rich2/cge-gen/cge-gen/cge99915/cge0365d15z | xppws | S=3 | 11/29/16 | 7:12 | Art: gen-2016-0101 | Input-1st disk, 2nd ??
Pagination not final (cite DOI) / Pagination provisoire (citer le DOI)

NRC Research Press

ARTICLE

Dissimilar evolutionary histories of two resistance gene families in the genus *Solanum*

Diana María Segura, Ricardo Williams Masuelli, and M. Virginia Sanchez-Puerta

Abstract: Genomic analyses have shown that most genes in eukaryotic lineages belong to families. Gene families vary in terms of number of members, nucleotide similarity, gene integrity, expression, and function. Often, the members of gene families are arranged in clusters, which contribute to maintaining similarity among gene copies and also to generate duplicates through replication errors. Gene families offer us an opportunity to examine the forces involved in the evolution of the genomes and to study recombination events and genomic rearrangements. In this work, we focused on the evolution of two plant resistance gene families, *Sw5* and *Mi-1*, and analyzed the completely sequenced nuclear genomes of potato and tomato. We first noticed that the potato genome carries larger resistance gene families than tomato, but all gene copies are pseudogenes. Second, phylogenetic analyses indicated that *Sw5* and *Mi-1* gene families had dissimilar evolutionary histories. In contrast to *Sw5*, *Mi-1* homologues suffered repeated gene conversion events among the gene copies, particularly in the tomato genome.

Key words: resistance, *Solanum*, evolution, *Mi-1*, *Sw5*.

Résumé : Des analyses génomiques ont montré que la plupart des gènes chez les eucaryotes appartiennent à des familles multigéniques. Ces familles varient en ce qui a trait au nombre de membres, à la similarité nucléotidique, à l'intégrité génique, à leur expression et à leur fonction. Souvent, les membres d'une famille sont disposés en suites, lesquelles contribuent à maintenir la similarité entre les copies de gènes et à générer des duplications via des erreurs de réplication. Les familles multigéniques offrent l'occasion d'examiner les forces derrière l'évolution des génomes et pour étudier les événements de recombinaison et les réarrangements génomiques. Dans ce travail, les auteurs se sont concentrés sur l'évolution de deux familles de gènes de résistance chez les plantes, *Sw5* et *Mi-1*. Les auteurs ont analysé ceux-ci au sein des génomes complètement séquencés de la pomme de terre et de la tomate. Les auteurs ont d'abord noté que le génome de la pomme de terre contient de plus grandes familles de gènes de résistance que celui de la tomate, mais que toutes les copies sont des pseudogènes. Ensuite, des analyses phylogénétiques ont indiqué que les familles de gènes *Sw5* et *Mi-1* présentaient des historiques évolutifs dissemblables. Au contraire de *Sw5*, les homologues de *Mi-1* ont connu de nombreux événements de conversion génique entre copies, particulièrement au sein du génome de la tomate. [Traduit par la Rédaction]

Mots-clés : résistance, *Solanum*, évolution, *Mi-1*, *Sw5*.

Introduction

New genes in a genome are commonly originated by duplication, chromosomal rearrangements, and subsequent divergence from pre-existing genes (Lawton-Rauh 2003). Gene duplication may occur by local or genome-wide events, such as polyploidization, and lead to the formation of a gene family. A gene family is a set of similar genes formed by duplication that generally share related biochemical functions. Most genes within a plant genome belong to gene families that originated as tandem duplicates, dispersed duplications, or genome-wide duplications, and are often arranged in clusters (Chen et al. 2007; Jia et al. 2015; Wu et al. 2014). A great number of gene families were described in plant genomes, including the phytochrome photoreceptor family in tomato (Alba et al. 2000), the apyrase family in legumes that plays an important role in nutrition, photomorphogenesis, and nodulation (Cannon et al. 2003), and the

Received 5 May 2016. Accepted 11 July 2016.

Corresponding Editor: B. Golding.

D.M. Segura and R.W. Masuelli. IBAM, Universidad Nacional de Cuyo, CONICET, Facultad de Ciencias Agrarias, Almirante Brown 500, M5528AHB, Chacras de Coria, Argentina.
M.V. Sanchez-Puerta. IBAM, Universidad Nacional de Cuyo, CONICET, Facultad de Ciencias Agrarias, Almirante Brown 500, M5528AHB, Chacras de Coria, Argentina; Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, 5500, Mendoza, Argentina.

Corresponding author: M. Virginia Sanchez-Puerta (email: mvsanchezpuerta@fca.uncu.edu.ar).

Published at www.nrcresearchpress.com/gen on 14 September 2016.

calmodulin family in Solanaceae, which codes for different calcium sensor isoforms (Zhao et al. 2013). In addition, most resistance genes belong to gene families and are found in clusters (Hulbert et al. 2001; Andolfo et al. 2013). In some cases, gene family members are highly similar and produce large amounts of a single product, usually needed in a demanding metabolic process (Walsh et al. 2001). These gene families are subject to concerted evolution, maintaining highly similar copies by repeated gene conversion. In contrast, other gene families present dissimilar copies that recombine and exchange fragments, generating new allelic forms and increased variability (Ashfield et al. 2012). In addition, gene families whose members are arranged in clusters can suffer genesize amplification by tandem duplications (due to replication slippage or unequal recombination), duplications of large regions (even the whole genome), transposition of DNA sequences, or retrotransposition of RNA transcripts (Thornton et al. 2000). The birth-and-death model of gene families implies that whereas some duplicated genes are maintained in the genome for a long time, others are deleted or inactivated through deleterious mutations (Nei and Rooney 2005). It has been suggested that disease resistance loci followed this model of evolution in plants (Michelmore and Meyers 1998; Nam et al. 2004). Their selective effects and population genetic parameters, such as effective population size, will determine the fate of these genomic changes.

