Molecular Plant Pathology

MOLECULAR PLANT PATHOLOGY (2013) **14**(7), 663–677 DOI: 10.1111/mpp.12037

Genomic analysis reveals MATH gene(s) as candidate(s) for *Plum pox virus* (PPV) resistance in apricot (*Prunus armeniaca* L.)

ELENA ZURIAGA[1,†], JOSÉ MIGUEL SORIANO[1,†], TETYANA ZHEBENTYAYEVA[2], CARLOS ROMERO[1], CHRIS DARDICK[3], JOAQUÍN CAÑIZARES[4] AND MARIA LUISA BADENES[1,*]

[1]*Instituto Valenciano de Investigaciones Agrarias (IVIA), Apartado Oficial, 46113 Moncada, Valencia, Spain*
[2]*Department of Genetics and Biochemistry, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA*
[3]*USDA-ARS Appalachian Fruit Research Station, 2217 Wiltshire Road, Kearneysville, WV 25430, USA*
[4]*Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain*

SUMMARY

Sharka disease, caused by *Plum pox virus* (PPV), is the most important viral disease affecting *Prunus* species. A major PPV resistance locus (*PPVres*) has been mapped to the upper part of apricot (*Prunus armeniaca*) linkage group 1. In this study, a physical map of the *PPVres* locus in the PPV-resistant cultivar 'Goldrich' was constructed. Bacterial artificial chromosome (BAC) clones belonging to the resistant haplotype contig were sequenced using 454/GS-FLX Titanium technology. Concurrently, the whole genome of seven apricot varieties (three PPV-resistant and four PPV-susceptible) and two PPV-susceptible apricot relatives (*P. sibirica* var. *davidiana* and *P. mume*) were obtained using the Illumina-HiSeq2000 platform. Single nucleotide polymorphisms (SNPs) within the mapped interval, recorded from alignments against the peach genome, allowed us to narrow down the *PPVres* locus to a region of ~196 kb. Searches for polymorphisms linked in coupling with the resistance led to the identification of 68 variants within 23 predicted transcripts according to peach genome annotation. Candidate resistance genes were ranked combining data from variant calling and predicted functions inferred from sequence homology. Together, the results suggest that members of a cluster of meprin and TRAF-C homology domain (MATHd)-containing proteins are the most likely candidate genes for PPV resistance in apricot. Interestingly, MATHd proteins are hypothesized to control long-distance movement (LDM) of potyviruses in *Arabidopsis*, and restriction for LDM is also a major component of PPV resistance in apricot. Although the PPV resistance gene(s) remains to be unambiguously identified, these results pave the way to the determination of the underlying mechanism and to the development of more accurate breeding strategies.

INTRODUCTION

Sharka disease, caused by *Plum pox virus* (PPV), was described for the first time infecting plums (*Prunus domestica* L.) in Bulgaria around 1917 (Atanasoff, 1932). Since then, it has spread into most temperate fruit crop-growing areas (Capote *et al*., 2006), and currently is the most important viral disease affecting *Prunus* species (Scholthof *et al*., 2011). The global cost of PPV worldwide in the last 30 years has been estimated as 10 000 million euros (Cambra *et al*., 2006b). PPV is transmitted by aphids in a nonpersistent manner, and therefore chemical treatments are not effective in preventing plant infection. Control measures are primarily focused on the use of certified healthy plants and the eradication of infected trees. However, this latter measure is inefficient because of the time lapse between inoculation and the appearance of symptoms, allowing the persistence of virus reservoirs, especially in endemic areas (Martínez-Gómez *et al*., 2000). Moreover, PPV can also infect ornamental and wild *Prunus* species that serve as potential reservoirs and sources of new inoculum, thereby hindering PPV eradication programmes (James and Thompson, 2006). In this context, although epidemiological studies (Cambra *et al*., 2006a; Labonne and Dallot, 2006) and improved PPV detection methods (Olmos *et al*., 2006) have contributed to better management of the disease, the growth of PPV-resistant varieties would be the ideal long-term solution.

At the beginning of the 1990s in those European countries severely affected by the disease, apricot (*Prunus armeniaca* L.) breeding programmes initiated efforts to introgress PPV resistance (reviewed by Badenes and Llácer, 2006; Bassi, 2006; Bassi and Audergon, 2006; Karayiannis, 2006; Rubio *et al*., 2004). After wide screening of *Prunus* germplasm, some resistant sources were identified in a handful of North American apricot cultivars (Brooks and Olmo, 1997). These have subsequently been used as PPV resistance donors, even though they are not well adapted to the various growth conditions of different European regions (Martínez-Gómez *et al*., 2000). With regard to other *Prunus* species, some genotypes showing tolerance or a hypersensitive response to PPV infection have been found in plum (Hartmann and Neumüller, 2006), but no resistance sources have yet been found in peach (*Prunus persica* L. Batsch) (Escalettes *et al*., 1998). In apricot, the introgression of PPV resistance into commercial cultivars has been accomplished, but breeding progress is still hampered by difficulties inherent to both fruit tree management and PPV resistance phenotyping (Egea *et al*., 2009; Martínez-Calvo *et al*., 2010).

**Correspondence*: Email: badenes_mlu@gva.es
†These authors contributed equally to this work.

The genetic control of PPV resistance in apricot has remained controversial; however, most inheritance studies agree on the importance of the involvement of one major dominant locus (the *PPVres* locus) located in the upper part of apricot linkage group 1 (Dondini *et al*., 2011; Lalli *et al*., 2008; Lambert *et al*., 2007; Marandel *et al*., 2009; Pilarova *et al*., 2010; Soriano *et al*., 2008). More recently, Vera-Ruiz *et al*. (2011) developed high-density simple sequence repeat (SSR) linkage maps and narrowed down the *PPVres* support intervals to 7.3 and 5.9 cM in 'Lito' and 'Goldrich' maps, respectively, with an estimated size according to the peach genome sequence of 2.16 Mb (Arús *et al*., 2012; Jung *et al*., 2004). From this point, Soriano *et al*. (2012) further refined the *PPVres* locus to 1.2-cM and 0.9-cM intervals in 'Lito' and 'Goldrich' maps, respectively, corresponding to an interval of ~270 kb in the peach genome. Moreover, Soriano *et al*. (2012) developed new markers tightly linked to PPV resistance, providing an efficient tool for marker-assisted selection in *Prunus* resistance breeding, and paving the way for future positional cloning by sequencing the *PPVres* locus. Similar strategies have been employed to identify resistance genes to different pathogens in other tree species. For instance, three clustered genes were identified as candidates for the *Rvi15* apple scab resistance gene after sequencing a bacterial artificial chromosome (BAC) clone spanning the resistance locus (Galli *et al*., 2010). Similarly, Parravicini *et al*. (2011) discovered *in silico* two genes as the most probable fire blight resistance genes in apple after sequencing a BAC clone of 189 kb.

In this article, overgo probes, designed from the SSRs spanning the *PPVres* locus defined by Soriano *et al*. (2012), have been hybridized against an apricot BAC library in order to construct a physical map for the *PPVres* region in the cultivar 'Goldrich'. BAC clones covering the resistant haplotype, as well as nine complete apricot genomes, were sequenced and analysed for polymorphisms within the mapped interval. The identification of segregating markers allowed us to narrow down the *PPVres* locus to ~196 kb according to the peach genome syntenic region. Polymorphisms linked in coupling with the resistance were interrogated, and putative functions for all transcripts comprised in this interval were inferred from sequence homology. On the basis of these analyses, candidate resistance genes were ranked and their possible role in apricot PPV resistance is discussed.

RESULTS

Construction of the BAC contig spanning the *PPVres* locus

The six high-density filters containing 101 376 clones of the 'Goldrich' BAC library were screened in two hybridization rounds. The first round was performed with a pool of six digoxigenin-labelled overgo probes designed from SSR markers: PGS1.20, PGS1.21, PGS1.22, PGS1.23, PGS1.24 and PGS1.252 (Soriano *et al*., 2012). Thirteen positive BAC clones were identified and assigned to the corresponding resistant or susceptible haplotype according to the SSR alleles determined by polymerase chain reaction (PCR) typing. Four clones were in coupling phase with the resistance, six in repulsion phase and three were considered as false positives because they could not be confirmed by PCR. In the second round of hybridization, 20 new overgo probes derived from the BAC-end sequences (BES) of the 10 previously confirmed positive clones were used, identifying 39 additional BACs. After PCR genotyping using BES and SSR primers, 14 clones amplified SSR alleles in coupling with the resistance, six were in repulsion phase and 19 were found to be false positives.

Resistant (R) and susceptible (S) BAC contigs were constructed (Fig. 1) covering the region of ~270 kb between PGS1.20 and PGS1.252 defined by Soriano *et al*. (2012) according to the peach genome syntenic region. The R-contig comprises a region of ~282 kb fully covered by six BAC clones ranging from ~24 to ~119 kb in size. The S-contig is covered by six BAC clones ranging from ~25 to ~80 kb in size and comprises a region of ~331 kb with an internal gap of ~58 kb.

Next-generation sequencing (NGS) data processing and mapping against the peach genome

Pyrosequencing of the six BACs from the R-contig produced a total of 127 802 reads (Table 1) ranging from 40 to 641 bp in length and averaging 348 bp, as expected from the 454/GS-FLX sequencing platform. The sequencing of apricot genomes with Illumina technology produced a total of 1 445 495 513 reads (Table 1), ranging from 91 443 246 for 'Harlayne' to 373 801 518 for 'Canino'. After removing low-quality regions, as well as vector and adaptor contaminants from all NGS data, 1 391 686 236 trimmed reads (96.27% of raw sequences), with an average length of 95.19 bp for Illumina and 344.79 bp for 454 sequences, were used for subsequent analyses.

Cleaned sequences were aligned onto the peach genome syntenic region from scaffold_1 corresponding to the interval between positions 7 986 205 and 8 281 900, encompassing the interval between PGS1.20 and PGS1.252 positions. As a result, 1 364 156 sequences were successfully mapped, representing 0.1% of the trimmed sequences, and the average read depth for all the bases was 294.14*x* (Table 1).

