different contigs of overlapping BAC clones agree with this hypothesis. However, due to potentially absent sequences in the used maize BAC library (approximately 4 × genome coverage) and the high stringency of hybridisation conditions (65 °C), the existence of further homologues to pic19 cannot be ruled out.

The function of RGAs flanking R genes may be various: they might display: (1) R genes conferring isolate-specific resistance, (2) pseudogenes representing an important source for the evolution of new resistance-specificities by recombination or gene conversion with functional genes, or (3) rudiments of already overcome resistances. This study demonstrates, as others before, that homologous RGAs can be tightly linked. But is also demonstrates that they can be simultaneously located at different loci. There is no evidence for duplicated loci between the two map positions of pic19 homologues in chromosomal bins 1.10 and 6.00/01 according to the autotetraploidy hypothesis (Helentjaris et al. 1995; Gaut and Doebley 1997). However, initial duplication followed by single nucleotide changes is apparently the driving force in the evolution of these loci. NBS regions appear to be under purifying selection consistent with its proposed but unproven effector function (Michelmore and Meyers 1998).

Clustering of RGA homologues, such as observed in this study, for pic19 is consistent with reports in maize as well as other crops. Studies on the *Pto, Cf* and *Dm* clusters in tomato and lettuce (Martin et al. 1993; Thomas et al. 1997; Meyers et al. 1998b) revealed the existence of numerous RGAs in the direct physical neighbourhood of R genes. Data from the Arabidopsis genome sequencing project uncovered 160 R genes (The Arabidopsis Genome Intiative 2000). The NBS sequences currently in the databases are located in approximately 21 genomic clusters and 14 isolated loci (Meyers et al. 1999). Besides the RGA clusters in maize identified by Collins et al. (1998), Zhao et al. (2001) noted the mapping of two maize R genes, *Rxo* and *Rpa*, to the same position on the short arm of maize chromosome 6 (Fig. 2) in the *Scmv1* region. Furthermore, the authors identified 5–6 RGAs in the direct neighbourhood of *Rxo* and *Rpa*. Therefore, clustering of homologous sequences within short chromosomal stretches seems also to be a common feature in the organisation of R genes and RGAs in maize.

Restriction analysis identified five contigs which are not overlapping with each other. Since contigs 1–3 and 4–5 are mapping to different chromosomes, overlapping can also be excluded. Genetic mapping identified two recombinants between contigs 1 and 2. Furthermore, restriction analysis clearly separates contigs 1 to 3. Therefore, the minimum size of the region covered by the three contigs on chromosome 6 estimated by the length of restriction fragments is 450 kbp. Repeated PCR using contig 3-specific primers did not amplify any product of the two genotypes recombinant between contigs 1 and 2. Therefore, it seems likely that these individuals lost the corresponding sequence. Pryor et al. (1987) and Bennetzen et al. (1988) reported spontanous mutations to susceptibility in different *rp1* maize rust R genes with frequencies of up to 0.5%. It has been proposed that the instability of this region is due to gene conversion or unequal crossing-over events between mispaired sequence repeats during meiosis (Sudupak et al. 1993; Hu and Hulbert 1994). The same mechanisms could explain the loss of contig 3-specific sequences in these two recombinant genotypes.

Xia et al. (1999) identified five QTLs for SCMV resistance in PopB. One major QTL mapped near *nor* on the short arm of chromosome 6 (*Scmv1*) and another major QTL near the centromere on the long arm of chromosome 3 (*Scmv2*). Three minor QTLs were located on chromosomes 1, 5 and 10. Quint et al. (2002) mapped pic19 in PopB to the *Scmv1* region on chromosome 6.00/01. Therefore, it was expected that at least one pic19 homologue maps to this region. Linkage of three BAC contigs containing pic19 homologues supports the hypothesis of a R gene cluster rather than one single pleiotropic R gene in this region. Especially interesting is the map position of contigs 4 and 5, which map to the same region as the QTL previously detected for SCMV resistance on chromosome 1 (Fig. 2). Therefore, these contigs might as well contain candidate genes underlying the QTL in this region.

Sequence comparison of pic19 homologues

The pic19 homologues which were amplified by PCR in four out of the five contigs displayed a very high sequence similarity of 94–98% at the nucleotide level. A comparison of 11 genes of the tomato Cf-4/9 cluster (Parniske et al. 1997) exhibits an equally high degree of overall sequence homology (92–99%). Shen et al. (1998) sequenced part of the NBS motif of cross-hybridising lettuce RGAs and showed that linked members of the same RGA families had an identity in the range of 54–98% at the deduced amino-acid level. The authors defined RGAs displaying at least 50% sequence similarity on the amino-acid level as members of the same RGA family. The pic19 RGA family includes members clustering at two different genome regions in contrast to the other discussed RGA families. However, the relationship between sequence similarity and the physical position of NBS-encoding RGAs is very complex and cannot be generalised because some clusters also contain quite diverse sequences (Meyers et al. 1999).