In this work, we studied the evolution of two resistance gene families in the genus *Solanum*, family Solanaceae. We focused on resistance gene families because they show a complex evolution in previous studies of the family Solanaceae (Bakker et al. 2011; Lozano et al. 2012). The resistance gene family *Sw5*, which was described in the wild tomato *Solanum peruvianum* (Spassova et al. 2001), is involved in the resistance to the Tomato Spotted Wilt Virus. In particular, the *Sw5-b* gene copy confers resistance to this virus, and it has been mapped to the long arm of chromosome 9 in the tomato genome (Chagué et al. 1996; Stevens et al. 1991, 1995). The gene *Mi-1.2* confers resistance to the potato aphid *Macrosiphum euphorbiae*, the whitefly *Bemisia tabaci*, and a group of root-knot nematodes of the genus *Meloidogyne* (Dropkin 1969; Nombela et al. 2003; Rossi et al. 2003; Jablonska et al. 2007). This gene belongs to the *Mi-1* gene family, which forms a cluster of seven copies located in the short arm of chromosome 6 and two copies located in chromosome 5 of the *Solanum lycopersicum* nuclear genome (Seah et al. 2004, 2007). Approximately 45 *Mi-1* homologues were identified in the genome of *Solanum tuberosum* diploid heterozygous line RH89-039-16—five times more than those in tomato (Sanchez-Puerta and Masuelli 2011). However, because the potato line analyzed was heterozygous, it was difficult to determine whether those homologues were actual genes or different alleles (Sanchez-Puerta and Masuelli 2011). The genomic arrangement, number of copies, integrity, and evolution of these gene families in solanaceous species remain unclear.

Recently, complete genomic sequences of three species of *Solanaceae* were published: the double monoploid line *S. tuberosum* Group Phureja DM1-3 516R44 (The Potato Genome Sequencing Consortium 2011), *S. lycopersicum* cultivar Heinz 1706 (The Tomato Genome Consortium 2012), and *Capsicum annuum* cultivar CM334 (Kim et al. 2013). The availability of these genomes enabled us to study gene families in three important and closely related crops by comparative genomics. To contrast the evolution of the multigene families *Sw5* and *Mi-1*, we included three smaller resistance genes families (*Hero*, *Hpa2*, and *Pto*) and three single-copy genes (*Waxy*, *Atg-1*, and *Ufm-1*) not related to pathogen resistance.

Materials and methods

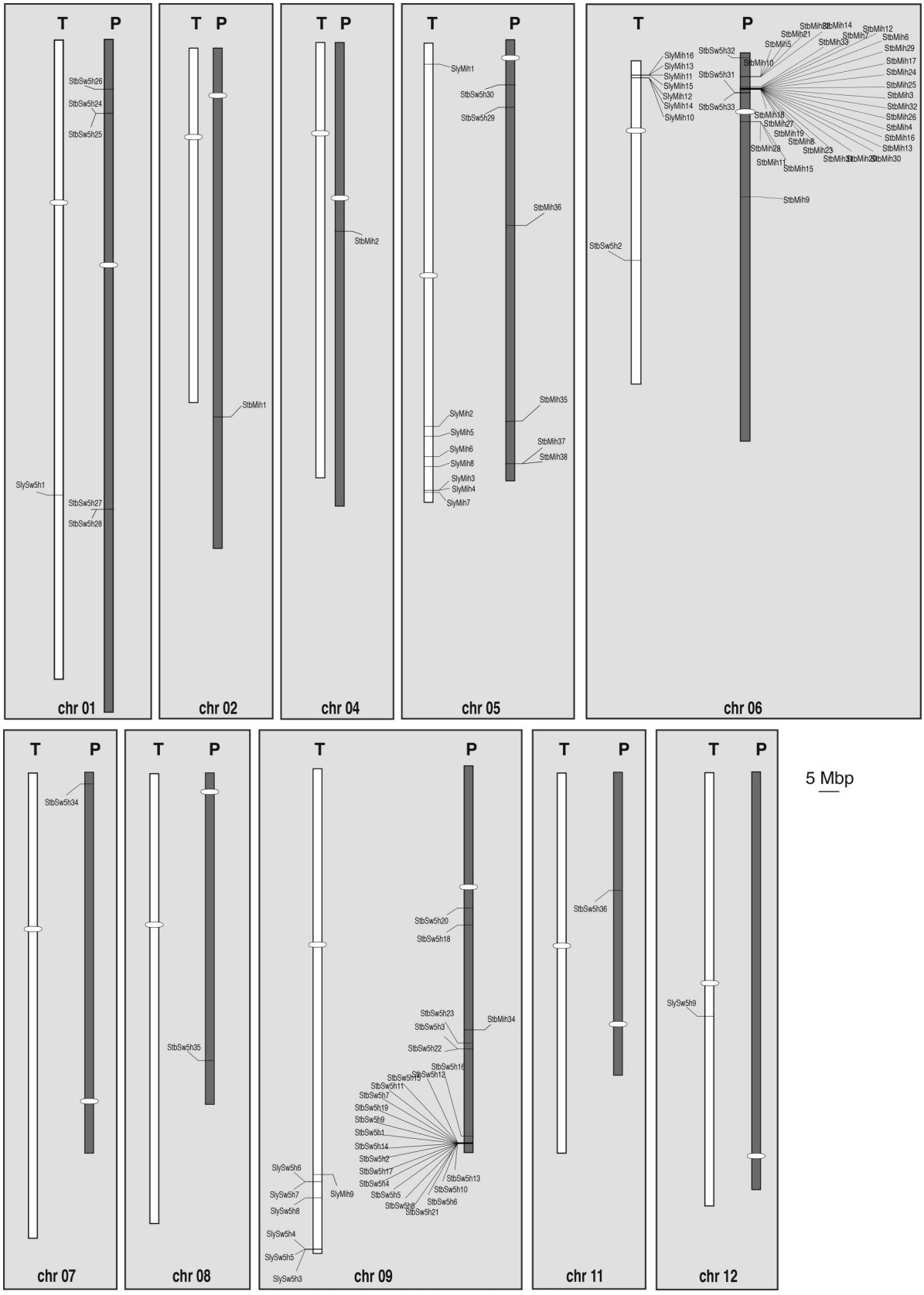
Nucleotide data

The genomes of *S. lycopersicum* cultivar Heinz 1706 and a double monoploid of *S. tuberosum* Group Phureja DM1-3 516R44 were analyzed with Blast. Query sequences included the *Sw5-b* gene from *S. lycopersicum* cultivar Stevens (AY007366; 3740 bp), the *Mi-1.2* gene from *S. lycopersicum* variety Motelle (AF091048, 3849 bp), *Hero* from *S. lycopersicum* (NM_001247066; 4280 bp), *Gpa2* from *S. tuberosum* (AF195939; 2710 bp), *Pto* from *S. peruvianum* (DQ019220; 915 bp), *Ufm-1* from *Arabidopsis thaliana* (NM_106420.4; 647 bp), *Waxy* from *S. tuberosum* (EU548081.1; 1320 bp), and *Agt1* from *A. thaliana* (NM_126925.4; 1540 bp). All hits with coverage >450 bp and E-value <$1e^{-50}$, including full length and truncated copies, were extracted from the genomic sequences. Coding sequences from *Capsicum annuum* cultivar CM334 genome (Kim et al. 2014) were analyzed with BLAST to identify all homologues to the genes *Sw5-b* and *Mi-1.2*. *Capsicum annuum* is closely related to the genus *Solanum* and can be helpful to analyze the evolutionary history of resistance gene families in potato and tomato. Chromosome maps were drawn with Vector NTI (Lu and Moriyama 2004) and edited with Adobe Illustrator. The approximate location of centromeres in the potato and tomato chromosomes was estimated from Park et al. (2007) and SolGenomics Network, respectively.

