Genomic characterization of putative allergen genes in peach/almond and their synteny with apple

Lin Chen1,2, Shuiming Zhang1,5, Eudald Illa3, Lijuan Song1, Shandong Wu2, Werner Howad3, Pere Arús3, Eric van de Weg4, Kunsong Chen1 and Zhongshan Gao*1,2

Address: 1Institute of Fruit Science, The State Agriculture Ministry Laboratory of Horticultural Plant Growth, Development and Biotechnology, Zhejiang University, Hangzhou, 310029, PR China, 2Allergy Research Center, Zhejiang University, Hangzhou, 310058, PR China, 3IRTA. Centre de Recerca en Agrigenòmica CSIC-IRTA-UAB, Carretera de Cabrils Km2; 08348 Cabrils (Barcelona), Spain, 4Plant Research International, Wageningen University and Research Centre, PO box 16, 6700AA, Wageningen, The Netherlands and 5Key Laboratory of Pomology, Anhui Agricultural University, Hefei 230036, PR China

Email: Lin Chen - d05chenlin@zju.edu.cn; Shuiming Zhang - zhangshuiming@zju.edu.cn; Eudald Illa - eudald.illa@irta.es; Lijuan Song - sjlj2006215.slij@yahoo.com.cn; Shandong Wu - wushandong@sohu.com; Werner Howad - werner.howad@irta.es; Pere Arús - pere.arus@irta.es; Eric van de Weg - eric.vandeweg@wur.nl; Kunsong Chen - akun@zju.edu.cn; Zhongshan Gao* - gaozhongshan@zju.edu.cn

* Corresponding author

Abstract

Background: Fruits from several species of the Rosaceae family are reported to cause allergic reactions in certain populations. The allergens identified belong to mainly four protein families: pathogenesis related 10 proteins, thaumatin-like proteins, lipid transfer proteins and profilins. These families of putative allergen genes in apple (Mal d 1 to 4) have been mapped on linkage maps and subsequent genetic study on allelic diversity and hypoallergenic traits has been carried out recently. In peach (Prunus persica), these allergen gene families are denoted as Pru p/du 1 to 4 and for almond (Prunus dulcis) Pru p/du 1 to 4. Genetic analysis using current molecular tools may be helpful to establish the cause of allergenicity differences observed among different peach cultivars. This study was to characterize putative peach allergen genes for their genomic sequences and linkage map positions, and to compare them with previously characterized homologous genes in apple (Malus domestica).

Results: Eight Pru p/du 1 genes were identified, four of which were new. All the Pru p/du 1 genes were mapped in a single bin on the top of linkage group 1 (G1). Five Pru p/du 2 genes were mapped on four different linkage groups, two very similar Pru p/du 2.01 genes (A and B) were on G3, Pru p/du 2.02 on G7, Pru p/du 2.03 on G8 and Pru p/du 2.04 on G1. There were differences in the intron and exon structure in these Pru p/du 2 genes and in their amino acid composition. Three Pru p/du 3 genes (3.01–3.03) containing an intron and a mini exon of 10 nt were mapped in a cluster on G6. Two Pru p/du 4 genes (Pru p/du 4.01 and 4.02) were located on G1 and G7, respectively. The Pru p/du 1 cluster on G1 aligned to the Mal d 1 clusters on LG16; Pru p/du 2.01A and B on G3 to Mal d 2.01A and B on LG9; the Pru p/du 3 cluster on G6 to Mal d 3.01 on LG12; Pru p/du 4.01 on G1 to Mal d 4.03 on LG2; and Pru p/du 4.02 on G7 to Mal d 4.02 on LG2.

Conclusion: A total of 18 putative peach/almond allergen genes have been mapped on five linkage groups. Their positions confirm the high macro-synteny between peach/almond and apple. The insight gained will help to identify key genes causing differences in allergenicity among different cultivars of peach and other Prunus species.
Background

In general, eating fruit is strongly recommended to improve health, but this can cause an allergic reaction in certain individuals. Most fruit crops in temperate and some subtropical zones belong to the Rosaceae family, mainly the subfamilies Maloideae and Prunoideae. Fruits from the Maloideae, like apple [1,2] and pear [3], and Prunoideae such as peach [4,5], sweet cherry [6,7], plum [8], apricot [9] and almond [10] have been reported to cause allergic reactions in Europe. The annual peach production in China has been around 6 million tons in recent years, and has lead to both an increase in peach consumption and more prevalent fruit allergy.

It is considered that general avoidance of fruit has a negative effect on health in allergic patients and also affects their quality of life. Therefore, the low allergic property of fruit is worth considering in new peach breeding programs. As cultivars of the same species differ in allergenicity [11], the selection and breeding of new, low-allergenic cultivars is becoming feasible by a multidisciplinary approach, such as in the EU-SAFE project [1], where doctors and plant geneticists identified the apple cultivar 'Santana' as low allergenic in the majority of Dutch patients with a history of allergy to apple [12]. A similar approach was initiated in peach in recent years through collaboration between China and Europe.

A variety of allergens from different fruits have been identified by means of experimental immunology and molecular biology, in particular by protein and gene identification and sequencing[13]. From the structural and biochemical point of view, fruit allergens identified belong to a limited number of protein families [13]. In the case of Rosaceae, the majority of allergens present in fresh fruit belong to four families: pathogenesis-related protein 10 (PR-10 protein, birch allergen Bet v 1 homologues), thaumatin-like proteins (TLP, PR-5 proteins), non-specific lipid transfer proteins (nsLTPs, PR-14 proteins) and profilins (PRF) [13]. Whole, unprocessed fruit containing PR-10 protein elicit allergy, while TLP, LTP and profilins are stable proteins because of intra-molecular disulfide bonds. These four allergen families in Prunus fruits are already on the international allergen list http://www.allergen.org. In peach they are denoted as Pru p 1 (PR10), Pru p 2 (TLP), Pru p 3 (LTP), and Pru p 4 (profilin), and in almond as Pru du 1, 2, 3, and 4, respectively. Due to the common structure and properties of each protein family over a wide range of plant species, genera and even families, allergy cross-reactivity has been frequently observed, for example between Pru p 1 and Mal d 1 proteins from peach and apple respectively with Bet v 1 proteins in birch pollen [14]. Peach allergy occurs more often in southern Europe, and two different peach allergy profiles have been observed in Spain. One of these is the oral allergy syndrome (OAS) caused by Pru p 1 and Pru p 4, the other is the systemic gastrointestinal symptom caused by Pru p 3 [15]. In China both profiles have also been reported after eating fresh peach [16]. In addition, some contact with peach fruit hair or even with pollen during the blooming season caused adverse reactions (from some patients’ claims).

Genomic characterization of the sequences and linkage map positions of the four apple allergen gene families were reported in 2005 [17-19], and subsequent associations between the genetic constitution and allergenicity of different cultivars gave the first indications on the isoallergens (genes) and protein variants (alleles) involved in low and high allergenic responses [20]. Comparison between apple (Malus domestica) and Prunus linkage maps suggests a high level of synteny of the linkage groups between these two genera [21]. For Prunus G1, the upper part contains several RFLP markers that aligned with apple LG13 and LG16 on which the majority of Mal d 1 genes are located, and the lower part aligned with LG8, where Mal d 4.03A is located. Prunus G3 aligns with apple LG9 where apple Mal d 2.01A/B and Mal d 4.01 were mapped. Therefore, some of the homologous allergen genes of Prunus fruit may be located at syntenic positions as compared with the apple genome.

