

## Narrowing down the apricot *Plum pox virus* resistance locus and comparative analysis with the peach genome syntenic region

ELSA MARÍA VERA RUIZ<sup>1,†</sup>, JOSÉ MIGUEL SORIANO<sup>1,†</sup>, CARLOS ROMERO<sup>1,†</sup>, TETYANA ZHEBENTYAYEVA<sup>2</sup>, JAVIER TEROL<sup>1</sup>, ELENA ZURIAGA<sup>1</sup>, GERARDO LLÁCER<sup>1</sup>, ALBERT GLENN ABBOTT<sup>2</sup> AND MARÍA LUISA BADENES<sup>1,\*</sup>

<sup>1</sup>Instituto Valenciano de Investigaciones Agrarias (IVIA), Apartado Oficial, 46113 Moncada, Valencia, Spain

<sup>2</sup>Department of Genetics and Biochemistry, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA

### SUMMARY

Sharka disease, caused by the *Plum pox virus* (PPV), is one of the main limiting factors for stone fruit crops worldwide. Only a few resistance sources have been found in apricot (*Prunus armeniaca* L.), and most studies have located a major PPV resistance locus (*PPVres*) on linkage group 1 (LG1). However, the mapping accuracy was not sufficiently reliable and *PPVres* was predicted within a low confidence interval. In this study, we have constructed two high-density simple sequence repeat (SSR) improved maps with 0.70 and 0.68 markers/cm, corresponding to LG1 of 'Lito' and 'Goldrich' PPV-resistant cultivars, respectively. Using these maps, and excluding genotype–phenotype incongruent individuals, a new binary trait locus (BTL) analysis for PPV resistance was performed, narrowing down the *PPVres* support intervals to 7.3 and 5.9 cm in 'Lito' and 'Goldrich', respectively. Subsequently, 71 overlapping oligonucleotides (overgo) probes were hybridized against an apricot bacterial artificial chromosome (BAC) library, identifying 870 single BACs from which 340 were anchored onto a map region of approximately 30–40 cm encompassing *PPVres*. Partial BAC contigs assigned to the two allelic haplotypes (resistant/susceptible) of the *PPVres* locus were built by high-information content fingerprinting (HICF). In addition, a total of 300 BAC-derived sequences were obtained, and 257 showed significant homology with the peach genome scaffold\_1 corresponding to LG1. According to the peach syntenic genome sequence, *PPVres* was predicted within a region of 2.16 Mb in which a few candidate resistance genes were identified.

### INTRODUCTION

*Plum pox virus* (PPV) is the causative agent of sharka disease, the most important disease affecting *Prunus* species. The eradication of infected foci is a very expensive and time-consuming method and is inefficient in terms of controlling the spread of the disease. Epidemiological studies (Cambra *et al.*, 2006; Labonne and Dalot, 2006) and improved detection methods at early infection stages (Olmos *et al.*, 2006) have contributed to a better management of the disease, but the best long-term solution is to grow new PPV-resistant varieties. Unfortunately, only a few resistance sources have been reported in apricot (*Prunus armeniaca* L.) (Martínez-Gómez *et al.*, 2000) and plum (*Prunus domestica* L.) (Hartmann and Neumüller, 2006), and no peach (*Prunus persica* L. Batsch) cultivar resistant to PPV has been found (Escalettes *et al.*, 1998).

Apricot breeding programmes aimed at introducing PPV resistance using resistant genitors were initiated in different countries severely affected by sharka at the beginning of the 1990s (Bassi and Audergon, 2006). However, conventional fruit breeding is burdened by difficulties inherent to the evaluation of trees, such as the long juvenile period and the high space requirement. Moreover, phenotyping for PPV resistance is based on a biological test that uses peach as graft-inoculated woody indicator (Moustafa *et al.*, 2001) and requires the analysis of a minimum of six plants per genotype, including symptom score, enzyme-linked immunosorbent assay (ELISA) and polymerase chain reaction (PCR), from two to four growing seasons. Standardization of the phenotyping for PPV resistance has also proven to be difficult because of the factors involved in symptom development, for instance, variability of PPV isolates, the physiological state of the host and the inoculation method (Llácer *et al.*, 2007). All these factors together make PPV resistance assessment the bottleneck for most breeding programmes. It is therefore of major importance to develop efficient tools for PPV resistance screening. The development of molecular markers for marker-assisted selection (MAS) appears to be a promising method for

\*Correspondence: Email: badenes\_mlu@gva.es

†These authors contributed equally to this work.

the early selection of PPV-resistant hybrids, avoiding the time-consuming phenotyping procedure that delays the release of resistant selections in the breeding programmes. Furthermore, PPV-linked markers may eventually be useful to develop a positional cloning strategy for PPV resistance gene(s) based on a physical map.

Recent studies on the segregation of different intraspecific apricot crosses have shown that PPV resistance is controlled by at least one major dominant locus (hereafter termed *PPVres*) located on the upper part of apricot linkage group 1 (LG1) (Dondini *et al.*, 2010; Lalli *et al.*, 2008; Lambert *et al.*, 2007; Marandel *et al.*, 2009a; Pilarova *et al.*, 2010; Soriano *et al.*, 2008). Nevertheless, mapping was not accurate in any case and map positions predicted for *PPVres* varied from one study to another. Moreover, other minor loci have been suggested to underlie PPV resistance in apricot (Lambert *et al.*, 2007; Pilarova *et al.*, 2010). The scenario becomes more complex when PPV resistance derived from the *Prunus davidiana* P1908 clone is analysed. In this case, interspecific crosses with *P. persica* suggest that a discrete number of loci scattered across the genome contribute to control the trait, including a quantitative trait locus (QTL) located in the upper part of LG1 (Marandel *et al.*, 2009b; Rubio *et al.*, 2010). As a whole, all evidence indicates that *PPVres* should be located in this region.

In this study, we have developed high-density simple sequence repeat (SSR) linkage maps for the putative region comprising *PPVres* in apricot. In addition, overgo probes designed from SSR clone sequences have been used to hybridize an apricot bacterial artificial chromosome (BAC) library (Vilanova *et al.*, 2003), and new SSRs found in BAC-derived sequences (BDSs) have been mapped. SSRs allowed us to narrow down the *PPVres* locus in apricot, and positive BACs were anchored onto a region encompassing the *PPVres* locus. BDSs were also used for homology searches against the peach genome sequence (peach v1.0) recently released by the International Peach Genome Initiative (IPGI): <http://www.rosaceae.org/peach/genome>. Synteny with peach has been studied and the comparative analysis has facilitated the identification of putative candidate resistance genes from the peach genome.

## RESULTS

### SSR marker development

As a first step to improving the mapping accuracy of the *PPVres* locus, we developed high-density maps based on SSR markers starting from the 'L × L-98' and 'G × C' genetic maps obtained previously by Soriano *et al.* (2008).

To increase the quality and resolution of these maps, different tasks were accomplished. Missing data and scoring errors detected in the original SSR markers were corrected and SSRs

available from the literature were incorporated (Table S1, see Supporting Information). The resulting maps (data not shown) were used to select markers encompassing the *PPVres* locus in order to develop single-copy overgo probes. These and additional probes (see Experimental Procedures) were hybridized against a BAC library of the apricot resistant parent 'Goldrich'. One hundred and two primer pairs flanking microsatellite repeat motifs (Gol-serie) were designed from the obtained BAC end sequences (BESs) (Table S2, see Supporting Information). The newly developed SSRs were subsequently tested in three segregating populations ('L × L-98', 'G × C' and 'G × Ca'). For the 'L × L-98' family, 45 SSRs were monomorphic, 21 did not amplify or produced complex patterns and 36 were polymorphic, 21 of which were mapped (Fig. 1). With regard to 'G × C' and 'G × Ca', 52 were monomorphic for 'Goldrich', 11 did not amplify or produced complex patterns and 39 were polymorphic (heterozygote SSRs in 'Goldrich'), 17 of which were mapped (Fig. 1).

