Mol Genet Genomics (2011) 285:207–218
DOI 10.1007/s00438-010-0596-6

ORIGINAL PAPER

Evolution of nematode-resistant *Mi-1* gene homologs in three species of *Solanum*

Maria Virginia Sanchez-Puerta · Ricardo Williams Masuelli

Received: 12 May 2010 / Accepted: 9 December 2010 / Published online: 5 January 2011

Abstract Plants have evolved several defense mechanisms, including resistance genes. Resistance to the root-knot nematode *Meloidogyne incognita* has been found in wild plant species. The molecular basis for this resistance has been best studied in the wild tomato *Solanum peruvianum* and it is based on a single dominant gene, *Mi-1.2*, which is found in a cluster of seven genes. This nematode attacks fiercely several crops, including potatoes. The genomic arrangement, number of copies, function and evolution of *Mi-1* homologs in potatoes remain unknown. In this study, we analyzed partial genome sequences of the cultivated potato species *S. tuberosum* and *S. phureja* and identified 59 *Mi-1* homologs. *Mi-1* homologs in *S. tuberosum* seem to be arranged in clusters and located on chromosome 6 of the potato genome. Previous studies have suggested that *Mi-1* genes in tomato evolved rapidly by frequent sequence exchanges among gene copies within the same cluster, losing orthologous relationships. In contrast, *Mi-1* homologs from cultivated potato species (*S. tuberosum* and *S. phureja*) seem to have evolved by a birth-and-death process, in which genes evolve mostly by mutations and interallelic recombinations in addition to sequence exchanges.

Keywords Potato · *Meloidogyne* · Resistance · Evolution · Birth-and-death model

Communicated by C. Gebhardt.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-010-0596-6) contains supplementary material, which is available to authorized users.

M. V. Sanchez-Puerta (✉)
Instituto de Ciencias Basicas, IBAM-CONICET and Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo, 5500 Mendoza, Argentina
e-mail: mvsanchezpuerta@fca.uncu.edu.ar
URL: http://sites.google.com/site/sanchezpuerta/mvsp

R. W. Masuelli
Instituto de Biología Agrícola Mendoza (IBAM), Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo, and INTA EEA La Consulta, 5500 Mendoza, Argentina

Introduction

Solanaceous plants, like many other plants, are attacked by a wide range of pathogens, including oomycetes, nematodes, insects, viruses, bacteria and fungi. Plants have evolved several defense mechanisms, including resistance genes (R genes) that belong to large families with dozens of genes. Root-knot nematodes, *Meloidogyne* species, are able to attack and seriously damage more than 2,000 plant species, including almost all the major crops (Vos et al. 1998; Williamson and Kumar 2006). Out of more than 70 known species of *Meloidogyne*, only four of them are of major economic importance worldwide: *M. incognita*, *M. hapla, M. arenaria* and *M. javanica* (Sasser and Carter 1985). Resistance to the root-knot nematode *Meloidogyne incognita* has been found in wild plant species, such as certain accessions of upland cotton, wild tomato, sweet potato and pepper (Williamson and Kumar 2006; Niu et al. 2007). The molecular basis for resistance against *M. incognita* has been best studied in the wild tomato *Solanum peruvianum*. Resistance found in some genotypes of *S. peruvianum* is effective against three root-knot nematodes (*M. arenaria, M. incognita* and *M. javanica*), the potato aphid *Macrosiphum euphorbiae* and the whitefly *Bemisia tabaci* (Dropkin 1969; Nombela et al. 2003; Rossi et al. 2003; Jablonska et al. 2007). Cultivated tomato varieties of *S. lycopersicum* are naturally susceptible to

infections by these nematodes; however, modern cultivars of tomato show high levels of resistance due to the introduction of a multigene locus called *Mi-1* that has been introgressed from *S. peruvianum* (Gilbert and McGuire 1956; Seah et al. 2007). Nematodes penetrate roots and reach the vascular system in both resistant and susceptible plants. However, in resistant plants, a localized tissue necrosis at the feeding site occurs, preventing pathogen multiplication. The resistance is based on a single dominant gene, *Mi-1.2*, localized on the short arm of chromosome 6 of tomato (Kaloshian et al. 1998; Milligan et al. 1998). The gene *Mi-1.2* is found in a cluster of seven homologous copies in both susceptible and resistant genotypes of tomato (Seah et al. 2007). The function of most of the *Mi-1* genes has not been determined.

Resistance genes share conserved elements and are divided into a few classes. Most R genes in plants contain a nucleotide-binding (NB) region and a C-terminal leucine-rich repeat (LRR) domain, and encode proteins named NB–LRR proteins (van Ooijen et al. 2007). A subset of these proteins contains an N-terminal domain with a predicted coiled-coil (CC) structure and this subgroup is referred to as CC–NB–LRR proteins. The seven homologs of the gene *Mi-1.2* belong to the CC–NB–LRR family of plant resistance genes (Fig. 1). The CC structure is a repeated heptad sequence with interspersed hydrophobic amino acid residues. It consists of two or more alpha helices that interact to form a supercoil and is implicated in protein–protein interaction; probably involved in signaling (Martin et al. 2003; van Ooijen et al. 2008). Between the NB and LRR domain, there is a well-conserved region defined as the ARC domain, which consists of two distinguishable subdomains: ARC1 and ARC2 (Albrecht and Takken 2006; Rairdan and Moffett 2006). The NB and ARC domains are contiguous and referred together as the NB–ARC domain, which might act as the molecular switch that controls the activation state of the protein, defined by the binding of ATP or ADP (Martin et al. 2003; Albrecht and Takken 2006; Rairdan and Moffett 2006). As it is typical for NB–LRR resistance proteins, the NB region is the most conserved among R genes (Williamson and Kumar 2006). The LRR domain is highly divergent in primary structure and number of repeats, consisting of 2–42 repeats, each comprising a β-sheet. It is implicated in protein–protein interactions, as a major determinant of recognition specificity (Michelmore and Meyers 1998; Martin et al. 2003). The C-terminal of the LRR domain is involved in ligand recognition and has been shown to be under diversifying selection (Michelmore and Meyers 1998; Williamson and Kumar 2006; van Ooijen et al. 2007).

Homologous genes to *Mi-1* in the genus *Solanum* have been reported in a single wild potato species, *S. bulbocastanum*. The gene *Rpi-blb2* from *S. bulbocastanum* confers resistance to the oomycete *Phytophthora infestans* and is the closest homolog to the gene *Mi-1.2* known in potato, sharing 81% amino acid sequence identity with the tomato gene *Mi-1.2* (van der Vossen et al. 2005). This gene is located in chromosome 6 of the *S. bulbocastanum* genome, in a syntenic position to the tomato locus *Mi-1* and also found in a cluster arrangement (van der Vossen et al. 2005; Seah et al. 2007). No resistance gene to *M. incognita* has been reported in tuber-bearing potatoes and no homologs of the gene *Mi-1.2* are known for the cultivated potato *S. tuberosum* or any other of the more than 200 wild potatoes known besides *S. bulbocastanum*. The genomic arrangement, number of copies, function and evolution of *Mi-1* homologs in potato species remain unknown. Here, we describe 59 *Mi-1* homologs from cultivated potato species (*S. tuberosum* ssp. *tuberosum* and *S. phureja*) and study the evolutionary history of this gene family in the genus *Solanum*, including duplication events, diversification rates and evidence of gene recombination.