In peach, differences in major allergen Pru p 1 (PR-10) and Pru p 3 (LTP) content among cultivars have been reported in recent years [11,22], which encouraged us to conduct research similar to that done in apple [17-19]. Advances in three aspects make the genomic research of Prunus allergen genes possible. First, a set of candidate genes is now available based on previous medical and molecular biology research [23-25]. Second, a large number of peach fruit ESTs are available in databases (mainly from researchers in Chile, Italy and USA), allowing us to find different members of the four putative allergen gene families. Third, a highly saturated genetic Prunus reference map T × E (based on the F2 progeny of 'Texas' almond and 'Earlygold' peach) and a simple bin-mapping strategy [26] provides an efficient and reliable way to map new genes.

Results

PCR-based cloning and individual clone sequencing with genomic DNA templates of the Prunus T × E F1 hybrid (MB1-73, the plant used for the generation of the T × E F2 population by selfing), produced 35 consensus sequences (GenBank: EU424239-EU424273) represent 18 different gene loci. Within a single gene locus, two putative alleles from peach and almond were distinguished through marker test and sequence comparison. The total length of each sequence and length of its open reading frames (ORFs), exon and intron are summarized in Table 1. Prim-
ers used for linkage mapping are presented in Table 2. Each gene family is described below in more detail.

**Pru p/du 1 (PR-10) genes**

According to four Pru p 1 cDNA reference sequences derived from peach fruit in the GenBank database, we obtained 15 different PR-10 DNA sequences [GenBank: EU424239-EU424253] from MB1-73 by PCR cloning, and classified them into eight distinct gene members based on their DNA sequence identity less than 98%, of which four were new (Pru p/du 1.05 and Pru p/du 1.06A-C) (Table 1) and formed a new group distinct from the initial reference sequence [GenBank:DQ251187] (Pru p 1.01). All Pru p/du 1 genes had a single intron with variable size, (98–233 nt) and two exons of 184 nt and 299 nt each, except in the second exon in Pru p/du 1.04 that was 296 nt long. Alleles derived from almond cv 'Texas' and those in apple cv 'Earlygold' had different intron sizes in four Pru p/du 1 members (Table 1), used to map these genes in polycrylamide gel electrophoresis (PAGE). DNA sequence divergence between allele from peach and almond is very low (1%–2.1%). Amino acid (aa) sequences from these genes were deduced from their coding DNA sequences (Figure 1). Most Pru p/du 1 had 160 aa, except for Pru p/du 1.04 with 159aa, with the common feature of a p-loop (GxG-GxGxxK). Overall aa sequence identity among different isoallergens was 71.9%–98.8%, and three Pru p 1.06 members (A, B, C) have about 98% aa sequence identity (Figure 1). Segments of identical aa sequence were found in several locations, especially surrounding the p-loop. Pru p/du 1.01 had a proline (P) residue in the p-loop, which is similar to Mal d 1.01 and 02 and all Bet v 1, while the other Pru p/du 1 isoallergens had a valine (V) in this position. The predicted protein molecular weights were in the range of 17.1–17.5 kDa, and pI values ranged from 4.9 to 6.0. Comparing the aa composition between alleles from peach and almond of each gene in Figure 1, we found that there was only one difference for Pru p/du 1.01, 02, 03, 04 and 06B, while Pru p/du 1.05 and Pru p/du 1.06A differed by 2 and 5 aa, respectively. A phylogenetic tree of the deduced amino acids from peach cv. 'Earlygold' alleles and those in apple (mainly cv. 'Fiesta') (Figure 2A) showed that at least one Pru p 1 isoallergen member is present in each of the four subfamilies of apple Mal d 1 genes (PR-10) identified in a previous study [17]. Furthermore, we can see in the subfamily III branch that Pru p 1.05 and three highly similar Pru p 1.06 members (A-C) clustered with three Mal d 1.06 members (A-C), while Pru-1.04 clustered together with Mal d 1.07, one of the apple intronless genes in subfamily IV. All eight Pru p/du 1 genes mapped to bin 1:14, at the top of G1 (Figure 3A, Table 1).

**Pru p/du 2 (TLP)**

Analysis of the peach TLP sequences in the NCBI identified four members of this gene family at the start. A single primer pair based on [GenBank:AF362988] resulted in four very similar sequences assigned as two separate genes...
(Pru p/du 2.01A and 2.01B) with minor intron size differences. Genome walking to upstream was employed in Pru p/du 2.02 to search for more polymorphisms and downstream to Pru p/du 2.03, since its initial reference sequence was partial. Five members of Prunus TLPs were identified after analysis of the genomic sequences [GenBank: EU424254-EU424263]: Pru p/du 2.01A and -01B, Pru p/du 2.02, -03 and -04 (Table 1), of which Pru p/du 2.01B is new. The Pru p/du 2 members had distinct genomic structure (Table 1). Three of them had one intron separating a short (58 or 61 nt) and long (671–773 nt) exon, while one member (Pru p/du 2.04) had two introns.

**Table 2: Markers developed and cloning and sequencing primers for mapping**

| Gene       | Primer sequence | Tm °C | Cycles | Allele length (nt) |
|------------|-----------------|-------|--------|--------------------|
| Pru du 1.01 | For-ATGGGTTGCTTCACATATGAGG | 58    | 35     | T-218              |
|            | Rev-CTGACACATATATCAGTTG |       |        |                    |
| Pru p 1.01  | For-TCATACAGCTACACCTGCTT | 58    | 35     | E-205              |
|            | Rev-CATAGGCAAGGGAGAATGGTTG |       |        |                    |
| Pru p/du 1.02 | For-GTTGGAACCATCAAGAGGAC | 58    | 35     | T-188              |
|            | Rev-GTTTACATGACAAGAATACAT |       |        | E-179              |
| Pru p/du 1.03 | For-GGTAAGGTGATGCTATTTGA | 58    | 35     | T-153              |
|            | Rev-GTTCCATACCTCCGTTCACAA |       |        | E-150              |
| Pru p/du 1.04 | For-CCTTTTGAGGTAAGTGTTTAG | 58    | 35     | T-157              |
|            | Rev-GTTTACATGACAAGAATACAT |       |        | E-164              |
| Pru p/du 1.05 | For-GCTTCCAATCATCAGAATACCT | 60    | 35     | T-179              |
|            | Rev-GTGGGAACTAGTGTTGCAAGCA |       |        | E-182              |
| Pru p/du 1.06A | For-TAGTTTAGGTGCTTGCAATGCT | 58    | 35     | 1871, sequencing   |
|            | Rev-GAAGTTCACAAAGTAGATGTCG |       |        |                    |
| Pru p/du 1.06B | For-ATCCCGAAGGTGTGATCCCTG | 60/58 | 5/35   | T-182              |
|            | Rev-GATTAGAATTTAAGAGCTTACTGT |       |        |                    |
| Pru p/du 1.06C | For-CTAATTAATACATAAATCTTTGGAACAGT | 56/54 | 5/35   | E-272              |
|            | Rev-TGTAGCTCACAATCCCTCGC |       |        |                    |
| Pru p 2.01A | For-CCTTAGATTACACAGCCCTTCCA | 60    | 35     | T-175              |
|            | Rev-GATGAGGTAGTCAGGGCAAGA |       |        |                    |
| Pru p 2.01A | For-CAATGTGCTCTGATCTGTTTTG | 60    | 35     | E-205              |
|            | Rev-ACGGAGCTCACAGAGATTGT |       |        |                    |
| Pru p 2.01B | For-CTGTTGATGCTTAAAGAGAC | 60    | 35     | T-187              |
|            | Rev-GCACTGGGTCTTAAAGAGAC |       |        |                    |
| Pru p 2.01B | For-CAATGTGCTCTGATCTGTTTTG | 60    | 35     | E-622              |
|            | Rev-GGTTGAGAGCTGTAGCTAGTG |       |        |                    |
| Pru p 2.02  | For-AAGAATCCATCAACTGAAATCC | 56/40 | 5/35   | T-271              |
|            | Rev-CTACTAGGGTCTTCATCATCCGA |       |        | E-315              |
| Pru p 2.03  | For-ACATTGCCCTTCTGGAGACTA | 60    | 30     | T-237              |
|            | Rev-CCATTTGCTTCACAGATTCAA |       |        | E-229              |
| Pru p 2.04  | For-TCATGTTGAGCTGTCATGTC | 60    | 35     | T-267              |
|            | Rev-CTCAGAGCTTAGAAGAGTCTG |       |        | E-235              |
| Pru p 3.01  | For-CACCTAGCTCATACCTAGCTGTG | 60    | 35     | 530, sequencing    |
|            | Rev-GTGGAACTACCTCTAAGGTG |       |        |                    |
| Pru p 3.02  | For-CCTTACACCCGCAAATGCTTACATC | 60    | 35     | 319, sequencing    |
|            | Rev-CCACTGAAAACCCACAGCAGA |       |        |                    |
| Pru p 3.03  | For-AGGACGCCCCTTACCTCTTGGC | 60    | 35     | 413, sequencing    |
|            | Rev-CCACATGCTTCCGGTGTATCAGA |       |        |                    |
| Pru p 4.01  | For-AGAAGAAATACAGAAGCAAC | 55    | 35     | T-276              |
|            | Rev1-CAAGATACCAACCCAGTTGCG |       |        | E-631              |
|            | Rev2-CTTTGTTGCACCAAAGAATAGA |       |        |                    |
| Pru p 4.02  | For-AAGATACATGCTCATCCAAAGCC | 58    | 35     | T-188              |
|            | Rev-GTCCCTGAAATTTGAGTGC |       |        |                    |
| Pru p 4.02  | For-GGCAACAACCTCTCTTCTCTCCTTC | 58    | 35     | T-188              |
|            | Rev-CAAGATACCAACCCAGTTGGCC |       |        |                    |