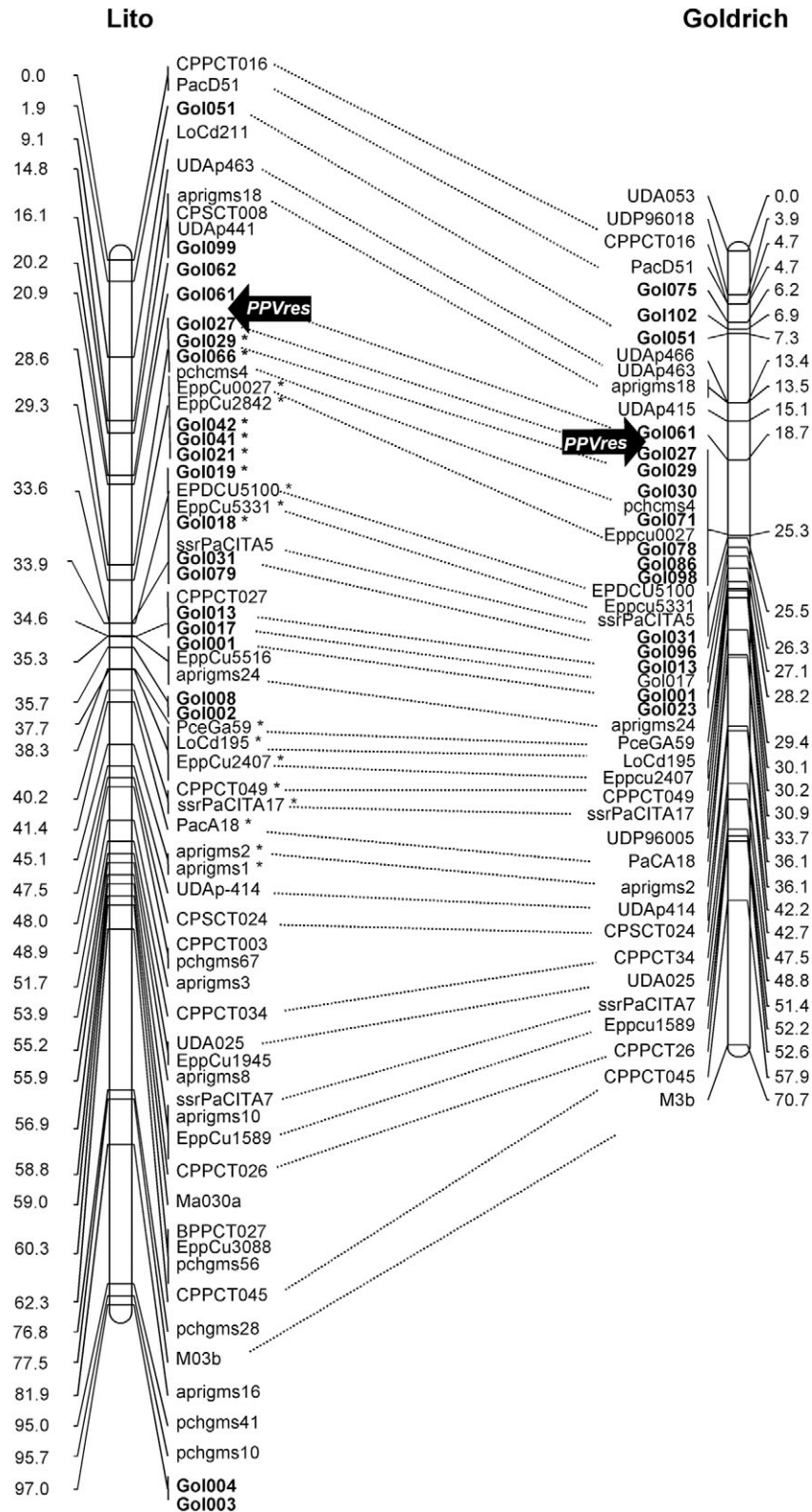
### Linkage maps

In the 'Lito' map (derived from the 'L × L-98' population), 50 of the total 187 SSRs tested were mapped into LG1, leading to a genetic map of 97.0 cm in length containing 68 SSRs (Fig. 1) and showing a marker density of 0.70 marker/cm. In 'Goldrich', 36 of the 187 SSRs were mapped into LG1, leading to a genetic map of 70.7 cm in length containing 48 SSRs (Fig. 1) and showing a marker density of 0.68 marker/cm. The construction of the 'Goldrich' map was based on individuals from two cross-pollination populations ('G × C' and 'G × Ca'). In the integrated 'Goldrich' map, 11 SSRs derived from the 'G × C' map, 15 originated from the 'G × Ca' map and 22 were shared in common by both maps (Fig. 1). Moreover, the 34 SSRs shared by 'Lito' and 'Goldrich' maps were shown to be completely collinear.

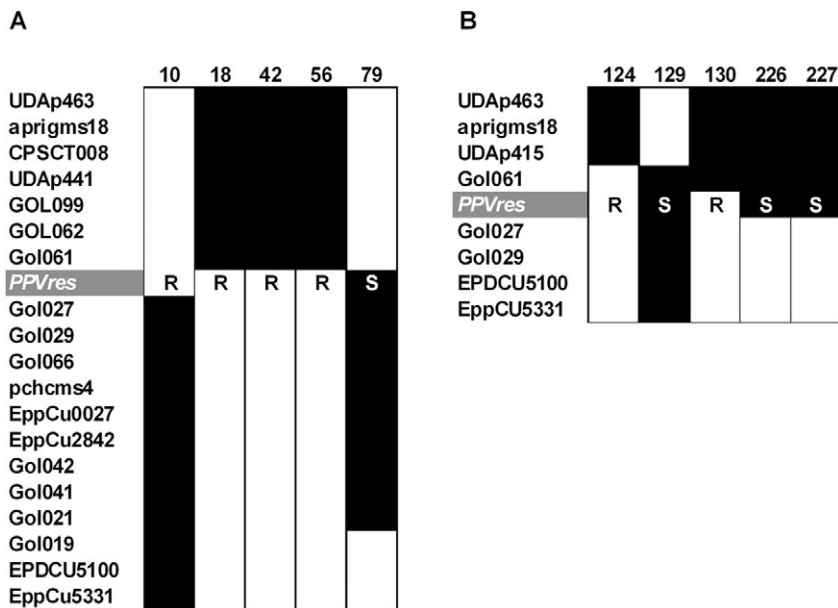
Quality improvement of the original maps was checked using different parameters. First, the marker density was increased by 10% in 'Lito' and doubled in 'Goldrich'. Second, the mean chi-squared contribution (as a measure of the goodness of fit for each mapped marker defined by JoinMap 3.0) decreased on average from 1.5 to 0.2 in 'Lito' and from 3.2 to 0.1 in 'Goldrich'. Furthermore, graphical ordering of genotype data from both maps enabled us to check the reasonable distribution of recombination breakpoints over the estimated map (Fig. 2).

### Association between SSRs and PPV resistance

The phenotype of PPV resistance of the parents and progenies used in this study was based on a bioassay using peach GF-305 as woody indicator and PPV Dideron strain 3.3 RB (Moustafa *et al.*, 2001). In this study, the scoring of PPV resistance phenotypes and marker genotypes within the binary trait locus (BTL)-associated genomic region defined by Soriano *et al.* (2008) was



**Fig. 1** High-density simple sequence repeat (SSR) maps of 'Lito' and 'Goldrich' LG1. The 'Goldrich' map resulted from the integration of 'G × C' and 'G × Ca' maps. Markers in bold came from 'G × C', those in green from 'G × Ca' and markers in black are common in both progenies. Distances in centimorgan (cM) are shown on the left in 'Lito' and on the right in 'Goldrich'. New SSRs are in bold, and asterisks indicate markers with distorted segregations at  $P < 0.01$ .



**Fig. 2** Graphical genotyping of the recombinant hybrids at the *PPVres* locus. (A) 'L × L-98' mapping population. (B) 'G × C' mapping population. The graphics show the region between markers UDAp463 and EppCu5331. Black vertical bars represent susceptible (S) and white bars resistant (R) chromosomal regions. Numbers at the top corresponded to the seedlings identified as recombinant hybrids in each population.

performed. As a result, 19 hybrids from the 'L × L-98' population showed genotype–phenotype incongruence (GPI) according to Gygax *et al.* (2004). These plants were classified as PPV susceptible, but genotyped as homozygotes or heterozygotes for the marker allele considered in coupling with PPV resistance originating from the resistant grandparent 'Stark Early Orange' ('SEO'). GPI was also found in 18 hybrids from the 'G × C' population and, in contrast with most cases, three were classified as PPV resistant, having their marker alleles in coupling with susceptibility. Susceptible or resistant hybrids classified as GPI had no recombination within the PPV resistance BTL-associated region. For accurate mapping, the phenotype of these plants was excluded from subsequent analysis, as reported by Patocchi *et al.* (1999).

Co-segregation analysis between PPV resistance and markers from the 'Goldrich' and 'Lito' LG1 new maps showed that only a few markers located on the upper part of the chromosome had a recombination frequency below 0.1 and a logarithm of the odds (LOD) > 5.0 (Table 1). A Kruskal–Wallis (KW) test was carried out to find associations between SSRs and PPV resistance in the new LG1 maps. In agreement with the co-segregation analysis, those markers shown to be strongly linked to PPV resistance exhibited the highest KW statistical values (Table 1). To confirm the KW test results, interval mapping (IM) was also performed in spite of the constraint caused by the binary phenotype distribution. IM results supported the detection of one BTL on both LG1 maps. In the 'Goldrich' map, the maximum LOD score (20.3) matched with the GoI061 marker, and the two-LOD support interval for the *PPVres* BTL was approximately 5.9 cm (Fig. 3A). In the 'Lito' LG1 map, a two-LOD support interval of approximately 7.3 cm was defined for this BTL, around the

maximum LOD value of 8.9 (Fig. 3B). Figure 3 also shows differences between KW and IM curves excluding and including GPI plants. When GPI hybrids were excluded, the maximum KW score values along the LG1 maps increased from 20.6 and 19.3 to 43.4 (GoI027) and 44.4 (GoI061) in 'Lito' and 'Goldrich', respectively (Table 1). The two-LOD support interval for the PPV resistance BTL decreased from 16.0 to 5.9 cm when GPI plants were excluded in the 'Goldrich' LG1 map, and could not even be defined in 'Lito' if GPI plants were not excluded (LOD < 3.0) (Fig. 3). Moreover, the maximum LOD score increased approximately three- to four-fold in both maps after the exclusion of GPI plants, whereas the peaks of the LOD scores did not shift significantly left or right, indicating that the *PPVres* core position was unchanged.

