ORIGINAL PAPER

Z. S. Gao · W. E. van de Weg · J. G. Schaart
H. J. Schouten · D. H. Tran · L. P. Kodde
I. M. van der Meer · A. H. M. van der Geest · J. Kodde
H. Breiteneder · K. Hoffmann-Sommergruber · D. Bosch
L. J. W. J. Gilissen

Genomic cloning and linkage mapping of the *Mal d 1* (*PR-10*) gene family in apple (*Malus domestica*)

Abstract Fresh apples can cause birch pollen-related food allergy in northern and central European populations, primarily because of the presence of Mal d 1, the major apple allergen that is cross-reactive to the homologous and sensitizing allergen Bet v 1 from birch. Apple cultivars differ significantly in their allergenicity. Knowledge of the genetic basis of these differences would direct breeding for hypoallergenic cultivars. The PCR genomic cloning and sequencing were performed on two cultivars, Prima and Fiesta, which resulted in 37 different *Mal d 1* gDNA sequences. Based on the mapping of sequence-specific molecular markers, these sequences appeared to represent 18 *Mal d 1* genes. Sixteen genes were located in two clusters, one cluster with seven genes on linkage group (LG) 13, and the other cluster with nine genes on the homoeologous LG 16. One gene was mapped on LG 6, and one remained unmapped. According to sequence identity, these 18 genes could be subdivided into four subfamilies. Subfamilies I–III had an intron of different size that was subfamily and gene-specific. Subfamily IV consisted of 11 intronless genes. The deduced amino acid sequence identity varied from 65% to 81% among subfamilies, from 82% to 100% among genes within a subfamily, and from 97.5% to 100% among alleles of one gene. This study provides a better understanding of the genetics of *Mal d 1* and the basis for further research on the occurrence of allelic diversity among cultivars in relation to allergenicity and their biological functions.

Z. S. Gao · W. E. van de Weg · L. J. W. J. Gilissen
Allergy Consortium Wageningen,
Wageningen University and Research Centre, P.O. Box 16,
6700AA Wageningen, The Netherlands

Z. S. Gao
Experimental Nursery of Shanxi Forestry Department,
Beiying Nan Lu 52, Taiyuan, Shanxi 030031,
People's Republic of China

Z. S. Gao · W. E. van de Weg (✉) · J. G. Schaart
H. J. Schouten · D. H. Tran · L. P. Kodde
I. M. van der Meer · A. H. M. van der Geest · J. Kodde
D. Bosch · L. J. W. J. Gilissen
Plant Research International,
Wageningen University and Research Centre, P.O. Box 16,
6700AA Wageningen, The Netherlands
E-mail: eric.vandeweg@wur.nl
Tel.: +31-317-477281
Fax: +31-317-418094

H. Breiteneder · K. Hoffmann-Sommergruber
Department of Pathophysiology, Medical University of Vienna,
AKH-EBO-3Q, Währinger Gürtel 18-20, 1090 Vienna, Austria

Present address: D. H. Tran
Department of Biology, Norwegian University of Science and
Technology, 7491 Trondheim, Norway

Introduction

Fresh apples can cause birch pollen-related food allergy in northern and central European populations. About 50–75% of birch pollen-sensitized patients suffer from an oral allergy syndrome after eating apples (Ebner et al. 1991; Pauli et al. 1996). This allergy results from IgE-mediated cross-reactivity between Bet v 1, the major birch allergen, and Mal d 1, the major apple allergen. Both Bet v 1 and Mal d 1 belong to a group of pathogenesis-related (PR)10 proteins (Breiteneder and Ebner 2000). The *PR-10* genes appear abundantly in plants (Wen et al 1997; Hoffmann-Sommergruber and Radauer 2004; Liu and Ekramoddoullah 2004). Expression of these genes is induced by stress, such as attack by plant pathogens, and occurs during ripening (Atkinson et al. 1996; Puehringer et al. 2000).