Narrowing down the *PPVres* locus

A selection of single nucleotide polymorphisms (SNPs) identified from the NGS alignment against peach and annotated between positions corresponding to markers PGS1.20–PGS1.21 and PGS1.24–PGS1.252 (Fig. 1) were screened. The identification of segregating SNPs (Table S1, Supporting Information) allowed us to

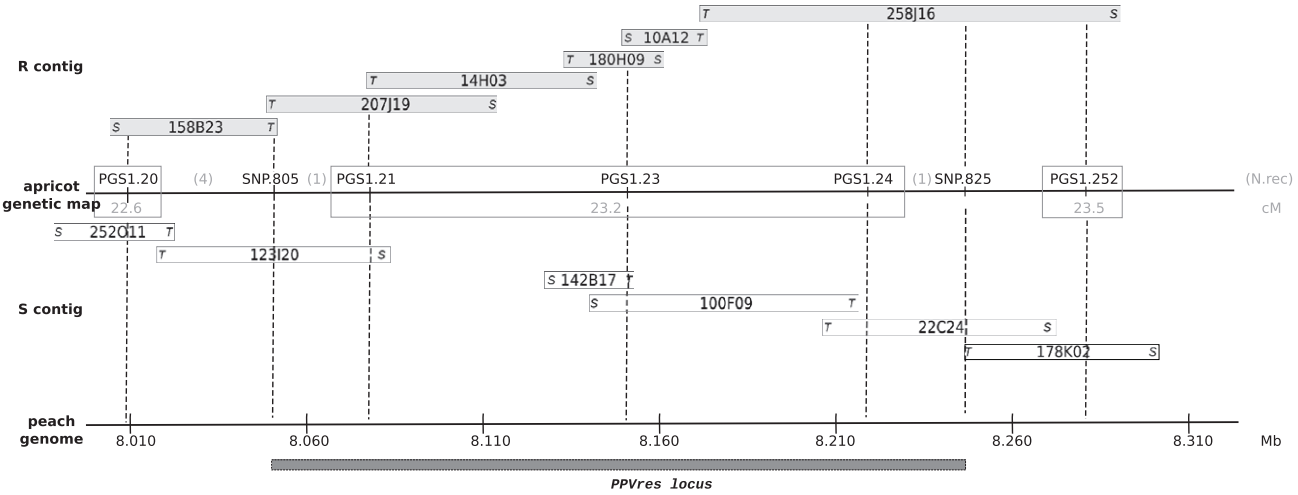

Fig. 1 Resistant (R) and susceptible (S) bacterial artificial chromosome (BAC) contigs spanning the *PPVres* locus in the apricot resistant cultivar 'Goldrich'. R-contig BACs (grey boxes) were in coupling phase with *Plum pox virus* (PPV) resistance and S-contig BACs (white boxes) were in repulsion phase. BAC-ends SP6 (S) and T7 (T) are indicated. A line in between these two contigs represents the 'Goldrich' genetic map for the *PPVres* locus. Markers and numbers of recombinants (N.rec) are indicated above the line, and distances in centimorgans (cM) are shown below the line. Markers, anchored into the 'Goldrich' genetic map, are connected by broken lines to their syntenic peach genomic positions in megabases (Mb) according to BLASTN results.

Table 1 Next-generation sequencing (NGS) data statistics. The total number of raw, cleaned and mapped sequences used from each sample is indicated. Coverage indicates the average number of aligned reads against the reference.

Sample	Phenotype	DNA source	NGS platform	No. raw sequences	Cleaned sequences		Mapped sequences		
					No,	Av. length (bp)	No.	%	Coverage
'Goldrich'	R	BACs	454	127802	109335	344.79	43605	39.88	
		BES, markers	Sanger	33	33	1033.88	21	63.64	
		gDNA	Illumina	137954275	131437990	95.18	105893	0.08	221.75
'Harlayne'	R	gDNA	Illumina	91443246	87739924	95.1	70326	0.08	143.12
'Stark Early Orange'	R	gDNA	Illumina	156657196	150085821	95.07	114894	0.08	224.79
'Canino'	S	gDNA	Illumina	373801518	365657436	96.40	462907	0.13	1163.02
'Krasnoshchekii'	S	gDNA	Illumina	142966212	137511644	95.48	114002	0.08	230.50
'Reale d'Imola'	S	gDNA	Illumina	161083342	152255001	95.37	121293	0.08	246.47
'Shalakh'	S	gDNA	Illumina	97418198	93361180	93.09	66927	0.07	143.41
Prunus mume	S	gDNA	Illumina	128834518	123740463	94.07	131072	0.11	132.95
Prunus sibirica	S	gDNA	Illumina	155337008	149787409	96.96	133216	0.09	141.23
Total				1445623348	1381686236		1364156	0.1	294.14

BAC, bacterial artificial chromosome; BES, BAC-end sequence.

redefine the *PPVres* locus by introducing them into the 'Goldrich' genetic map (Fig. 1). Taking into account the new recombination breakpoints (Fig. 2), the *PPVres* locus was shrunk to the interval flanked by SNP-805 (peach genome position: scaffold_1: 8 050 804) and SNP-825 (peach genome position: scaffold_1: 8 247 059) with three SSR markers co-segregating with the resistance (PGS1.21, PGS1.23 and PGS1.24) (Fig. 1). As a result, the *PPVres* locus was predicted to comprise ~196 kb according to the peach genome syntenic region.

Screening for candidate resistance genes in the *PPVres* locus

The annotation of the peach genome sequence by the International Peach Genome Initiative (IPGI), available from the Genome Database for Rosaceae (GDR), showed 31 predicted transcripts in the ~196-kb region (Table 2). Three corresponded to alternative transcripts of the same locus (ppa006833m, ppa006841m and ppa006846m). Moreover, three other transcripts (ppb020867m, ppa008960m and ppa017827m), comprised in a ~22-kb region (between positions 8 173 840 and 8 196 599), seemed to be partially or completely lost in apricot genomes. Only ppa008960m was present in *P. mume* and *P. sibirica*, but showed a low alignment coverage. Variant calling to detect SNPs or deletion/insertion polymorphisms (DIPs) (Tables S2 and S3, see Supporting Information) was performed using the alignment of the NGS data against the peach genome. As a result, between 8022 and 11 272 putative SNPs per genotype and between 2990 and 3242 putative DIPs per genotype were discovered (Table 3). Among the different genic

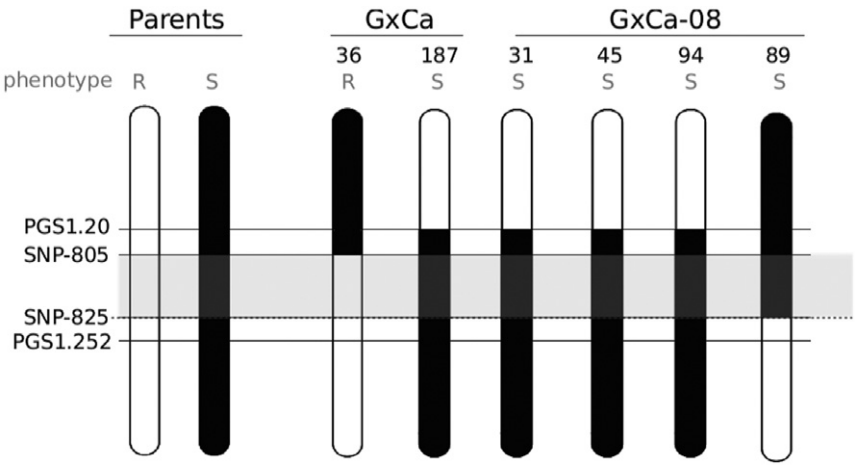

Fig. 2 Graphical genotyping of recombinant hybrids belonging to 'Goldrich' × 'Canino' populations ('G× Ca' and 'G×Ca-08') at the *PPVres* locus. Black vertical bars represent susceptible (S) and white bars represent resistant (R) chromosomal regions. Grey box represents the position of the *PPVres* locus. Recombinant hybrids are numbered at the top.

compartments, introns showed greater variant frequency (average: 1 SNP/24 bp and 1 DIP/78 bp) than exons (average: 1 synonymous SNP (sSNP)/140 bp, 1 nonsynonymous SNP (nsSNP)/168 bp and 1 DIP/630 bp) (Table 3). Overall, around 44%–47% of exonic SNPs were nonsynonymous.

In the screening for candidate resistance genes within the *PPVres* locus, three premises based on prior knowledge were taken into account. First, PPV resistance is generally accepted to be controlled by a major dominant locus (Dondini *et al*., 2011; Lalli *et al*., 2008; Lambert *et al*., 2007; Marandel *et al*., 2009; Pilarova *et al*., 2010; Soriano *et al*., 2008). Second, the allelic composition of the *PPVres* locus suggests a common PPV-resistant ancestor for the resistant cultivars used in this study ('Goldrich', 'Harlayne' and 'Stark Early Orange') (Pilarova *et al*., 2010; Soriano *et al*., 2012; Zhebentyayeva *et al*., 2008). Third, segregation ratios recorded from different apricot intraspecific populations confirm that these resistant varieties are heterozygous for PPV resistance (Karayiannis *et al*., 2008; Soriano *et al*., 2012).

According to these premises, three filters were sequentially applied (Table 4) to discriminate SNPs/DIPs associated with resistance from all detected variants after mapping against the susceptible peach (Table S2). In the first, variants preserved in all resistant or all susceptible varieties were screened. In addition, variants preserved in both 'Goldrich' and 'Harlayne', or only in 'Stark Early Orange', but not in any susceptible variety, were also screened in order to detect mutations affecting the same responsible gene that could be different in 'Stark Early Orange' and the other two accessions. As a result, 87 variants within 23 predicted transcripts were detected, and therefore five transcripts (ppa020498m, ppa007310m, ppa021666m, ppa024143m and ppa018241m) were discarded (Table 4). The second filter selected heterozygous variants present in resistant varieties, discarding two of the original 87 (Table 4). Lastly, the third filter selected variants in coupling with the resistance by checking their presence in the BAC R-contig sequences (Table S3), and reduced the number of variants to 68 located within 23 predicted transcripts, ranging from 1 to 17 per transcript (Table 4). Twelve of these 23 transcripts had additional variants predicted to produce early stop codons, resulting in putative truncated proteins in one or more PPV-susceptible varieties (Table 4). These 68 variants can be roughly classified into nonsynonymous (12 lead to amino acid changes and one produces a frameshift) and synonymous [16 in exons, 28 in intronic regions and 11 in untranslated regions (UTRs)] (Table 4). Two transcripts show only a single nonsynonymous mutation, 14 show only synonymous mutations (ranging from one to three each) and seven have both (ranging from one to four nonsynonymous and 1–14 synonymous in each). A total of 33 variants are located within just three transcripts.

For most of the transcripts present in the *PPVres* locus, highly similar *Arabidopsis* proteins were found, but a few showed moderate/low similarity values with *Arabidopsis* peptides (Table 2). This was the case for ppb020867, ppa021666m and ppa018241m, all having id < 50%, overlap length < 55% and E values > 1e-34. Others, such as ppa000690, ppa000792 and ppa008960, showed id < 50%, or an overlap length < 50%, such as ppa020498. In addition, ppa013116 showed homology with an 'unknown protein' according to The Arabidopsis Information Resource (TAIR) database (Table 2). Significant similarity degrees were also found between some annotated transcripts in peach. Three were alternative transcripts of the same locus (ppa006841m, ppa006833m and ppa006846m) sharing three variants, but differing in the beginning of the intronic region within the 5′-UTR. Another two were adjacent transcripts, ppa000690m and ppa000792m, which showed a high similarity (id = 81.4%). Finally, six meprin and TRAF-C homology domain (MATHd)-containing genes clustered together in apricot, between positions 8 150 688 and 8 206 582, whereas, in the peach genome, nine MATHd genes were present in the cluster. All six MATHd genes had variants linked in coupling with the resistance, but ppa022254m and ppa023061 had only a 1-bp insertion within the fourth and fifth introns, respectively. Therefore, there is a low probability that they have any impact on protein function. The

Table 2 Gene content of the *PPVres* locus in the peach syntenic region. Position and length of the transcripts annotated by the International Peach Genome Initiative (IPGI), as well as the first BLASTP match on The Arabidopsis Information Resource (TAIR) database, are shown. Overlap length (amino acids), percentage id and E value are indicated for each *Prunus*/*Arabidopsis* gene pair. Reciprocal BLASTP results are also detailed for those pairs being best reciprocal hits (*At* position in bold). Positions of simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers are included as reference. Peach genes not present in apricot genomes are indicated with a grey background.