### Primary BAC contigs

A set of primary BAC contigs surrounding the *PPVres* locus was anchored onto the apricot LG1 by hybridization-based screening of an apricot BAC library (Vilanova *et al.*, 2003) with 71 overgo probes. Twelve of these probes gave weak or no hybridization signals and were discarded from the study. The remaining 59 probes identified a total of 870 single positive BACs but, after removal of those probably related to duplicated sequences (numerous groups) and those not confirmed by PCR, only 340 could be assigned unambiguously to the set of overgo probes used. A set of BACs identified with the same probe or different probes sharing at least one BAC clone or one marker made up a primary contig when confirmed by PCR (Table 2). SSRs linked to the first-round overgo probes and new ones derived from BESSs were used to confirm primary contigs by PCR (Tables S1, S2 and



**Table 1** Results of co-segregation, Kruskal–Wallis (KW) and interval mapping (IM) analysis for *Plum pox virus* (PPV) resistance on 'Lito' and 'Goldrich' maps excluding genotype–phenotype incongruent plants.

| Map        | Marker    | Pos* | Co-segregation |       |       | KW test        |       | Interval mapping            |             |  |
|------------|-----------|------|----------------|-------|-------|----------------|-------|-----------------------------|-------------|--|
|            |           |      | <i>r</i> †     | LOD‡  | KW§   | <i>P</i> (KW)¶ | IM**  | <i>R</i> <sup>2</sup> (%)†† | <i>a</i> ‡‡ |  |
| Lito       | UDAp463   | 14.8 | 0.08           | 6.20  | 27.12 | <0.0001        | 5.78  | 33.6                        | 0.379       |  |
|            | aprigms18 | 16.1 | 0.08           | 6.22  | 27.21 | <0.0001        | 5.85  | 34.0                        | 0.374       |  |
|            | UDAp441   | 16.1 | 0.06           | 7.46  | 31.14 | <0.0001        | 6.62  | 37.5                        | 0.392       |  |
|            | Gol 062   | 20.2 | 0.07           | 7.51  | 31.43 | <0.0001        | 6.93  | 38.3                        | 0.375       |  |
|            | Gol 061   | 20.9 | 0.07           | 7.53  | 31.51 | <0.0001        | 7.04  | 39.3                        | 0.372       |  |
|            | Gol 027   | 28.7 | 0.01           | 11.05 | 43.39 | <0.0001        | 6.77  | 38.1                        | 0.357       |  |
|            | EPPCU0027 | 29.3 | 0.03           | 9.21  | 38.42 | <0.0001        | 6.20  | 35.5                        | 0.348       |  |
|            | Gol 019   | 33.6 | 0.05           | 7.90  | 33.79 | <0.0001        | 5.22  | 30.9                        | 0.338       |  |
| Goldrich§§ | EPPCU5331 | 34.0 | 0.07           | 6.89  | 29.46 | <0.0001        | 4.74  | 28.5                        | 0.328       |  |
|            | UDAp463   | 28.1 | 0.11           | 6.85  | 32.04 | <0.0001        | 9.00  | 47.8                        |             |  |
|            | aprigms18 | 28.1 | 0.09           | 8.08  | 35.99 | <0.0001        | 12.23 | 59.4                        |             |  |
|            | UDAp415   | 31.3 | 0.05           | 9.73  | 43.03 | <0.0001        | 16.95 | 78.3                        |             |  |
|            | Gol 061   | 34.0 | 0.05           | 10.25 | 44.37 | <0.0001        | 20.26 | 78.6                        |             |  |
|            | Gol 027   | 37.2 | 0.07           | 9.70  | 41.97 | <0.0001        | 15.25 | 66.6                        |             |  |
|            | EPPCU5331 | 43.9 | 0.17           | 6.22  | 26.31 | <0.0001        | 7.51  | 41.8                        |             |  |

\*Position in centimorgan (cM) on linkage group 1 (LG1).

†Recombination frequency between markers and PPV resistance trait estimated by JoinMap 3.0.

‡Logarithm of odds (LOD) score for co-segregations.

§Kruskal–Wallis test statistical values.

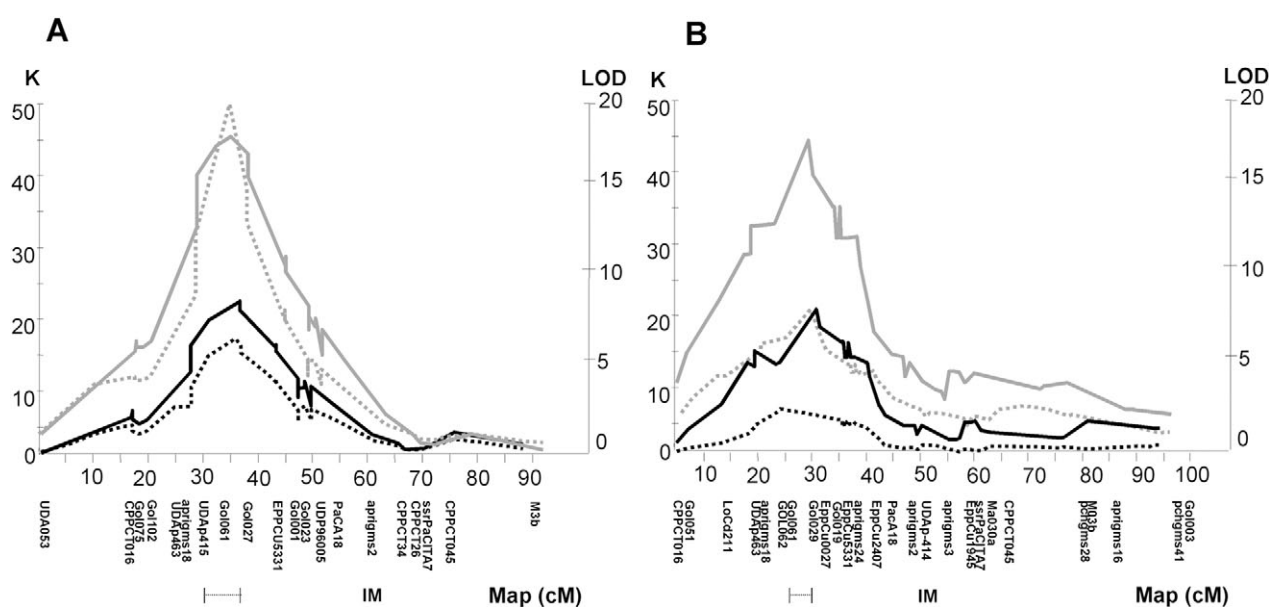
¶Probability associated with the KW value.

\*\*LOD score under IM.

††Percentage of the contribution to the total variance.

‡‡Additive effect.

§§Statistics were only calculated from the 'G × C' population.



**Fig. 3** Kruskal–Wallis (KW) statistical values (full line) and interval mapping (IM) logarithm of odds (LOD) score (dotted line) at markers on linkage group 1 (LG1). (A) 'Goldrich' map. (B) 'Lito' map. Grey lines correspond to the analyses excluding and black lines including genotype–phenotype incongruent plants. Bars at the bottom indicate the IM two-LOD support interval excluding genotype–phenotype incongruent plants.

2). BACs identified using mapped SSR-associated probes were directly anchored onto the apricot genetic map. BACs identified using unmapped SSRs or peach BESs could be anchored in some cases, mapping new SSRs developed from BESs such as those

associated with EPDCU3122 (Gol051, Gol075 and Gol102), UDAp444 (Gol099), M16a (Gol061 and Gol062), AG51ssr (Gol027, Gol029, Gol030, Gol066 and Gol071), EPPB4232 (Gol021) and AG116 (Gol086) (Tables S2 and 2).