Mal d 1 has been identified as a 17- to 18-kDa protein of 158–159 amino acids encoded by 480–483 nucleotides [(nt) Vieths et al. 1995; Schoning et al. 1996; Hoffmann-Sommergruber et al. 1997]. Over 30 *Mal d 1* DNA

sequences obtained from cDNA and gDNA of leaves and fruits of various apple cultivars (Vanek-Krebitz et al. 1995; Hoffmann-Sommergruber et al. 1997; Son et al. 1999; Holm et al. 2001; Ziadi et al. 2001) were published in the database GenBank at the beginning of this study. Their phylogenetic tree indicated four distinct groups with seven members (Puehringer et al. 2003). Recently, Beuning et al. (2004) sequenced cDNA libraries of cultivar Royal Gala and identified 12 *Mal d 1*-related genes, five of which were new. The *Mal d 1* family seemed to consist of at least 15 members (Atkinson et al. 1996), which means that at least three *Mal d 1* homologous genes have not been cloned and sequenced. To date, neither is there a conclusive answer to the total number of members in the *Mal d 1* gene family in the apple genome, nor an answer about their positions on the linkage map.

Apple cultivars differ considerably in allergenicity, e.g., Golden Delicious is highly allergenic as experienced by many apple-allergy patients, whereas Gloster causes only mild reactions (Vieths et al. 1994; Hsieh et al. 1995). It is likely that genetic factors are involved. For breeding hypoallergenic apple cultivars, it is essential to understand these differences at the qualitative and quantitative level of Mal d 1 isoallergens and their variants. As a first step, we examined the number of *Mal d 1* genes present in the genome and their locations on molecular marker linkage maps of apple. For this purpose, we first cloned and sequenced *Mal d 1* genes from genomic DNA of two parental cultivars, Prima and Fiesta. Then we developed sequence-specific markers for mapping. This strategy enabled us to characterize genomic sequences and to determine 18 *Mal d 1* genes and their map positions.

Materials and methods

Plant materials

The PCR cloning and sequencing was conducted on two cultivars, Prima (PM) and Fiesta (FS). These two cultivars were chosen because they are used as parent in three available mapping populations: PM × FS (Maliepaard et al. 1998), Jonathan (JO) × PM, and FS × Discovery (DS). With regard to the allergenicity of four parental cultivars, skin-prick tests on Dutch apple allergenic patients revealed FM and DS to be high and PM and JO to be intermediate allergenic (W.E. van de Weg et al., unpublished data).

PCR primers

Primer pairs were designed using the software program Primer Designer, version 2.0 (Scientific and Educational Software, Cary, N.C., USA) based on all available sequence information of *Mal d 1* genes. Preferably, the forward and reverse primers covered the whole length of the gene-coding sequences, but primers for getting the middle part of a gene were also tried if the first primers failed to amplify the expected product or sequences. Both conserved and specific primers were used to obtain all possible *Mal d 1* sequences. Some of the cloning primers were adjusted after the first round of cloning and sequencing. The *Mal d 1* cloning primers were first derived from the sequences in the GenBank. Then new cloning primers were added when new *Mal d 1*-like sequences were obtained from unpublished research on gene regulation in apple stock root induced by wounding and auxin at Plant Research International, Wageningen, The Netherlands. Primers for genome walking were designed according to new sequences obtained in this research. (All these primer sequences are listed in Table 1.)