Phylogenetic analyses

Sw5 and *Mi-1* homologues identified in potato, tomato, and hot pepper were aligned individually using MacClade (Maddison and Maddison 2000). Each nucleotide alignment was subjected to manual editing taking into account the encoded amino acid sequences to identify positional homology. 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 (4 categories). One hundred bootstrap replicates were performed.

Fig. 1. Physical location of *Mi-1* and *Sw5* homologues of tomato and potato. Tomato and potato chromosomes are represented in white (T) and grey (P), respectively. *Sw5* and *Mi-1* homologues are indicated on the right and left of each chromosome, respectively. Missing chromosomes do not contain any *Mi-1* or *Sw5* homologues. The approximate location of centromeres in the potato and tomato chromosomes was estimated from Park et al. (2007) and SolGenomics Network, respectively.



Results

Sw5 gene family is four times greater in potato than in tomato

BLAST searches of *Sw5-b* against the nuclear genomes of *S. lycopersicum* and *S. tuberosum* yielded 9 (*SlySw5h*) and 36 (*StbSw5h*) homologues, respectively (Table S1)[1]. Only one homologue (*SlySw5h3* from tomato) covered the entire gene length (3740 bp) and the rest were variable in size (due to indels), ranging from 490 to 4572 bp (Table S1)[1]. Out of 45 *Sw5* homologues identified in both species of *Solanum*, 44 were classified as pseudogenes because they were truncated, contained premature stop codons, or had indels that led to frameshifts. The identity between potato homologues and the gene *Sw5-b* was 73%–94%; and between tomato homologues and *Sw5-b* was 77%–97% (Table S1)[1]. We identified 26 *Sw5* homologues in *C. annuum*, which were located in chromosomes 3, 7, and 10 (Table S3)[1]. The *Sw5* tomato homologues were found in four chromosomes (chr. 1, 6, 9, and 12), while those in the potato genome were identified in seven different chromosomes (chr. 1, 5, 6, 7, 8, 9, and 11) (Fig. 1). There were several non-syntenic *Sw5* homologues between potato and tomato genomes, such as those found in the potato chromosomes 1, 5, 6, 7, 8, and 11 and those in chromosomes 6 and 12 of tomato (Fig. 1). Most *Sw5* homologues were clustered in chromosomes 9 of both tomato and potato genomes. A particularly dense area encompassed 17 *Sw5* homologues clustered within ~360 kb of chromosome 9 of *S. tuberosum*.

Mi-1 homologues were mainly pseudogenes and clustered in chromosome 6 of tomato and potato

A BLAST search of *Mi-1.2* homologues against the complete nuclear genomes of *S. lycopersicum* and *S. tuberosum* found 16 (*SlyMih*) and 38 (*StbMih*) gene copies, respectively, although most of them were not full-length (Table S2)[1]. The functional gene copy *Mi-1.2* of *S. lycopersicum* has three exons, two of which (exons 2 and 3) are protein-coding exons. The first two exons and introns encompass 1532 bp and were identified only in 5 homologues of the tomato genome under analysis (*SlyMih10*, *SlyMih11*, *SlyMih12*, *SlyMih13*, and *SlyMih14*), and in 8 of the 38 homologues in the potato genome (*StbMih3*, *StbMih6*, *StbMih7*, *StbMih8*, *StbMih12*, *StbMih14*, *StbMih17*, and *StbMih18*). Exon 3 sequences were recognized in all homologues of tomato and potato showing variable lengths (Table S2)[1]. Out of the 54 *Mi-1* homologues identified in both genomes, 51 were classified as pseudogenes because they were truncated, contained premature stop codons, or had indels that led to frameshifts. Pseudogenes were variable in size, ranging from 717 to 5077 bp due to the occurrence of indels. The identity of each potato gene sequence to the tomato *Mi-1.2* copy ranged between 76% and 92%. The three full-length *Mi-1* copies (*SlyMih11*, *SlyMih13*, and *SlyMih14*) were found in the tomato genome and no complete genes were found in the potato genome. The *Mi-1* gene family in *Capsicum annuum* had 71 members located in chromosomes 2, 5 and 6 (Table S3). *Mi-1* homologues were found in three tomato chromosomes (chr 5, 6 and 9), and five chromosomes of potato (chr 2, 4, 5, 6 and 9) (Fig. 1). Tomato *Mi-1* homologues present in chromosome 6 clustered in two loci, while those in potato clustered in five loci of the same chromosome (Fig. 1). The three potato-specific clusters in chromosome 6, along with the homologues *StbMih1* (in chromosome 2), *StbMih2* (in chromosome 4), and several copies in chromosome 5 have no syntenic copies in tomato (Fig. 1).

Table 1. Number of gene copies in the tomato and potato genomes.

Gene name	Chromosome number[a]	No. of gene copies in *Solanum lycopersicum*	No. of gene copies in *Solanum tuberosum*
Mi-1.2 (3732 bp)	2	0	1
	4	0	1
	5	8	4
	6	7	31
	9	1	1
Sw5-b (3741 bp)	1	1	5
	5	0	2
	6	1	3
	7	0	1
	8	0	1
	9	6	23
	12	1	0
Hero (4280 bp)	3	1	1
	4	8	11
	6	1	1
	9	1	1
Gpa2 (3208 bp)	10	0	1
	12	3	15
Pto (915 bp)	3	0	1
	5	5	6
	6	1	1
	11	1	1
Waxy (1320 bp)	8	1	1
Agt-1 (1090 bp)	12	1	1
Ufm-1 (283 bp)	12	1	1

[a]Chromosomes that are not listed do not contain gene copies for the corresponding gene family.