**Table 2** Bacterial artificial chromosomes (BACs) and primary contigs anchored onto the apricot 'Lito' genetic map.

| Primary contigs | Overgo probes*          | No. of hybridization-positive BACs | No. of BDSs | No. of anchored BACs confirmed by PCR§ |
|-----------------|-------------------------|------------------------------------|-------------|--|
| 1               | EPDCU3122†              | 18                                 | 36          | 20                                     |
|                 | EPPCU3062†              | 2                                  |             |  |
| 2               | UDAp463                 | 32                                 | 0           | 18                                     |
| 3               | aprigms18               | 16                                 | 0           | 15                                     |
| 4               | CPST008                 | 4                                  | 10          | 22                                     |
|                 | UDAp444†                | 9                                  |             |  |
|                 | UDAp441                 | 17                                 |             |  |
| 5               | M16a†                   | 26                                 | 10          | 10                                     |
|                 | P101L17_T7 (MDL1)†      | 10                                 |             |  |
|                 | CPPCT10                 | 16                                 |             |  |
| 6               | AG51ssr† (2)‡           | 28                                 | 51          | 19                                     |
| 7               | EPPCU0027 (2)           | 75                                 | 46          | 73                                     |
|                 | EPPCU2842               | 4                                  |             |  |
|                 | P003L11 (pchcms4)       | 11                                 |             |  |
|                 | EPPB4232† (3)           | 43                                 |             |  |
|                 | N089B04_T7†             | 13                                 |             |  |
| 8               | N012N20_T7† (AG116) (1) | 19                                 | 23          | 14                                     |
|                 | N094009_SP6†            | 8                                  |             |  |
| 9               | UDAp435†                | 54                                 | 23          | 30                                     |
|                 | EPDCU5100 (4)           | 52                                 |             |  |
|                 | EPPCU5331               | 10                                 |             |  |
|                 | N002_B09†               | 18                                 |             |  |
|                 | N012D6_SP6†             | 10                                 |             |  |
|                 | N036C13_T7†             | 46                                 |             |  |
| 10              | ssrPaCITA5 (5)          | 30                                 | 39          | 32                                     |
|                 | L012N14_T7†             | 13                                 |             |  |
|                 | CPPCT027 (1)            | 22                                 |             |  |
| 11              | EPPCU5516               | 2                                  | 62          | 52                                     |
|                 | aprigms24 (6)           | 203                                |             |  |
|                 | N021110_SP6†            | 1                                  |             |  |
|                 | EPPCU2407               | 3                                  |             |  |
| 12              | Cd83N12                 | 15                                 | 0           | 27                                     |
|                 | L009D02_T7†             | 10                                 |             |  |
|                 | UDP96-005               | 22                                 |             |  |
| 13              | UDP96-005               | 22                                 | 0           | 0                                      |
| 14              | PacA18                  | 8                                  | 0           | 8                                      |

The numbers of hybridization-positive BACs and BAC-derived sequences (BDSs) obtained from each group of overgo probes are also indicated.

\*Groups of overgo probes were named according to linked markers previously mapped in *Prunus* spp. and hybridized in the first round.

†Markers or probes indirectly mapped in the apricot maps.

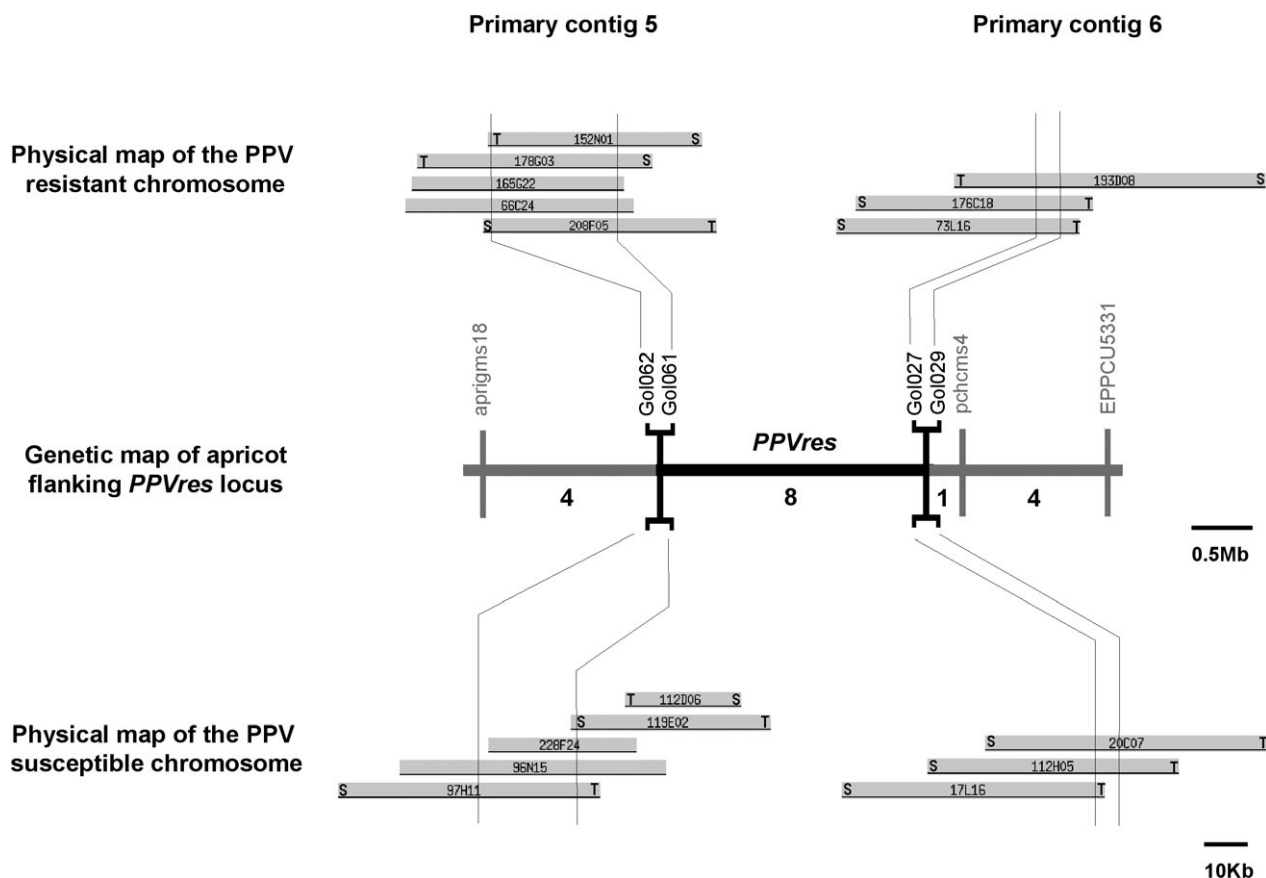
‡The numbers of new overgo probes developed from BAC end sequences (BESs) are shown in parentheses.

§Polymerase chain reactions (PCRs) were performed with simple sequence repeat (SSRs) linked to overgo probes and SSRs developed from BDSs.

As a result, 14 primary contigs confirmed by PCR were found in the upper part of LG1, covering a region of approximately 30 cm in 'Goldrich' and approximately 40 cm in 'Lito' from EPDCU3122 (Gol051–Gol075) to PacA18 (Table 2 and Fig. 1). These contigs comprised a total of 340 anchored BAC clones ranging from eight at the primary contig 13 (PacA18) to 73 at the primary contig 7 (EPPCU0027). PCRs were also performed with the aim to assemble the different primary contigs, but no overlapping was found. BACs anchored onto the *PPVres* locus (Gol061–Gol027) were firstly assigned by PCR screening to their corresponding haplotype (resistant vs. susceptible) within the heterozygous genome of the *P. armeniaca*-resistant cultivars. These BACs were then assembled into contigs by

high-information content fingerprinting (HICF) with a cut-off value of  $1E-20$ . The contigs obtained covered two regions (of approximately 70–120 kb) flanking a gap with an estimated size according to the peach genome sequence of 2.16 Mb (Fig. 4).

Hybridization results identified putative duplications or multiple-copy loci within the apricot LG1. For instance, CPPCT10 (unmapped) seems to have three copies in different LG1 primary contigs (5, 7 and 10). Similarly, aprigms24-positive hybridizations were found in the LG1 primary contigs 11 (map position) and 1 (LG1 top). In addition, two SSRs developed from the primary contig 11 (aprigms24) mapped at the bottom of LG1 (Gol003 and Gol004), also suggesting possible duplications.



**Fig. 4** Bacterial artificial chromosome (BAC) contigs within the *PPVres* locus region anchored to the apricot 'Lito' genetic map. Markers flanking the *PPVres* locus in the 'Lito' map are shown in grey and numbers correspond to the distances measured in centimorgan (cM). Scale bars for the apricot BAC contigs (10 kb) and for LG1 (0.5 Mb) are shown on the right. BAC ends from which markers were developed are indicated as T (T7)/S (SP6).