	GDR/JGI							TAIR		BLASTP peach against TAIR			BLASTP At against GDR		
Molecular marker	Peach gene id	Transcript start	Transcript stop	Genomic length	Protein length	At locus	Protein length	Description	At position	Overlap length	%id	E value	Overlap length	%id	E value
SNP-805		8050804													
	ppa020498m	8054332	8056455	2124	264	AT2G30350.2	368	Excinuclease ABC, C subunit, N-terminal	chr2:12934237–12935742	164	71	4.00E-067			
	ppa007310m	8059267	8062957	3691	373	AT3G17430.1	375	Nucleotide-sugar transporter family protein	**chr3:5966597–5968962**	376	85	e-168	376	80	e-169
PGS1.21	ppa004549m	8078825	8081487	2663	503	AT1G56720.3	492	Protein kinase superfamily protein	chr1:21263630–21265559	504	72	0.0			
	ppa007758m	8082833	8086119	3287	356	AT3G17430.1	364	Protein kinase superfamily protein	**chr3:5956601–5958882**	347	79	e-163	346	79	e-159
	ppa000690m	8093363	8099939	6577	1035	AT2G47070.1	881	Squamosa promoter binding protein-like 1	chr2:19337144–19340552	643	49	e-167			
	ppa000792m	8101127	8107079	5953	1002					1020	44	0.0	1016	47	0.0
	ppa020554m	8114349	8116135	1787	485	AT3G06430.1	486	Tetratricopeptide repeat (TPR)-like superfamily protein	**chr3:1956658–1958240**	418	76	0.0	445	76	0.0
	ppa023381m	8117825	8125610	7786	1392	AT3G53480.1	1450	Pleiotropic drug resistance 9	chr3:19825366–19831644	1455	62	0.0			
	ppa000229m	8128589	8135634	7046	1429					1416	61	0.0			
	ppa006833m	8141239	8143517	2279	393	AT3G17390.1	393	*S*-Adenosyl-methionine synthetase	**chr3:5952484–5953665**	393	90	0.0	393	90	0.0
	ppa006841m														
	ppa006846m														
	ppa013518m	8148425	8150167	1743	118	AT3G15580.1	115	Ubiquitin-like superfamily protein	**chr3:5274075–5275000**	115	76	1.00E-046	115	76	3.00E-049
PGS1.23	ppa022254m	8150688	8152509	1822	314	AT3G17380.1	309	TRAF-like family protein	**chr3:5950240–5952124**	306	51	3.00E-089			
	ppb022195m	8156175	8157899	1725	319					305	55	3.00E-095			
	ppa008951m	8159042	8160589	1548	313					303	57	e-106	303	57	e-108
	ppa024696m	8168103	8172400	4298	319					308	50	6.00E-087			
	ppb020867m	8173007	8175838	2832	336					157	42	2.00E-030			
	ppa008960m	8176411	8178193	1783	312					305	49	1.00E-079			
	ppa017827m	8192292	8194960	2669	314					310	51	2.00E-089			
	ppa023061m	8201953	8203266	1314	296					291	51	3.00E-082			
	ppa019595m	8205205	8206582	1378	299					298	50	9.00E-079			
	ppa013116m	8208726	8209814	1089	139	AT1G48170.1	153	Unknown protein	**chr1:17788810–17789635**	101	72	3.00E-040	101	72	4.00E-039
	ppa000137m	8210743	8217757	7015	1160	AT5G18525.1	1639	Protein serine/threonine kinases; protein tyrosine kinases; ATP binding; protein kinases	**chr5:6146932–6153658**	1680	58	0.0	1678	57	0.0
PGS1.24	ppa006199m	8218685	8221818	3134	423	AT1G09430.1	424	ATP-citrate lyase A-3	**chr1:3042135–3044978**	420	88	0.0	420	88	0.0
	ppa021666m	8223846	8224599	754	251	AT3G56530.1	319	NAC domain-containing protein 64	chr3:20948911–20950045	162	29	5.00E-008			
	ppa024143m	8225331	8225750	420	139	N/A									
	ppa018241m	8228143	8230197	2055	167	AT2G30280.1	346	RNA-directed DNA methylation 4	chr2:12909804–12911860	191	43	2.00E-034			
	ppa021156m	8232169	8233179	1011	156	AT4G04780.1	139	Mediator 21	chr4:2432665–2433397	156	59	4.00E-031			
	ppa010512m	8233367	8235295	1929	247	AT3G17365.1	239	*S*-Adenosyl-L-methionine-dependent methyltransferases superfamily protein	**chr3:5947144–5948766**	230	66	3.00E-090	268	53	2.00E-075
	ppa000912m	8238644	8244925	6282	964	AT2G22300.2	1032	Signal responsive 1	chr2:9471599–9476472	631	43	e-141			
SNP-825	8247059														

Table 3 Summary of variants recorded from the alignment of the nine apricot genomes and the bacterial artificial chromosome (BAC) contig sequences against the 196-kb region of the peach genome containing the *PPVres* locus. Distribution in different genic compartments is shown for each sample and on average.

'Goldrich'	8022	30	1618	530/459	69	5316	2990	3	592	139	20	2236
'Harlayne'	11070	38	1937	743/584	70	7698	3081	5	555	148	19	2354
'Stark Early Orange'	10949	30	1957	727/587	72	7576	3037	3	577	167	18	2272
'Canino'	8187	31	1609	501/436	70	5540	3098	6	571	136	25	2360
'Krasnoshchekii'	10826	31	1960	732/627	73	7403	3057	5	605	148	19	2280
'Reale d'Imola'	10754	33	1925	725/578	71	7422	3050	5	570	163	19	2293
'Shalakh'	10678	40	1942	728/583	68	7317	3065	5	580	166	16	2288
Prunus mume	11272	45	1944	755/666	72	7790	3242	4	565	154	16	2503
Prunus sibirica	11225	52	2014	779/674	74	7632	3075	9	616	162	21	2267
Average	10331.44	36.7	1878.4	691/577	71.0	7077.1	3077.2	5.0	581.2	153.7	19.2	2317.0
%		0.4	18.2	6.7/5.6	0.7	68.5		0.2	18.9	5.0	0.6	75.3
Length region	196255	806	45303	96860	2552	50734	196255	806	45303	96860	2552	50734
Frequency (1 variant/*X* bp)	19.0	22.0	24.1	140.2/167.9	35.9	7.2	63.8	161.2	77.9	630.3	132.8	21.9

other four transcripts showed variants that might have stronger effects. ppa008951m had one nsSNP within the first MATH domain; ppa024696m and ppa019595m had one nsSNP within the second MATH domain and also additional frameshift mutations in some varieties; ppb022195m had a 5-bp deletion in the three resistant varieties within the second exon that should produce a truncated protein having just 42 amino acids, leading to a loss of both MATH domains.

The 23 transcripts selected during the filtering process can be roughly classified into two groups according to their homology with proteins involved in virus resistance (see Discussion section and Table 2). The first group comprises six transcripts (ppa023381m, ppa000229m, ppa006199m, ppa013518m, ppa020554m, ppa013116m) putatively coding for proteins unrelated to pathogen resistance according to the accumulated evidence. In the second group, eight transcripts (ppa006833m, ppa006841m, ppa006846m, ppa021156m, ppa010512m, ppa00912m, ppa000690, ppa000792m) are somehow related to biotic stress resistance (and, particularly, in some cases with virus resistance), but not to cell-to-cell or long-distance viral movement. In addition, nine transcripts (three putatively coding for serine/threonine kinases and six for MATHd proteins) seem to be closely associated with potyvirus resistance and, particularly, with PPV resistance in *Arabidopsis* regarding MATHd proteins.

Microsynteny of peach vs. *Arabidopsis*

The *sha3* locus has been reported recently to control PPV infection in *Arabidopsis thaliana* (Cosson *et al*., 2010; Pagny *et al*., 2012). In order to identify putative orthologous genes between *Prunus* and *A. thaliana*, BLASTP analysis of the 31 predicted proteins within the *PPVres* locus was performed against the TAIR database (E value cut-off < 1e-6) (Table 2 and Fig. S1, see Supporting Information). Thus, *Prunus*/*Arabidopsis* gene pairs that are BLASTP best-reciprocal hits (BRHs) were considered as putative orthologues. Under this criterion, seven peach transcripts were found to be putative orthologues of *Arabidopsis* transcripts located in the upper part of chromosome 3 (ppa020554m at position ~1.96 Mb and the rest between ~5.27–5.97 Mb), two were located in chromosome 1, one at the distal end of chromosome 2 and another on the upper part of chromosome 5.

Putative orthologues were also supported by the conserved gene order between both species along the defined chromosome region (Nozawa and Nei, 2007; Zheng *et al*., 2005). Interestingly, five of the seven *Arabidopsis* transcripts located in chromosome 3 preserved the same order, but inverted relative to their putative peach orthologues (Table 2). In addition, all nine peach transcripts putatively coding for TRAF-like family proteins showed high homology with one *Arabidopsis* transcript (At3g17380), but ppa008951m was the only BLASTP BRH with a 55.41% identity along 88.15% of the QS length and an E value of 5e-95. BLASTing At3g17380.1 (located in scaffold_3: 5950200–5952235) against the peach genome predicted peptides showed high-similarity E values (ranging from 7e-108 to 1e-80) with all nine TRAF-like genes present in the cluster.

In addition, BLASTP analysis of the restricted *Tobacco etch virus* (TEV) movement 3 (RTM3) protein (coded by the At3g58350.1 transcript located within the *sha3* locus in scaffold_3: 21 591 452–21 592 962) against the peach predicted peptides showed a 29.76% similarity along 96% of the QS length (with an E value of 5e-23) with ppa024552m, a ubiquitin carboxyl-terminal hydrolase 12 protein (containing one MATH domain) located in the peach scaffold_2: 4 372 558–4 374 144.

DISCUSSION

Physical mapping and variant calling from NGS data

Fine mapping of the *PPVres* locus in apricot was hampered for years by the difficulties inherent to fruit tree management, the lack of large segregating populations and the limited efficiency of phenotypic resistance assays (Llácer *et al*., 2007). Recently, Soriano *et al*. (2012) restricted the *PPVres* locus to a short interval of ~270 kb by combining data from two PPV resistance sources ('Goldrich' and 'Stark Early Orange'). Here, we have constructed a detailed physical map of this region by assembling 'Goldrich' BAC

Table 4 Filtering process of variants potentially associated with *Plum pox virus* (PPV) resistance. Filter 1 selected variants conserved in all resistant or all susceptible varieties indicating their type, peach genome position and genic compartments. Filter 2 identified variants present in heterozygosity. Filter 3 found variants present in the R-haplotype of 'Goldrich'. Variants putatively leading to frameshift mutations in some genotypes are indicated in the last column. Boundaries of the *PPVres* locus and peach genes not present in apricot genomes are indicated with a grey background.