1. Designed mismatching nucleotides are in bold and underlined, added pigtail bases in bold and italics.
2. Annealing temperature (Tm)
3. Number of PCR cycles. In the case of two values, a touch-down PCR was performed in two steps: the first number of the "Tm" and "Cycles" column refers to the Tm and number of cycles, respectively, of the first step.
4. T and E refer to parental source from 'Texas' almond and 'Earlygold' peach, respectively. Genes mapped by direct sequencing are indicated by 'sequencing'.
the deduced precursor of *Prunus* TLPs with the reference TLP amino acid sequences in the Rosaceae and other plants, we assigned the putative signal peptides to deduce the mature protein (Figure 4). Pru p/du 2.02 had a shorter signal peptide, 23 aa in peach and 21 aa in almond, while the other four members had a 24 aa signal peptide. They differed considerably in protein size, with Pru p/du 2.03 and Pru p/du 2.04 having longer peptides in the C-terminal than Pru p/du 2.01 and 02. Amino acid sequences identity from alleles from peach and almond for five Pru p/du 2 members are different, they are 84.4% for Pru p/du 2.02, about 96% for Pru p/du 2.01A and -01B, 98% for Pru p/du 2.03, and 99.4% for Pru p 2.04 p/du 2.04.

Based on the percentage amino acid identity among five Pru p 2 isoallergens, Pru p 2.01A, 01B and 2.02 were clustered with four apple Mal d 2 (01A, 01B, 02 and 03) (Figure 2B) proteins. Pru p 2.01A and 2.01B had 96.7% identity in their aa sequence, their closest homologous members being apple Mal d 2.01A and B. Pru p 2.03 and 2.04 are quite distant from 2.01 and 2.02 (Figure 2B) and no close counterparts were found in apple. The overall sequence identity level in amino acids among the four Pru p 2 proteins was around 40% (Figure 2B), but a long sequence region (DFYDVSLVDG [F/Y]NLPM) around position 130, and the typical 16 cysteine (C) to form 8 disulfide bonds, were conserved in all Pru p/du 2 members. Mature Pru p 2.01 and 2.02 had a molecular weight of about 23 kDa. Mature Pru 2.03 differed in length and weight with mature Pru du 2.03, with the molecular weight being 25.23 kDa and 27.46 kDa respectively. Mature Pru 2.04 had the highest number of aa (307) with a calculated molecular weight of 30.7 kDa. Unlike the Pru p 1 genes that were placed in a single bin, these five Pru p/du 2 members mapped on four linkage groups: Pru p 1 du 2.01A and B on G3, -2.02 on G7, -2.03 on G8, and -2.04 on G1 (Figure 3A, C, D, E).
Phylogenetic trees derived from the deduced PR-10, TLP, LTP and profilin (PRF) families in peach (Pru p 1–4) and apple (Mal d 1–4) based on our genomic sequences and ESTs (in bold) retrieved from the DNA NCBI Database (see their GenBank accession numbers below).

A: PR-10 family. Peach: Pru p 1.01 [GenBank:EU424240], Pru p 1.02 [GenBank:EU424248], Pru p 1.03 [GenBank:EU424252], Pru p 1.04 [GenBank:EU424255], Pru p 1.05 [GenBank:EU424257], Pru p 1.06A [GenBank:EU424259], Pru p 1.06B [GenBank:EU424250], Pru p 1.06C [GenBank:EU424253]. Apple: Mal d 1.01 [GenBank:AY789238]; Mal d 1.02 [GenBank:AY789241]; Mal d 1.03A [GenBank:AY789244]; Mal d 1.03B [GenBank:AY789244], Mal d 1.03C [GenBank:AY789266]; Mal d 1.03D [GenBank:AY789267], Mal d 1.03E [GenBank:AY789269], Mal d 1.03F [GenBank:AY789271], Mal d 1.03G [AY789274], Mal d 1.04 [GenBank:AY789244], Mal d 1.05 [GenBank:AY789246], Mal d 1.06A [GenBank:AY789249], Mal d 1.06B [GenBank:AY789251], Mal d 1.06C [GenBank:AY789254], Mal d 1.07 [GenBank:AY789257], Mal d 1.08 [GenBank:AY789260], Mal d 1.09 [GenBank:AY789262].

B: TLP family. Peach: Pru p 2.01A [GenBank:EU424257], Pru p 2.01B [GenBank:EU424259], Pru p 2.02 [GenBank:EU424255], Pru p 2.03 [GenBank:EU424261], Pru p 2.04 [GenBank:EU424263]. Apple: Mal d 2.01A [GenBank:AY792599], Mal d 2.01B [GenBank:AY792603]; Mal d 2.02 (consensus sequence derived from ESTs, [GenBank:CO9044477, CV082311, CO723595, CN445021] and Mal d 2.03 (consensus of 8 ESTs: [GenBank:CO866703, CO901275, CN495042, CV084040, CO866711, CN444138, CO866347, CN491711] [18]).

C: LTP family: Peach: Pru p 3.01 [GenBank:EU424265], Pru p 3.02 [GenBank:EU424257]. Apple: Mal d 3.01 [GenBank:AE792599], Mal d 3.02 [GenBank:AE792603]; Mal d 3.03 [GenBank:AE792599], Mal d 3.04 [GenBank:AE792603].

D: Profilin Family: Peach: Pru p 4.01 [GenBank:EU424271], Pru p 4.02 [GenBank:EU424273]. Apple: Mal d 4.01 [GenBank:AE792608], Mal d 4.02A [GenBank:AE792610], Mal d 4.02B (consensus of 3 ESTs, [GenBank:CN578860, CN992900, CV082623], Mal d 4.03A [GenBank:AE792616], Mal d 4.03B (consensus of 3 ESTs, [GenBank:CO722645, CO052497, CO418275].