Comparison with other resistance gene families and single copy genes

Blast searches of additional resistance gene families, *Hero*, *Gpa2*, and *Pto*, against the nuclear genomes of potato and tomato also yielded a greater number of gene

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2016-0101.

Fig. 2. Analyses of *Sw5* homologues. (A) Maximum Likelihood phylogenetic tree of 71 sequences from *Capsicum annuum* (CA), *Solanum lycopersicum* (SlySw5h), and *Solanum tuberosum* (StSw5h) based on 3741 bp. Numbers on branches correspond to bootstrap support values >50% from 100 bootstrap replicates. The chromosome location (C*n*) is shown next to the gene name. (B) Graphical representation of *Sw5* homologues located in chromosome 9 of *S. lycopersicum* (above) and *S. tuberosum* (below). Well-supported evolutionary relationships are indicated by solid lines. Arrows with numbers indicate *Sw5* homologues and letters represent different loci.

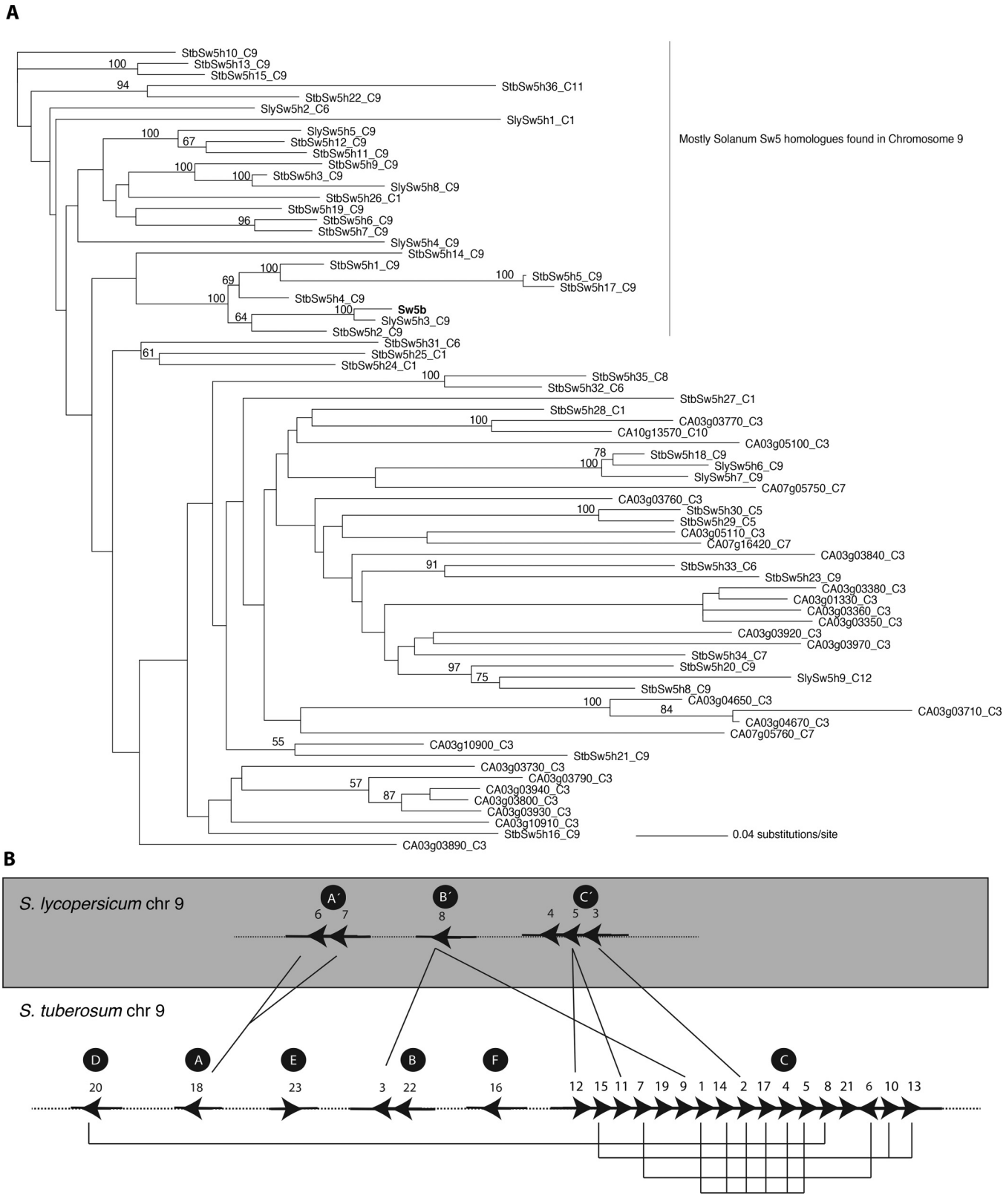


Fig. 3. Analyses of *Mi-1* homologues. (A) Maximum Likelihood phylogenetic tree of 126 sequences from *Capsicum annuum* (CA), *Solanum lycopersicum* (SlyMih), and *Solanum tuberosum* (StbMih) based on 3876 bp. Numbers on branches correspond to bootstrap support values >50% from 100 bootstrap replicates. The chromosome location (C*n*) is shown next to the gene name. (B) Graphical representation of *Mi-1* homologues from chromosome 6 from *S. lycopersicum* (above) and *S. tuberosum* (below). Well-supported evolutionary relationships are indicated by solid lines. Arrows with numbers indicate *Mi-1* homologues and letters represent different loci.

copies in *S. tuberosum* than in *S. lycopersicum*, and those copies were distributed in a higher number of chromosomes of the potato genome (Table 1). In contrast, low-copy genes (*Ufm-1*, *Waxy*, and *Agt1*) had only one copy of each gene in the nuclear genomes of potato and tomato, and they were in syntenic locations (Table 1).

Evolutionary analyses of *Sw5* homologues and their genomic distribution

The phylogenetic tree based on *Sw5* sequences showed that tomato, potato, and hot pepper *Sw5* homologues did not form species-specific monophyletic groups (Fig. 2A). In most cases, homologues found in the same chromosomes in both potato and tomato did not group together in the tree. The distribution of the *Sw5* in chromosomes 9 of potato and tomato was compared to their evolutionary relationships (Fig. 2B). Along chromosome 9, tomato *Sw5* homologues were found in three loci (A′, B′ and C′), and potato homologues in six, where loci D, E, and F were not present in tomato. In addition, 17 potato homologues formed a large cluster (Fig. 2B, cluster C). Based on the phylogenetic tree (Fig. 2A), most of these clustered homologues were closely related to each other, except for *StbSw5h8* and *StbSw5h21*. Finally, we observed a strongly supported sister relationship between the gene copy used as query for the Blast analyses (*Sw5b*) and the only putatively functional copy of *Sw5* identified in this study, *SlySw5h3* (Fig. 2A).