		Filter 1					Filter 2	Filter 3	
		Variants preserved in all R or S but different between R/S					Variants heterozyg.	Variants in R-haplotype	Additional variants
Molecular marker	Peach gene ID	Genotypes showing the variant	Peach genome position	Gene position	SNP/DIP	Amino acid change			
SNP-805									
	ppa020498m	None							
	ppa007310m	None							
PGS1.21	ppa004549m								
		All 3 R	8080436	Exon	1122G>A		+	+	
		All 3 R	8081456	3′-UTR	*133T>A		+	+	
									several genotypes: Asn361fs
	ppa007758m								
		All 3 R	8082853	3′-UTR	*159A>G		+	+	
		All 3 R	8085641	5′-UTR Intron	-11-136G>C		+	–	
	ppa000690m								
		All 3 R	8093993	Intron	581+49A>G		+	–	
		All 3 R	8094042	Intron	581+98A>T		+	–	
		All 3 R	8096099	Intron	1419+268T>C		+	–	
		All 3 R	8096392	Intron	1420-97A>G		–	+	
		All 3 R	8097512	Exon	1914C>T		+	+	
		All 3 R	8097830	Intron	2013+219G>A		+	+	
		All 3 R	8098172	Intron	2014-154G>T		+	–	
		All 3 R	8098234	Intron	2014-92A>T		+	+	
		All 3 R	8098913	Exon	2601C>T		+	+	
		All 3 R	8099359	Exon	2813C>T	Thr938Ile	+	+	
		All 3 R	8099700	3′-UTR	*47T>C		+	+	
		All 3 R	8099757	3′-UTR	*104G>A		+	+	
									All: Gly858del
	ppa000792m								
		All 3 R	8101366	Exon	239C>T	Pro80Leu	+	+	
		All 3 R	8101955	Intron	567-8A>G		+	-	
		All 3 R	8102060	Exon	664A>G	Thr222Ala	+	+	
		All 3 R	8102560	Exon	753C>T		+	+	
		All 3 R	8102720	Exon	913G>T	Asp305Tyr	+	+	
		All 3 R	8103277	Intron	1344+126C>T		+	–	
		All 3 R	8103354	Intron	1345-152C>A		+	–	
		All 3 R	8103424	Intron	1345-82C>A		+	–	
		All 3 R	8104026	Intron	1645+25_1645+30delCCCCCT		+	+	
		All 3 R	8105026	Intron	1939-56G>A		+	+	
		All 3 R	8105607	Exon	2464G>T	Val822Phe	+	+	
		All 3 R	8106308	Intron	2557-48T>C		+	+	
		All 3 R	8106978	3′-UTR	*171C>T		+	–	
		All 3 R	8107025	3′-UTR	*218C>A		+	–	
		:							Shalakh: Arg215fs (putative) *P. mume*: Gln194fs (putative)
	ppa020554m								
		All 3R:	8115155	Exon	474C>T		+	+	
									Canino: Leu203fs *P. sibirica*: Val229fs
	ppa023381m								
		All 3R:	8120816	Exon	1665T>A		+	+	
	ppa000229m								
		All 3R:	8133376	Exon	1216G>A	Ala406Thr	+	+	
									Almost all: His475fs *P. sibirica*: Leu464*
	ppa006841m								
		All 3 R	8142119	Exon	699C>T		+	–	
		All 3 R	8143183	5′-UTR Intron	-10-357_-10-358insG		+	+	
		All 3 R	8143185	5′-UTR Intron	-10-359_-10-360insCC		+	+	
									P. sibirica: Gln392fs
	ppa013518m								
		All 3 R	8148875	Intron	97-187T>C		+	+	
		All 3 R	8149092	Exon	127C>T		+	+	
		All 3 R	8149803	Intron	268+12A>G		+	+	
									P. mume: Pro54fs, Ala90fs
PGS1.23	ppa022254m								
		All 3R:	8151317	Intron	618-170_618-169insA		+	+	
	ppb022195m								
		All 3 R	8156185	Exon	951G>A		+	+	
		All 3 R	8156301	Exon	835C>A		+	+	
		All 3 R	8156345	Intron	803-12_803-11nsTCA		+	+	

Table 4 Continued.

Molecular marker	Peach gene ID	Filter 1: Variants preserved in all R or S but different between R/S — Genotypes showing the variant	Peach genome position	Gene position	SNP/DIP	Amino acid change	Filter 2: Variants heterozyg.	Filter 3: Variants in R-haplotype	Additional variants
		All 3 R	8156357	Intron	803-24C>T		+	+	
		All 3 R	8156361	Intron	803-27T>C		+	–	
		All 3 R	8156365	Intron	803-32delA		–	–	
		All 3 R	8156383	Intron	803-50T>A		+	+	
		All 3 R	8156414	Intron	802+31C>T		+	+	
		All 3 R	8156440	Intron	802+5A>G		+	+	
		All 3 R	8156451	Exon	796A>G	Thr266Ala	+	+	
		All 3 R	8156515	Exon	732A>G		+	+	
		All 3 R	8156707	Intron	627-87A>G		+	–	
		All 3 R	8156890	Exon	621G>A		+	+	
		All 3 R	8156931	Exon	580G>A	Glu195Lys	+	+	
		All 3 R	8157378	Exon	260_261insA		+	–	
		All 3 R	8157384	Exon	254delG		+	–	
		All 3 R	8157588	Intron	167+11_167+12delTT		+	+	
		All 3 R	8157652	Exon	111_115delCAAAC	Ser37fs	+	+	
		All 3 R	8157671	Exon	96G>A		+	+	
		All 3 R	8157704	Exon	63A>C		+	+	
		All 3 R	8157738	Intron	38-9delG		+	–	
		All 3 R	8157754	Intron	38-25_38-24insCTT		+	+	
		All 3 R	8157762	Intron	38-33T>A		+	+	
	ppa008951m								
		All 3 R:	8159891	Exon	370G>A	Asp124Asn	+	+	
	ppa024696m								
		All 4 S + sibirica	8168482	Intron	158+137delT		–	–	
		All 3 R	8170400	Intron	159-934_159-933insA		+	–	
		All 3 R	8172048	Exon	714T>G		+	+	
		All 3 R	8172050	Exon	716T>G	Val239Gly	+	+	
									All: Thr2del Stark Early Orange and krasnos: Asn221fs
	ppb020867m								
		All 3 R	8173027	Exon	992A>G/T	Ile331Ala, Ille331Ser	+	+ (992A>T; Ille331Ser)	
		All 3 R	8173129	Exon	890A>T	Gln297Leu	+	+	
									Partially deleted
	ppa008960m	Absent							
	ppa017827m	Absent							
	ppa023061m								
		All 3 R	8202124	Intron	752-31T>C		+	+	
		All 4 S	8203123	Intron	128+17A>G		–	–	
		All 4 S	8203215	Exon	53A>G	Glu18Gly	–	–	
									5 genotypes: Tyr95fs
	ppa019595m								
		All 3 R	8205669	Exon	589A>G	Lys197Glu	+	+	
		All 3 R	8206170	Intron	298+6C>T		+	+	
									Canino, Krasnos and Shalakh: Gly208fs and Lus209fs
	ppa013116m								
		All 3 R	8209518	Intron	195+76A>G		+	+	
		All 3 R	8209807	5′-UTR	-20G>A		+	+	
	ppa000137m								
		All 3 R	8210884	Exon	4851T>C		+	+	
		All 3 R	8211235	Intron	4697+114C>T		+	+	
		All 3 R	8211269	Intron	4697+80A>T		+	+	
		All 3 R	8211502	Exon	4614C>T		+	+	
		All 3 R	8215860	Exon	1612C>T	Arg538Cys	+	+	
PGS1.24	ppa006199m								
		All 3 R:	8220099	Intron	768+20T>A		+	+	
									Stark Early Orange, Krasnos and *P. sibirica*: Ser51* All Glu116*
	ppa021666m	None							
	ppa024143m	None							
	ppa018241m	None							
	ppa021156m								
		All 4 S	8232620	Intron	394+67_394+69delTTG		–	–	
		All 3 R	8232674	Intron	394+15C>T		+	+	
		All 3 R	8232889	Intron	264+28C>A		+	+	

Table 4 Continued.

Molecular marker	Peach gene ID	Filter 1: Variants preserved in all R or S but different between R/S: Genotypes showing the variant	Peach genome position	Gene position	SNP/DIP	Amino acid change	Filter 2: Variants heterozyg.	Filter 3: Variants in R-haplotype	Additional variants
		Stark Early Orange	8232989	Exon	192T>C	Ile64Arg			
		Stark Early Orange	8233002	Exon	179G>T	Thr60Asn			
		Stark Early Orange	8233036	Exon	147C>G	Ala49Pro			
		Stark Early Orange	8233045	Exon	137C>T	Ala46Thr			
	ppa010512m								
		All 4 S	8233736	Exon	204G>C		–	–	
		All 3 R	8234739	Intron	557-9T>C		+	+	
	ppa000912m								
		All 3 R	8238840	Intron	132+64G>T		+	+	
		All 3 R	8241759	Intron	[844-149C>T], [844-149C>G]		+	+ (844-149C>T)	
SNP-825									

clones into two separate haplotype contigs and by performing deep sequencing of both the R-haplotype contig and the entire genome of nine apricot accessions. Segregating SNPs found in the NGS data alignment against the peach genome were used to narrow down the *PPVres* locus to ~196 kb. In agreement with the high degree of map collinearity found among *Prunus* spp. (Dirlewanger *et al*., 2004; Dondini *et al*., 2007; Lambert *et al*., 2004; Olmstead *et al*., 2008), over 93% of similarity was detected in the *PPVres* locus between peach and *P. armeniaca*, *P. mume* and *P. sibirica*. However, discrepancies were also found: several misalignments (gaps) were detected (data not shown), and three peach MATH domain-containing transcripts were found to be partially or completely lost in apricot. Moreover, the presence of additional genes in apricot could not be fully discarded.

Comparison of resistant/susceptible genotypes allowed us to screen for differential variants at the *PPVres* locus by considering the genetic evidence supporting the proposed model for PPV resistance in apricot (Soriano *et al*., 2008; Vera-Ruiz *et al*., 2011). Thus, only those variants in coupling with the resistance and present in heterozygosity in all the resistant varieties or, alternatively, variants in repulsion with the resistance and present in homozygosity in all the susceptible varieties, were considered to be relevant. A total of 68 variants within 23 putative transcripts were found to fulfil these requirements, but none belonged to the second type. From these, 23% were nonsynonymous variants, leading to amino acid or frameshift changes.

It is generally accepted that most nonsilent mutations are not effectively neutral, leading to functional changes in the affected protein in hominids, *Drosophila* and enteric bacteria (Eyre-Walker and Keightley, 2007). However, a recent survey across human SNP–disease associations showed that, although DNA variants pursued for functional validation are usually those predicted to lead to significant amino acid changes, sSNPs show a similar likelihood and effect size for disease association (Chen *et al*., 2010). Interestingly, variants causing a premature stop codon are most likely to be associated with disease and have a higher effect size than other SNP types (Chen *et al*., 2010). As a whole, none of the 23 transcripts can be fully rejected according to the reported variants. In this study, we have combined information derived from the type of mutation detected in every transcript and their putative function inferred from sequence homology to classify them with regard to their likelihood of being candidate genes.