Amino acid substitutions (x100)
Four peach and almond nsLTP members were traced from the NCBI DNA database, three of them have complete coding sequences, the fourth (GenBank:BQ641139) is a partial almond cDNA sequence distant from the other three. Genomic cloning and sequencing was done with the first three members Pru p/du 3.01-03 [GenBank:EU424264-EU424269]. Pru p/du 3 gene sequences have
Alignment of amino acid sequences of Pru p/du 2 members identified in the Prunus 'Texas' almond × 'Earlygold' peach hybrid MB1-73.

Figure 4
Alignment of amino acid sequences of Pru p/du 2 members identified in the Prunus 'Texas' almond × 'Earlygold' peach hybrid MB1-73. Four levels of identity are shown: (1) black, 100%; (2) grey, 80%; (3) light grey, 60%; and (4) white. The signal peptides were deduced by comparison with other TLPs in Rosaceae fruits or other plants. 16 conserved cysteine residues are arrowed.
two exons and one intron (Table 1). The size of the second exon was conserved and extremely short (10 nt), while the first was variable (344, 362 and 341 nt). Sequences from peach and almond LTPs are highly conserved in exon and intron size, with only one base pair length difference in introns of Pru p/du 3.02 and 3.03 (Table 1).

The deduced pre-mature Pru p 3 proteins had different signal peptide lengths of 26 aa, 30 aa and 25 aa for Pru p/du 3.01–03, respectively (Figure 5). In peach and almond (E and T), Pru p/du 3.01 alleles had identical signal amino acids, but their mature proteins differed in 5 aa, of which 4 are at the C-terminal end. Pru p/du 3.02 had identical aa, and Pru p/du 3.03, only one amino acid difference P/T at position 88. The three Prunus LTPs isoallergens differed greatly in their signal peptide region, with about 40–80% sequence identity (Figure 2C), their alignments were according to the eight conserved cysteine residues (Figure 5). All mature Pru p/du 3 were alkaline (pI = 8.5–9.22) and their molecular weight was around 9 kDa. Pru p 3.01 grouped closely with apple Mal d 3.01 and 02 (Figure 2C), and Pru p 3.02 with the putative Mal d 3.03 (e-cloning analysis assigned). Pru p 3.03 is distant from Pru p 3.01 and -02, and shares about 50% aa identity with the others. Interestingly, all three Pru p 3 genes mapped on bin 6:74 (Figure 3B). One of them, Pru p 3.03 (as Ltp2), had previously been mapped using the whole TxE population and was located at 72.5 cM from the top of G6 [19].

**Pru p/du 4 (profilin, PRF)**

Genomic sequences for two profilin genes Pru p/du 4.01 and Pru p/du 4.02 were obtained based on their cDNA data [GenBank: EU424270-EU424278]. Both genes have three exons with conserved length, 123, 138 and 135 nt, in agreement with that found in most plant species. DNA polymorphisms in Pru p/du 4.01 occurred mainly in the first intron, but in Pru p/du 4.02 they were in the first and second exon. These two genes gave the same protein sequences for peach and almond (Figure 6), although a few synonymous polymorphisms existed at the DNA level. Molecular weights were 14.06 and 14.19 kDa for Pru p 4.01 and 4.02, respectively, and both were acidic (calculated pI 4.5 and 4.6). Compared with the six known Mal d 4 isoallergens, Pru p 4.01 was similar to Mal d 4.03A and B (93.1% and 93.9% identity), while Pru p 4.02 was similar to Mal d 4.02A and B (95.4% and 96.9% identity) (Figure 2D). Pru p 4.01 was mapped on G1 and Pru p 4.02 on G7 (Figure 3A and 3D).

**Discussion**

This study provides basic genomic sequences and map positions of four allergen families in *Prunus persica* (peach) and *Prunus dulcis* (almond). This information is a prerequisite for unravelling the genetic basis of allergenic characteristics of *Prunus* fruits, and is likely to be similar for other *Prunus* species due to the high level of synteny of the species within this genus [21].

**Gene family and relevance of individual genes in allergenicity**

Gene family is becoming a prevailing topic in genomic research. It is thought that the pressure to conserve protein sequence and structure is associated with an increase in copy number of certain genes during evolution [27]. Although food allergens belong to a limited number of protein families [28], each protein family in a single spe-

---

**Figure 5**

Alignment of amino acid sequences of LTPs members identified in the *Prunus 'Texas' almond × 'Earlygold' peach hybrid MB1-73*. Eight conserved cysteine residues are arrowed. The lipid-binding motifs are shown by green oval.
cies can be putatively encoded (experimentally demonstrated in some cases) by multiple genes in specific tissues, thus making the in vivo allergic reaction very complex. There is a need to identify genes implicated in the allergic reaction from the many putative allergen genes existing. Searching the updated Arabidopsis thaliana DNA sequence database, we found 29 putative Pru p 1-homologous genes scattered over the five chromosomes (Table 3). More than half of these (17/29) are located on chromosome 1 in four clusters, each with 3 to 7 genes. Nineteen TLP genes are present on four chromosomes. LTP genes form a large gene family with 98 members, more or less equally distributed over all five chromosomes, and profilin is a small gene family with 5 members mapping to three chromosomes. It is probable that, in peach and apple, only a small fraction of PR-10, TLP and LTP gene members are relevant to allergenic trait.

A limited number of putative allergen genes may be expressed in consumed fruits, their expression profiles at mRNA and protein level are essential to allergenicity. Evidence of cDNA present in the ripe fruit is not sufficient. The need to qualify and quantify their individual or combined allergenicity is a major challenge for further investigation, where the application of new experimental and computational tools in genomics and proteomics will be crucial. Results in this area are starting to emerge and integrate, such as differential expression of LTP 1 (Pru p 3.01) and LTP2 (Pru p 3.02) [29], the expression studies of Pru p 3.01 (Pp-LTP1 in two peach genotypes [23], where major differences were detected in expression at mRNA and protein accumulation levels between them, and the research on IgE binding capacity [30], where specific IgE-binding protein regions of Pru p 3 were found for two types of allergy (OAS and systemic) by the mimotope mapping technique. Of four Pru p 2 isoallergen genes, a clone PpAz44 [GenBank:AF362988] matching to Pru p 2.01 was abundantly expressed in ripe peach fruit, while clone PpAz8 [GenBank:AF362987] (Pru p 2.02) was undetected by northern blot analysis [24]. Pru p 2.03 and Pru p 2.04 were identified from the fruit ESTs, the latter may be in large amount of transcript, since 29 ESTs can be assembled into one consensus gene. A very low level of Pru p 4 (profilin) was has been identified as pollen-associated allergen, and two isoallergen genes were isolated from fruit tissue [25]. Different Pru p 1 genes expression was not reported previously, our preliminary expression study was performed by means of real-time PCR for six Pru p 1 isoallergen genes on full ripe fruits of two cultivars. The result showed that the most abundantly expressed member is Pru p 1.01, followed by Pru p 1.06, then very low expression level for Pru p 1.04 and -1.05, while Pru p 1.02 and -0.3 were hardly detected (data not shown).