Evolutionary analyses of *Mi-1* homologues

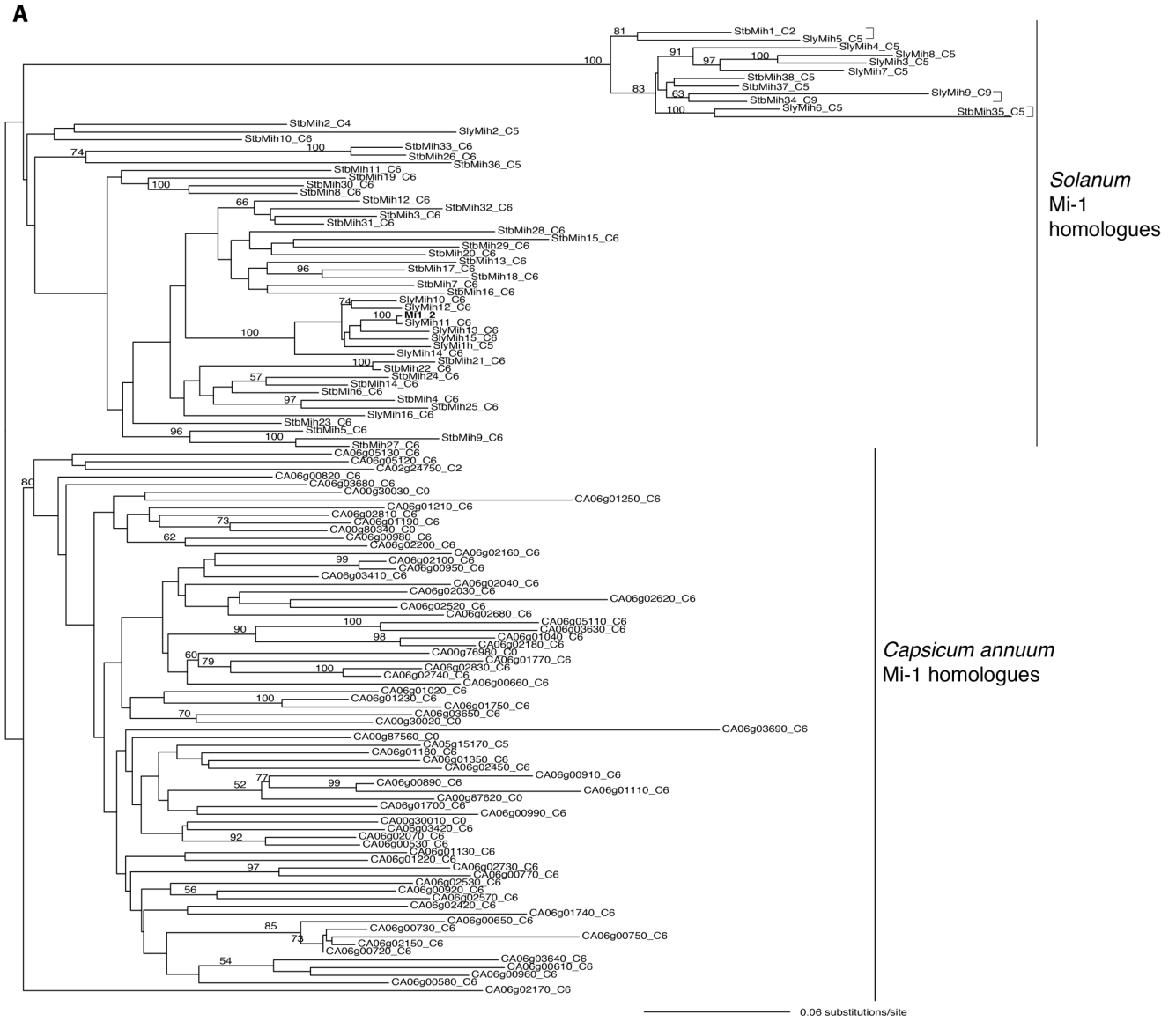
A phylogenetic analysis of *Mi-1* homologues showed that potato and hot pepper sequences were more numerous and diverse than those from tomato (Fig. 3A). All homologues from the genus *Solanum* formed a monophyletic clade divided in two main groups: (*i*) one clade formed by all sequences from chromosome 6 of either tomato or potato, the *SlyMih1* and *SlyMih2* genes from tomato chromosome 5, and the genes *StbMih2* and *StbMih36* from potato chromosomes 4 and 5, respectively; and (*ii*) the other clade included the majority of homologues located in chromosomes 2, 5, and 9 of either species (Fig. 3A). All but one tomato homologue from chromosome 6 clustered in a single highly supported group (BS = 100%) embedded within a group of potato sequences from chromosome 6. In contrast, tomato *Mi-1* homologues from chromosomes 5 and 9 were interspersed with potato homologues from chromosomes 2, 5, and 9. Individual tomato and potato sequences were not found as sister taxa, except for three pairs of homologues from different chromosomes (Fig. 3A, brackets). In addition, we observed a strongly supported sister relationship between the gene copy used as query for the Blast analyses (*Mi-1.2*) and one of the functional copies of *Mi-1* identified in this study, *SlyMih11* (Fig. 2A).

Comparing the physical location of *Mi-1* homologues in chromosome 6 of both species of *Solanum*, we observed that *Mi-1* copies formed clusters in two and five loci in tomato and potato, respectively (Fig. 3B). Cluster C in potato was the most highly populated. All homologues found in the two tomato clusters, except for *SlyMih16*, formed a well-supported clade in the evolutionary tree, closely related to the potato homologues found in chromosome 6 (Fig. 3A). Comparative analyses of the potato and tomato clusters revealed the presence of highly similar intergenic regions between clusters A and A′ from potato and tomato, respectively, indicating they were homologous. However, gene copies found in clusters A and A′ did not group together in the phylogenetic tree (Fig. 3A).