Materials and methods

Sequence data acquisition

Sequence similarity searches were performed with BLAST (Altschul et al. 1990) against the nr/nt (non-redundant nucleotide collection), GSS (genome survey sequences) and HTGS (high throughput genomic sequences) GenBank databases, restricted to the genus *Solanum*, and against the sequenced chromosomes of the tomato genome from Sol Genomics Network (Mueller et al. 2009), using the sequence from the gene *Mi-1.2* from *S. lycopersicum* as the query sequence. Blast hits with an expect-value $<1e^{-20}$ were considered putative *Mi-1* homologs and included in subsequent analyses. Table S1 includes GenBank accession numbers for all sequences used in this study. *StMi1h*

Fig. 1 Gene (**a**) and protein (**b**) structure of the tomato gene *Mi-1.2*. Filled exons are translated into protein. *CC* coiled-coil structure, *NB* nucleotide-binding region, *LRR* leucine-rich repeat domain

sequences belong to a diploid heterozygous line RH89-039-16 (RH) of the cultivated potato *Solanum tuberosum* ssp. *tuberosum* and *SpMi1h* genes to a doubled monoploid DM1-3 516R44 (DM) of the potato *S. phureja* (Visser et al. 2009). Protein motif prediction was performed using InterProScan from the European Bioinformatics Institute (Quevillon et al. 2005).

Phylogenetic analyses

Potential homologs of the gene *Mi-1* obtained from GenBank databases were aligned and edited manually using MacClade (Maddison and Maddison 2000). Maximum likelihood analyses were performed with Garli 0.951 (Zwickl 2006) under the General Time Reversible model with parameters for invariant sites and gamma-distributed rate heterogeneity (GTR + I+Γ 4; four rate categories). This substitution model was supported by hierarchical likelihood ratio tests done using Modeltest v.3.7 (Posada and Crandall 1998). Ten independent runs were conducted using the automated stopping criterion or for up to 5,000,000 generations to ensure convergence to a similar topology and likelihood score. One hundred bootstrap replicates were performed.

Alternative topology test

The approximately unbiased (AU) test aids to assess whether phylogenetic incongruences are significant. The CONSEL package (Shimodaira 2002) was used to calculate the AU *p* values for unconstrained and constrained trees. Based on the individual gene region (N-terminal, NBS and LRR) data sets, we compared the unconstrained tree topology for a gene region (i.e., the best ML tree found with this data set) to the topology constrained according to the phylogeny obtained with the other data sets (i.e., the topology found using another gene region). We repeated the same process for each gene region. The most likely tree under each constraint was determined by searching for the best tree compatible with that constraint using PAUP* (Swofford et al. 2002). The site likelihoods for these trees and for the best tree in the unconstrained analysis were exported from PAUP* and the AU *p* values were calculated from these data using CONSEL.

Tests of gene recombination

Putative gene conversion events during *Mi-1* evolution in the genus *Solanum* were analyzed with the pairwise homoplasy index (PHI) test (Bruen et al. 2006) as implemented in the program SplitsTree4 (Huson and Bryant 2006). Sequence exchanges were also identified in the *Mi-1* alignment using Geneconv (Sawyer 1989); no mismatch was allowed (gscal = 0) and only events with $p < 0.05$ were reported.

Selection analysis

The Molecular Evolutionary Genetics Analysis (MEGA 4.0) software package (Tamura et al. 2007) was used to analyze the selective pressures on gene pairs by calculating the dN/dS (dN: number of non-synonymous substitutions per non-synonymous site; dS: number of synonymous substitutions per synonymous site) ratio between homologous genes. The ratios were calculated between pairs of genes as well as each gene region (i.e., N-terminal, NB–ARC or LRR regions). Comparisons of the dN/dS ratios were done to evaluate the selection pressures of each gene region independently. For testing positive or purifying selection, the method of Nei and Gojobori (1986) was used as implemented in MEGA 4.0.

Results

Characterization of *Mi-1* homologs in the cultivated potato *S. tuberosum*

We retrieved nucleotide sequences from BAC clones from *Solanum tuberosum* deposited in GenBank (Visser et al. 2009). We searched for homologs of the tomato gene *Mi-1.2* by BLASTn against the HTGS (High Throughput Genomic Sequences) database and found 59 hits to *Mi-1*-like genes from *S. tuberosum* with an *E* value $< 1e^{-19}$ (Fig. 2). These hits corresponded to 16 partially sequenced genomic BAC clones from *S. tuberosum* cultivar RH89-039-16 (Visser et al. 2009). Out of the 16 BAC clones, 11 were previously anchored to chromosome 6 of the *S. tuberosum* nuclear genome (Visser et al. 2009), while the others had not been localized yet (Fig. 2; Table S1). We assembled all fragments that belonged to these 16 BAC clones using the program Sequencher and found that several BAC sequences were overlapping with one another: AC237930 and AC232390; AC238200 and AC231973; AC237846 and AC238283; AC238108 and AC238211; and AC238274 with AC237844. Based on this analysis, two additional BACs (AC237846 and AC237844) were localized to chromosome 6 (Fig. 2; Table S1). Among the 59 hits, 14 sequences (Fig. 2) were identical to at least one other sequence from *S. tuberosum* (Figure S1) and were removed from the data set to avoid redundancy, leaving 45 different potato *Mi-1*-like sequences (*StMi1h* genes) that were included in further analyses.

An alignment of all 45 *Mi-1* homologs of *S. tuberosum* was prepared. The *Mi-1* homolog from *S. bulbocastanum* (*Rpi-blb2*) was also added to this alignment. Some