DNA isolation, PCR cloning, and sequencing

Genomic DNA was extracted using the CTAB-based, large-scale nuclei-isolation method (Roche et al. 1997). The PCR cloning and sequencing procedures have been described previously (Gao et al. 2005). Here we mention only the key points and some changes. The PCR was performed in two steps with *Pfu* polymerase (Stratagene, La Jolla, Calif., USA) and super *Taq* (HT Biotechnology, UK) using a PTC-200 machine (MJ Research, Waltham, Mass., USA). The amplified fragments were purified by Qiaquick Gel Extraction Kit (Qiagen, Germany) and ligated into the pGEM-T easy vector (Promega, Madison, Wis., USA) and used to transform XL1 Blue competent cells (Stratagene) according to the protocols of the manufacturer. For each fragment, eight to ten white colonies were selected for the isolation of plasmid DNA using the Qiaprep Turbo BioRot Kit (Qiagen) by a Bio Robot 9600 (Qiagen, USA). The DNA sequencing was performed on a 96-capillary system (ABI 3700; Applied Biosystems, Foster City, Calif., USA). If one primer pair produced more than two different sequences, then additional clones were sequenced, or PCR cloning was performed again in order to obtain enough replicated sequences.

Genome-walking approach

Genome walking was applied to gDNA of cvs. PM and FS using the Universal Genome walker kit (Clontech, Palo Alto, Calif., USA) to get precise sequences at the two ends of a gene and its flanking region. For each cultivar, four libraries were constructed using *Dra*I, *Eco*RV, *Pvu*II, and *Stu*I enzymes to digest 2.5 μg of the gDNA. Adaptors were ligated to the digested DNA fragments. Four groups of gene-specific primers (Table 1) were designed and used for nested PCR together with two adapter primers (AP1 and AP2). The product obtained from one of the four libraries was excised from the gel and subsequently purified, ligated, transformed, and sequenced as already described.

Sequence analysis

The DNA sequences and single nucleotide polymorphisms (SNPs) were analyzed using the SEQMAN program (DNAstar, Madison, Wis., USA). Intron sizes were deduced by comparing the genome sequences with known cDNA sequences or by putative splicing patterns. The phylogenetic tree was created and sequence identity percentages were calculated using the Clustal W by Megalign program (DNAstar). Multiple DNA and amino acid sequence alignments were performed with the GeneDoc program (http://www.psc.edu/biomed/genedoc).

Designing and testing of sequence-specific markers

Two types of molecular markers were used to distinguish a specific sequence or allele in the context of the PM × FS or the JO × PM population: single nucleotide amplification polymorphism (SNAP) markers (Drenkard et al. 2000; Gao et al. 2005) and simple sequence repeat (SSR) markers. The SNAP markers were tested first on gDNA of PM, FS, and eight individuals of their population to confirm PCR conditions, expected product size and segregation pattern. Then the well-working markers were tested on the entire population. Some markers were also applied to JO × PM to map several Mal d 1 genes that did not segregate in PM × FS or to confirm the results of PM × FS.

The SSR primers were designed for regions flanking the repetitive stretch. The reverse primers had so called pig-tails, i.e., GTTT at the 5′ end according to Brownstein et al. (1996). Primer labeling, PCR amplification, and gel electrophoresis were performed as described previously (Gianfranceschi et al. 1998).

Mapping genes on molecular linkage groups

Two molecular marker linkage maps, PM × FS (population size $n = 144$) and JO × PM (population size $n = 196$), were used to map sequence-specific molecular markers. Grouping and mapping were performed with JoinMap, version 3.0 (van Ooijen and Voorrips 2001), using the Kosambi mapping function. The LOD and recombination threshold was 4 and 0.45, respectively. Final drawings of the marker maps were generated with MapChart (Voorrips 2001).

## Nomenclature of the Mal d 1 sequences

After consulting the Allergen Nomenclature Committee, we differentiated isoallergen genes when their protein sequences displayed less than 95% identity and denoted these genes according to current allergen nomenclature (King et al. 1995). Genes with more than 95% DNA sequence identity that are also clustered on the same

linkage group (LG) were denoted by adding a capital letter to their isoallergen name, such as Mal d 1.03A, Mal d 1.03B, etc. Previous denotations of Mal d 1 isoallergens and variants (Mal d 1.01 to Mal d 1.04, http://www.allergen.org//isoall) were maintained as much as possible. Finally, we denoted silent mutations according to Gao et al. (2005).