Table 3: Members of four gene families related to allergenicity, pathogenesis related 10 proteins (PR-10), thaumatin-like proteins (TLP), lipid transfer proteins (LTP) and profilin (PRF) on the five Arabidopsis thaliana chromosomes traced from the GenBank database

| Gene | Chrom-1 | Chrom-2 | Chrom-3 | Chrom-4 | Chrom-5 | Total |
|------|---------|---------|---------|---------|---------|-------|
| PR-10 | 17      | 3       | 2       | 3       | 4       | 29    |
| TLP   | 8       | 2       | 2       | 6       | 3       | 19    |
| LTP   | 19      | 14      | 16      | 25      | 24      | 98    |
| PRF   | 2       | 2       | 1       | 5       |         | 5     |
In apple, a very recent comprehensive gene expression of four classes of putative allergens in apple fruit has gained insight of the genetic and environmental factors affecting the allergenic potential of apples [31]. They showed that some putative members are statistically different in transcript accumulation in ripe fruit, while others were not detectable. This was a very nice reference to think about comparative transcription and protein expression for all set of peach putative allergen genes among diverse peach cultivars.

We found that the putative genes encoding the major allergens reported in peach fruit, Pru p 1 and Pru p 3 [15], were located on two linkage groups, G1 and G6. Based on their bin map positions, we can use SSR markers tightly linked or flanking these two major Prunus allergen genes for a diversity survey of a large number of genotypes to search for potential candidates for further sequencing of the allergen genes. Clinical evaluation of allergenicity for different peach cultivars is needed for association analysis to find low allergenic alleles.

**Comparative mapping between Prunus and Malus using the sequence and position of allergen genes**

Dirlewanger et al. [21] compared the Prunus and apple genomes using a limited number of common markers. They found that three Prunus linkage groups (G1, G3 and G4) were syntenic to seven apple linkage groups, G1 to LG13 and LG16 at the top and LG8 at the bottom, G3 to LG9 and LG17, and G4 to LG5 and LG10. Information is also available on the correspondence between homoeologous linkage groups within the apple genome obtained initially with RFLPs by Maliepaard et al. [32], to which various allergen genes and SSRs were recently added by Gao et al. [17-19] and Celton et al. [33], respectively. With the availability of genomic sequences and the linkage map positions of the four allergen-families in apple (26 apple allergen genes on eight linkage groups) [17-19] and peach (18 genes on 5 linkage groups, this study), we have additional information useful for updating the synteny between these two species.

Three of the eight Prunus bins where we located putative allergen genes corresponded to known syntenic and colinear regions of the apple genome (Figure 3): the Pru p/du 1 on G1, corresponding to regions with clusters of homologous genes (Mal d 1) of apple on LG13 and LG16, Pru p/du 4.01 on the lower part of G1 and Mal d 4.03 on LG8, and Pru p/du 2.01A/B on G3 and Mal d 2.01A/B on LG9. Three more syntenic regions were detected in this study: Pru p/du 4.02 on G7 and the region of LG2 where Mal d 4.02 maps, Pru p/du 3.01-03 on G6 corresponding to Mal d 3.01 on LG12 and Mal d 3.02 on LG4; and Pru p/du 4.01 which, in addition to the region of LG8 already mentioned, is syntenic with the upper part of LG15 where Mal d 4.03B maps. No counterparts were found to the Pru p/du 2.02,-03, and -04 genes, which were mapped on G7, G8 and G1, respectively, on the apple genome. The positions of these three genes in the Prunus genome indicate syntenic regions of apple where additional Mal d 2 genes may be located. Because Mal d 1 on apple LG 13 and LG16 are clustered in a region of about 2 cM, so we can deduce our mapped Pru p 1 genes are also in a cluster.

There are two hypotheses on the origin of the two genomes that generated the apple amphidiploid genome (2n = 34). One proposes that these were from members of two Rosaceae subfamilies, one from the Amygdaloideae (thus closer to Prunus with x = 8) and another from the Spiraeoideae (x = 9) [34]. The other hypothesis proposes that both ancestral genomes had a basic chromosome number of x = 9 at least one being Spiraeoideae (x = 9) [35]. Combining map position and sequence comparison, the cluster of Pru p 1 genes mapped on Prunus G1 is more similar to that of the Mal d 1 genes of LG16 than to LG13: a) Mal d 1 isoallergen genes were classified into four subfamilies and members of sub-families were distributed asymmetrically to LG13 and LG16 (Figures 2, 3), i.e., LG13 contains Mal d 1 members from subfamilies I and IV, but LG16 contains representative members of the four subfamilies; b) Pru p 1 genes fall into the four subfamilies as defined in apple, and had a higher deduced amino acid sequence similarity to the Mal d 1 members on LG16 (Figure 2 and data of percentage amino acid identity in the Results section). These results favor the two primitive genomes for apple and one is closer to Prunus than the other.

**Bin mapping feasibility and effectiveness**

The bin mapping approach provides a fast and cheap way to map new genes [26] that has worked very well in our hands. Once we cloned and sequenced the two putative alleles for each gene from MB1-73, we either created an allele-specific marker and genotyped the eight plants of the bin set, or cloned the fragments from the individuals of the bin set and sequenced them to determine their genotypes. One of the advantages of using the T × E population for bin mapping is that, due to its inter-specific nature, polymorphism is very high [26]. The disadvantage of the bin mapping approach is that there is a loss of precision in marker location (8–25 cM), but this can be solved for markers of particular interest by running the whole population to establish a more precise position or narrow down a region where closely related genes or common RFLP markers in apple has been fine mapped. For example, we estimate Pru p/du 1 genes (in bin 1:14) are in a cluster and likely around 13 cM from the top based on location of Mal d 1 genes and anchor marker MC001 (Figure 3A). A very large number of ESTs from peach fruit are on public databases. This information is essential to find
candidate genes to be mapped, and their positions compared to that of major genes or quantitative trait loci of interest mapped using phenotypic data. Bin mapping, as in this case, can facilitate the positioning of these candidate genes. Exchange of DNA and genotype data of the T × E bin set in the peach research community is helping candidate gene mapping, with genetic analysis and breeding applications following on.

Conclusion
Eighteen putative Prunus allergen genes were located on five of the eight linkage groups of this genus. Genomic characterization of Prunus allergens provided further evidence for synteny between Prunus and Malus. This information will be directed towards identifying the key genes or isoforms causing the differences in allergenicity among different cultivars by additional gene and protein expression and genetic association studies. These results are a first step towards the understanding of allergenicity caused by peach fruit and for the development of new cultivars with enhanced health properties and subject to specific recommendations on fruit consumption.

Methods

Plant materials

PCR cloning and sequencing for targeting allergen genes to identify different gene members and alleles was mainly on the 'Texas' (T) almond and 'Earlygold' (E) peach hybrid (MB1-73) used to generate the reference mapping population (T × E) and it is now used by the international Prunus community. Therefore, two putative alleles, one from 'Texas' (T) and another from 'Earlygold' (E), will be simultaneously amplified with a pair of primers based on a reference sequence. For T and E allele-specific marker tests, DNA from T and E were used as controls. Eight plants were used for the bin set (the set of plants used for bin mapping): the grandparent E, the parent MB1-73, and six plants from the F2 progeny (# 5, 12, 23, 30, 34, 83) [26]. DNA samples were isolated and provided by IRTA (Institut de Recerca i Tecnologia Agrària, Terrassa, Barcelona, Spain) for gene sequences that were not cover the ORF or to avoid the mistakes induced by primer sequence, or to explore polymorphisms in the flanking regions. Four libraries were constructed using Dra I, EcoR V, Pvu II and Stu I enzymes to digest 2.5 μg of gDNA from MB1-73. Adaptors were ligated to the digested DNA fragments. Two gene specific primers (GSP, Table 4) and two adapters were used for nest PCR. The product from two of the four libraries was excised from the gel and subsequently purified, ligated, transformed and sequenced as described in the above section.

Sequence analysis
DNA sequences were analysed and single nucleotide polymorphisms (SNPs) were identified using the SEQUAM program (Lasergene 6, Madison, Wis.). Intron sizes were deduced by comparing the genome sequences with cDNA (ESTs) sequences or by putative conserved splicing patterns in plants. The phylogenetic tree was created and percentage sequence identity calculated using Clustal W by Megalign (Lasergene 6). The GenDoc program [http://www.psc.edu/biomed/genedoc](http://www.psc.edu/biomed/genedoc) was used for multiple DNA and amino acid sequence alignment.