Consequences for gene isolation and development of closely linked molecular markers

Our study shows that RGAs mapping to the target region can be successfully used to identify further-linked candidate sequences. This approach is useful with regard to map-based cloning, which 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).

A second benefit is the possibility to develop closely linked molecular markers for application of MAS in practical plant breeding. Furthermore, RGAs were shown to be highly polymorphic and the rate of success of converting RGAs into codominant CAPS or dominant SNP markers is high compared to other genomic sequences (Quint et al. 2002).

However, the high sequence similarity of these linked RGA sequences reveals potential problems for the use of RGAs as molecular markers. Discrimination of homologous sequences by PCR might be problematic because of simultaneous amplification of several homologues (Fig. 3c). Therefore, these indistinguishable and genetically linked homologues might be mapped as a single-copy 'ghost' marker analogous to a 'ghost' QTL (Martinez and Curnow 1992) as illustrated by an example in Fig. 3. Suppose a RGA was sequenced from one end in two genotypes 1 and 2, and a SNP (SNP1) is identified between them (in Fig. 3a) and used for mapping in a segregating population. Suppose further that the two real homologous copies RGA1 and RGA2, which can be distinguished in both genotypes by a second SNP (SNP2), remain unrecognised. Hence, scoring the polymorphism between the parental genotypes 1 and 2 at SNP1 results in a single-copy 'ghost' marker representing RGA1 and RGA2. Using a codominant marker system in a segregating progeny of the parental genotypes 1 and 2 results in deviations between the marker genotypes of the 'ghost' RGA and those of the real RGAs (Fig. 3c, d). The expected genotype frequencies will be shifted towards heterozygotes. Instead of estimating the correct map positions of RGA1 and RGA2 relative to linked marker loci, the position of the fictive 'ghost' marker is estimated. As a consequence, the position of the 'ghost' marker may change the positions and even the order of other marker loci and target genes or QTLs (Fig. 3c, for details see Frisch et al. 2003). Map distances will be overestimated and the positions of the fictive 'ghost' marker might deviate greatly from the true marker loci RGA1 and RGA2. The divergence of the map positions of the 'ghost' marker and its real underlying loci grows with the map distance between the real loci of RGA1 and RGA2. Wrong estimates of map positions of 'ghost' marker loci might have fatal effects for map-based cloning or MAS. Relative to the position of this 'ghost' marker and other markers, one also arrives at completely wrong estimates of map positions of the hypothetical target genes.

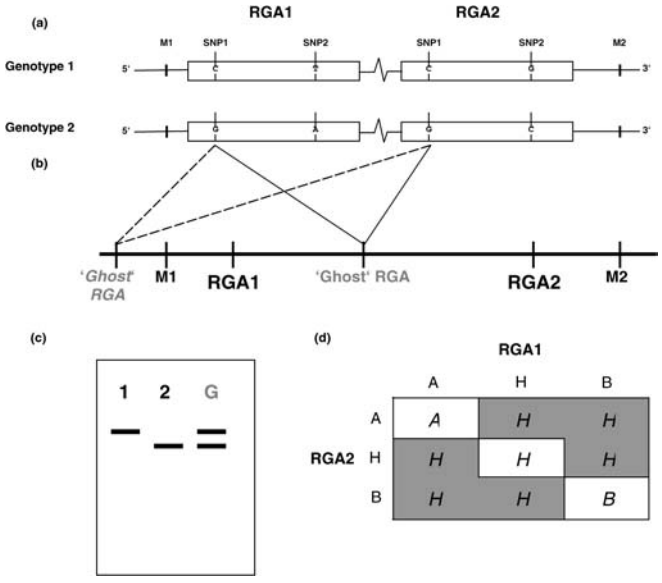

Fig. 3a–d Occurrence of 'ghost' markers. (**a**) Physical organisation of two homologous linked sequences RGA1 and RGA2 and two flanking markers M1 and M2; the RGAs can be distinguished between two genotypes by SNP1 and amongst each other by SNP2; (**b**) using SNP1 for scoring the RGA genotype in a segregating population results in the mapping of both homologous sequences to one single-copy 'ghost' RGA; the map position of this 'ghost' RGA may either be in between RGA1 and RGA2 but can also be outside this interval, and may even result in different orders of marker loci ('ghost' RGA in *italics*). (**c**) banding pattern of RGA1 (1), RGA2 (2) and the resulting 'ghost' marker (*G*), simultaneous amplification of 1 and 2 would result in banding pattern G; (**d**) the resulting scoring of RGA1 and RGA2 with a codominant marker system; scores for the 'ghost' marker are shown in *italics*, deviations of scores between the 'ghost' marker and one of the real RGAs are indicated by a *shaded background*

To avoid these scenarios, the implementation of a genomic library screening, including genetic mapping of potential homologues as reported in this paper, seems necessary for safe application of RGA markers for gene isolation and MAS. Especially for R genes, it has been demonstrated that clustering of homologous sequences is a common feature in the plant kingdom. However, these complications in the application of RGA sequences as molecular markers are not restricted to RGAs. They are existent for all DNA markers derived from potentially duplicated sequences, such as ESTs from large gene families, AFLPs or other markers located in recombination hotspots.