Results

Generating and grouping Mal d 1 sequences

Fifteen primer pairs (including four for genome walking) were used for PCR amplification using gDNA of two cultivars, PM and FS (Table 1), which resulted in about 300 raw Mal d 1 sequences. These sequences could be aligned into 43 different sequences by the SEQMAN program. Most of our sequences have been identified from at least two clones. Those derived from a single clone had crosschecks to the GenBank or were confirmed by SNAP marker tests. Six sequences with PCR errors were excluded. Ultimately, 37 correct sequences could be deduced. One out of these 37 sequences represented a
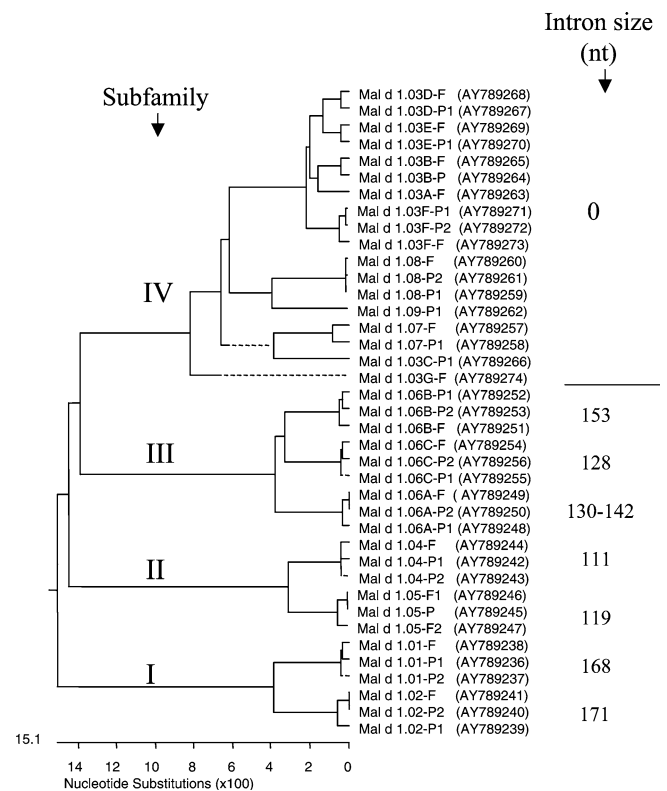
Fig. 1 Phylogenetic tree of 39 coding DNA sequences of the Mal d 1 gene family of the cultivars Prima (PM) and Fiesta (FS) and their intron size. The P and F behind the gene name indicate the source cultivar, P PM, F FS; 1 and '2 after the cultivar symbols P and F refer to different alleles of the same gene. Mal d 1.02-P2 sequence is the same as Mal d 1.02-F, Mal d 1.05-P is the same as Mal d 1.05-F1, and Mal d 1.06A-P2 is the same as Mal d 1.06A-F