Allele-specific markers and bin mapping approach

Three types of molecular markers were used: 1) two dominant single nucleotide amplification polymorphism (SNAP) markers [18] representing T and E respectively, 2) based on these sequences. Preferably, the forward and reverse primers covered the whole region of the gene encoding sequences (Table 4). Detailed PCR cloning and sequencing procedures have been described previously [19], here we mention only the key points and changes.

First, the melting temperature for amplification was optimized by gradient PCR with hybrid (MB1-73) DNA, then PCR was performed in two steps with Pfu polymerase (Stratagene, La Jolla, Calif.) and Taq (Takara Biotechnology, Dalian, China) using an Eppendorf Mastercycler. The amplified fragments were purified by Qiaquick Gel Extraction Kit (Qiagen, Germany), ligated into the pGEM-T easy vector (Promega, Madison, Wis.) and used to transform DH5α competent cells (Takara Biotechnology, Dalian, China) according to the manufacturer's protocol. For each purified fragment, 6–8 white colonies were selected for plasmid isolation and sequencing by Invitrogen (Shanghai, China) using an ABI 3730 (Applied Biosystems, Foster City, Calif.) platform. If there are more than two distinct sequences with initial one primer pair, additional colonies were sequenced to explore new gene members. All primers were synthesised by Invitrogen (Shanghai, China).

Genome walking

Genome walking was applied to gDNA of MB1-73 using the Universal Genome Walker kit (Clontech, Palo Alto Calif.) for gene sequences that were not cover the ORF or to avoid the mistakes induced by primer sequence, or to explore polymorphisms in the flanking regions. Four libraries were constructed using Dra I, EcoR V, Pvu II and Stu I enzymes to digest 2.5 μg of gDNA from MB1-73. Adaptors were ligated to the digested DNA fragments. Two gene specific primers (GSP, Table 4) and two adaptors were used for nest PCR. The product from two of the four libraries was excised from the gel and subsequently purified, ligated, transformed and sequenced as described in the above section.
Table 4: PCR primer pairs used for genomic cloning of the different peach/almond allergen genes and for (bin) mapping the genes

| Primers group | Reference Sequence | Primer name | Primer sequence (5' – 3') | Tm °C/cycles | Product (nt) | Gene       |
|---------------|-------------------|-------------|---------------------------|--------------|-------------|------------|
|               |                   | Pfu Taq     |                           |              |             |            |
| 1             | DQ251187          | PpPr10f     | ACCATGGGCTTCTCACATA        | 56/30        | 586/623     | Pru p/du 1.01/05 |
|               |                   | PpPr10r     | AATTTAGTTAGGCACTCGG        | 53/30        | 762/765     | Pru p/du 1.02 |
| 2             | DY646736          | Pp1-2f      | GTTTGCGTTTAAACTCTCC        | 53/30        | 624         | Pru p/du 1.03 |
|               |                   | Pp1-2r      | GATTAGTTAGGCACTCGG         | 53/30        | 624         | Pru p/du 1.03 |
| 3             | DY647300          | Pp1-3f      | CTTTGATCGTTTCCCCAAGT      | 53/30        | 615/622     | Pru p/du 1.04 |
|               |                   | Pp1-3r      | TTAGGTGACTAGCATGCGG        | 53/30        | 615/622     | Pru p/du 1.04 |
| 4             | DY653062          | Pp1-4f      | TTTACAGAACTAGGTTGTG        | 53/30        | 615/622     | Pru p/du 1.04 |
|               |                   | Pp1-4r      | TTAGGTGACTAGCATGCGG        | 53/30        | 615/622     | Pru p/du 1.04 |
| 5             | GW1                | Pp1-1A-GSP1 | AACCTTCACCAAAAGCTGTTCTTGGT | 57/30        | 1100        | Pru p/du 1.05 |
|               |                   | Pp1-1A-GSP2 | GTCAGATCCAGAACAAGGCTTTGAAA | 58/30        | 1546        | Pru p/du 1.06A/C |
| 6             | GW2                | P1-1A3-GSP1 | GGGTCATCGTCTGTGTTAACCT    | 58/30        | 1546        | Pru p/du 1.06A/C |
|               |                   | P1-1A3-GSP2 | CCACCTCACTACCCCCCAACAGA   | 58/30        | 1546        | Pru p/du 1.06A/C |
| 7             |                   | P1-1-6      | CACCTCCAGTCTCATAGC         | 57/30        | 720         | Pru p/du 1.05 |
|               |                   | PpPr10r     | See above                 | 57/30        | 593         | Pru p/du 1.06A/B |
| 8             |                   | P1-7f       | ATGGGCTCTTCACATCTCAG       | 57/30        | 593         | Pru p/du 1.06A/B |
|               |                   | P1-7r       | TTGATTTAGGCTAGCATGCTG     | 57/30        | 593         | Pru p/du 1.06A/B |
| 9             |                   | P1-8f       | TAGTTATAGGCTGTTGCAATGCT    | 58/30        | 1346        | Pru p/du 1.06C |
|               |                   | P1-8r2      | GGCGATCGTTCTGTGTTAACCT    | 58/30        | 1346        | Pru p/du 1.06C |
| 10            |                   | P1-9f       | See above                 | 58/30        | 1908        | Pru p/du 1.06A |
|               |                   | P1-9r       | GAAGCTTCAACTACATGTC        | 58/30        | 1908        | Pru p/du 1.06A |
| 11            | AF362988           | PpTL2f      | ACACGCAAGCCTAATTAGCAGA    | 56/30        | 879/892     | Pru p/du 2.01A/B |
|               |                   | PpTL2r      | GCTATGCGAGAGAATGTAGGAGG    | 56/30        | 879/892     | Pru p/du 2.01A/B |
| 12            | AF362987           | PpTL1f      | ATGGGATGACCGCTTGAGAGCG    | 56/30        | 974         | Pru p/du 2.02 |
|               |                   | PpTL1r      | GTCGAGAGCACTCCTGATATTT    | 56/30        | 974         | Pru p/du 2.02 |
| 13            | GW3                | PTL1-GSP1   | GATAATCGAGATGTTCAAGAGAGGC   | 52/30       | 500/1500    | Pru p/du 2.02 |
|               |                   | PTL1-GSP2   | AGAAGTTGGAAGCTGAGAGAGGACTG | 52/30       | 500/1500    | Pru p/du 2.02 |
| 14            | BF717226,          | PpTLP3f     | AAGTTAAAGGCTTCTTGGCCTG     | 53/30        | 659         | Pru p/du 2.03 |
|               | DW349103           | PpTLP3r     | CACGTGCCACCAAGACCATC       | 53/30        | 659         | Pru p/du 2.03 |
|               | GW4                | PruTLP3-GSP1 | AGTACTCACCCACCTCCTCCTCGCTG   | 53/30       | 800/1000    | Pru p/du 2.03 |
|               |                   | PruTLP3-GSP2 | GCTACCTTGAAGCTTGATCGGTC     | 53/30       | 800/1000    | Pru p/du 2.03 |
| 15            | Consensus of 29 ESTs | PpTLP4f     | ACTTTGCTGAATCTGGCTCGG     | 58/30        | 1800/1825   | Pru p/du 2.04 |
|               |                   | PpTLP4r     | GAAGTTGGAAGAGCGAGAGCGG     | 58/30        | 1800/1825   | Pru p/du 2.04 |
| 16            |                   | AY072050    | ATCATGATCAAGAGAGATGG       | 52/30        | 645         | Pru p/du 3.01 |
|               |                   | X96716      | CTGACAAGCTGAGAGAGAGAGC     | 52/30        | 645         | Pru p/du 3.01 |
| 17            | AY093699           | PpLTP2f     | TACCTGATGTTCTTGGAGAGGA    | 58/30        | 513/514     | Pru p/du 3.02 |
|               |                   | PpLTP2r     | CTGACTTCTTAGAAGAAGCTC     | 58/30        | 513/514     | Pru p/du 3.02 |
| 18            | AY002185           | PpLTP3f     | CCCACGGAGCAAGACACCATGA    | 58/30        | 527/528     | Pru p/du 3.03 |
|               |                   | PpLTP3r     | CATCTGACACTTCTCTTCCA      | 58/30        | 527/528     | Pru p/du 3.03 |
| 19            | AJ491881           | Pp4.01f     | AGAAGAAAACTAACAGAGCAAAC   | 51/30        | 1041        | Pru p/du 4.01 |
|               | AJ491882           | Pp4.01r     | AAAAGTCACTGCGAGCAAT       | 51/30        | 1041        | Pru p/du 4.01 |
|               |                   | Pp4.02f     | CAGCAAAACAACAAAGAATTG     | 53/30        | 754         | Pru p/du 4.02 |
|               |                   | Pp4.02r     | TCTAGAAGACATCTCCTCAAC     | 53/30        | 754         | Pru p/du 4.02 |