Ranking candidate genes within the *PPVres* locus

In recent years, a number of dominant and recessive virus resistance genes have been identified and characterized in plants (Maule *et al*., 2007). Dominant resistance is generally associated with so-called *R* genes that confer resistance to bacteria, fungi, nematodes and viruses (Soosaar *et al*., 2005). In particular, *R* genes conferring dominant virus resistance mostly fall into the nucleotide-binding site-leucine-rich repeat class (Maule *et al*., 2007). However, there are a few exceptions, such as the *RTM* resistance genes. Interestingly, the first three *RTM* genes were initially found to restrict long-distance movement (LDM) of TEV in *A. thaliana*, but recent evidence has allowed their effects to be extended to other potyviruses, such as PPV (Decroocq *et al*., 2006). Recently, two new *RTM* loci have been genetically identified in *A. thaliana* (Cosson *et al*., 2012), indicating that at least five proteins are involved in the complex RTM resistance process. However, recessive resistance has been commonly related to mutations of components of the eukaryotic translation initiation complex, mostly affecting the eukaryotic initiation factor 4E necessary for virus replication (Robaglia and Caranta, 2006). Their implication in PPV resistance has not been fully demonstrated, but a striking correlation was found by Marandel *et al*. (2009) in *Prunus davidiana* (Carrière) Franch. Nevertheless, in spite of these significant advances, no factor involved in the PPV resistance mechanism has been identified to date in its natural host, *Prunus* spp.

Fine mapping and NGS variant analysis led us to restrict the number of candidate genes for PPV resistance in apricot to 23. However, implication in PPV resistance of 12 of these seems quite

unlikely: six have little or no direct connection to genes known to confer resistance to pathogens, and the other six have been linked to virus resistance, but the reported phenotypes differ substantially from that which has been observed for PPV resistance in apricot. Within this latter group, ppa006833m, ppa006841m and ppa006846m represent alternative transcripts of a unique gene sharing as BRH an *Arabidopsis S*-adenosyl-methionine (AdoMet) synthetase. Recently, Użarowska *et al*. (2009) identified an AdoMet synthetase as a candidate gene for resistance to two potyviruses in maize. However, other candidate genes were also proposed by these authors, and therefore a direct correlation with this phenotype should be further confirmed. Moreover, the role of the apricot transcript does not seem plausible as two DIPs within the intronic region of the 5′-UTR should explain the PPV resistance phenotype. Three additional transcripts putatively encode for peptides highly similar to transcriptional regulatory proteins. ppa021156m is similar to the MED21 subunit (At4g04780) of the mediator multiprotein complex, a coactivator required by DNA-binding transcription factors for transcriptional activation of polymerase II-transcribed genes (Sato *et al*., 2004). Reduced expression of *Arabidopsis* MED21 has been associated with disease susceptibility and embryo-lethal phenotypes (Dhawana *et al*., 2009). The first intron has been suggested to play an important regulatory role in many genes, most probably involved in transcriptional control (Majewski and Ott, 2012). This apricot transcript has one SNP within the first intron in the resistant varieties, but its location at the beginning of the intron could imply a low plausibility of its effect (Chen *et al*., 2010). ppa010512m is the BRH of the At3g17365 protein, an AdoMet-dependent methyltransferase. Recently, an AdoMet methyltransferase was described as an innate defence protein against potexvirus accumulation in *Nicotiana benthamiana* (Cheng *et al*., 2009). However, the implication of AdoMet methyltransferases has not been reported for potyvirus infection. Moreover, resistant varieties have just one SNP within the seventh intron, and therefore a role for this gene in PPV resistance seems unlikely. Lastly, ppa000912m is partially similar to the *Arabidopsis* signal responsive 1 (SR1) protein, which has been described as a negative regulator of plant defences (Qiu *et al*., 2012). SR1 acts as a negative regulator of salicylic acid (SA)-mediated plant immunity by repressing the transcription of two SA signaling-positive regulators (Du *et al*., 2009; Nie *et al*., 2012). ppa000912m has two SNPs predicted within an exonic splicing enhancer of the first intron and within the fifth intron, respectively. Moreover, dominant loss-of-function SR1 mutants have been shown to display enhanced resistance to fungal and bacterial pathogens (Nie *et al*., 2012). As a whole, this evidence would support a possible implication of this gene in the apricot PPV resistance. However, the type of constitutive disease resistance (including the hypersensitive response) associated with SR1 mutants (Du *et al*., 2009) does not fit with the observed symptoms during PPV infection in apricot.

Candidate genes for PPV resistance in apricot

The compiled evidence suggests that PPV resistance in apricot is more closely linked to the restriction of the virus LDM than to specific host resistance responses (Dicenta *et al*., 2003; Ion-Nagy *et al*., 2006). Thus, the most compelling candidate genes within the *PPVres* locus include three types of genes encoding proteins with known functions associated with the control of virus movement. In the first type, ppa000690m and ppa000792m have a plant-specific DNA-binding domain and also ankyrin repeats. Both transcripts are homologous to the squamosa promoter-binding-like protein 1 belonging to a family of DNA-binding proteins and putative transcription factors (Cardon *et al*., 1999). Ankyrin repeats have been related to cell-to-cell tobamovirus movement (Ueki *et al*., 2010); however, no similarity between the apricot transcripts and the ankyrin repeat-containing protein described by Ueki *et al*. (2010) was observed (data not shown). Moreover, the ankyrin repeat has been found in many proteins, spanning a wide range of functions, typically involved in the mediation of specific protein–protein interactions (Mosavi *et al*., 2004). Altogether, the implication of these two proteins in PPV resistance in apricot seems quite unlikely.

The second class includes three transcripts showing similarity to serine/threonine protein kinases (STKs), which have been reported to be important factors for defence signal transduction (Xu and Deng, 2010). Moreover, glycosylation and phosphorylation have been suggested to play, in a broad sense, crucial roles in the regulation of potyviral capsid protein (CP) functions (Fernández-Fernández *et al*., 2002), among which is involvement in the movement of infectious viral RNA from cell to cell and over long distances (Dolja *et al*., 1994). ppa004549m shows high similarity to three *Arabidopsis* STK proteins (At1g56720, At5g18500 and At2g42960). Its R-haplotype has one SNP within the 3′-UTR that could have an important effect on gene expression regulation, because it is located within a predicted hsa-miR-223 target site. Interaction between microRNA (miRNA) and its target mRNA appears to result in translational repression or, in some cases, cleavage of cognate mRNAs, causing partial or full silencing of the respective protein-coding genes (Laios *et al*., 2008). We could hypothesize that the expression of this STK protein could not be repressed in apricot-resistant varieties, therefore having an active phosphorylation function, and, conversely, phosphorylation of viral movement proteins has been implicated in plant virus movement (Ivanov *et al*., 2003). ppa000137m is the putative orthologue of the *Arabidopsis* At5g18525, a DDB1-CUL4-associated factor 1 (DCAF1) protein which could be a substrate receptor for the CUL4 RING ubiquitin ligase complex (Lee *et al*., 2008). Human DCAF1 acts by providing human immunodeficiency virus type 1 with the equipment for the degradation of specific host proteins and by counteracting its proteasome targeting by another cellular E3 ubiquitin ligase (Le Rouzic *et al*., 2008). However, in plants,

mutant analyses have shown that DCAF1 has an essential role in plant embryogenesis and is involved in multiple developmental pathways (Zhang *et al*., 2008). ppa007758m, a putative orthologue of the *Arabidopsis* At3g17410 STK protein, is highly similar to the *Solanum lycopersicum* L. Pto kinase interactor 1 (Pti1) (id = 87%). It has one SNP within a predicted selenocysteine insertion sequence element in the 3′-UTR, a region thought to be implicated in protein–protein interactions (Martin *et al*., 1998). Moreover, At3g17410 expression was induced by infections of PPV and other positive-sense RNA viruses (Babu *et al*., 2008). However, Pti has been reported to be involved in the Pto-mediated signalling pathway, leading to the hypersensitive response against fungi (Zhou *et al*., 1995), a type of response that does not fit with the observed phenotype in the case of PPV resistance in apricot.

The third class of *PPVres* locus candidate genes is formed by a cluster of six MATHd-containing genes, two of which have synonymous variants with putatively low impact, three of which have at least one amino acid change within a MATH domain, and one of which has a 5-bp deletion that should produce a truncated protein lacking both MATH domains. Taking into account that the mutated allele is dominant and shows resistance in heterozygosity, the implication of these apricot transcripts could be more plausibly explained by a gain-of-function or a dominant negative mutation, where either the mutant protein inhibits directly the activity of the wild-type protein through dimerization, or competes with the wild-type protein for another protein that is required for normal function (Fay and Spencer, 2005).

Interestingly, other MATHd-containing proteins have been described to be involved in virus LDM. The *RTM3* gene, which encodes a MATHd-only protein (Cosson *et al*., 2010), is one of the dominant *RTM* genes underlying potyvirus resistance previously described in *Arabidopsis* (Chisholm *et al*., 2001; Decroocq *et al*., 2006; Mahajan *et al*., 1998). However, a mutation in any one of the three first identified *RTM* genes completely abolishes the resistance (Cosson *et al*., 2010). Although compelling, the genetics of RTM resistance seems to be different from that of PPV in apricot, as the mutated protein is predicted to confer PPV resistance in heterozygosity. Interestingly, the *sha3* locus, which encompasses the *RTM3* gene and includes another seven *RTM3*-like TRAF domain-containing genes, has been described recently to restrict PPV LDM in *Arabidopsis* in a recessive fashion (Pagny *et al*., 2012). However, apricot MATH cluster genes show highest similarity to At3g17380 in *Arabidopsis*, which is located on the upper part of chromosome 3, whereas the *sha3* locus maps at the bottom of the same chromosome. It has been suggested that most of the MATH genes of the MATHd-only protein family are organized in clusters in the *Arabidopsis* genome (Cosson *et al*., 2010). Similarly, in the apricot *PPVres* locus and in the syntenic peach region, clusters of six and nine TRAF-like genes, respectively, were observed. However, their putative *Arabidopsis* orthologous gene (At3g17380) seems to be a single-copy gene, suggesting that duplication events probably occurred after the diversification of these species. Finally, the two new genetically identified *RTM* loci are not in *PPVres* locus syntenic locations, as *RTM4* maps in *Arabidopsis* chromosome 1 and *RTM5* in chromosome 2 (Cosson *et al*., 2012).

In conclusion, this is the first report in which whole-genome sequencing (WGS) (seven apricot varieties and two wild relatives) and the physical map of the *PPVres* locus have been combined. Our data support member(s) of a cluster of MATHd-containing genes as the most promising candidate gene(s) for PPV resistance in apricot. The functions of the MATHd-like proteins are not well known, although currently available evidence suggests that they might link specific protein substrates to ubiquitin ligase complexes (Zapata *et al*., 2007). Functional analyses are currently in progress in *Arabidopsis* and plum, taking advantage of the high transformation efficiency of the latter species (Petri *et al*., 2008), in order to elucidate the role of TRAF-like proteins in PPV resistance. The results presented here will contribute to the implementation of new strategies to improve breeding for resistance, including the introgression of resistance monitored by marker-assisted selection. They will also contribute to increase the knowledge of the natural resistance to PPV in the main plant–pathogen systems of PPV.