Acknowledgements The present study was supported by a grant from the Deutsche Forschungsgemeinschaft, Grant No. LU601/3. The authors gratefully acknowledge the skilled technical assistance of Elisabeth Kokai-Kota. We would also like to thank Michael D. McMullen for linkage analysis and Matthias Frisch for his helpful suggestions on the issue of "ghost" markers.

References

Bennetzen JL, Qin MM, Ingels S, Ellingboe AH (1988) Allele-specific and mutator-associated instability at the *Rp1* disease-resistance locus of maize. Nature 332:369–370

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

Davis G, Musket T, Melia-Hancock S, Duru N, Qu J, Sharopova N, Schultz L, McMullen MD, Woodman WL, Long MJ, Lee M, Vogel JM (2000) A high-resolution genetic map of the B73 × Mo17 population. Maize Genetics Conference Abstracts 42:P79

Ellis J, Lawrence GJ, Finnegan EJ, Anderson PA (1995) Contrasting complexity of two rust resistance loci in flax. Proc Natl Acad Sci USA 92:4185–4188

Frisch M, Quint M, Lübberstedt T, Melchinger AE (2003) Duplicate marker loci can result in incorrect locus order on linkage maps. (submitted)

Gaut BS, Doebley JF (1997) DNA sequence evidence for the segmental allotetraploid origin of maize. Proc Natl Acad Sci USA 94:6809–6814

Hammond-Kosack KE, Jones JDG (1997) Plant disease resistance genes. Annu Rev Plant Physiol Plant Mol Biol 48:575–607

Helentjaris T (1995) Atlas of duplicated sequences. Maize Genet Coop Newslett 69:67

Hu G, Hulbert S (1994) Evidence for the involvement of gene conversion in meiotic instability of the *Rp1* rust resistance genes in maize. Genome 37:742–746

Kesseli R, Witsenboer H, Stanghellini M, Vandermark G, Michelmore RW (1993) Recessive resistance to *Plasmopara lactucae-radicis* maps by bulked segregant analysis to a cluster of dominant resistance genes in lettuce. Mol Plant-Microbe Interact 6:722–728

Liu SS, Kowalsky SP, Lan TH, Feldmann KA, Peterson AH (1996) Genome-wide high-resolution mapping by recurrent intermating using *Arabidopsis thaliana* as a model. Genetics 142:247–258

Lander ES, Green P, Abraramson J, Baslow A, Daly MJ, Lincoln SE, Newburg L (1987) MapMaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–184

Landry BS, Kesseli R, Farrara B, Michelmore RW (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331–337

Leipoldt M, Schmidtke J (1982) In: Dover G, Flavell R (eds) Genome evolution. Academic Press, New York, pp 219–236

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262:1432–6

Martinez O, Curnow RN (1992) Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor Appl Genet 85:480–488

Meyers BC, Chin, DB, Shen KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, Michelmore RW (1998a) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell 10:1817–1832

Meyers BC, Shen KA, Rohani P, Gaut BS, Michelmore RW (1998b) Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell 11:1833–1846

Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J 20:317–332

Michelmore RW, Meyers BC (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res 8:1113–1130

Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BB, Jones JD (1997) Novel disease resistance specifities results from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. Cell 91:821–832

Pryor T (1987) Stability of alleles of *Rp* (resistance to *Puccinia sorghi*). Maize Genet Newslett 61:37–38

Quint M, Melchinger AE, Dußle CM, Lübberstedt T (2000) Breeding for virus resistance in maize. Genetika 32:529–545

Quint M, Mihaljevic R, Dußle CM, Xu ML, Melchinger AE, Lübberstedt T (2002) Development of RGA-CAPS markers and genetic mapping of candidate genes for SCMV resistance in maize. Theor Appl Genet 105:355–363

Saxena KMS, Hooker AL (1968) On the structure of a gene for disease resistance in maize. Proc Natl Acad Sci USA 61:1300–1305

Shen KA, Meyers BC, Islani-Faridi MN, Chin DB, Stelly DM, Michelmore RW (1998) Resistance gene candidates indetified by PCR with degenerate oligonnucleotide primos map to clusters of resistance genes in lettuce. Mol Plant-Microbe Interact 8:815–823

Sudupak MA, Bennetzen JL, Hulbert SH (1993) Unequal exchange and meiotic instability of disease-resistance genes in the *Rp1* region on maize. Genetics 133:119–125

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

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 *Cladiosporum fulvum* identifies sequences that determine recognitial specifity in *Cf-4* and *Cf-9*. Plant Cell 9:2209–2224

Witsenboer H, Kesseli R, Fortin MG, Stanghellini M, Michelmore RW (1995) Sources and genetic fine structure of a cluster of genes for resistance to three pathogens of lettuce Theor Appl Genet 91:178–188

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

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 IX Conference, p 90