Discussion

The availability of genomic sequences from *S. tuberosum* and *S. lycopersicum* allowed us to study the evolutionary history of resistance genes in two closely related species of the family Solanaceae. Most resistance genes are organized in tandem arrays, clusters, and super-clusters (Chen et al. 2007; Andolfo et al. 2013). In this study, tomato and potato sequences of two resistance gene families were often found in clusters, with exceptional members that were distantly located possibly due to duplications followed by translocations. Large resistance gene clusters provide a huge potential for the generation of sequence variation through recombination, facilitating the creation of new alleles or genes upon which pathogen selective pressures act. The investigation of the spatial arrangement of resistance clusters is useful in reconstructing the history of the chromosomal rearrangement that shaped the genome architecture of the ancestral species.

The comparative study we undertook indicated that all gene families analyzed were more numerous in *S. tuberosum* than in *S. lycopersicum*, even though their genomes are roughly equal in size (705 and 759 Mb, respectively). This phenomenon may be due to successive duplications in the potato genome or repeated gene losses in tomato. In previous surveys, syntenic comparisons between the tomato and potato genomes indicated that tomato suffered the deletion of roughly one-third of the redundant proteome and transposable elements in euchromatic regions (The Tomato Genome Consortium 2012). This study also revealed that the tomato genome


A
StbMih1_C2
SlyMih5_C5
SlyMih4_C5
SlyMih8_C5
SlyMih3_C5
SlyMih7_C5
StbMih38_C5
StbMih37_C5
SlyMih9_C9
StbMih34_C9
SlyMih6_C5
StbMih35_C5
StbMih2_C4
SlyMih2_C5
StbMih10_C6
StbMih33_C6
StbMih26_C6
StbMih36_C5
StbMih11_C6
StbMih19_C6
StbMih30_C6
StbMih8_C6
StbMih12_C6
StbMih32_C6
StbMih3_C6
StbMih31_C6
StbMih28_C6
StbMih15_C6
StbMih29_C6
StbMih20_C6
StbMih13_C6
StbMih17_C6
StbMih18_C6
StbMih7_C6
StbMih16_C6
SlyMih10_C6
SlyMih12_C6
Mi1_2
SlyMih11_C6
SlyMih13_C6
SlyMih15_C6
SlyMih1h_C5
SlyMih14_C6
StbMih21_C6
StbMih22_C6
StbMih24_C6
StbMih14_C6
StbMih6_C6
StbMih4_C6
StbMih25_C6
SlyMih16_C6
StbMih23_C6
StbMih5_C6
StbMih9_C6
StbMih27_C6
CA06g05130_C6
CA06g05120_C6
CA02g24750_C2
CA06g00820_C6
CA06g03680_C6
CA00g30030_C0
CA06g01250_C6
CA06g01210_C6
CA06g02810_C6
CA06g01190_C6
CA00g80340_C0
CA06g00980_C6
CA06g02200_C6
CA06g02160_C6
CA06g02100_C6
CA06g00950_C6
CA06g03410_C6
CA06g02040_C6
CA06g02030_C6
CA06g02620_C6
CA06g02520_C6
CA06g02680_C6
CA06g05110_C6
CA06g03630_C6
CA06g01040_C6
CA06g02180_C6
CA00g76980_C0
CA06g01770_C6
CA06g02830_C6
CA06g02740_C6
CA06g00660_C6
CA06g01020_C6
CA06g01230_C6
CA06g01750_C6
CA06g03650_C6
CA00g30020_C0
CA06g03690_C6
CA00g87560_C0
CA05g15170_C5
CA06g01180_C6
CA06g01350_C6
CA06g02450_C6
CA06g00910_C6
CA06g00890_C6
CA06g01110_C6
CA00g87620_C0
CA06g01700_C6
CA06g00990_C6
CA00g30010_C0
CA06g03420_C6
CA06g02070_C6
CA06g00530_C6
CA06g01130_C6
CA06g01220_C6
CA06g02730_C6
CA06g00770_C6
CA06g02530_C6
CA06g00920_C6
CA06g02570_C6
CA06g02420_C6
CA06g01740_C6
CA06g00650_C6
CA06g00730_C6
CA06g00750_C6
CA06g02150_C6
CA06g00720_C6
CA06g03640_C6
CA06g00610_C6
CA06g00960_C6
CA06g00580_C6
CA06g02170_C6
0.06 substitutions/site
Solanum
Mi-1
homologues
Capsicum annuum
Mi-1 homologues
B
S. lycopersicum chr 6
A´
B´
2,3 Mb
2,7 Mb
16 13 11 15
12 14 10
S. tuberosum chr 6
A
B
C
D
E
2,7 Mb
2,8Mb
4,3Mb
4,5Mb
4,8Mb
9,7Mb
21,4Mb
10 5 21 22 14
33 7
12 6 29 17 24 25 3 32 26 4 16 13 30 20 31 23 8 19 27 18
15 11 28
9

has approximately an equal number of resistance genes as *A. thaliana*, but fewer than grape, rice, and the potato genomes (The Tomato Genome Consortium 2012). Plant resistance gene family evolution has probably been influenced by polyploidization, genome size variation, natural and artificial selection including domestication breeding and cultivation, and gene family interactions (Jia et al. 2015; Wu et al. 2014; Zhong et al. 2015). Given that the family Solanaceae evolved largely in the absence of polyploidization, except for the cultivated potato, but this event is insignificant in evolutionary times, any increase in the size of resistance gene families should be due to segmental duplications (Wu et al. 2010). Cross-hybridization experiments in *Solanum* showed that resistance genes have an ancient origin that predated speciation events in the Solanaceae (Quirin et al. 2012). After that, tomato and potato genomes suffered small and large rearrangements, along with duplications and deletions that obscured orthologous relationships because neither gene content nor order is strictly conserved (Andolfo et al. 2013).

The evolutionary histories of *Sw5* and *Mi-1* gene families were markedly different. Individual tomato *Sw5* homologues maintained a close phylogenetic relationship with potato and hot pepper sequences. This suggests that several gene copies were present in the ancestor of these *Solanaceae* due to early segmental duplications. In contrast, sister relationships between tomato, potato, and hot pepper *Mi-1* sequences were not observed. Instead, tomato homologues formed a highly supported monophyletic group, possible due to recurring gene conversion events within the tomato (Seah et al. 2007) and potato (Sanchez-Puerta and Masuelli 2011) genomes, respectively.