Table 1 The PCR primer pairs used for cloning of the different *Mal d 1* genes

Primer pairs[a]	Reference Sequence[a]	Primer name[b]	Primer sequence (5'–3')	T_m (°C/no. cycles)[c]		Product (nucleotides)	Intron (nucleotides)	Isoallergen gene member
				Pfu	*Taq*			
1	AY026910	M1a-For M1a/b/c-Rev	ATCTCCAACACAATACTCTCAAC AAAGCCACACAACCTTCGAC	58/25	60/2	800	168	*Mal d 1.01*
2	L42952 AF020542	Mald1b-For M1b-Rev	CATCCTTGGTAGTTGCTTTC ACCATAGAAACATATTAATTTAGT	52/30	54/2	723	171	*Mal d 1.02*
3	Z48969	M1a/b/c-For	TCATGGGTGTCTACACATTTG	50/35	50/2	439; 483	0	*Mal d 1.03B;* *Mal d 1ps1*
4	Z72425–Z72428	M1c-Rev *HB20-For*[b]	ATTAATTTAGTTGTAGGCATCC ATGGGGTGTTTCAATTACGAAACTGAG	54/30	56/2	480	0	*Mal d 1.03A;* *Mal d 1.08;* *Mal d 1.09*
5	AY026908 AY026909	*HB22-Rev*[b] M1.03-For1	TTAGTTGTAGGCATCGGAGTGTGCCAAGAGG ATGGGGTGTTTTCACATACGAA	55/25	57/2	486	0	*Mal d 1.07;* *Mal d 1.08;* *Mal d 1.09;* *Mal d 1.03C,D, E, G*
6	Z72426–Z72428	M1.03-Rev1 M1.04-s27 M1.04-as496	AAGAATTTAGTTGTAGGCATC GWYCACCTCTGTCATCCCTG TCTCAACATCACCCTTAGCA	56/30	58/2	470	111	*Mal d 1.04*
7	Z72426–Z72428	M1.04-s62	AGGCYTTTATCCTTGATGGC	56/30	58/2	492; 500	111; 119	*Mal d 1.04;* *Mal d 1.05*
8	Based on 6	M1.04-as553 M1.04UP-GSP1 M1.04UP-GSP2	GAAGCTTGAAGAGACCGGAA CTCCTTCAATCATAGAGTAACTGTATGT CCAAAGGTAACTTCTTGTGGTTCC			747		*Mal d 1.04*
9	Based on 6	M1.04DW-GSP1 M1.04DW-GSP2	ACATACAGTTACTCTATGATTGAAGGA CATCAAAACCACTAGCCACTACCATGC			830		*Mal d 1.04*
10	Based on 7	M1.04UP-GSP1 M1.05UP-GSP2	See above GAGAGAATATTGAGGGTTGTCTAGCTAA			747		*Mal d 1.05*
11	Based on 7	M1.04DW-GSP1 M1.05DW-GSP2	See above ATTGACCGCATCTCCTGATGG			477		*Mal d 1.05*
12[c]	3262-PRI 3425-PRI	M1.06-For M1.06-Rev	CATGGGTGTCCTCACATACGAAAC CTAGTTGTAGGCATCAGGATTG	55/25	57/2	609–634	128–142	*Mal d 1.06A– C*
13	Based on 12	M1.07-fw M1.07-rev	ATGGGGTGTCCTCACATACGAAACT TTAGTTGTAGGCATCAGGATTGGCCACAAGGTG	62/30	64/2	608	128	*Mal d 1.06C*
14[c]	3076-PRI	M1.03-For2	ATCATGGGTGTTTTCACATAT	55/25	57/2	491	0	*Mal d 1.03A;* *Mal d 1.03B;* *Mal d 1.03F*
15	Based on 14	M1.03-Rev2 M1.03-For2	CATATAATTTAGTTGTAGGCGTCT See above	55/25	57/2	491	0	*Mal d 1.03A;* *Mal d 1.03B;* *Mal d 1.03F*
		M1.03-Rev3	CATATAATTTAGTTGTAGGCGT					

[a] Primer pairs 1–7 were based on *Mal d 1* sequences in the GenBank; 8–11 are gene-specific for genome walking to upstream (*UP*) and downstream (*DW*); 12–15 were derived from unpublished *Mal d 1* sequences at Plant Research International (*PRI*), Wageningen

[b] HB20-For and HB22-Rev are original primers of Hoffmann-Sommergruber et al. (1997)

[c] Annealing temperature

Table 2 Classification of all *Mal d 1* sequences according to the phylogenetic tree and map positions