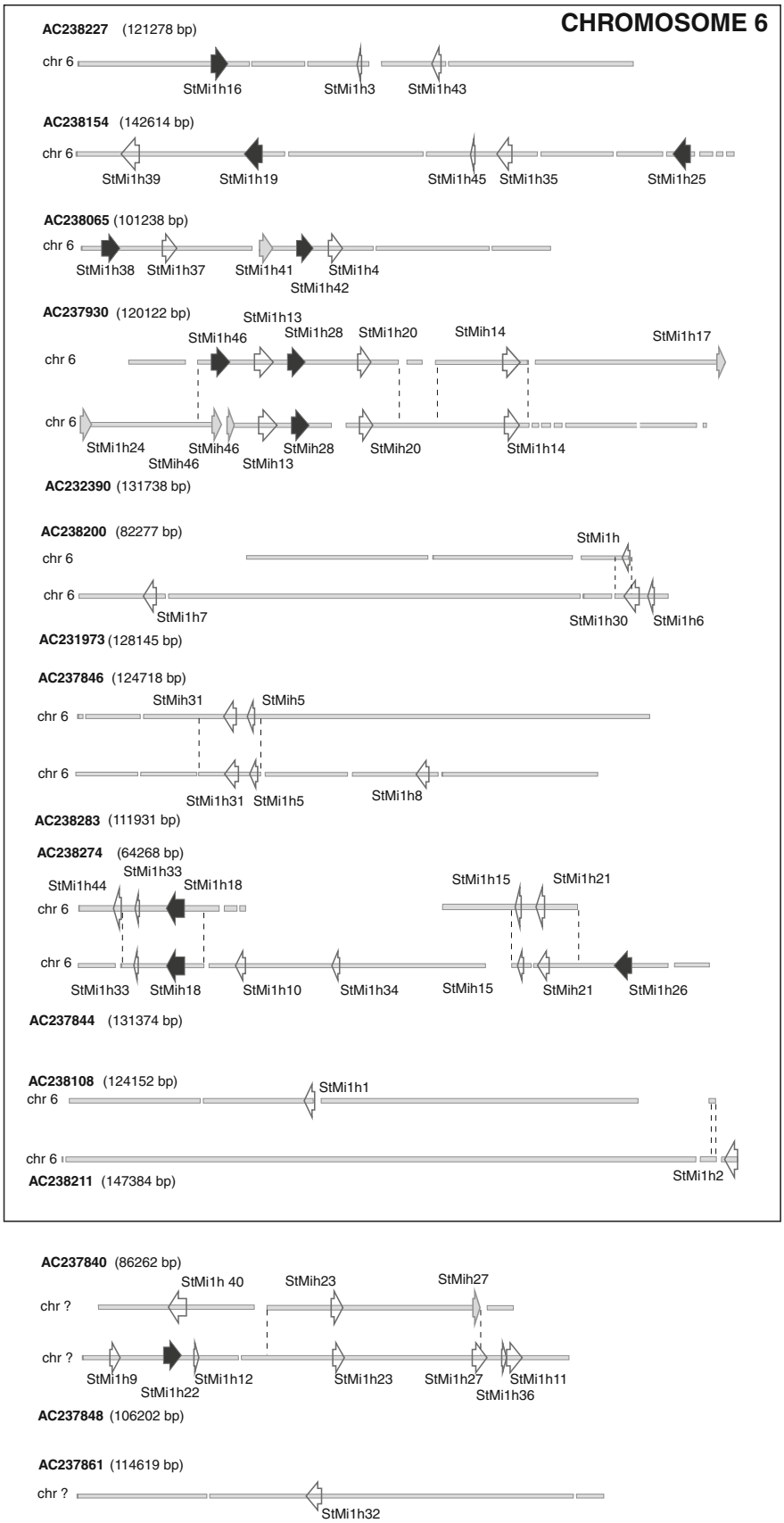

Fig. 2 BAC clones from *Solanum tuberosum* RH89-039-16 containing *Mi-1* homologs. Length of each BAC clone is shown next to the GenBank accession number. *Gray bars* represent unordered fragments within each clone. Relative size of *bars* refer to the relative length of sequenced fragments. Thirteen BAC clones were anchored to chromosome 6. *Arrows* represent *Mi-1* homologs and their direction indicates the DNA strand in which they are encoded. *Black arrows* are full-length genes, *gray arrows* are partially sequenced genes and *white arrows* are pseudogenes. *Dashed lines* indicate that linked BAC clones are overlapping

sequences of *Mi-1* homologs in *S. tuberosum* were partial sequences either missing the 5′ or 3′ end of the gene due to incomplete sequencing of BAC clones (Table S1). Out of the 45 *Mi-1* homologs from *S. tuberosum*, 32 were labeled as pseudogenes because they were truncated or contained premature stop codons or indels that led to frameshifts (Table S1). Ten potato sequences were full length, compared to the gene *Mi-1.2* from tomato and could translate into a full-length protein with no frameshifts: *StMi1h_16, StMi1h_18, StMi1h_19, StMi1h_22, StMi1h_25, StMi1h_26, StMi1h_28, StMi1h_38, StMi1h_42* and *StMi1h_46* (Table S1, in boldface). These ten genes encode potentially functional proteins.

Characterization of *Mi-1* homologs in the cultivated tomato *S. lycopersicum*

Seven and nine *Mi-1* homologs are known from resistant (*Mi-1.1–Mi-1.7*) and susceptible (*Mi-1.A–Mi-1.I*) *S. lycopersicum* cultivars, respectively (Seah et al. 2004; Seah et al. 2007). The genes *Mi-1.1–Mi-1.7* belong to a genomic region that was introgressed from the wild tomato *S. peruvianum* (Gilbert and McGuire 1956; Seah et al. 2007). Southern blots using an *Mi-1* probe showed that resistant and susceptible tomatoes contained ~7–9 copies of *Mi-1* genes each (Milligan et al. 1998; Seah et al. 2004). Furthermore, we performed BLAST searches against the fully sequenced chromosomes from *S. lycopersicum* for homologs of the tomato gene *Mi-1.2* in the databases from Sol Genomics Network. Blast hits longer than 500 nt included nine sequences: seven located in chromosome 6 and two in chromosome 5.

Phylogenetic analysis of all 23 sequenced *Mi-1* homologs from *S. lycopersicum* and all 59 sequences from *S. tuberosum* is shown in Figure S1. The nine tomato sequences identified in the sequenced tomato genome correspond to the seven *Mi-1* copies from the susceptible tomato cultivar previously reported (*MiA-G*; Seah et al. 2007). In addition, two *Mi-1* homologs were identified in chromosome 5 from the tomato nuclear genome. One of the sequences appeared embedded within potato sequences, while the other sequence was located within the tomato clade (Figure S1).

Structure of *Mi-1* genes

Mi-1 homologs from *S. lycopersicum* present three exons, two of which (exon 2 and 3) are protein-coding exons (Fig. 1). Exon 1 belongs to the 5′ UTR (untranslated region) of the gene (Milligan et al. 1998). Sequences of exon 2 from *S. tuberosum* were available from 11 *StMi1h* genes and they were short and conserved in size (42 nt long), same as in tomato *Mi-1* homologs (Figure S2). Complete exon 3 sequences were available from 14 potato *StMi1h* genes and they were variable in size, ranging from 3,726 to 4,065 nt. The size of exon 3 was overall conserved among homologs in *S. lycopersicum* and the few sequences available from *S. bulbocastanum* and *Capsicum annuum*, but several potato *Mi-1* genes contained long insertions, the longest of which was 186 nt. Two insertions (9- and 24-nt long) were shared exclusively by all potato *StMi1h* sequences and *S. bulbocastanum Rpiblb2*, but not by tomato *Mi-1* genes. Potato sequences (including pseudogenes) shared 79–97% nucleotide identity with each other. The average identity of exon 3 sequences with tomato *Mi-1* homologs was 75–94%.

Two introns are known for *Mi-1* homologs in tomato and they are localized near the 5′ end of the gene (Fig. 1). Intron 1 is longer and interrupts the 5′ UTR, while intron 2 interrupts the coding region (Milligan et al. 1998). Intron 2 was also identified in all potato *StMi1h* sequences that were available toward the 5′ end of the gene. Intron 2 in *S. tuberosum* was located at a conserved position between exon 2 and 3, as described in the tomato gene *Mi-1.2*, but was slightly longer, ranging in length from 82 to 91 nt (Figure S2).

Evolution of *Mi-1* homologs in potato and tomato

Based on the *Mi-1* alignment of potato and tomato sequences, a number of phylogenetic analyses were performed. A maximum likelihood (ML) tree based on exon 3 sequences is shown in Fig. 3. Even though some sequences are partial, the length of most sequences is greater than 70% of the total length of the alignment. It has been shown that a small proportion of missing data in an alignment does not affect the phylogenetic analysis significantly (Philippe et al. 2004; Wiens 2006).

What first stands out is that potato sequences are more numerous and diverse than tomato homologs. Surprisingly, tomato and potato sequences do not form sister clades; instead, the tomato sequences mostly cluster together and are located among the potato sequences. Potato *StMi1h* genes represent the great majority and form small clades of less than ten potato sequences with variable support and a broad range of branch lengths. The ML tree also shows that 11 *Mi-1* homologs (clade that includes *Mi1F* to *Mi1_6*) from tomato form a highly supported monophyletic group embedded within *S. tuberosum* sequences (Fig. 3). Another two sequences from tomato (*Mi1_3* and *Mi1_5*) form a strongly supported clade and are sister to two sequences from *S. tuberosum* (*StMi1h_32* and *StMi1h_33*) with high bootstrap support. The other tomato sequence (*Mi1A*) is not robustly placed. Full-length *StMi1h* genes (bold-faced names in Fig. 3) neither clustered together, nor were they found to be sister to the tomato clade (Fig. 3). Most of the *StMi1h* genes were pseudogenes (Table S1) and showed