The tomato and potato *Mi-1* homologues are distributed in less than perfectly matching positions across the two genomes, suggesting that conserved gene order on chromosomes is a mere vestige of common ancestry. *SlyMih* sequences of chromosome 6 are clustered in two loci, while *StbMih* sequences are found in five loci in the same chromosome. So, three potato clusters (C, D, and E) of chromosome 6, *StbMih1* from chromosome 2, *StbMih2* from chromosome 4, and some genes from chromosome 5 in potato, have no syntenic position in tomato. Novel clusters in potato may have been originated through translocation events. For example, a translocation of a segment of the ancestral cluster B of the potato chromosome 6 could have given rise to the first members of cluster C (Fig. 3B). Evidence for this hypothesis includes the close relationship of *StbMih33-StbMih26* and *StbMih7-StbMih32* in the phylogenetic tree and the similarity of the intergenic sequences between these two pair of genes. Additional clusters of chromosome 6 of potato (clusters D and E) may have also originated by the duplication and translocation of an ancestral copy found in clusters C or A, respectively (Fig. 3B). The great number of copies in cluster C was also puzzling. The increase in size of this cluster may have occurred by successive unequal crossing over events. The intimate phylogenetic relationships of some of the genes from this group (*StbMih12-StbMih3-StbMih31*, *StbMih6-StbMih25-StbMih4*, *StbMih17-StbMih18*, and *StbMih30-StbMih8*) were congruent with this hypothesis.

In contrast to the inferred evolution of the two resistance gene families, single copy genes did not suffer gene duplications in any of the genomes analyzed, indicating that gene clusters are more prone to duplication and translocation through unequal crossing overs or errors in the replication or recombination processes (Krebs et al. 2013).

The distribution of homologous sequences in the genomes analyzed and the discovery of redundant genes provide insight into how novel pathogen-resistance genes can evolve via subsequent gene duplication events, unequal crossing over, translocation, and diversifying selection (Michelmore and Meyers 1998; Roth et al. 2007; Long et al. 2003). In conclusion, comparisons of gene arrangements between related plant genomes offer insight on their genome evolution. The complete sequencing of the tomato and potato genomes permitted an unprecedented view of *Solanum* resistance gene families, which may also impact in molecular breeding efforts.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported in part by SecTyP grant #06/M053 from Universidad Nacional de Cuyo and grant PICT-277 from ANCyPT. R.W.M. and M.V.S.-P. are researchers of CONICET.

References

Alba, R., Kelmenson, P., Cordonnier-Pratt, M., and Pratt, L. 2000. The phytochrome gene family in tomato and the rapid differential evolution of this family in angiosperms. Mol. Biol. Evol. **17**(3): 362–373. doi:10.1093/oxfordjournals.molbev.a026316. PMID:10723737.

Andolfo, G., Sanseverino, W., Rombauts, S., Van de Peer, Y., Bradeen, J.M., Carputo, D., et al. 2013. Overview of tomato (*Solanum lycopersicum*) candidate pathogen recognition genes reveals important *Solanum R* locus dynamics. New Phytol. **197**: 223–237. doi:10.1111/j.1469-8137.2012.04380.x. PMID:23163550.

Ashfield, T., Egan, A., Pfeil, B., Chen, N., Podicheti, R., Ratnaparkhe, M., et al. 2012. Evolution of a complex disease resistance gene cluster in diploid *Phaseolus* and tetraploid *Glycine*. Plant Physiol. **159**: 336–354. doi:10.1104/pp.112.195040. PMID:22457424.

Bakker, E., Borm, T., Prins, P., van der Vossen, E., Uenk, G., Arens, M., et al. 2011. A genome-wide genetic map of NB-LRR disease resistance loci in potato. Theor. Appl. Genet. **123**: 493–508. doi:10.1007/s00122-011-1602-z. PMID:21590328.

Cannon, S.B., McCombie, W.R., Sato, S., Tabata, S., Denny, R., Palmer, L., et al. 2003. Evolution and microsynteny of the apyrase gene family in three legume genomes. Mol. Genet. Genomics, **270**: 347–361. PMID:14598165.

Chagué, V., Mercier, J.C., Guénard, M., de Courcel, A., and Vedel, F. 1996. Identification and mapping on chromosome 9 of RAPD markers linked to *Sw-5* in tomato by bulked seg-

regant analysis. Theor. Appl. Genet. **92**: 1045–1051. doi:10.1007/BF00224047. PMID:24166634.

Chen, G., Pan, D., Zhou, Y., Lin, S., and Ke, X. 2007. Diversity and evolutionary relationship of nucleotide binding site-encoding disease-resistance gene analogues in sweet potato (*Ipomoea batatas* Lam.). J. Biosci. **32**(4): 713–721. doi:10.1007/s12038-007-0071-7. PMID:17762144.

Dropkin, V. 1969. Cellular responses of plants to nematode infections. Annu. Rev. Phytopathol. **7**: 101–122. doi:10.1146/annurev.py.07.090169.000533.

Hulbert, S., Webb, C., Smith, S., and Sun, Q. 2001. Resistance gene complexes: evolution and utilization. Annu. Rev. Phytopathol. **39**: 285–312. doi:10.1146/annurev.phyto.39.1.285. PMID:11701867.

Jablonska, B., Ammiraju, J., Bhattarai, K., Mantelin, S. Martinez, de Ilarduya, O., Roberts, P., and Kaloshian, I. 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. PMID:17172289.

Jia, Y.X., Yuan, Y., Zhang, Y., Yang, S., and Zhang, X. 2015. Extreme expansion of NBS-encoding genes in *Rosaceae*. BMC Genet. **16**(1): 48. PMID:25935646.

Kim, S., Park, M., Yeom, S., Kim, Y., Lee, J., Lee, H., et al. 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nature, **46**: 270–279. doi:10.1038/ng.2877.

Krebs, J., Lewin, B., Goldstein, E.S., and Kilpatrick 2013. Lewin's essential genes. Jones & Bartlett Publishers. pp. 847.

Lawton-Rauh, A. 2003. Evolutionary dynamics of duplicated genes in plants. Mol. Phylogenet. Evol. **29**: 396–409. doi:10.1016/j.ympev.2003.07.004. PMID:14615182.

Long, M., Beltrán, E., Thornton, K., and Wang, W. 2003. The origin of new genes: glimpses from the young and old. Nat. Rev. **4**: 865–865. doi:10.1038/nrg1204.