1GW-genome walking, only gene specific primers (GSP) shown, according to sequences obtained with previous primer group, representative reference EST accessions for group 16: [GenBank:DY634569; DY644567; DY651848; DY634415]
co-dominant markers based on length differences due to Indels and variation in the number of simple sequence repeats (SSRs) and 3) SNP markers based on direct sequencing of the PCR products from 8 bin samples. Markers 1) and 2) were used for the mapping of Pru p 1, 2, and 4 genes and 3) for Pru p 3 genes.

With SNAP markers, the annealing temperature is usually critical. It was optimised by gradient PCR using gDNAs of ‘Texas’ almond and ‘Earlygold’ peach (the two parents of MB1-73) as contrasting templates, and then applied to genomic DNA of 8 bin samples. Indels and SSR markers were revealed by 6% PAGE and silver staining. For direct sequencing, after PCR amplification (GoTaq, Promega), an aliquot of the PCR reaction was loaded on an agarose gel to check for specific amplification. If the PCR produced a single fragment, the remaining PCR reaction was purified using Roche High Pure PCR Product purification kit (Germany). An aliquot of 0.5 μl to 5 μl of purified PCR was used as template for sequencing with forward and/or reverse primers. SNPs and/or indels were detected using Sequencher software, version 4.7.

For mapping, we followed the strategy of “selective” or ‘bin’ mapping [26,37] where, based on the data of a detailed map already constructed (in this case the TxE map), a few plants are selected having a high number of recombination breakpoints and an approximately even distribution of these breakpoints throughout the genome. The joint genotype of these plants for a new marker defines a single map position, allowing its mapping with less time and cost, and with a moderate loss of precision compared with the use of the whole population. For Prunus, a set of only eight plants: one of the parents (‘Earlygold’) the F1 hybrid (MB1-37) and six plants (# 5, 12, 23, 30, 34 and 83) from the TxE progeny, was selected by Howad et al. [26] allowing to establish the position of any marker to 67 fragments of the genome (“bins”) of an average size of 7.8 cM. The identification of each bin is unequivocal except for two pairs (2:45/3:04 and 5:41/8:30), which have an identical genotype and require the genotype of an additional F2 plant (#27) to separate them. By comparing the genotype appeared in H and E, we were able to distinguish T allele from that of E. Marker tests were performed on these eight individuals of the bin set that were genotyped as homozygous for the ‘Texas’ allele (noted as A), homozygous for the ‘Earlygold’ allele (B) or heterozygous as MB1-73 (H). The resulting segregation pattern was compared to the defined data for each bin, and the matching bin was identified [26].

Nomenclature of the putative allergen sequences

We followed the systematic allergen nomenclature [38] for the four different protein classes and encoding genes. Generally, allergens are designated by the first three letters of the name of the genus, followed by a space and the first one or two letters of the species name. These letters are usually followed by a space and an Arabic number according to the order in which the allergens were identified, whereby the same number is used to designate homologous allergens of related species. Additional numbers and letters can be added to distinguish different isoallergens and variants [19]. Accordingly, PR-10, TLP, LTP and PRF proteins are denoted as Pru p 1 to 4 in peach and Pru du 1 to 4 in almond. Amino acid sequences with more than 5% dissimilarity usually belong to a different isoallergen and are distinguished by a two-digit code. If there are several genes within isoallergen group (on the DNA sequence identity, usually more than 95%), they are further distinguished by adding a capital letter after their isoallergen name, such as Pru p 2.01A and Pru p 2.01B. This procedure has been employed in apple [17] and birch [39]. The plant material in this study is an interspecies hybrid between peach and almond; we assigned two alleles of the same gene locus by its original species source, from Prunus persica as Pru p and from Prunus dulcis as Pru du. Therefore, Pru p/du refers to a gene in interspecies hybrid parent MB1-37.

References

1. Hoffmann-Sommergruber K., the SAFE consortium: The SAFE project: ‘plant food allergies: field to table strategies for reducing their incidence in Europe’ an EC-funded study. Allergy 2005, 60(4):436-442.
2. Fernandez-Rivas M, Bolhaar S, Gonzalez-Mancebo E, Asero R, van Leeuwen A, Bohle B, Ebner C, Rigby N, Sancho AI, Miles S, Zuidmeer L, Knulst A, Breiteneder H, Mills C, Hoffmann-Sommergruber K, van Ree R: Apple allergy across Europe: How allergen sensitization profiles determine the clinical expression of allergies to plant foods. Journal of Allergy and Clinical Immunology 2006, 118(2):481-488.
3. Karamlou F, Scheurer S, Wangorsch A, May S, Haustein D, Vieths S: Pyr c 1, the major allergen from pear (Pyrus communis), is a new member of the Bet v 1 allergen family. J Chromatogr B Biomed Sci Appl 2001, 756(1-2):281-293.
4. Fernandez-Rivas M, van Ree R, Cuevas M: Allergy to Rosaceaee fruits without related pollinosis. Journal of Allergy and Clinical Immunology 1997, 100(6):728-733.
5. Pastorella EA, Farioli L, Pravettoni V, Bertoli E, Conti A: The major allergen of peach (Prunus persica) is a lipid transfer protein. Journal of Allergy and Clinical Immunology 1999, 103(3):520-526.

6. Scheurer S, Metzner K, Haustein D, Vieths S: Molecular cloning, expression and characterization of Pru av 1, the major cherry allergen. Molecular Immunology 1997, 34(8-9):619-629.

7. Scheurer S, Lauter I, Faustok K, Mutschler W, Harcz C, Enrique L, Lidsell J, Cistero-Bahina A, Vieths S: Strong allergenicity of Pru av 3, the lipid transfer protein from cherry, is related to high stability against thermal processing and digestion. Journal of Allergy and Clinical Immunology 2004, 114(4):900-907.

8. Pastorella EA, Farioli L, Pravettoni V, Giuffrida MG, Ortolani C, Fortunato D, Trambiaoli C, Scibola E, Calamari AM, Robino AM, Conti A: Characterization of the major allergen of plum as a lipid transfer protein. J Chromatogr B Biomed Sci Appl 2001, 756(1-2):123.