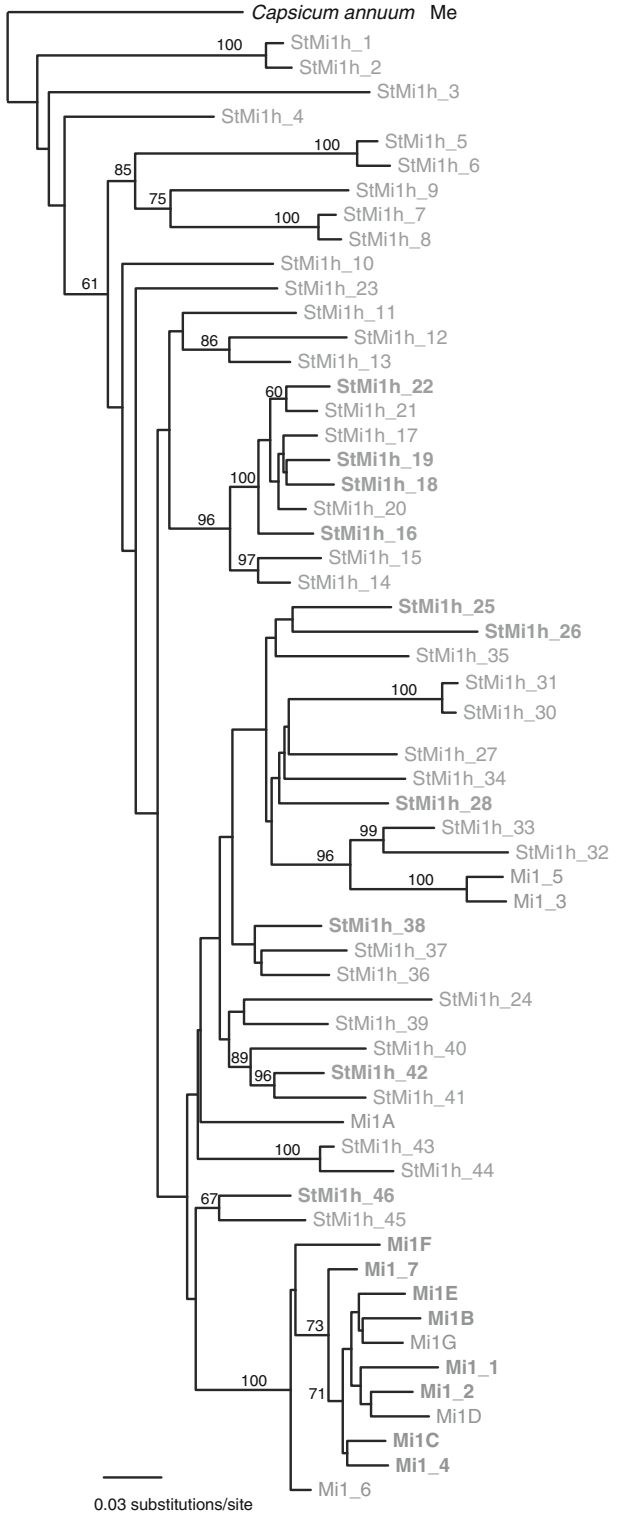

Fig. 3 Arbitrarily rooted-tree of 60 *Mi-1* homologs from *S. lycopersicum* (*Mi1* genes) and *S. tuberosum* (*StMi1h* genes). Maximum likelihood analysis based on 3,876 exonic positions using GTR+I+Γ4 model of substitution and rooted with *Capsicum annuum* gene *Me*. Numbers on branches correspond to bootstrap support values >59% from 100 bootstrap replicates. Names in *boldface* indicate that the genes have a complete ORF

Table 1 Average dN/dS values from the three gene regions of full-length *StMi1h* gene pairs from *Solanum tuberosum* and gene pairs from putative functional *Mi-1* genes from *S. lycopersicum*

Gene pair	N-terminal region	NBS region	LRR region
Mi1.2/Mi1.1	**0.57***	**0.50***	**1.75***
Mi1.1/Mi1.4	**0.50***	0.60	1.13
Mi1.4/Mi1.7	**0.50***	1.00	0.67
Mi1.7/Mi1.2	**0.50***	0.56	1.00
Mi1B/Mi1C	**0.20***	0.50	1.20
Mi1C/Mi1E	1.00	**0.50***	1.00
Mi1E/Mi1F	**0.31***	**0.45***	1.00
Mi1F/Mi1B	**0.31***	0.75	1.17
StMih16/StMih18	**0.67***	0.50	0.90
StMih18/StMih19	**0.67***	1.00	1.75
StMih19/StMih22	**0.40***	2.00	**1.60***
StMih22/StMih25	0.90	**0.56***	0.92
StMih25/StMih26	**0.44***	0.67	1.09
StMih26/StMih28	**0.56***	0.75	0.92
StMih28/StMih38	**0.50***	**0.50***	1.00
StMih38/StMih42	**0.56***	1.67	0.73
StMih42/StMih46	**0.33***	**0.50***	0.80
StMih46/StMih16	**0.27***	**0.59***	0.85

* Values of dN/dS significantly different from 1 ($p < 0.05$). When value is >1 or <1, it suggests that the sequence is under positive or purifying selection, respectively

longer branch lengths due to a faster evolutionary rate than putative functional genes.

Analyses of different gene regions of *Mi-1* homologs

Potato *StMi1h* genes, along with tomato *Mi-1* homologs, encode CC–NB–LRR proteins that can be divided into three regions: N-terminal (including the CC structure), NBS (nucleotide-binding site including NB, ARC1 and ARC2 motifs) and LRR, separated by short transition regions (Fig. 1). We performed analyses on these three regions independently. The similarity among potato *StMi1h* sequences varied depending on the gene region. Nucleotide identity of the N-terminal region among potato sequences was 60–100%, 81–100% for the NBS region and 75–100% for the LRR region. In addition, we studied the nucleotide identity and selection pressures over individual gene regions of all ten full-length (putative functional) *StMi1h* homologs. Nucleotide identity of N-terminal, NBS and LRR regions among full-length *StMih* genes was 87–98%, 82–98% and 84–96%, respectively. We calculated the dN/dS ratios between gene pairs and found that dN/dS values for the LRR regions in tomato and potato sequences were higher than for the other two regions (Table 1). Tests of selection showed that the

N-terminal and NBS regions from tomato *Mi-1* and potato *StMih* genes were under neutral or purifying selection (Z test $p < 0.05$; Table 1). In contrast, the LRR region of the tomato *Mi1.2/Mi1.1* and potato *StMi1h19/StMih22* genes were under positive selection (Z test $p < 0.05$), while the LRR region of the rest was under neutral selection (Z test $p < 0.05$; Table 1).

Separate phylogenetic analyses of each region are depicted in Fig. 4. The three trees show an overall similarity to each other with reduced support values compared to the tree based on the full alignment (Fig. 3). However, we can identify some robust incongruences among the trees supported by bootstrap values >60% (marked with an asterisk on Fig. 4). The incongruences occurred among potato or tomato sequences independently. To evaluate the statistical significance of those incongruences, we performed statistical analyses on the tomato and potato data sets separately. First, we performed phylogenetic analyses on each of the three gene regions of *StMi1h* and *Mi-1* homologs independently, keeping an identical taxon sampling for each gene region (Figures S3 and S4, respectively). These trees contain a smaller taxon sampling than the analyses in Fig. 4 because it only includes those gene sequences that are long enough to cover the three gene regions. Looking at Figures S3 and S4, we found that incongruences among the trees based on different gene regions were evident. AU tests rejected ($p < 0.01$) the hypothesis that the trees of potato *StMi1h* sequences based on N-terminal, NBS and LRR regions were equally likely (Figure S3 B). This indicates that the evolutionary history of a particular gene region has not been identical to another gene region of the same gene. Given that the *StMi1h* genes belong to a heterozygotic diploid nuclear genome from *S. tuberosum*, it suggests that crossing over and/or gene conversion events among alleles or gene copies have taken place. Putative gene conversion events during *StMi1h* gene evolution in *S. tuberosum* were analyzed with statistical measures: PHI test and Geneconv. The PHI test found statistically significant evidence for the occurrence of recombination ($p = 0.0003$). A total of five sequence exchanges were detected between 26 *StMi1h* genes using Geneconv ($p < 0.05$).