### Syntenic analysis

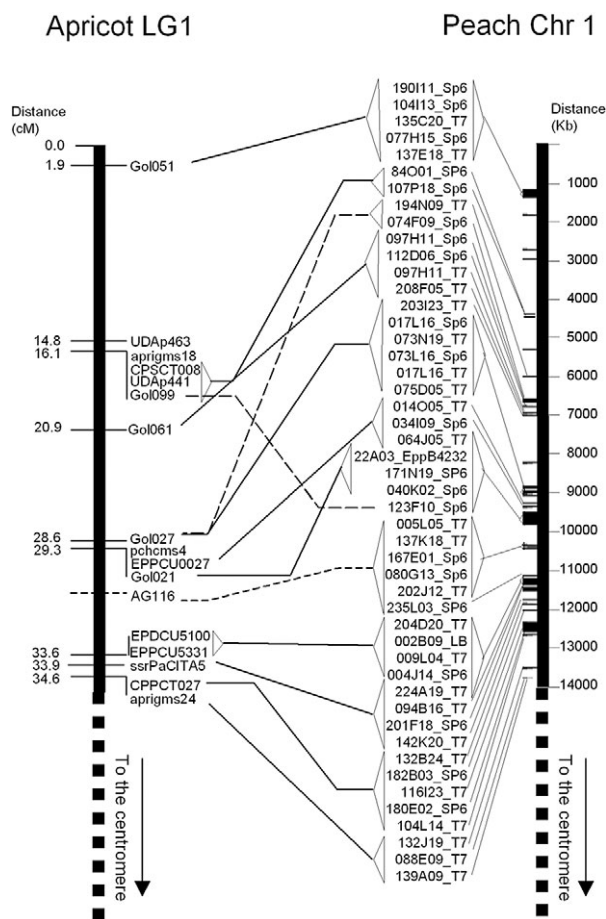
A total of 135 BAC clones, of the 340 BACs identified with single-copy overgo probes and anchored onto the region surrounding the *PPVres* locus, were partially sequenced, obtaining 270 BAC end reads and 30 BAC internal sequences (together termed 'BDS' for 'BAC-derived sequence'). Eighty-six BDSs were assembled into 37 sequence contigs, ranging from two to four sequences each, and 214 remained as singletons, leading to a total of 251 single query sequences.

The nonrepetitive fraction of the apricot BDSs was used in a BLASTN search (Altschul *et al.*, 1990) against the complete nucleotide genome sequence of *P. persica* (IPGI), with a cut-off value of  $1E-14$ . The genomic sequences were displayed with chromosomes as single searchable FASTA sequences. In order to map the BDSs unambiguously on the heterologous complete genomes, only those sequences producing single significant hits were taken into account. BLASTN analysis revealed that 257 of the 300 obtained BDSs (35 contigs and 177 singletons) showed significant homology with the peach genome scaffold\_1

sequence corresponding to LG1 (Table S3, see Supporting Information, and Fig. 5). The rest matched other scaffolds or were removed because of low sequence quality. All showed identity over 80% and E values below  $1E-10$ , and the high-scoring pair (HSP) length was, on average, 60% of the query sequence length (Table S3). Following the approach of Lai *et al.* (2006), we used forward and reverse BDS read pairs, separated by the approximate length of BAC clone inserts, to analyse the synteny between *P. armeniaca* and *P. persica*. To be considered as potentially collinear with the target genome, the apricot mate pairs had to map in the heterologous genome into a region comprising between 10 and 300 kb, and also be oriented properly. The analyses of the sequences identified 112 apricot BAC end pairs that met these criteria in peach. Furthermore, the majority of these BDS pairs mapped on the peach genome at a distance similar to the insert size of the *P. armeniaca* library (64 kb on average), suggesting a high level of synteny between apricot and peach. As a whole, the apricot BDSs mapped on the peach genome were shown to be collinear with the apricot genetic map obtained (Fig. 5). Furthermore, the BAC contigs produced by the

HICF analysis still in progress show the same organization as that inferred from the peach genome mapping, supporting the high degree of conservation between the two genomes.

The high level of synteny observed between *P. persica* and *P. armeniaca* allowed us to analyse the genomic landscape of the



**Fig. 5** Graphical representation of some apricot BAC-derived sequence (BDS) hits on the peach linkage group 1 (LG1) physical map (scaffold\_1). Genetic distances in centimorgan (cM) are shown on the left of the apricot 'Lito' LG1 map. Physical distances are shown in kb on the right of the peach scaffold\_1. Broken grey lines correspond to noncollinear BDSs and the broken black line to an unmapped marker (AG116).

**Table 3** Candidate resistance genes in the peach genome scaffold\_1 sequence located in the region comprising the markers Gol061 and Gol027.

| Peach gene ID | Transcript start | Transcript stop | Organism                       | Description                                       | E value | Percentage identity |
|---------------|------------------|-----------------|--------------------------------|---|---------|---------------------|
| ppa015992m    | 6855590          | 6858994         | <i>Arabidopsis thaliana</i>    | LRR receptor-like serine/threonine protein kinase | 4E-63   | 44.52               |
| ppa015042m    | 6860841          | 6863059         | <i>A. thaliana</i>             | LRR receptor-like serine/threonine protein kinase | 2E-73   | 45.14               |
| ppa017831m    | 7535331          | 7539680         | <i>A. thaliana</i>             | LRR receptor-like serine/threonine protein kinase | 3E-60   | 29.19               |
| ppa003371m    | 7570354          | 7574507         | <i>A. thaliana</i>             | LRR receptor-like serine/threonine protein kinase | 0       | 75.04               |
| ppa004549m    | 8078825          | 8081487         | <i>A. thaliana</i>             | Probable receptor-like protein kinase             | 0       | 64.4                |
| ppa007758m    | 8082833          | 8086119         | <i>Lycopersicon esculentum</i> | Pto-interacting protein 1                         | 9E-174  | 87.36               |

LRR, leucine-rich repeat.

peach region homologous to the apricot *PPVres* locus, in order to find putative candidate genes responsible for the resistance to PPV. The genomic region delimited by the apricot markers comprised 2.16 Mb, containing a total of 251 transcription units, as annotated by IPGI. We performed a further characterization of the open reading frames (ORFs) and their predicted proteins using BLASTX. We also obtained the gene ontology (GO) terms associated with each gene using BLAST2GO (Conesa *et al.*, 2005). These analyses showed that six peach transcription units were significantly homologous to resistance (*R*) genes already characterized in other species (Table 3). However, no significant homologies with disease resistance protein sequences deposited in the GenBank database were found after BLASTX analysis (Altschul *et al.*, 1990) for those apricot BDSs located within the *PPVres* locus supporting interval.

## DISCUSSION

### PPV resistance mapping

The genetic control underlying apricot PPV resistance has been under discussion for some time, and most recent studies point to a major locus (*PPVres*) located on LG1 as the dominant factor (Marandel *et al.*, 2009a). Unfortunately, *PPVres* mapping accuracy is not sufficiently reliable for map-based cloning approaches, and new improvements are required to accomplish this task. In this regard, we focused on two points: to increase the quality and resolution of the available maps and to devise an alternative mapping strategy for PPV resistance.

Starting from the apricot LG1 'L × L-98' and 'Goldrich' linkage maps developed by Soriano *et al.* (2008), we first increased significantly the marker density (especially in 'Goldrich') by adding new SSRs. Most were incorporated within the wide interval defined by *aprigms18* and *ssrPaCITA17*, where *PPVres* was proposed to be located according to Soriano *et al.* (2008). In this region, the marker density doubled that of the whole LG1 map, reaching 1.5 and 1.3 markers/cm in 'Lito' and 'Goldrich', respectively. Only reliable SSR markers were used for map construction, and the goodness-of-fit measurement, as well as the complete collinearity found between both maps, gave evidence of the high quality of the new maps.



With regard to the second point, PPV resistance fine mapping is known to be hindered for different reasons, such as the strong environmental dependence of PPV resistance scoring (Decroocq *et al.*, 2005) and the difficulties in carrying out large-scale experiments with woody species (Llácer *et al.*, 2007). In our particular case, PPV resistance was scored as a binary trait (resistance vs. susceptibility), as reported by Soriano *et al.* (2008), because intermediate phenotypes of PPV susceptibility, such as those described by Decroocq *et al.* (2005) for *P. davidiana*, have not yet been reliably defined in *P. armeniaca*. However, discrepancies between scored phenotype and genotypes for the *PPVres* locus surrounding markers were found in some individuals, termed 'genotype–phenotype incongruence' (GPI) according to Gygax *et al.* (2004). As molecular marker data are more trustworthy than phenotyping data for PPV resistance, we removed phenotyping data from GPI plants following the strategy successfully used by Patocchi *et al.* (1999) to map the apple scab (*Venturia inaequalis*) resistance gene *Vf*. The usefulness of this procedure was further confirmed by several studies mapping other apple scab resistance genes, such as *Vbj* (Gygax *et al.*, 2004), *Vb* (Erdin *et al.*, 2006) and *Vd3* (Soriano *et al.*, 2009). GPI plants were explained by Patocchi *et al.* (1999) as the result of double recombination events or incorrect classification. The expected frequency of double recombinants between flanking markers in coupling with the *PPVres* locus (Gol061 and Gol027) is much less than the observed frequency of GPI (data not shown). Therefore, incorrect classification seems to be a more reasonable hypothesis in this case. Resistant seedlings might have been misclassified as susceptible for different reasons, such as the latent PPV resistance already documented in apricot (Karayiannis, 2006) or the presence of modifier factors affecting the major resistance gene (Erdin *et al.*, 2006). Moreover, susceptible seedlings misclassified as resistant might be caused by PPV inoculation escapes or the accumulation of PPV resistance-related minor QTLs.