9. Pastorella EA, Purello D Ambrosio F, Pravettoni V, Farioli L, Giuffrida G, Monza M, Ansaloni R, Fortunato D, Scibola E, Rivolta F, Incorvaia C, Bengtsson A, Conti A, Ortolani C: Evidence for a lipid transfer protein as the major allergen of apricot. Journal of Allergy and Clinical Immunology 2000, 105(2):371-377.

10. Tawde P, Venkatesh YP, Wang F, Teuber SS, Sathe SK, Roux KH: Cloning and characterization of profilin (Mal d 4) from almond (Prunus dulcis) cultivars and dependence of their content on fruit ripening. Journal of Agricultural and Food Chemistry 2004, 52(26):7977-8000.

11. Kosters SS, Villeg-Boerstra BJ, Dubois AEJ: Assessment of the reduced allergenic properties of the Santana apple. Ann Allergy Asthma Immunol 2007, 99(4):522-525.

12. Breiteneder H, Ebner C: Molecular and biochemical classification of plant-derived food allergens. Journal of Allergy and Clinical Immunology 2000, 106(1):27-36.

13. Bohle B: The impact of pollen-related food allergens on pollen allergy. Allergy 2007, 62(1):3-10.

14. Gamboa PM, Caceres O, Antepara I, Sanchez-Monge R, Duffort O, Barber D, Salcedo G: Two different profiles of peak allergy in the north of Spain. Allergy 2007, 62(4):408-414.

15. Wen ZM, Ye ST: A report of 50 patients with artemisia polleinosis and plant food allergy. National Medical Journal of China 2002, 82(9):626-629. (in Chinese with English abstract)

16. Gao Z, Zheng W, Wang Z, Liu Q, Schenkel MF, Gilissen LJW: Genomic cloning and linkage mapping of the Mal d 1 (PR-10) gene family in apple (Malus domestica). Theoretical and Applied Genetics 2004, 109(1):171-183.

17. Gao Z, Zheng W, Wang Z, Liu Q, Schenkel MF, Gilissen LJW: Genomic characterisation and linkage mapping of the apple allergen genes Mal d 2 (thiamin-like protein) and Mal d 4 (profilin). Theoretical and Applied Genetics 2003, 107(6):1087-1097.

18. Gao Z, Zheng W, Wang Z, Liu Q, Schenkel MF, Gilissen LJW: Linkage map positions and allelic diversity of two Mal d 3 (non-specific lipid transfer protein) genes in the cultivated apple (Malus domestica). Theoretical and Applied Genetics 2005, 110(1):181.

19. Gao Z, Zheng W, Wang Z, Liu Q, Schenkel MF, Gilissen LJW: Linkage map positions and allelic diversity of two Mal d 3 (non-specific lipid transfer protein) genes in the cultivated apple (Malus domestica). Theoretical and Applied Genetics 2005, 110(3):479-491.

20. Gao Z, Zheng W, Wang Z, Liu Q, Schenkel MF, Gilissen LJW: Linkage map positions and allelic diversity of two Mal d 3 (non-specific lipid transfer protein) genes in the cultivated apple (Malus domestica). Theoretical and Applied Genetics 2005, 110(3):479-491.

21. Gao Z, Zheng W, Wang Z, Liu Q, Schenkel MF, Gilissen LJW: Linkage map positions and allelic diversity of two Mal d 3 (non-specific lipid transfer protein) genes in the cultivated apple (Malus domestica). Theoretical and Applied Genetics 2005, 110(3):479-491.

22. Ahrazem O, Jimeno L, Lopez-Torrejon G, Herrero M, Espada JL, Sanchez-Monge R, Duffort O, Barber D, Salcedo G: Assessing allergen levels in peach and nectarine cultivars. Allergy Asthma Immunol 2007, 99(1):42-47.

23. Pastorella EA, Vegera M, Delso J, Pasini G, Ramina A, Gemignani C, Marcer G, Pravettoni V, Tonutti P: Different expression of Pp-LTP1 and accumulation of Pru p 3 in fruits of two Prunus persica L. Batsch genotypes. Plant Science 2006, 171(1):106-113.

24. Breiteneder H, Scibola E, Ansaloni R, Fortunato D, Calamari AM: Multiplex mapping with a few plants: Using selective mapping for microsatellite saturation of the Prunus reference map. Genetics 2005, 171(3):1303-1309.

25. Buttolon A, Beghldo M, Rassor A, Bonghi C, Tonutti P: Differential expression of two lipid transfer protein genes in reproductive organs of peach (Prunus persica L. Batsch). Plant Science 2002, 163(5):993-1000.

26. Piazzia LF, Tordessilas L, Cuesta-Herranz J, Compes E, Sanchez-Monge R, Palacin A, Salcedo G, Diaz-Perales A: Mimotope mapping as a complementary strategy to define allergen IgE-epitopes. Peach Pru p 3 allergen as a model. Molecular Immunology 2008, 45(8):2269-2276.

27. Ahrazem O, Barcaccia G, Rami A: Genetic and environmental factors affecting allergen-related gene expression in apple fruit (Malus domestica L. Borkh). Journal of Agricultural and Food Chemistry 2008, 56(7):6707-6716.

28. Botton A, Lezzer P, Dorigoni A, Barcaccia G, Rami A: Evidence for a lipid transfer protein as the major allergen of apricot. Journal of Allergy and Clinical Immunology 2000, 105(2):371-377.

29. Botton A, Beghldo M, Rassor A, Bonghi C, Tonutti P: Differential expression of two lipid transfer protein genes in reproductive organs of peach (Prunus persica L. Batsch). Plant Science 2002, 163(5):993-1000.

30. Piazzia LF, Tordessilas L, Cuesta-Herranz J, Compes E, Sanchez-Monge R, Palacin A, Salcedo G, Diaz-Perales A: Mimotope mapping as a complementary strategy to define allergen IgE-epitopes. Peach Pru p 3 allergen as a model. Molecular Immunology 2008, 45(8):2269-2276.

31. Botton A, Lezzer P, Dorigoni A, Barcaccia G, Rami A: Genetic and environmental factors affecting allergen-related gene expression in apple fruit (Malus domestica L. Borkh). Journal of Agricultural and Food Chemistry 2008, 56(7):6707-6716.

32. Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dune-maart F, Guifford P, Gardiner S, Guilford P, Van Heusden AW, Ju et al. 2008, 653-640.

33. Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink van Ginkel M, King GJ: Aligning male and female linkage maps of apple (Malus pumila Mill.) using multiple-allele markers. Theoretical and Applied Genetics 1998, 97(1-2):60-73.

34. Cleton JM, Tustin S, Chapnê D, Gardiner S: Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from Malus ESTs and Pyrus genomic sequences. Theoretical and Applied Genetics 2008, 16:407-420.

35. Phipps JB, Robertson KR, Rohrer JR: Origins and evolution of subfam. Maloideae (Rosaceae). Systematic Botany 1991, 16:303-332.

36. Viruel MA, Messeguer R, de Vicente MC, Garcia-Mas J, Pujol-Monech P, Vargas F, Arus P: A linkage map with RFLP and isozyme markers for almond. Theoretical and Applied Genetics 1995, 91:964-971.

37. Vision T, Brown DG, Shmoys DB, Durrett RT, Tanksley SD: Selective mapping: a strategy for optimizing the construction of high-density linkage maps. Genetics 2000, 155(1):407-420.

38. King TP, Hoffman D, Lowenstein H, Marsh DG, Platts TAE, Thomas W: Allergen nomenclature. Journal of Allergy and Clinical Immunology 1995, 95:5-14.

39. Schenk MF, Gilissen L, Esselink GD, Smulders MIM: Seven different genes encode a diverse mixture of isoforms of Bet v 1, the major birch pollen allergen. BMC Genomics 2006, 7:168.