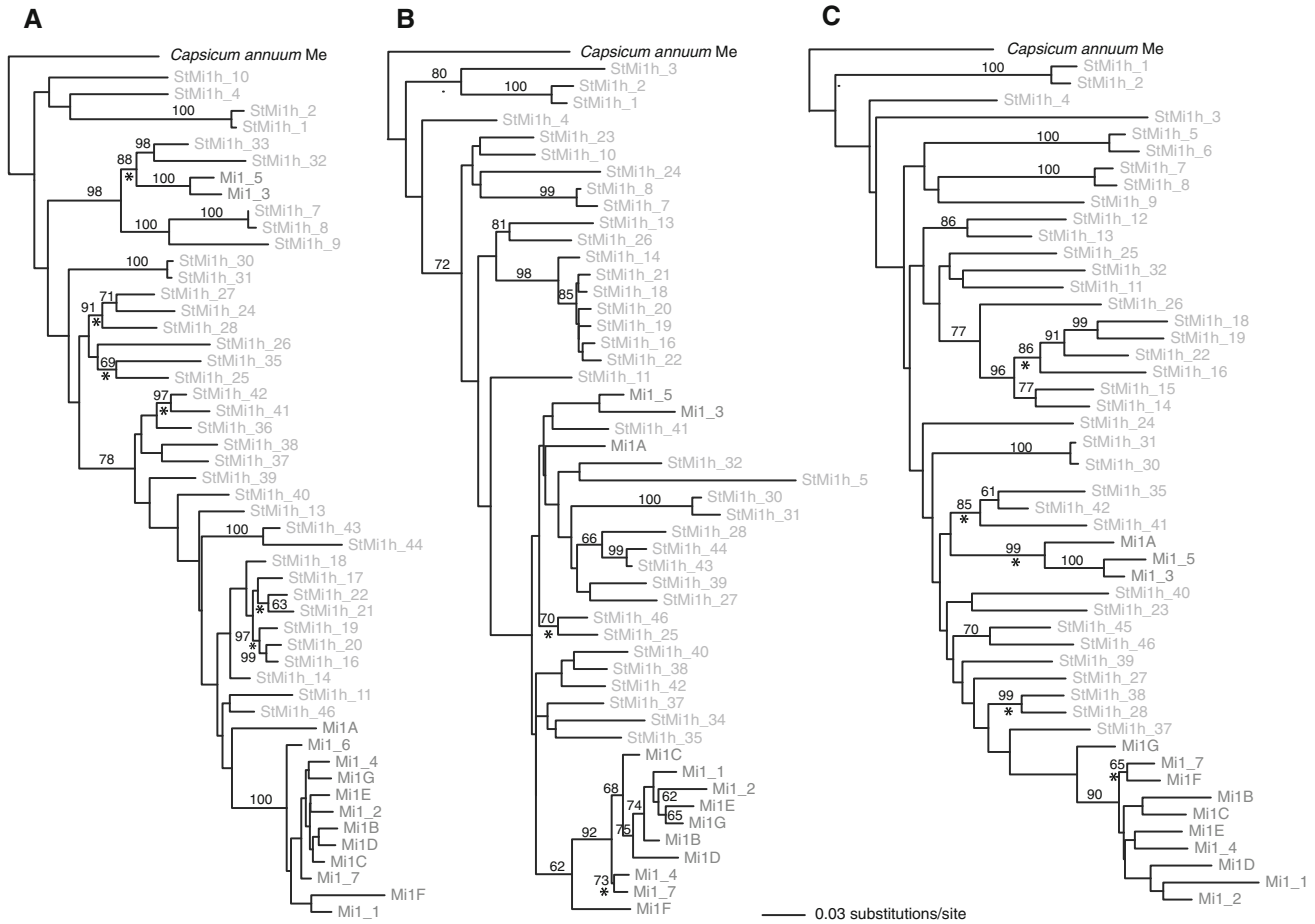

Fig. 4 Phylogenetic analyses of different regions of the *Mi-1*-like genes: **a** N-terminal (2,088 nt long); **b** NBS (873 nt long); and **c** LRR (1,068 nt long). Best trees based on nucleotides under maximum likelihood (*ML*) using the program Garli. Trees are plotted on a common scale. *Numbers* on branches correspond to bootstrap support values >59% from 100 bootstrap replicates. *Mi-1* and *StMi1h* refer to *S. lycopersicum* and *S. tuberosum* genes, respectively

In the case of tomato *Mi-1* genes, the trees based on individual gene regions were significantly different to each other (AU $p < 0.01$; Figure S4). This result suggests an incongruent evolution of each gene region in these two tomato lineages (resistant and susceptible cultivars). It is important to note that resistant cultivars carry genes that are introduced from the wild tomato *S. peruvianum*. Therefore, we should think of the *Mi-1* genes from the resistant tomato cultivars (*Mi-1.1–Mi-1.7*) as sequences from *S. peruvianum*. In this case, gene conversion may have taken place before the divergence of these two lineages or during the evolution of each tomato lineage. A total of six gene conversion events were identified among the 13 *Mi-1* genes from tomato using Geneconv ($p < 0.05$).

Comparison of *Mi-1* homologs in another potato species

We wished to understand the evolution of *Mi-1* genes among potato lineages; thus, we took advantage of a sequencing project of *S. phureja* DM1-3 (Visser et al. 2009). By using BLASTn and selecting the genome survey sequence (GSS) GenBank database, which contains ~350,000 entries of *S. phureja*, we retrieved 17 *S. phureja* sequences. The sequences were 618–939 nt in length (Table S1) and were included in the overall alignment of *Mi-1* homologs. In addition, two *Mi-1* homologs from *Capsicum annuum* were included (*CaMi* and *Me*). Phylogenetic analyses were performed on this larger data set and the cladogram was arbitrarily rooted (Fig. 5). The tree showed that *S. phureja* sequences (in boldface in Fig. 5) do not clade together, but form eight small groups with *StMi1h* sequences from *S. tuberosum*, most of which are strongly supported. No sequence from *S. phureja* is sister to any tomato *Mi-1* sequence. Highly similar sequences from *S. phureja* might be partial sequences of the same gene in the *S. phureja* genome. For example, the sequences *SpMi1h_11*, *SpMi1h_12* and *SpMi1h_13* seem to be overlapping reads of a single gene and form a 1,465-bp contig. Similarly, the sequences *SpMi1h_4* and *SpMi1h_5* form a 1,149-bp contig. The gene *CaMi* from *C. annuum* is surprisingly similar to the gene *Mi1-2* from tomato (Chen et al. 2007) and cluster together with 100% bootstrap support.