PPV resistance mapping reported by Soriano *et al.* (2008) revealed the presence of a putative single BTL in the upper part of LG1. In this work, we performed the same analyses, but using the new high-density maps and excluding GPI plants. KW tests and IM confirmed the presence of a single BTL in apricot LG1. The IM two-LOD support intervals defined in both maps were smaller than those obtained previously by Soriano *et al.* (2008), and the maximum LOD and KW statistical values, and consequently the detection accuracy, increased dramatically. Furthermore, according to the maximum values detected, it can be concluded that the BTL position shifted slightly upwards in LG1 after these new estimations. Together, BTL mapping and graphical genotyping suggest that *PPVres* is located within the interval defined by Gol027 and Gol061 markers. This position is roughly consistent with those previously suggested in other *P. armeniaca* genetic backgrounds (Marandel *et al.*, 2009a; Pilarova *et al.*,

2010) and in *P. davidiana* (Marandel *et al.*, 2009b; Rubio *et al.*, 2010), and particularly coincident with those proposed by Lambert *et al.* (2007) and Dondini *et al.* (2010). Following mapping, apricot BAC clones identified by overgo probe hybridization were anchored onto the genomic region comprising the *PPVres* locus.

### BAC anchoring to the genetic map encompassing the *PPVres* locus

In *Prunus*, a physical map-based cloning strategy has already been used to isolate the *Ma* gene for root-knot nematode resistance from Myrobalan plum (Claverie *et al.*, 2004). Similarly, in this work, we have initiated the first steps towards map-based cloning of the *PPVres* gene(s) in apricot. New recently released tools will facilitate this task: the peach physical map (Zhebentyayeva *et al.*, 2008) anchored to the *Prunus* reference map (Aranzana *et al.*, 2003) and the peach genome sequence already available (IPGI): <http://www.rosaceae.org/peach/genome>.

As a first milestone, an apricot BAC library (Vilanova *et al.*, 2003) was hybridized using overgo probes to identify BAC clones located in the region comprising the *PPVres* locus. As a rough average, 14.7 clones per probe were detected; however, once unconfirmed BACs had been rejected, only 340 BACs could be assigned to 59 overgo probes, averaging 5.8 clones per probe. Some of the removed BACs were probably detected by unspecific hybridization of repetitive sequences (i.e. Ap139A09\_T7, Ap211002\_SP6 and UDAp435 probes detected 170, 57 and 54 BACs, respectively). The BAC library coverage was predicted to be 22 haploid genome equivalents, but the first observed coverage after screening with restriction fragment length polymorphism (RFLP) probes was only around eight genome equivalents (Vilanova *et al.*, 2003). Our results are consistent with these observations. The total of 340 BACs was distributed into 14 primary contigs partially covering a wide region of approximately 30–40 cm on the upper part of LG1 (from Gol051–Gol075 to PacA18).

Several putative duplications along LG1 were found by the analysis of hybridization results. In agreement with this finding, Lambert *et al.* (2004) observed that 33% of the RFLPs analysed in the apricot population 'Polonais' × 'SEO' detected two or more loci, suggesting the presence of genomic duplicated regions. Comparative mapping across *Prunus* also revealed divergences that have been attributed to multilocus RFLP or SSR markers (Dirlewanger *et al.*, 2004). More recently, Zhebentyayeva *et al.* (2008) developed a framework physical map for peach, finding that 683 of 2636 markers hybridized to multiple contigs. Duplication of genomic regions has been suggested, among other reasons, to explain this result.

BAC clones anchored onto the *PPVres* locus were assembled into contigs by HICF. *Prunus* heterozygosity led to the separation

of allelic BACs into two different contigs (susceptible and resistant) by PCR screening. These contigs covered two regions with an estimated size of approximately 70–120 kb, flanking an uncovered gap whose size, inferred from comparative genomics with the peach genome, was around 2.16 Mb. This physical distance corresponds to 7.7 and 4.4 cm in the 'Lito' and 'Goldrich' genetic maps, respectively (from Gol061 to Gol027). According to the estimated sizes of the peach genome (~290 Mb) and the *Prunus* general map (519 cm) (Zhebentyayeva *et al.*, 2008), the relationship between the physical and genetic distances is close to 0.56 Mb/cm on average. The ratios estimated for the apricot *PPVres* locus were lower than this (0.27 and 0.48 Mb/cm in 'Lito' and 'Goldrich', respectively), but this disagreement is probably a result of differences in the saturation of the genetic maps.

### Syntenic analysis

Apricot BDSs were obtained from nine of the 14 primary contigs anchored to the genetic maps. Most query sequence contigs and singletons (>66%) were shown to be collinear when compared with peach scaffold\_1, according to the marker order established in the apricot genetic maps. Together, apricot collinear BDSs matched peach sequences distributed within a region of approximately 12.7 Mb (located between ~1.2 and 13.9 Mb starting from the top) of the total 46.9 Mb of scaffold\_1. However, a significant part of the BDSs matched noncollinear positions within this interval (9%) or downward on the chromosome (25%). Moreover, a significant number of hits were found in other scaffolds. These results support the presence of duplicated regions previously suggested by the hybridization experiments. However, other factors already reported might contribute to explain multiple hits (BDSs containing repetitive sequences, transposons or genes belonging to multigene families) or non-collinear matches (BDSs derived from false-positive BACs).

Apricot primary contigs partially covered peach sequence fragments varying in size from approximately 0.03 Mb for contig 4 (CPST008) to 1.61 Mb for contig 11 (CPPCT027). Gaps between primary contigs also varied in size, from the shortest value of 0.07 Mb between contigs 8 (AG116) and 9 (EPDCU5331) to the longest value of 2.16 Mb between contigs 5 (M16) and 6 (AG51) comprising the *PPVres* locus. The sizes of these gaps, inferred from the peach scaffold\_1 sequence, explain why no overlapping was detected between primary contigs by PCR. Soriano *et al.* (2008) predicted that the *PPVres* locus was located within a region surrounding *ssrPaCITA5*, approximately between the primary contigs 8 (AG116) and 11 (CPPCT027). Therefore, overgo probes were designed to cover mainly this area, and all gaps between contigs 7 and 10 were shorter than 0.5 Mb. However, new mapping performed in this work revealed that the most probable position was slightly upwards between contigs 5 (M16) and 6 (AG51), where coverage was less dense.

Additional efforts will be required to develop a complete physical map covering the *PPVres* locus.

Homology searches, carried out against the GenBank database, revealed no homologies with disease resistance genes for the BDSs located within the *PPVres* locus confidence interval (Gol061–Gol027). Meanwhile, a rough analysis of the corresponding peach sequence has revealed the presence of genes coding for R proteins belonging to different classes, such as leucine-rich repeat (LRR) receptor-like and pto-like serine/threonine protein kinases (Martin *et al.*, 2003), which might be considered as potential candidates for PPV resistance gene(s). However, it is important to take into account that no PPV resistance sources have been found to date in peach (Escalettes *et al.*, 1998) and, consequently, the peach *PPVres* locus syntenic region is predicted to lack the major PPV resistance gene or to contain a nonfunctional version of this gene. Therefore, further work will be necessary to analyse in depth the possible involvement of the apricot putative orthologous genes in PPV resistance.

## EXPERIMENTAL PROCEDURES

### Plant material

Three populations segregating for PPV resistance were used for mapping. An F<sub>2</sub> derived from selfing of the PPV-resistant cultivar 'Lito' ('L × L-98') ( $n = 81$ ), originally obtained from the cross 'SEO × Tyrinthos', and two F<sub>1</sub> from the crosses 'Goldrich × Currot' ('G × C') ( $n = 81$ ) and 'Goldrich × Canino' ('G × Ca') ( $n = 171$ ) (Soriano *et al.*, 2008). 'SEO' and 'Goldrich' were the PPV-resistant parents used in each case.

### DNA isolation

DNA was extracted from 50 mg of young leaves following the method described by Doyle and Doyle (1987). DNA quantification was performed by comparison with  $\lambda$  DNA (Promega, Madison, WI, USA).

### SSR marker analysis

Eighty-five SSR markers located on *Prunus* LG1 genetic maps and available from the GDR website (Genome Database for Rosaceae: <http://www.rosaceae.org>) were analysed (Table 1). In addition, 102 new primer pairs flanking microsatellite repeat motifs were designed using Primer3 (Rozen and Skaletsky, 2000) from positive apricot BACs identified by hybridization (Table S1; and see below). All SSRs were tested in the three apricot populations.