EXPERIMENTAL PROCEDURES

Plant material

Nine apricot genotypes were selected for WGS, representing three PPV-resistant cultivars ('Goldrich', 'Stark Early Orange' and 'Harlayne'), four PPV-susceptible cultivars ['Canino', 'Shalakh', 'Krasnoshchekii (syn. 'Hungarian Best') and 'Reale d'Imola'] and two PPV-susceptible nondomesticated species: *Prunus mume* (Sieb.) Sieb & Zucc. and *Prunus sibirica* L. var. *davidiana* (Carrière) (Table 5). For WGS, all cultivars and species, except 'Canino', came from the repository at Nikita Botanical Garden, Crimea, Ukraine, and have been genotyped previously in Zhebentyayeva *et al*. (2008). The Spanish genotype 'Canino' is maintained as part of the germplasm collection at IVIA, Valencia, Spain.

DNA extraction

Genomic DNA from 'G×Ca' and 'G×Ca-08' recombinant hybrids and their parents ('Goldrich' and 'Canino') was isolated from young leaves using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol, and stored at −20 °C until use. Genomic DNA from the Nikita Botanical Garden samples was extracted and stored at −20 °C according to the previously published protocol by Zhebentyayeva *et al*. (2008). DNA integrity was checked on agarose gel and quantification was performed using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA).

Construction of the BAC contig spanning the *PPVres* locus

BAC clone identification was carried out using digoxigenin-labelled overgo probes hybridized in pools (Madishetty *et al*., 2007) against an apricot BAC

Table 5 Apricot cultivars used in this study. Geographical origin, pedigree and *Plum pox virus* (PPV) resistance phenotype are indicated.

Cultivar	Geographical area	Country of origin	Pedigree	PPV resistance	Reference
'Goldrich'	North America	USA	Sunglo × Perfection*	R	Dosba *et al*. (1992)
'Harlayne'	North America	Canada	[(Reliable × o.p.) × o.p.] × Sunglo*	R	Dosba *et al*. (1992)
'Stark Early Orange'	North America	USA	Unknown*	R	Syrgiannidis (1980)
'Canino'	Western Europe	Spain	Unknown†	S	Avinent *et al*. (1993)
'Krasnoshchekii'	Eastern Europe	Ukraine	Unknown‡	S	Evaluated at IVIA
'Reale d'Imola'	Western Europe	Italy	Unknown†	S	Tradafirescu and Topor (1999)
'Shalakh' (syn. Erevan)	Irano-Caucasian	Armenia	Unknown‡	S	Karayiannis (1989)
Prunus mume	Japan and China§			S	James and Thompson (2006)
Prunus sibirica	Eastern Siberia, Manchuria and northern China§			S	James and Thompson (2006)

R, resistant; S, susceptible.
*Brooks and Olmo (1997).
†Della Strada *et al*. (1989).
‡Zhebentyayeva *et al*. (2008).
§Layne *et al*. (1996).

library developed from the PPV-resistant cultivar 'Goldrich' and printed on six high-density filters (Vilanova *et al*., 2003). Overgo probes were designed using the Overgo 1.02i software (http://www.mouse-genome.bcm.tmc.edu/webovergo/OvergoInput.asp), prepared as described by Hilario *et al*. (2007) and hybridized following the manufacturer's instructions (Roche, Basle, Switzerland). In a first hybridization round, six overgo probes were designed using, as templates, the flanking sequences of the SSR markers developed for the fine mapping of the *PPVres* locus by Soriano *et al*. (2012). Positive BAC clones were assigned to their corresponding R- or S-haplotype by PCR screening. BES were obtained from all positive clones, and primer pairs were designed using Primer3 software (Rozen and Skaletsky, 2000) to verify overlapping BACs by PCR. Twenty additional overgo probes were designed from BES for a second hybridization round to cover those gaps in which clones did not overlap. BES were BLASTed against the peach v1.0 genome sequence (IPGI, http://www.rosaceae.org/species/prunus_persica/genome_v1.0) using a stand-alone version of BLAST (Altschul *et al*., 1990) in order to indicate the orientation of the BAC clones and to provide an indirect estimation of their size.

Next-generation sequencing

BAC clones of the R-haplotype contig cv. 'Goldrich' were massively parallel pyrosequenced as a pool using 454 GS-FLX Titanium NGS technology (Roche), commercially conducted by Macrogen Inc. (Seoul, South Korea). WGS of the nine apricot genomes was conducted on an Illumina HiSeq2000 platform, using 100-bp paired-end reads. Sequencing of cv. 'Canino' was commercially conducted by Macrogen Inc. The remaining six cultivars and two nondomesticated species were sequenced at genomic facilities at DHMRI (David H. Murdock Research Institution, Kannapolis, NC, USA; http://www.dhmri.org).

NGS data analysis

All raw reads were processed using the ngs_backbone pipeline (Blanca *et al*., 2011) with the configuration file 'backbone.conf' (File S1, see Supporting Information). After removing the low-quality regions as well as vector and adaptor contaminants, cleaned reads were aligned to the peach genome syntenic region using CLC Genomics Workbench software v. 5.1 (Aarhus, Denmark). The FASTA sequence of the syntenic peach genome region was downloaded from the GDR website (http://www.rosaceae.org) (Jung *et al*., 2004) and used as reference. Variant calling to detect SNPs or DIPs was performed with CLC Genomics Workbench software v. 5.1 using the following parameters: DIPs: minimum (min) coverage, 4; min variant frequency, 35%; maximum (max) expected variation, 2; SNPs: window length, 11; max number of gaps and mismatches, 3; min average of surrounding bases, 15; min quality of central base, 20; min coverage, 4; min variant frequency, 35%; max expected variation, 2.

SNP linkage map

A selection of SNPs identified from the NGS data and annotated between positions corresponding to markers PGS1.20–PGS1.21 and PGS1.24–PGS1.252 were screened in 'Goldrich', 'Canino' and the six recombinant hybrids described by Soriano *et al*. (2012) in order to narrow down the *PPVres* locus. Primers (Table S1) were developed using Primer3 software (Rozen and Skaletsky, 2000). PCR amplifications were performed in a GeneAmp® PCR System 9700 thermal cycler (Perkin–Elmer, Freemont, CA, USA) in a final volume of 20 μL, containing 75 mM Tris-HCl, pH 8.8, 20 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.1 mM of each deoxynucleoside triphosphate (dNTP), 20 ng of genomic DNA and 1 U of Taq polymerase (Invitrogen, Carlsbad, CA, USA). The temperature profile was: 10 min at 95 °C, followed by 40 cycles of 20 s at 94 °C, 45 s at 52 °C and 1 min at 72 °C, finishing with 10 min at 72 °C. After amplification, PCR products were purified following the manufacturer's instructions for the High Pure PCR Product Purification Kit (Roche). DNA sequencing was performed using an ABI Prism 3130 genetic analyser, following the manufacturer's instructions for the BigDye terminator v.3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA). Each PCR fragment was sequenced using both forward and reverse primers. Sequences were edited and assembled using the PREGAP4 and GAP4 modules of the Staden software package (Staden *et al*., 2000).

Filtering of SNPs and DIPs in coupling with PPV resistance

Peach genome annotation data available from the GDR were used as a reference to identify associated polymorphisms within the predicted transcripts.

In order to discriminate the SNPs/DIPs linked in coupling with the resistance from all the variants found against peach, three filters were sequentially applied: (i) selected variants must be conserved within the group of resistant or susceptible varieties, being different between groups; (ii) resistant cultivars must be heterozygous for the variant; (iii) variants must be present in the resistant haplotype and therefore in the 454 BAC contig sequences. Complete transcripts with variants fulfilling these three requirements were interrogated in order to detect additional changes that might discard them as candidate genes for PPV resistance. All the routine of filtering and mathematical/statistical calculations was performed using OpenOffice Calc 3.0.1.

Microsynteny analysis between peach and *A. thaliana*

BLASTP analysis of the 31 proteins encoded by the predicted transcripts comprised within the peach syntenic region was performed against the TAIR database (version TAIR10 proteins), using the TAIR BLAST 2.2.8 tool (http://www.arabidopsis.org/Blast/index.jsp), with an E value cut-off < 1e-6, in order to predict gene functions based on homology (Table 2). This table also indicates those *Prunus*/*Arabidopsis* gene pairs that are BLASTP BRHs (BLASTing *Arabidopsis* proteins against the peach predicted peptides annotated by IPGI) identifying putative orthologues. The same procedure was used to predict putative orthologues in *Prunus* of some *A. thaliana* genes present in the *sha3* locus (Pagny *et al*., 2012).

ACKNOWLEDGEMENTS

This research was supported by grants from the Ministerio de Educación y Ciencia of Spain (Research Projects AGL2007-60709 and AGL2010-20595), the European Union Seventh Framework Programme (FP7/2007–2013, grant agreement: 204429), the European Regional Development Fund (ERDF), a grant from the US Department of Agriculture (USDA) Foreign Agricultural Service (FAS) Technical Assistance for Specialty Crops (TASC) program and a cooperative agreement between USDA-ARS/Clemson University.

We acknowledge Dr Albert Glenn Abbott for his comments on the manuscript.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Arús, P., Verde, I., Sosinski, B., Zhebentyayeva, T. and Abbott, A. (2012) The peach genome. *Tree Genet. Genomes*, **8**, 531–547.

Atanasoff, D. (1932) *Plum Pox. A New Virus Disease*. Bulgaria: Yearbook University of Sofia 1932/33, Faculty of Agronomy.

Avinent, L., Hermoso de Mendoza, A., Llácer, G. and García, S. (1993) Transmisión del virus de la sharka y sensibilidad varietal en albaricoquero. In: *Actas del II Congreso Ibérico de Ciencias Hortícolas, Zaragoza, Spain (Acta nº 9)*, pp. 200–206. Madrid: Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria.

Babu, M., Griffiths, J.S., Huang, T.S. and Wang, A. (2008) Altered gene expression changes in Arabidopsis leaf tissues and protoplasts in response to Plum pox virus infection. *BMC Genomics*, **9**, 325.

Badenes, M.L. and Llácer, G. (2006) Breeding for resistance: breeding for Plum pox virus resistant apricots (*Prunus armeniaca* L.) in Spain. *EPPO Bull.* **36**, 323–326.

Bassi, D. (2006) Breeding for resistance: breeding for resistance to *Plum pox* virus in Italy. *EPPO Bull.* **36**, 327–329.

Bassi, D. and Audergon, J.M. (2006) Apricot breeding, update and perspectives. *Acta Hort.* **701**, 279–294.

Blanca, J.M., Pascual, L., Ziarsolo, P., Nuez, F. and Cañizares, J. (2011) Ngs_backbone: a pipeline for read cleaning, mapping and SNP calling using Next Generation Sequence. *BMC Genomics*, **12**, 285.

Brooks, R.M. and Olmo, H.P. (1997) *The Brooks and Olmo Register of Fruit and Nut Varieties*, 3rd edn. Alexandria: ASHS Press.

Cambra, M., Capote, N., Cambra, M.A., Llácer, G., Botella, P. and López-Quílez, A. (2006a) Epidemiology of sharka disease in Spain. *EPPO Bull.* **36**, 271–275.

Cambra, M., Capote, N., Myrta, A. and Llácer, G. (2006b) Plum pox virus and the estimated costs associated with sharka disease. *EPPO Bull.* **36**, 202–204.