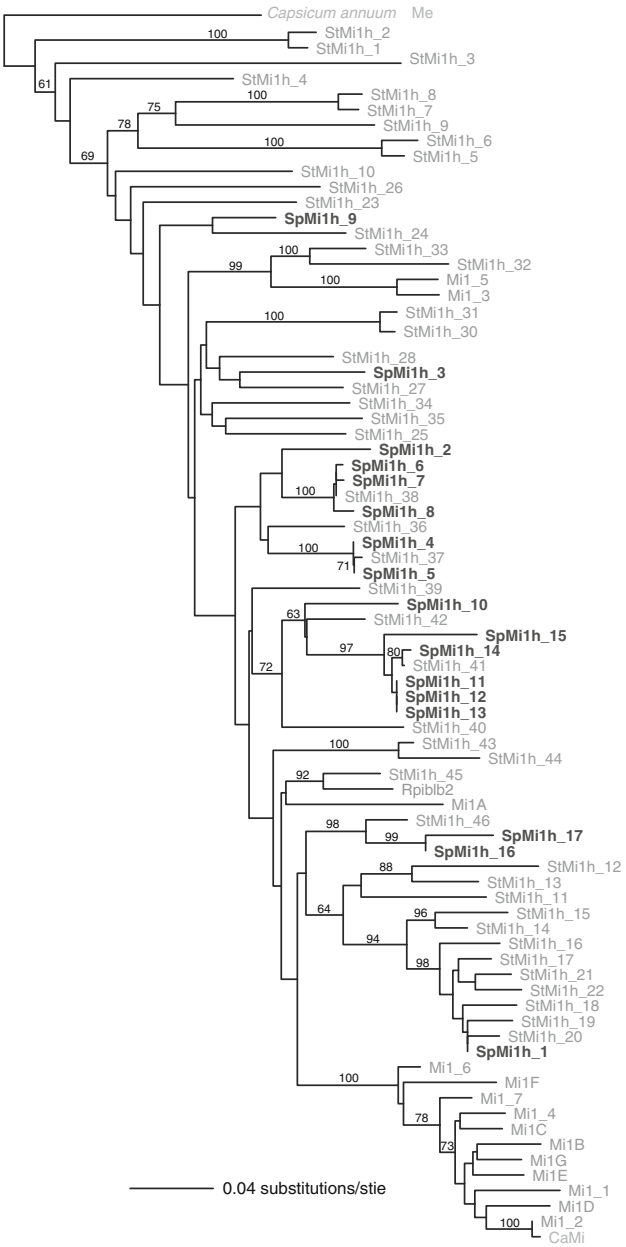

Fig. 5 Phylogenetic analyses of *Mi-1* homologs from *S. lycopersicum* (*Mi1* genes), *S. tuberosum* (*StMi1h* genes) and *S. phureja* (*SpMi1h* genes). Maximum likelihood tree based on 3,888 exonic positions using GTR+I+ Γ4 model of substitution and arbitrarily rooted with *Capsicum annuum* gene *Me*. *Numbers* on branches correspond to bootstrap support values >59% from 100 bootstrap replicates

Discussion

Organization of *Mi-1* homologs in *Solanum*

The cultivated tomato (*S. lycopersicum* L.) and potato (*S. tuberosum* L.) belong to closely related but distinct sections (sections *Lycopersicum* and *Petota*, respectively) of the genus *Solanum* (Fig. 6). The *S. tuberosum* genome is similar in size to the *S. lycopersicum* genome (850 Mb vs. 1,000 Mb, respectively) with a high level of macrocolinearity and microsinteny (Tanksley et al. 1992; Bryan and Hein 2008) and similar gene content, sequence and order (Zhu et al. 2008). Comparative analyses revealed that resistance genes were present in homologous locations in the genomes of potato and tomato (Grube et al. 2000; Gebhardt and Valkonen 2001). In addition, resistant genes in plant

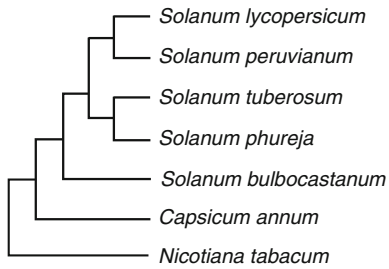

Fig. 6 Cladogram showing the phylogentic relationships among representative solanaceous species

genomes are often arranged in clusters (Michelmore and Meyers 1998; Gebhardt and Valkonen 2001; van der Linden et al. 2004). The clusters comprise tandem arrays of genes that determine resistance to multiple pathogens as well as to multiple variants of a single pathogen (Kuang et al. 2004).

The short arm of chromosome 6 is particularly interesting to study in the plant family Solanaceae, given that it is a hotspot for R genes. All seven *Mi-1* genes (*Mi-1.1–Mi-1.7*) in the resistant cultivated tomato *S. lycopersicum* are contained within a 650-kb region introgressed from *S. peruvianum* and are grouped in two clusters separated by 300 kb (Seah et al. 2007). Unsurprisingly, most *Mi-1* homologs in *S. tuberosum* (*StMi1h* genes) are also located on chromosome 6 of the potato genome (Visser et al. 2009) and arranged in clusters (Fig. 2). Even though the BAC clones analyzed here are unfinished and not yet assembled, we found 4–7 *StMi1h* genes in each of seven BAC clones from *S. tuberosum*, indicating that *StMi1h* genes form clusters (Fig. 2). In addition, some BAC clones from *S. tuberosum* overlap indicating that bigger *Mi-1* gene clusters are formed. For example, BAC clones AC238274 and AC237844 overlap and contain a total of eight *StMih* genes in less than 180 kb. Gene duplications and transposon insertions may have contributed to the generation of these gene clusters and diversity of gene copy number in *Solanum*, as suggested for other gene clusters (Michelmore and Meyers 1998; Kuang et al. 2005; Friedman and Baker 2007). A previous study suggested that the *Rpiblb2* gene belongs to a cluster in the wild potato *S. bulbocastanum* that has expanded in comparison to the seven *Mi-1* gene cluster described in tomato (van der Vossen et al. 2005).

Number and diversity of *Mi-1* genes in the genus *Solanum*

Analyses of partial genome sequences of *S. tuberosum* and *S. phureja* indicated the presence of a great number of *Mi-1* homologs in the nuclear genome of these tuber-bearing species (~42–45 homologs in *S. tuberosum* and 14 in *S. phureja*). The difference in number of *Mi-1* homologs in the two potato species may be due to sampling error given the limited sequences available from *S. phureja* in sequence databanks (~680,000 nucleotide sequences available from *S. phureja* and ~1,400,000 nucleotide sequences from *S. tuberosum*). In addition, a higher number of different *StMi1h* sequences of *S. tuberosum* may be due to the heterozygotic nature of this genome (Visser et al. 2009). In such case, distinct sequences may represent different alleles and not different gene copies. For example, BAC clones AC238283 and AC231973 may carry different alleles of the same genes (*StMi1h6-5, StMi1h7-8* and *StMi130-31*) given that they are syntenic (Fig. 2) and the pairs of genes are closely related in the phylogenetic tree (Fig. 3). We can rule out that they represent two sequences of exactly the same portion of the genome, because they show sufficient differences to distinguish them. The observed differences are much greater than the estimated sequencing error, which is <0.3 errors per 1,000 bp. The sequencing error was calculated by comparing all sections of BAC clones that were sequenced twice.

Analyses of the *S. lycopersicum* genome, by BLAST of the sequenced genome (this study), Southern blots (Milligan et al. 1998; Seah et al. 2004) and RFLP studies (Seah et al. 2004), revealed that susceptible *S. lycopersicum* contained nine *Mi-1* homologs in its nuclear genome: seven in chromosome 6 and two in chromosome 5. Considering that the sequenced tomato is a homozygous diploid, there are nine *Mi-1* homologs per haplotype. Following the same reasoning, *S. tuberosum* RH is a heterozygous diploid, and thus contains 21–23 (42–45 divided by two) homologs per haplotype, while *S. phureja* DM is a doubled haploid with 14 homologs per haplotype Therefore, the estimated number of tomato *Mi-1* homologs per haplotype is about one-third the estimated number of *Mi-1* homologs in *S. tuberosum*. The significant expansion of this gene family in potato could respond to several factors, including their differences in reproductive behavior: the cultivated tomato reproduce by self-fertilization and potatoes by outcrossing (Cipar et al. 1964; Frankel and Galun 1977).