SSR amplifications were performed in a GeneAmp® PCR System 9700 thermal cycler (Perkin–Elmer, Fremont, CA, USA) in a final volume of 20  $\mu$ L, containing 75 mM Tris-HCl, pH 8.8,

20 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.5 mM MgCl<sub>2</sub>, 0.1 mM of each deoxynucleoside triphosphate (dNTP), 20 ng of genomic DNA and 1 U of Taq polymerase (Invitrogen, Carlsbad, CA, USA). Each PCR was performed using the procedure of Schuelke (2000) with three primers: the specific forward primer of each microsatellite with the M13(-21) tail at its 5' end at 0.4 μM, the sequence-specific reverse primer at 0.8 μM and the universal fluorescent-labelled M13(-21) primer at 0.4 μM. PCR conditions were performed as described by Soriano *et al.* (2008). Allele lengths were determined using an ABI Prism 3130 Genetic Analyzer with the aid of GeneMapper software, version 4.0 (Applied Biosystems, Foster City, CA, USA).

### Linkage analysis and QTL identification

Linkage analysis was carried out using JoinMap 3.0 software (Van Ooijen and Voorrips, 2001) with the Kosambi mapping function (Kosambi, 1944) used to convert recombination units into genetic distances. Linkage groups were established using a minimum LOD threshold of 8.0 and a recombination frequency below 0.4. The 'Lito' map was constructed using only SSR co-dominant markers segregating in the 'L × L-98' population and setting the F<sub>2</sub> data type. The final 'Goldrich' map was obtained by integrating the two 'Goldrich' maps derived from the 'G × C' and 'G × Ca' populations using the JoinMap 'map integration' function. These 'Goldrich' genetic maps were constructed following the 'two-way pseudo-test-cross' model of analysis (Grattapaglia and Sederoff, 1994) and setting 'cross-pollinator' data type in both populations. All SSRs heterozygous for 'Goldrich' were scored as dominant for mapping, except those showing segregation <hk × hk> (as per JoinMap 3.0), which were removed. Segregation of the markers was analysed one by one to correct possible mistakes in the JoinMap 3.0 output.

The KW rank-sum test (Lehmann, 1975) was applied using MapQTL version 4.0 software (Van Ooijen, 2000) with a threshold value of  $P < 0.005$  to test for associations between markers and PPV resistance. IM analysis (Lander and Botstein, 1989; Van Ooijen, 1992) was performed to support the detection of putative QTLs by the KW test. The LOD chromosome-wide significance threshold to decide on the presence or absence of a QTL for IM (Van Ooijen, 1999) was determined with a 5% significance level using permutation tests (Churchill and Doerge, 1994) carried out on LG1, and corresponded to a value of 3.0 in both maps. A confidence interval around the position of the largest LOD was indicated by a two-LOD support interval (Van Ooijen, 1992).

### BAC library hybridization

BAC identification was made using radioactively labelled overgo probes hybridized in pools (Madishetty *et al.*, 2007; Ross *et al.*, 1999) against an apricot BAC library developed from the PPV-resistant cultivar 'Goldrich' (Vilanova *et al.*, 2003).

Seventy-one overgo probes were designed. Forty-seven were hybridized in a first round, 34 came from clone sequences containing SSR markers previously mapped on the upper part of *Prunus* LG1 (Howad *et al.*, 2005; Lalli *et al.*, 2008; Soriano *et al.*, 2008) and 13 originated from peach BESs corresponding to BACs anchored onto the peach physical map by Zhebentyayeva *et al.* (2008). Subsequently, 24 additional overgo probes, developed from apricot BESs identified in the first round, were hybridized (Table 2).

Positive BACs were verified and assigned to individual probes by re-hybridization to colony dot blots. Positive BACs were reconfirmed by PCR using SSR markers. Overgo probes were designed using Overgo1.02i software (Cai *et al.*, 1998), selecting sequences without homology with repetitive motifs with CENSOR software (Kohany *et al.*, 2006) and intron-exon junctions predicted with GENSCAN (Burge and Karlin, 1997) when possible, following the website: <http://www.mouse-genome.bcm.tmc.edu/webovergo>.

### BAC end sequencing

BAC clones were inoculated into 96-deep-well microplates and grown for 20 h at 37 °C. Cells were harvested by centrifugation and BACs were purified in 96-well plates by a standard alkaline lysis protocol. BAC DNA was precipitated with isopropanol and washed with 70% ethanol. Sequencing was carried out on an ABI3730 equipment with the 'Dye Terminator' process using an ABI kit version 3.1 (Applied Biosystems). BAC internal sequences were obtained using overgo primers LA/LB.

The software Phred was used for base calling and Crossmatch for vector masking (Ewing and Green, 1998). Repetitive DNA was identified with RepeatMasker software (Smit *et al.*, 1996), using the Viridiplantae section of the RepBase Update (Jurka *et al.*, 2005) as database. Assembly was performed with CAP3 (Huang and Madan, 1999), using read quality and default parameters. Similarity searches were performed with the standalone version of BLAST (Altschul *et al.*, 1990) against the National Center for Biotechnology Information (NCBI) nonredundant protein database and the peach genome (peach v1.0, IPGI 2010: <http://www.rosaceae.org/peach/genome>). Parsing of the BLAST results was performed with the Bio::SearchIO module from the Bioperl package (Stajich *et al.*, 2002). Coding sequences were annotated with GO terms using BLAST2GO (Conesa *et al.*, 2005). SPUTNIK (Abajian, 1994) was used to identify SSRs. Apricot BDSs were compared with peach using BLASTN (Altschul *et al.*, 1990).

### Fingerprinting reactions and fingerprinted contigs (FPC) assembly

Fingerprinting reactions were executed using five restriction enzymes, *Bam*HI, *Eco*RI, *Hae*III, *Xba*I and *Xho*I (New England

BioLabs, Ipswich, MA, USA), and the ABI PRISM® SNaPShot® restriction fragment labelling kit according to Luo *et al.* (2003).

BAC fingerprint profiles, including peak areas, peak heights and fragment sizes, were collected by the instrument-implemented program ABI Data Collection v2.0 and automatically scored using the GeneMapper v4.0 software package (Applied Biosystems). An ABI sequencer-compatible package, GenoProfiler v2.0 (You *et al.*, 2007), was used for fragment analysis, data filtering from background noise and the removal of repetitive and vector bands.

The HICF-compatible version of FPC v8.9 (<http://www.agcol.arizona.edu/software/fpc>) was used to assemble BAC clones into contigs (Soderlund *et al.*, 1997). FPC build was constructed at a tolerance of six with a cut-off of 1E-20, followed by the automatic Dqer function to minimize the number of false-positive merges.

## ACKNOWLEDGEMENTS

This research was supported by grants from the Ministerio de Educación y Ciencia of Spain, Research Project AGL2007-60709, and by the European Union Seventh Framework Programme FP7-KBBE-2007-1. EMVR was funded by a fellowship from the Ministerio de Educación y Ciencia of Spain.