Lozano, R., Ponce, O., Ramirez, M., Mostajo, N., and Orjeda, G. 2012. Genome-wide identification and mapping of NBS encoding resistance genes in *Solanum tuberosum* Group Phureja. PloS ONE, **7**(4): e34775. doi:10.1371/journal.pone.0034775. PMID:22493716.

Lu, G., and Moriyama, E.N. 2004. Vector NTI, a balanced all-in-one sequence analysis suite. Brief. Bioinform. **5**(4): 578–588. doi:10.1093/bib/5.4.378. PMID:15606974.

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

Michelmore, R.W., and Meyers, B.C. 1998. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. **8**: 1113–1130. PMID:9847076.

Nam, J., de Pamphilis, C.W., Ma, H., and Nei, M. 2004. Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. Proc. Natl. Acad. Sci. U.S.A. **101**: 1910–1915. doi:10.1073/pnas.0308430100. PMID:14764899.

Nei, M., and Rooney, A.P. 2005. Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. **39**: 121–152. doi:10.1146/annurev.genet.39.073003.112240. PMID:16285855.

Nombela, G., Williamson, V., and 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. doi:10.1094/MPMI.2003.16.7.645. PMID:12848430.

Park, T.-H., Kim, J.-B., Hutten, R.C.B., van Eck, H.J., Jacobsen, E., and Visser, R.G.F. 2007. Genetic positioning of centromeres using half-tetrad analysis in a 4x–2x cross population of potato. Genet. Soc. Am. **176**: 85–94. doi:10.1534/genetics.107.070870.

Philippe, H., Snell, E., Bapteste, E., Lopez, P., Holland, P., and Casane, D. 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Mol. Biol. Evol. **21**: 1740–1752. doi:10.1093/molbev/msh182. PMID:15175415.

Quirin, E., Mann, H., Meyer, R., Traini, A., Chiusano, M.L., Litt, A., and Bradeen, J. 2012. Evolutionary meta-analysis of *Solanaceous* resistance gene and *Solanum* resistance gene analog sequences and a practical framework for cross-species comparisons. Mol. Plant-Microbe Interact. **25**(5): 603–612. doi:10.1094/MPMI-12-11-0318-R. PMID:22352721.

Rossi, M., Goggin, F., Milligan, S., Kaloshian, I., Ullman, D., and Williamson, V. 2003. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc. Natl. Acad. Sci. U.S.A. **95**: 9750–9754. PMID:9707547.

Roth, C., Rastogi, S., Arvestad, L., Dittmar, K., Light, S., Ekman, D., and Liberles, D. 2007. Evolution after gene duplication: models, mechanisms, sequences, systems and organisms. J. Exp. Zool. **308B**: 58–73. doi:10.1002/jez.b.21124.

Sanchez-Puerta, M.V., and Masuelli, R.W. 2011. Evolution of nematode-resistant *Mi-1* gene homologues in three species of *Solanum*. Mol. Genet. Genomics, **285**: 207–218. doi:10.1007/s00438-010-0596-6.

Sasser, J., and Carter, C. 1985. An advance treatise on *Meloidogyne*. NC State University Graph, Raleigh, N.C.

Seah, S., Yaghoobi, J., Rossi, M., Gleason, C.A., Williamson, V.M. 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. doi:10.1007/s00122-004-1594-z. PMID:14963654.

Seah, S., Telleen, A., and 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. doi:10.1007/s00122-007-0519-z. PMID:17318492.

Spassova, M., Prins, T., Folkertsma, R., Klein-Lankhorst, R., Hille, J., Goldbach, R., and Prins, M. 2001. The tomato gene *Sw5* is a member of the coiled coil, nucleotide binding, leucine-rich repeat class of plant resistance genes and confers resistance to TSWV in tobacco. Mol. Breed. **7**: 151–161. doi:10.1023/A:1011363119763.

Stevens, M.R., Scott, S.J., and Gergerich, R.C. 1991. Inheritance of a gene for resistance to tomato spotted wilt virus (TSWV) from *Lycopersicon peruvianum* Mill. Euphytica, **59**: 9–17. doi:10.1007/BF00025356.

Stevens, M.R., Lamb, E.M., and Rhoads, D.D. 1995. Mapping the *Sw-5* locus for tomato spotted wilt virus resistance in tomatoes using RAPD and RFLP analyses. Theor. Appl. Genet. **90**: 451–456. PMID:24173937.

The Potato Genome Sequencing Consortium. 2011. Genome sequence and analysis of the tuber crop potato. Nature, **475**: 189–195. doi:10.1038/nature10158. PMID:21743474.

The Tomato Genome Consortium 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature, **485**: 635–641. doi:10.1038/nature11119. PMID:22660326.

Thornton, J., and DeSalle, R. 2000. Gene family evolution and homology: genomics meets phylogenetics. Annu. Rev. Genom. Hum. Genet. **1**: 41–73. doi:10.1146/annurev.genom.1.1.41.

Walsh, B., and Stephan, W. 2001. Multigene families: Evolution. Encyclopedia of Life Sciences.

Wiens, J. 2006. Missing data and the design of phylogenetic analyses. J Biomed Inform. **39**: 34–42. doi:10.1016/j.jbi.2005.04.001. PMID:15922672.

Wu, P., Shao, Z.Q., Wu, X.Z., Wang, Q., Wang, B., Chen, J.Q., et al. 2014. Loss/retention and evolution of NBS encoding genes upon whole genome triplication of *Brassica rapa*. Gene, **540**(1): 54–61. doi:10.1016/j.gene.2014.01.082. PMID:24576745.

Wu, F., and Tanksley, S.D. 2010. Chromosomal evolution in the

plant family *Solanaceae*. BMC Genomics, **11**: 182. doi:10.1186/1471-2164-11-182. PMID:20236516.

Zhao, Y., Liu, W., Xu, Y., Cao, J., Braam, J., and Cai, X. 2013. Genome-wide identification and functional analyses of calmodulin genes in *Solanaceous* species. BMC Plant Biol. **13**: 70. doi:10.1186/1471-2229-13-70. PMID:23621884.

Zhong, Y., Yin, H., Sargent, D.J., Malnoy, M., and Cheng, Z.M. 2015. Species-specific duplications driving the recent expansion of NBS-LRR genes in five Rosaceae species. BMC Genomics, **16**(1): 77. doi:10.1186/s12864-015-1291-0. PMID:25759136.

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