The phylogenetic tree in Fig. 5 shows that repeated duplications and divergence might have occurred early in the evolution of *Mi-1* homologs in tuber-bearing species. Several *Mi-1* homologs were maintained through the speciation of *S. tuberosum* and *S. phureja* and are evidenced by the sister relationships of *Mi-1* homologs from *S. tuberosum* and *S. phureja* observed in the phylogeny. It is possible that *Mi-1* duplications occurred in the ancestor of *S. tuberosum* and *S. phureja* and also during the independent evolution of these two lineages. The evolution of the *Mi-1* gene family parallels the evolution of the R3 complex in the *Solanum* species. The R3 complex locus located in chromosome 11 from potato is involved in oomycete resistance and has evolved after divergence from tomato

and this cluster has suffered an impressive diversification and multiplication in potatoes without disrupting the flanking colinearity (Huang et al. 2005).

Evolution of *Mi-1* homologs

Plant R gene clusters contain paralogs that confer resistance to distinct pathogens or variants of a single pathogen, and also pseudogenes and gene family members with unknown functions, as observed in *S. lycopersicum* and *S. tuberosum Mi-1*-like clusters. These clusters may evolve by sequence exchanges (unequal crossovers, gene conversions) between homologs in the same cluster (i.e., paralogs), resulting in diverse, chimeric genes (Kuang et al. 2005; Friedman and Baker 2007). Based on the frequency of these sequence exchanges, R genes display contrasting evolutionary patterns (Michelmore and Meyers 1998; Kuang et al. 2005; Friedman and Baker 2007).

Previous studies suggested that *Mi-1* genes in tomato have evolved rapidly by gene duplications and high frequent sequence exchanges among gene copies within the same cluster, erasing orthologous relationships and resulting in highly similar *Mi-1* genes (Friedman and Baker 2007; Seah et al. 2007). Phylogenetic analyses of individual gene regions of ten *Mi-1* genes from susceptible and resistant tomato cultivars found incongruent trees, consistent with putative sequence exchanges among gene copies (Seah et al. 2007). In addition, comparisons of indels in homologous sequences suggested that a patchwork of sequence exchanges had occurred (Seah et al. 2007). No statistical analyses were provided in such studies. Here, we re-analyzed the tomato *Mi-1* sequences and performed phylogenetic analyses including all 14 *Mi-1* genes. We found that trees based on individual gene regions were significantly incongruent with each other (AU $p < 0.01$). Furthermore, recombination tests confirmed that 6 gene conversion events among 13 *Mi-1* genes from tomato had occurred (Geneconv $p < 0.05$). These data support the hypothesis that tomato *Mi-1* genes underwent frequent recombination and gene conversion events that produced chimeric and homogenous *Mi-1* sequences. Numerous sequence exchanges between paralogs in R gene clusters have also been reported in several other plants (Parniske et al. 1997; Michelmore and Meyers 1998; Cooley et al. 2000; Dodds et al. 2001).

Similar analyses with *Mi-1* homologs from *S. tuberosum* showed slightly different evolutionary patterns. Significant differences were found in the evolution of the three gene regions of *StMi1h* genes (AU $p < 0.01$), although overall congruence was observed. Only five gene conversion events (Geneconv $p < 0.05$) were identified among the 26 *StMi1h* genes from *S. tuberosum* analyzed. Furthermore, it is important to note that the recombination events detected could have happened between alleles of the same gene (i.e., interallelic recombination), given that the *S. tuberosum* sequences may include both different gene family members and alleles. In addition, we observed that several sequences from *S. tuberosum* were sister to homologous genes from *S. phureja*, suggesting that orthologous relationships had not been lost. However, this could be an artifact of the origin of *S. tuberosum* RH. The clone RH89-039-16 originated from a cross between clones BC1034 and SUH2293. In turn, BC1034 descended from a cross between clones USW5295.7 and USW5337.3, both of which were derived from a cross between *S. phureja* and haploid *S. tuberosum* (Ramanna 1983; Rouppe van der Voort et al. 1997). Given this pedigree, it is expected to see sister relationships of gene copies of *S. tuberosum* and *S. phureja*, which could be remnants of the *S. phureja* genome in *S. tuberosum* RH. The amount of sister relationships suggest that at least some of them are due to their common ancestry and conservation of orthologous relationships. To test these hypotheses, additional sequences from other potato species will be highly useful.

Capsicum annuum carries an almost identical copy of the gene *Mi-1.2*

It is worth noting that one copy of the *Mi-1* homologs from *C. annuum* (*CaMi*) is extremely similar (98% nucleotide identity) to the gene *Mi-1.2* from tomato (Chen et al. 2007). The gene *CaMi* shares its function with *Mi-1.2* conferring resistance to the nematode *M. incognita* in *CaMi*-containing pepper or tomato plants (Chen et al. 2007). It is most puzzling that no *Mi-1* homolog sequenced from potatoes (*S. tuberosum* or *S. phureja*) is as similar as *CaMi* to the gene *Mi-1.2*. This could be the result of sampling error; i.e., such a gene has not yet been sequenced in potatoes, but it does exist. Under this scenario, we may find a *Mi-1* homolog highly similar to the gene *Mi-1.2* in wild potato species that are resistant to *M. incognita* (Nirula et al. 1969; Berthou et al. 2003). If a gene similar to *Mi-1.2* did not exist in potatoes, an explanation could entail the loss of such an *Mi-1* gene in both potato lineages (*S. bulbocastanum* and *S. tuberosum* + S. *phureja*), but maintained in both tomatoes and peppers (Fig. 6). Another scenario could be that the gene *Mi-1.2* originated in the tomato lineage and was laterally transferred from tomato to *C. annum*. Horizontal gene transfer among closely related plant species has been increasingly recognized (Keeling and Palmer 2008; Sanchez-Puerta et al. 2008). Out of all these possibilities, we believe that it is most likely that some potatoes do carry such a gene, but has not been identified yet.

Conclusions

Mi-1 homolog genes in the genus *Solanum* show a variety of evolutionary patterns. *Mi-1* genes from susceptible and

resistant cultivars of *Solanum lycopersicum* (resistant cultivars contain seven genes introduced from the wild tomato *S. peruvianum*) underwent concerted evolution by experiencing unequal crossing over and gene conversion events. In consequence, phylogenetic analyses show that orthologous relationships of *Mi-1* genes with *Mi-1* homologs from other *Solanum* species are blurred and paralogs are more similar than orthologs. This conclusion is strengthened by frequent recombination events that were detected by statistical analyses. In contrast, *Mi-1* homologs from cultivated potato species (*S. tuberosum* and *S. phureja*) may have evolved by a birth-and-death process, in which genes evolve mostly by mutations and interallelic recombinations. The "birth-and-death model" of evolution of a multigene family postulates that new genes are originated by gene duplication and duplicates are either maintained in the genome for a long time, deleted or become pseudogenes by accumulating deleterious mutations (Nei et al. 1997). Evidence for this type of evolution in *StMi1h* and *SpMi1h* genes include: (a) putative conservation of orthologous relationships among these genes; (b) the origin of additional gene copies by duplication events in the potato genome; (c) presence of pseudogenes; and (d) evidence of occasional recombination events between different gene copies.

Sequencing *Mi-1* homologs from a wide taxonomic range of potato-like species will increase our understanding of the evolution of genes in the *Mi-1* cluster in the genus *Solanum* and the impressive diversification of *Mi-1* genes in potato in comparison to those in tomato. In particular, it would allow identifying more easily those genes that could confer resistance or tolerance to pathogens.

Acknowledgments We would like to thank the anonymous reviewers for their highly valuable suggestions. This study was supported by grants from ANPCyT (PICT-2008-277) and Universidad Nacional de Cuyo (Sectyp 06/M022) to MVSP and RWM, who are researchers from CONICET.