## REFERENCES

- Abajian, C. (1994) SPUTNIK. Computer program. Available at <http://espressoftware.com/sputnik/index.html>.
- 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.
- Aranzana, M.J., Pineda, A., Cosson, P., Dirlwanger, E., Ascasibar, J., Cipriani, G., Ryder, C.D., Testolin, R., Abbott, A., King, G.J., Iezzoni, A.F. and Arús, P. (2003) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. *Theor. Appl. Genet.* **106**, 819–825.
- Bassi, D. and Audergon, J.M. (2006) Apricot breeding, update and perspectives. *Acta Hortic.* **701**, 279–294.
- Burge, C. and Karlin, S. (1997) Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94.
- Cai, W., Reneker, J., Chow, C., Vaishnav, M. and Bradley, A. (1998) An anchored framework BAC map of mouse chromosome 11 assembled using multiplex oligonucleotide hybridization. *Genomics*, **54**, 387–397.
- Cambra, M., Capote, N., Cambra, M.A., Llácer, G., Botella, P. and López-Quílez, A. (2006) Epidemiology of sharka disease in Spain. *EPPO Bull.* **36**, 271–275.
- Churchill, G.A. and Doerge, R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics*, **138**, 963–971.
- Claverie, M., Dirlwanger, E., Cosson, P., Bosselut, N., Lecouls, A.C., Voisin, R., Kleinhentz, M., Lafargue, B., Caboche, M., Chalhoub, B. and Esmenjaud, D. (2004) High-resolution mapping and chromosome landing at the root-knot nematode resistance locus *Ma* from Myrobalan plum using a large-insert BAC DNA library. *Theor. Appl. Genet.* **109**, 1318–1327.
- Conesa, A., Gotz, S., García-Gómez, J.M., Terol, J., Talón, M. and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.
- Decroocq, V., Foulouge, M., Lambert, P., Le Gall, O., Martin, C., Pascal, T., Schurdi-Levraud, V. and Kervella, J. (2005) Analogues of virus resistance genes map to QTLs for resistance to sharka disease in *Prunus davidiana*. *Mol. Genet. Genomics*, **272**, 680–689.
- Dirlwanger, 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.
- 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. (2010) Identification of QTL for resistance to plum pox virus strains M and D in Lito and Harcot apricot cultivars. *Mol. Breed.* DOI: 10.1007/s11032-010-9431-3.
- Doyle, J.J. and Doyle, J.L. (1987) A rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15.
- Erdin, N., Tartarini, S., Broggin, G.A.L., Gennari, F., Sansavini, S., Gessler, C. and Patocchi, A. (2006) Mapping of the apple scab resistance gene *Vb*. *Genome*, **49**, 1238–1245.
- Escalettes, V., Dosba, F., Lansac, M. and Eyquard, J.P. (1998) Genetic resistance to Plum pox potyvirus in peaches. *Acta Hortic.* **465**, 689–698.
- Ewing, B. and Green, P. (1998) Basecalling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194.
- Grattapaglia, D. and Sederoff, R.R. (1994) Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudotest-cross strategy and RAPD markers. *Genetics*, **137**, 1121–1137.
- Gygax, M., Gianfranceschi, L., Liebhard, R., Kellerhals, M., Gessler, C. and Patocchi, A. (2004) Molecular markers linked to the apple scab resistance gene *Vbj* derived from *Malus baccata* jackii. *Theor. Appl. Genet.* **109**, 1702–1717.
- 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.
- Howad, W., Yamamoto, T., Dirlwanger, E., Testolin, R., Cosson, P., Cipriani, G., Monforte, A.J., Georgi, L., Abbott, A.G. and Arús, P. (2005) Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. *Genetics*, **171**, 1305–1309.
- Huang, X. and Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res.* **9**, 868–877.
- Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O. and Walichiewicz, J. (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462–467.
- Karayiannis, I. (2006) Breeding for resistance to PPV in apricots in Greece. *EPPO Bull.* **36**, 319–322.
- Kohany, O., Gentles, A.J., Hankus, L. and Jurka, J. (2006) Annotation, submission and screening of repetitive elements in repbase: repbase submitter and censor. *BMC Bioinformatics*, **7**, 474.
- Kosambi, D.D. (1944) The estimation of map distance from recombination values. *Ann. Eugen.* **12**, 172–175.
- Labonne, G. and Dallot, S. (2006) Epidemiology of sharka disease in France. *EPPO Bull.* **36**, 268–270.
- Lai, C.W.J., Yu, Q., Hou, S., Skelton, R.L., Jones, M.R., Lewis, K.L.T., Murray, J., Eustice, M., Guan, P., Agbayani, R., Moore, P.H., Ming, R. and Presting, G.G. (2006) Analysis of papaya BAC end sequences reveals first insights into the organization of a fruit tree genome. *Mol. Genet. Genomics*, **276**, 1–12.
- 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., Arus, 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.
- Lander, E.S. and Botstein, D. (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**, 185–199.



- Lehmann, E.L. (1975) *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco, CA: Holden and Day.
- Llácer, G., Badenes, M.L. and Romero, C. (2007) Problems in the determination of inheritance of plum pox virus resistance in apricot. *Acta Hort.* **781**, 263–267.
- Luo, M.C., Thomas, C., You, F.M., Hsiao, J., Ouyang, S., Buell, C.R., Malandro, M., McGuire, P.E., Anderson, O.D. and Dvorak, J. (2003) High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labeling kit and sizing of restriction fragments by capillary electrophoresis. *Genomics*, **82**, 378–389.
- 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**, 1.
- Marandel, G., Salava, J., Abbott, A., Candresse, T. and Decroocq, V. (2009a) 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.
- Marandel, G., Pascal, T., Candresse, T. and Decroocq, V. (2009b) Quantitative resistance to Plum pox virus in *Prunus davidiana* P1908 linked to components of the eukaryotic translation initiation complex. *Plant Pathol.* **58**, 425–435.
- Martin, G.B., Bogdanove, A.J. and Sessa, G. (2003) Understanding the functions of plant disease resistance proteins. *Annu. Rev. Plant Biol.* **54**, 23–61.
- 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.
- Moustafa, T.A., Badenes, M.L., Martínez-Calvo, J. and Llácer, G. (2001) Determination of resistance to sharka (plum pox) virus in apricot. *Sci. Hort.* **91**, 57–70.
- Olmos, A., Capote, N. and Candresse, T. (2006) Detection and characterization of Plum pox virus: molecular methods. *EPPO Bull.* **36**, 262–266.
- Patocchi, A., Vinatzer, B.A., Gianfranceschi, L., Tartarini, S., Zhang, H.B., Sansavini, S. and Gessler, C. (1999) Construction of a 550kb BAC contig spanning the genomic region containing the apple resistance gene *Vf*. *Mol. Gen. Genet.* **262**, 884–891.
- 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' × 'Vestiar'. *Tree Genet. Genomes*, **6**, 467–475.
- Ross, M.T., LaBrie, S., McPherson, J. and Stanton, V.P. (1999) Screening large-insert libraries by hybridization. In: *Current Protocols in Human Genetics* (Boyl, A., ed.), pp. 561–565. New York: Wiley.
- Rozen, S. and Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (Krawetz, S. and Misener, S., eds), pp. 365–386. Totowa, NJ: Humana Press.
- Rubio, M., Pascal, T., Bachellez, A. and Lambert, P. (2010) Quantitative trait analysis of Plum pox virus resistance in *Prunus davidiana* P1908: new insights on the organization of genomic resistance regions. *Tree Genet. Genomes*, **6**, 291–304.
- Schuelke, M. (2000) An economic method for the fluorescent labelling of PCR fragments. *Nat. Biotechnol.* **18**, 233–234.
- Smit, A.F.A., Hubley, R. and Green, P. (1996) RepeatMasker Open-3.0. *Computer Program*. Available at <http://www.repeatmasker.org>.
- Soderlund, C., Longden, I. and Mott, R. (1997) FPC: a system for building contigs from restriction fingerprinted clones. *Comput. Appl. Biosci.* **13**, 523–535.
- Soriano, J.M., Vera-Ruiz, E.M., Vilanova, S., Martínez-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., Joshi, S., van Kaauwen, M., Noordijk, Y., Groenwold, R., Henken, B., van de Weg, W.E. and Schouten, H.J. (2009) Identification and mapping of the novel apple scab resistance gene *Vd3*. *Tree Genet. Genomes*, **5**, 475–482.
- Stajich, J.E., Block, D., Boulez, K., Brenner, S.E., Chervitz, S.A., Dagdigian, C., Fuellen, G., Gilbert, J.G.R., Korf, I., Lapp, H., Lehväslaiho, H., Matsalla, C., Mungall, C.J., Osborne, B.I., Pocock, M.R., Schattner, P., Senger, M., Stein, L.D., Stupka, E., Wilkinson, M.D. and Birney, E. (2002) The Biperl toolkit: perl modules for the life sciences. *Genome Res.* **12**, 1611–1618.
- Van Ooijen, J.W. (1992) Accuracy of mapping quantitative trait loci in autogamous species. *Theor. Appl. Genet.* **84**, 803–811.
- Van Ooijen, J.W. (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity*, **83**, 613–624.
- Van Ooijen, J.W. (2000) *MapQTL TM Version 4.0: User Friendly Power in QTL Mapping*. Addendum to the manual of version 3.0. Wageningen, The Netherlands: Plant Research International B.V.
- Van Ooijen, J.W. and Voorrips, R.E. (2001) *JoinMap@3.0, Software for the Calculation of Genetic Linkage Maps*. Wageningen, The Netherlands: Plant Research International B.V.
- 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.
- You, F.M., Luo, M.C., Gu, Y.Q., Lazo, G.R., Deal, K., Dvorak, J. and Anderson, O.D. (2007) GenoProfiler: batch processing of high throughput capillary fingerprinting data. *Bioinformatics*, **23**, 240–242.
- Zhebentyayeva, T.N., Swire-Clark, G., Georgi, L.L., Garay, L., Jung, S., Forrest, S., Blenda, A.V., Blackmon, B., Mook, J., Horn, R., Howad, W., Arús, P., Main, D., Tomkins, J.P., Sosinski, B., Baird, W.V., Reighard, G.L. and Abbott, A.G. (2008) A framework physical map for peach, a model Rosaceae species. *Tree Genet. Genomes*, **4**, 745–756.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Table S1** Simple sequence repeat (SSR) markers from the literature tested for 'Lito' and 'Goldrich' map construction.

**Table S2** Apricot simple sequence repeat (SSR) primers developed from the 'Goldrich' bacterial artificial chromosome (BAC) library. Primer sequence, primary contig, BAC-derived sequence (BDS) origin, size of amplified products and annealing temperature are indicated.

**Table S3** Results of similarity searches between apricot BAC-derived sequence (BDS) and peach scaffold\_1 sequence using BLASTN. Primary contigs (P) from which the BDS (query) were obtained and sequence contigs (S) are indicated. Query and hit HSP (high-scoring pair) records are also indicated.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.