Capote, N., Cambra, M., Llácer, G., Petter, F., Platts, L.G., Roy, A.S. and Smith, I.M. (2006) Current status of *Plum pox* virus and sharka disease worldwide. *EPPO Bull.* **36**, 205–218.

Cardon, G., Höhmann, S., Klein, J., Nettesheim, K., Saedler, H. and Huijser, P. (1999) Molecular characterisation of the Arabidopsis SBP-box genes. *Gene*, **237**, 91–104.

Chen, R., Davydov, E.V., Sirota, M. and Butte, A.J. (2010) Non-synonymous and synonymous coding SNPs show similar likelihood and effect size of human disease association. *PLoS ONE*, **5**, e13574.

Cheng, C.W., Hsiao, Y.Y., Wu, H.C., Chuang, C.M., Chen, J.S., Tsai, C.H., Hsu, Y.H., Wu, Y.C., Lee, C.C. and Meng, M. (2009) Suppression of bamboo mosaic virus accumulation by a putative methyltransferase in *Nicotiana benthamiana*. *J. Virol.* **83**, 5796–5805.

Chisholm, S.T., Parra, M.A., Anderberg, R.J. and Carrington, J.C. (2001) Arabidopsis RTM1 and RTM2 genes function in phloem to restrict long-distance movement of Tobacco Etch Virus. *Plant Physiol.* **127**, 1667–1675.

Cosson, P., Sofer, L., Le, H., Leger, V., Schurdi-Levraud, V., Whitham, S.A., Yamamoto, M.K., Gopalan, S., Le Gall, O., Candresse, T., Carrington, J.C. and Revers, F. (2010) RTM3 which controls long distance movement of potyviruses is a member of a new plant gene family encoding a meprin and TRAF homology (MATH) domain-containing protein. *Plant Physiol.* **154**, 222–232.

Cosson, P., Schurdi-Levraud, V., Le, Q.H., Sicard, O., Caballero, M., Roux, F., Le Gall, O., Candresse, T. and Revers, F. (2012) The RTM resistance to potyviruses in *Arabidopsis thaliana*: natural variation of the RTM genes and evidence for the implication of additional genes. *PLoS ONE*, **7**, e39169.

Decroocq, V., Sicard, O., Alamillo, J.M., Lansac, M., Eyquard, J.P., García, J.A., Candresse, T., Le Gall, O. and Revers, F. (2006) Multiple resistance traits control *Plum pox virus* infection in *Arabidopsis thaliana*. *Mol. Plant–Microbe Interact.* **19**, 541–549.

Della Strada, G., Pennone, F., Fideghelli, C., Monastra, F. and Cobianchi, D. (1989) *Monografia di cultivar di albicocco. Ministerio dell'Agricoltura e delle Foreste. Direzione Generale della Produzione Agricola*. Rome: Instituto Sperimentale per la Frutticolture.

Dhawana, R., Luoa, H., Foersterb, A.M., AbuQamara, S., Duc, H.N., Briggsc, S.D., Scheidb, O.M. and Mengistea, T. (2009) HISTONE MONOUBIQUITINATION1 interacts with a subunit of the mediator complex and regulates defense against necrotrophic fungal pathogens in Arabidopsis. *Plant Cell*, **21**, 1000–1019.

Dicenta, F., Martínez-Gómez, P., Rubio, M. and Audergon, J.M. (2003) Localisation and movement of *Plum pox virus* in apricot stem tissues. *Ann. Appl. Biol.* **142**, 99–105.

Dirlewanger, E., Graziano, E., Joobeur, T., Garriga-Calderé, F., Cosson, P., Howad, W. and Arús, P. (2004) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proc. Natl. Acad. Sci. USA*, **101**, 9891–9896.

Dolja, V.V., Haldeman, R., Robertson, N.L., Dougherty, W.G. and Carrington, J.C. (1994) Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. *EMBO J.* **13**, 1482–1491.

Dondini, L., Lain, O., Geuna, F., Banfi, R., Gaiotti, F., Tartarini, S., Bassi, D. and Testolin, R. (2007) Development of a new SSR-based linkage map in apricot and analysis of synteny with existing Prunus maps. *Tree Genet. Genomes*, **3**, 239–249.

Dondini, L., Lain, O., Vendramin, V., Rizzo, M., Vivoli, D., Adami, M., Guidarelli, M., Gaiotti, F., Palmisano, F., Bazzoni, A., Boscia, D., Geuna, F., Tartarini, S., Negri, P., Castellano, M., Savino, V., Bassi, D. and Testolin, R. (2011) Identification of

QTL for resistance to plum pox virus strains M and D in Lito and Harcot apricot cultivars. *Mol. Breed.* **27**, 289–299.

Dosba, F., Orliac, S., Dutranoy, F., Maison, P., Massonie, G. and Audergon, J.M. (1992) Evaluation of resistance to plum pox virus in apricot trees. *Acta Hort.* **309**, 211–219.

Du, L., Ali, G.S., Simons, K.A., Hou, J., Yang, T., Reddy, A.S. and Poovaiah, B.W. (2009) Ca^{2+}/calmodulin regulates salicylic-acid-mediated plant immunity. *Nature*, **457**, 1154–1158.

Egea, J., Campoy, J., Dicenta, F., Burgos, L., Patino, J.L. and Ruiz, D. (2009) 'Estrella' and 'Sublime' apricot Cultivars. *Hortscience*, **44**, 469–470.

Escalettes, V., Dosba, F., Lansac, M. and Eyquard, J.P. (1998) Genetic resistance to *Plum pox* potyvirus in peaches. *Acta Hort.* **465**, 689–698.

Eyre-Walker, A. and Keightley, P.D. (2007) The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8**, 610–618.

Fay, D. and Spencer, A. (2005) Genetic mapping and manipulation: Chapter 8—dominant mutations in WormBook. In: *WormBook* (**The C. elegans Research Community**, eds.). doi: 10.1895/wormbook.1.7.1. http://www.wormbook.org [accessed on March 28, 2013].

Fernández-Fernández, M.R., Camafeita, E., Bonay, P., Méndez, E., Albar, J.P. and García, J.A. (2002) The capsid protein of a plant single-stranded RNA virus is modified by O-linked N-acetylglucosamine. *J. Biol. Chem.* **277**, 135–140.

Galli, P., Patocchi, A., Broggini, G.A.L. and Gessler, C. (2010) The *Rvi15* (*Vr2*) apple scab resistance locus contains three TIR-NBS-LRR genes. *Mol. Plant–Microbe Interact.* **23**, 608–617.

Hartmann, W. and Neumüller, M. (2006) Breeding for resistance: breeding for *Plum pox* virus resistant plums (*Prunus domestica* L.) in Germany. *EPPO Bull.* **36**, 332–336.

Hilario, E., Bennell, T.F. and Rikkerink, E. (2007) Screening a BAC library with nonradioactive overlapping oligonucleotide (overgo) probes. *Methods Mol. Biol.* **353**, 79–91.

Ion-Nagy, L., Lansac, M., Eyquard, J.P., Salvador, B., Garcia, J.A., Le Gall, O., Hernould, M., Schurdi-Levraud, V. and Decroocq, V. (2006) PPV long-distance movement is occasionally permitted in resistant apricot hosts. *Virus Res.* **120**, 70–78.

Ivanov, K.I., Puustinen, P., Gabrenaite, R., Vihinen, H., Rönnstrand, L., Valmu, L., Kalkkinen, N. and Mäkinen, K. (2003) Phosphorylation of the potyvirus capsid protein by protein kinase CK2 and its relevance for virus infection. *Plant Cell*, **15**, 2124–2139.

James, D. and Thompson, D. (2006) Hosts and symptoms of *Plum pox* virus: ornamental and wild *Prunus* species. *EPPO Bull.* **36**, 222–224.

Jung, S., Jesudurai, C., Staton, M., Du, Z., Ficklin, S., Cho, I., Abbott, A., Tomkins, J. and Main, D. (2004) GDR (genome database for Rosaceae): integrated web resources for Rosaceae genomics and genetics research. *BMC Bioinformatics*, **5**, 130.

Karayiannis, I. (1989) Susceptibility of apricot cultivars to plum pox virus in Greece. *Acta Hort.* **235**, 271–274.

Karayiannis, I. (2006) Breeding for resistance: conventional breeding for *Plum pox* virus resistant apricots (*Prunus armeniaca* L.) in Greece. *EPPO Bull.* **36**, 319–322.

Karayiannis, I., Thomidis, T. and Tsaftaris, A. (2008) Inheritance of resistance to *Plum pox virus* in apricot (*Prunus armeniaca* L.). *Tree Genet. Genomes*, **4**, 143–148.

Labonne, G. and Dallot, S. (2006) Epidemiology of sharka disease in France. *EPPO Bull.* **36**, 267–270.

Laios, A., O'Toole, S., Flavin, R., Martin, C., Kelly, L., Ring, M., Finn, S.P., Barrett, C., Loda, M., Gleeson, N., D'Arcy, T., McGuinness, E., Sheils, O., Sheppard, B. and O'Leary, J. (2008) Potential role of miR-9 and miR-223 in recurrent ovarian cancer. *Mol. Cancer*, **7**, 35.

Lalli, D.A., Abbott, A.G., Zhebentyayeva, T.N., Badenes, M.L., Damsteegt, V., Polák, J., Krška, B. and Salava, J. (2008) A genetic linkage map for an apricot (*Prunus armeniaca* L.) BC1 population mapping plum pox virus resistance. *Tree Genet. Genomes*, **4**, 481–493.

Lambert, P., Hagen, L.S., Arús, P. and Audergon, J.M. (2004) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas × peach Earlygold reference map for Prunus. *Theor. Appl. Genet.* **108**, 1120–1130.

Lambert, P., Dicenta, F., Rubio, M. and Audergon, J.M. (2007) QTL analysis of resistance to sharka disease in the apricot (*Prunus armeniaca* L.) 'Polonais' × 'Stark Early Orange' F1 progeny. *Tree Genet. Genomes*, **3**, 299–309.

Layne, R.E.C., Bailey, C.H. and Hough, L.F. (1996) Apricots. In: *Fruit Breeding. Vol. II. Tree and Tropical Fruits* (Janick J. and Moore J.N., eds), pp. 79–111. New York, NY: John Wiley & Sons.

Le Rouzic, E., Morel, M., Ayinde, D., Belaïdouni, N., Letienne, J., Transy, C. and Margottin-Goguet, F. (2008) Assembly with the Cul4A-DDB1DCAF1 ubiquitin ligase protects HIV-1 Vpr from proteasomal degradation. *J. Biol. Chem.* **283**, 21 686–21 692.

Lee, J.H., Terzaghi, W., Gusmaroli, G., Charron, J.B., Yoon, H.J., Chen, H., He, Y.J., Xiong, Y. and Deng, X.W. (2008) Characterization of *Arabidopsis* and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. *Plant Cell*, **20**, 152–167.

Llácer, G., Badenes, M.L. and Romero, C. (2007) Problems in the determination of inheritance of plum pox virus resistance in apricot. *Acta Hortic.* **781**, 263–267.

Madishetty, K., Condamine, P., Svensson, J.T., Rodriguez, E. and Close, T.J. (2007) An improved method to identify BAC clones using pooled overgos. *Nucleic Acids Res.* **35**, e5.