References

Albrecht M, Takken F (2006) Update on the domain architectures of NLRs and R proteins. Biochem Biophys Res Commun 339:459–462

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Berthou F et al (2003) Enhancing the resistance of the potato to Southern root-knot nematodes by using *Solanum sparsipilum* germplasm. Euphytica 132:57–65

Bruen T, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. Genetics 172:2665–2681

Bryan G, Hein I (2008) Genomic resources and tools for gene function analysis in potato. Int J Plant Genomics 2008:1–9

Chen R, Li H, Zhang L, Zhang J, Xiao J, Ye Z (2007) *CaMi*, a root-knot nematode resistance gene from hot pepper (*Capsicum annuum* L.) confers nematode resistance in tomato. Plant Cell Rep 26:895–905

Cipar M, Peloquin S, Hougas R (1964) Variability in the expression of self-incompatibility in tuber-bearing diploid *Solanum* species. Am Potato J 41:155–162

Cooley M, Pathirana S, Wu H, Kachroo P, Klessig D (2000) Members of the *Arabidopsis HRT/RPP8* family of resistance genes confer resistance to both viral and oomycete pathogens. Plant Cell 12:663–676

Dodds P, Lawrence G, Ellis J (2001) Contrasting modes of evolution acting on the complex *N* locus for rust resistance in flax. Plant J 27:439–453

Dropkin V (1969) Cellular responses of plants to nematode infections. Annu Rev Phytopathol 7

Frankel R, Galun E (1977) Pollination mechanisms, reproduction and plant breeding. Springer, Berlin

Friedman A, Baker B (2007) The evolution of resistance genes in multi-protein plant resistance systems. Curr Opin Genet Dev 17:493–499

Gebhardt C, Valkonen J (2001) Organization of genes controlling disease resistance in the potato genome. Annu Rev Phytopathol 39:79–102

Gilbert J, McGuire D (1956) Inheritance of resistance to severe root-knot, *Meloidogyne incognita*, in commercial type tomatoes. Proc Am Soc Hortic Sci 68:437–442

Grube R, Radwanski E, Jahn M (2000) Comparative genetics of disease resistance within the Solanaceae. Genetics 155:873–887

Huang S et al (2005) Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. Plant J 42:251–261

Huson D, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267

Jablonska B et al (2007) The *Mi-9* gene from *Solanum arcanum* conferring heat-stable resistance to root-knot nematodes is a homolog of *Mi-1*. Plant Physiol 143:1044–1054

Kaloshian I et al (1998) Genetic and physical localization of the root-knot nematode resistance locus *Mi* in tomato. Mol Gen Genet 257:376–385

Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev 9:1–14

Kuang H, Woo S, Meyers B, Nevo E, Michelmore R (2004) Multiple genetic processes results in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. Plant Cell 16:2870–2894

Kuang H et al (2005) The *R1* resistance gene cluster contains three groups of independently evolving, type I *R1* homologues and shows substantial structural variation among haplotypes of *Solanum demissum*. Plant J 44:37–51

Maddison W, Maddison P (2000) MacClade version 4: analysis of phylogeny and character evolution. Sinauer Associates, Sunderland

Martin G, Bogdanove A, Sessa G (2003) Understanding the functions of plant disease resistance proteins. Annu Rev Plant Biol 54:23–61

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

Milligan S, Bodeau J, Yaghoobi J, Kaloshian I, Zabel P (1998) The root knot nematode resistance gene Mi from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. Plant Cell 10:1307–1320

Mueller L, Lankhorst R, Tanksley S et al (2009) A snapshot of the emerging tomato genome sequence. The Plant Genome 2:78–92

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nei M, Gu X, Sitnikova T (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci USA 94:7799–7806

Nirula K, Khushu C, Raj B (1969) Resistance in tuber-bearing *Solanum* species to root knot nematode, *Meloidogyne incognita*. Am Potato J 46:251–253

Niu C, Hinchliffe D, Cantrell R, Wang CY, Roberts P, Zhang J (2007) Identification of molecular markers associated with root-knot nematode resistance in upland cotton. Crop Sci 47:951–960

Nombela G, Williamson V, Muniz M (2003) The root-knot nematode resistance gene Mi-1.2 of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. Mol Plant Microbe Interact 16:645–649

Parniske M et al (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

Philippe H, Snell E, Bapteste E, Lopez P, Holland P, Casane D (2004) Phylogenomics of eukaryotes: impact of missing data on large alignments. Mol Biol Evol 21:1740–1752

Posada D, Crandall K (1998) ModelTest: testing the model of DNA substitution. Bioinformatics 14:817–818

Quevillon E et al (2005) InterProScan: protein domains identifier. Nucleic Acids Res 33:116–120

Rairdan G, Moffett P (2006) Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation. Plant Cell 18:2082–2093

Ramanna M (1983) First division restitution gametes through fertile desynaptic mutants of potato. Euphytica 32:337–350

Rossi M, Goggin F, Milligan S, Kaloshian I, Ullman D, Williamson V (2003) The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc Natl Acad Sci USA 95:9750–9754

Sanchez-Puerta MV, Cho Y, Mower JP, Alverson AJ, Palmer JD (2008) Frequent, phylogenetically local horizontal transfer of the cox1 group I intron in flowering plant mitochondria. Mol Biol Evol 25:1762–1777

Sasser J, Carter C (1985) An advance treatise on *Meloidogyne*. NC State Univ Graph, Raleigh

Sawyer S (1989) Statistical tests for detecting gene conversion. Mol Biol Evol 6:526–538

Seah S, Yaghoobi J, Rossi M, Gleason C, Williamson V (2004) The nematode-resistance gene, *Mi-1*, is associated with an inverted chromosomal segment in susceptible compared to resistant tomato. Theor Appl Genet 108:1635–1642

Seah S, Telleen A, Williamson V (2007) Introgressed and endogenous *Mi-1* gene clusters in tomato differ by complex rearrangements in flanking sequences and show sequence exchange and diversifying selection among homologues. Theor Appl Genet 114:1289–1302

Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. Syst Biol 51:492–508

Swofford D, Olsen G, Waddell P, Hillis D (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Tanksley S et al (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141–1160

van der Linden G, Wouters D, Mihalka V, Kochieva E, Smulders M, Vosman B (2004) Efficient targeting of plant disease resistance loci using NBS profiling. Theor Appl Genet 109:384–393

van der Rouppe Voort J et al (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol Gen Genet 255:438–447

van der Vossen E et al (2005) The *Rpi-blb2* gene from *Solanum bulbocastanum* is an *Mi-1* gene homolog conferring broad-spectrum late blight resistance to potato. Plant J 44:208–222

van Ooijen G, van den Burg H, Cornelissen B, Takken F (2007) Structure and function of resistance proteins in Solanaceous plants. Annu Rev Phytopathol 45:43–72

van Ooijen G, Mayr G, Kasiem M, Albrecht M, Cornelissen B, Takken F (2008) Structure–function analysis of the NB–ARC domain of plant disease resistance proteins. J Exp Botany 59:1383–1397

Visser RGF et al (2009) Sequencing the potato genome: outline and first results to come from the elucidation of the sequence of the world's third most important food crop. Am J Potato Res 86:417–429

Vos P et al (1998) The tomato *Mi-1* gene confers resistance to both root-knot nematodes and potato aphids. Nat Biotechnol 16:1365–1369

Wiens J (2006) Missing data and the design of phylogenetic analyses. J Biomed Inform 39:34–42

Williamson V, Kumar A (2006) Nematode resistance in plants: the battle underground. Trends Genet 22:396–403

Zhu W et al (2008) Analysis of 90 Mb of the potato genome reveals conservation of gene structures and order with tomato but divergence in repetitive sequence composition. BMC Genomics 9:286

Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. University of Texas, Austin