Mahajan, S.K., Chisholm, S.T., Whitham, S.A. and Carrington, J.C. (1998) Identification and characterization of a locus (RTM1) that restricts long-distance movement of tobacco etch virus in *Arabidopsis thaliana*. *Plant J.* **14**, 177–186.

Majewski, J. and Ott, J. (2012) Distribution and characterization of regulatory elements in the human genome. *Genome Res.* **12**, 1827–1836.

Marandel, G., Salava, J., Abbott, A., Candresse, T. and Decroocq, V. (2009) Quantitative trait loci meta-analysis of *Plum pox* virus resistance in apricot (*Prunus armeniaca* L.): new insights on the organization and the identification of genomic resistance factors. *Mol. Plant Pathol.* **10**, 347–360.

Martin, G.W., III, Harney, J.W. and Berry, M.J. (1998) Functionality of mutations at conserved nucleotides in eukaryotic SECIS elements is determined by the identity of a single nonconserved nucleotide. *RNA*, **4**, 65–73.

Martínez-Calvo, J., Llácer, G. and Badenes, M.L. (2010) Rafel and Xelva, two apricot varieties resistant to sharka. *HortScience*, **45**, 1904–1905.

Martínez-Gómez, P., Dicenta, F. and Audergon, J.M. (2000) Behaviour of apricot (*Prunus armeniaca* L.) cultivars in the presence of sharka (*Plum pox* potyvirus): a review. *Agronomie*, **20**, 407–422.

Maule, A.J., Caranta, C. and Boulton, M.I. (2007) Sources of natural resistance to plant viruses: status and prospects. *Mol. Plant Pathol.* **8**, 223–231.

Mosavi, L.K., Cammett, T.J., Desrosiers, D.C. and Peng, Z.-Y. (2004) The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci.* **13**, 1435–1448.

Nie, H., Zhao, C., Wu, G., Wu, Y., Chen, Y. and Tang, D. (2012) SR1, a calmodulin-binding transcription factor, modulates plant defense and ethylene-induced senescence by directly regulating NDR1 and EIN3. *Plant Physiol.* **158**, 1847–1859.

Nozawa, M. and Nei, M. (2007) Evolutionary dynamics of olfactory receptor genes in Drosophila species. *Proc. Natl. Acad. Sci. USA*, **104**, 7122–7127.

Olmos, A., Capote, N. and Candresse, T. (2006) Detection and characterization of *Plum pox* virus: molecular methods. *EPPO Bull.* **36**, 262–266.

Olmstead, J.W., Sebolt, A.M., Cabrera, A., Sooriyapathirana, S.S., Hammar, S., Iriarte, G., Wang, D., Chen, C.Y., van der Knaap, E. and Iezzoni, A.F. (2008) Construction of an intra-specific sweet cherry (*Prunus avium* L.) genetic linkage map and synteny analysis with the Prunus reference map. *Tree Genet. Genomes*, **4**, 897–910.

Pagny, G., Paulstephenraj, P.S., Poque, S., Sicard, O., Cosson, P., Eyquard, J.-P., Caballero, M., Chague, A., Gourdon, G., Negrel, L., Candresse, T., Mariette, S. and Decroocq, V. (2012) Family-based linkage and association mapping reveals novel genes affecting *Plum pox virus* infection in *Arabidopsis thaliana*. *New Phytol.* **196**, 873–886.

Parravicini, G., Gessler, C., Denancé, C., Lasserre-Zuber, P., Vergne, E., Brisset, M.N., Patocchi, A., Durel, C.E. and Broggini, G.A.L. (2011) Identification of serine/threonine kinase and nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes in the fire blight resistance quantitative trait locus of apple cultivar 'Evereste'. *Mol. Plant Pathol.* **12**, 493–505.

Petri, C., Webb, K., Hily, J.M., Dardick, C. and Scorza, R. (2008) High transformation efficiency in plum (*Prunus domestica* L.): a new tool for functional genomics studies in *Prunus* spp. *Mol. Breed.* **22**, 581–591.

Pilarova, P., Marandel, G., Decroocq, V., Salava, J., Krka, B. and Abbott, A.G. (2010) Quantitative trait analysis of resistance to *Plum pox* virus in the apricot F1 progeny 'Harlayne' × 'Vestar'. *Tree Genet. Genomes*, **6**, 467–475.

Qiu, Y., Xi, J., Du, L., Suttle, J.C. and Poovaiah, B.W. (2012) Coupling calcium/calmodulin-mediated signaling and herbivore-induced plant response through calmodulin-binding transcription factor AtSR1/CAMTA3. *Plant Mol. Biol.* **79**, 89–99.

Robaglia, C. and Caranta, C. (2006) Translation initiation factors: a weak link in plant RNA virus infection. *Trends Plant Sci.* **11**, 40–45.

Rozen, S. and Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols* (Misener, S. and Krawetz, S.A., eds), pp. 365–386. Totowa, NJ: Humana Press.

Rubio, M., Martinez-Gomez, P., Dicenta, F. and Audergon, J.M. (2004) Testing of genetic control hypothesis for Plum pox resistance in apricot. *Acta Hort.* **663**, 265–267.
Sato, S., Tomomori-Sato, C., Parmely, T.J., Florens, L., Zybailov, B., Swanson, S.K., Banks, C.A., Jin, J., Cai, Y., Washburn, M.P., Conaway, J.W. and Conaway, R.C. (2004) A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol. Cell*, **14**, 685–691.
Scholthof, K.-B.G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., Hemenway, C. and Foster, G. (2011) Top 10 plant viruses in molecular plant pathology. *Mol. Plant Pathol.* **12**, 938–954.
Soosaar, J.L.M., Burch-Smith, T.M. and Dinesh-Kumar, S.P. (2005) Mechanisms of plant resistance to viruses. *Nat. Rev. Microbiol.* **3**, 789–798.
Soriano, J.M., Vera-Ruiz, E.M., Vilanova, S., Martinez-Calvo, J., Llácer, G., Badenes, M.L. and Romero, C. (2008) Identification and mapping of a locus conferring *plum pox* virus resistance in two apricot improved linkage maps. *Tree Genet. Genomes*, **4**, 391–402.
Soriano, J.M., Domingo, M.L., Zurigaga, E., Romero, C., Zhebentyayeva, T., Abbott, A.G. and Badenes, M.L. (2012) Identification of SSR markers tightly associated with PPV resistance in apricot. *Mol. Breed.* **30**, 1017–1026.
Staden, R., Beal, K.F. and Bonfield, J.K. (2000) The Staden Package, 1998. Computer methods in molecular biology. In: *Bioinformatics Methods and Protocols* (Misener, S. and Krawetz, S.A., eds), pp. 115–130. Totowa, NJ: Humana Press.
Syrgiannidis, G. (1980) Selection of two apricot varieties resistant to sharka virus. *Acta Phytopathol. Sci. Hung.* **15**, 85–87.
Tradafirescu, M. and Topor, E. (1999) Investigation on the susceptibility of some apricot to plum pox (Sharka) virus. *Acta Hort.* **488**, 787–791.
Ueki, S., Spektor, R., Natale, D.M. and Citovsky, V. (2010) ANK, a host cytoplasmic receptor for the *Tobacco mosaic virus* cell-to-cell movement protein, facilitates intercellular transport through plasmodesmata. *PLoS Pathog.* **6**, e1001201.
Użarowska, A., Dionisio, G., Sarholz, B., Piepho, H.P., Xu, M., Ingvardsen, C.R., Wenzel, G. and Lübberstedt, T. (2009) Validation of candidate genes putatively associated with resistance to SCMV and MDMV in maize (*Zea mays* L.) by expression profiling. *BMC Plant Biol.* **9**, 15.
Vera-Ruiz, E.M., Soriano, J.M., Romero, C., Zhebentyayeva, T., Terol, J., Zuriaga, E., Llácer, G., Abbott, A.G. and Badenes, M.L. (2011) Narrowing down the apricot *Plum pox* virus resistance locus and comparative analysis with the peach genome syntenic region. *Mol. Plant Pathol.* **12**, 535–547.
Vilanova, S., Romero, C., Abernathy, D., Abbott, A.G., Burgos, L., Llácer, G. and Badenes, M.L. (2003) Construction and application of a bacterial artificial chromosome (BAC) library of *Prunus armeniaca* L. for the identification of clones linked to the self-incompatibility locus. *Mol. Gen. Genet.* **269**, 685–691.
Xu, Q. and Deng, X. (2010) Cloning and phylogenetic analyses of serine/threonine kinase class defense-related genes in a wild fruit crop 'chestnut rose'. *BMC Res. Notes*, **3**, 202.
Zapata, J.M., Martínez-García, V. and Lefebvre, S. (2007) Phylogeny of the TRAF/MATH domain. *Adv. Exp. Med. Biol.* **597**, 1–24.
Zhang, Y., Feng, S., Chen, F., Chen, H., Wang, J., McCall, C., Xiong, Y. and Deng, X.W. (2008) *Arabidopsis* DDB1-CUL4 ASSOCIATED FACTOR1 forms a nuclear E3 ubiquitin ligase with DDB1 and CUL4 that is involved in multiple plant developmental processes. *Plant Cell*, **20**, 1437–1455.
Zhebentyayeva, T.N., Reighard, G.L., Lalli, D., Gorina, V.M., Krška, B. and Abbott, A.G. (2008) Origin of resistance to *plum pox* virus in apricot: what new AFLP and targeted SSR data analyses tell. *Tree Genet. Genomes*, **4**, 403–417.
Zheng, X.H., Lu, F., Wang, Z.-Y., Zhong, F., Hoover, J. and Mural, R. (2005) Using shared genomic synteny and shared protein functions to enhance the identification of orthologous gene pairs. *Bioinformatics*, **21**, 703–710.
Zhou, J., Loh, Y.T., Bressan, R.A. and Martin, G.B. (1995) The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. *Cell*, **83**, 925–935.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1 Comparative analysis of the *Prunus PPVres* locus with the *Arabidopsis* genome. Peach transcripts, annotated by the International Peach Genome Initiative (IPGI), are represented by black boxes. Full lines connect *Prunus*/*Arabidopsis* gene pairs that are best-reciprocal BLASTP hits (E value cut-off < 1e-6). Broken lines correspond to the best unidirectional BLASTP hit when no best-reciprocal hit was obtained. Distances in kilobases are shown on the left of the peach physical map and in megabases on the right of the *Arabidopsis* chromosomes. Markers defining the *PPVres* locus are represented on the left of the peach map as a reference.
File S1 Configuration parameters for cleaning sequence reads with ngs_backbone pipeline (backbone.conf).
Table S1 Overgo and polymerase chain reaction (PCR) primers employed.
Table S2 Variants [single nucleotide polymorphisms (SNPs)/deletion/insertion polymorphisms (DIPs)] recorded from the alignment of the nine apricot genomes and bacterial artificial chromosome (BAC) contig sequences against the peach genome.
Table S3 Variants [single nucleotide polymorphisms (SNPs)/deletion/insertion polymorphisms (DIPs)] recorded from the alignment of the bacterial artificial chromosome (BAC) contig sequences from the R-haplotype of 'Goldrich' against the peach genome.