Subfamily	Gene	Previous names[a]	New sequences (alleles) Prima	Fiesta	Reference sequences[b] (isoallergen and variants names)
I	*Mal d 1.01*	Mal d 1a/c/d; PR-10c, Mal d 1.01	AY789236 (Mal d 1.0105.01), AY789237 (Mal d 1.0105.03)	AY789238 (Mal d 1.0105.02)	X83672 (Mal d 1.0101), Z48969 (Mal d 1.0102), AF124823 (Mal d 1.0103), AF124829 (Mal d 1.0104), AF124830 (Mal d 1.0105), AF124831 (Mal d 1.0106), AF124832 (Mal d 1.0107), AF126402 (Mal d 1.0108), AY026910 (Mal d 1.0109), AJ417551, **AY428579**
	Mal d 1.02	Mal d 1b; PR-10a; Mal d 1.02	AY789239 (Mal d 1.0201.09), AY789240 (Mal d 1.0201.01)	AY789241 (Mal d 1.0201.01)	L42952 (Mal d 1.0201), AF020542 (Mal d 1.0206), AF074721, AF124822 (Mal d 1.0202), AF124824 (Mal d 1.0203), AF125825 (Mal d 1.0204), AF124826, AF124827, AF124828, AF124833, AF124834, AF124835 (Mal d 1.0205), AF124836, AY026911 (Mal d 1.0207), AJ488060 (Mal d 1.0208), **AY428578**
II	*Mal d 1.04*	Mal d 1.04	AY789242 (Mal d 1.0404.01), AY789243 (Mal d 1.0404.02a)	AY789244 (Mal d 1.0404.02b)	Z72426 (Mal d 1.0401), Z72427 (Mal d 1.0402), Z72428 (Mal d 1.0403)
	Mal d 1.05		AY789245 (Mal d 1.0501)	AY789246 (Mal d 1.0501), AY789247 (Mal d 1.0502)	
III	*Mal d 1.06A*	Mal d 1e; Mal d 1f	AY789248 (Mal d 1.06A02.01), AY789250 (Mal d 1.06A01.01b)	AY789249 (Mal d 1.06A01.01a)	**AY428580, AY428581**
	Mal d 1.06B	Mal d 1g; Mal d 1h	AY789252 (Mal d 1.06B02), AY789253 (Mal d 1.06B03)	AY789251 (Mal d 1.06B01)	**AY428582, AY428583**
	Mal d 1.06C		AY789255 (Mal d 1.06C02), AY789256 (Mal d 1.06C03)	AY789254 (Mal d 1.06C01)	
IV	*Mal d 1.07*	PR-10d; Mal d 1k; Mal d 1.03	AY789258 (Mal d 1.0702)	AY789257 (Mal d 1.0701)	AY026909 (Mal d 1.0303), **AY428586**
	Mal d 1.08		AY789259 (Mal d 1.0801.01), AY789261 (Mal d 1.0801.02)	AY789260 (Mal d 1.0801.01)	
	Mal d 1.09		AY789262 (Mal d 1.0901)		
	Mal d 1.03A	Mal d 1.03		AY789263 (Mal d 1.03A01)	Z72425 (Mal d 1.0301)
	Mal d 1.03B	Mal d 1.03	AY789264 (Mal d 1.03B01)	AY789265 (Mal d 1.03B02)	AY186248 (Mal d 1.0304)
	Mal d 1.03C	Mal d 1 11	AY789266 Mal d 1.03C01		**AY428587**
	Mal d 1.03D	PR-10b, Mal d 1.03	AY789267 (Mal d 1.03D02)	AY789268 (Mal d 1.03D01)	AY026908 (Mal d 1.0302)
	Mal d 1.03E	Mal d 1i	AY789270 (Mal d 1.03E01)	AY789269 (Mal d 1.03E02)	**AY428584**
	Mal d 1.03F	Mal d 1j	AY789271 (Mal d 1.03F01), AY789272 (Mal d 1.03F01)	AY789273 (Mal d 1.03F02)	**AY428585**
	Mal d 1.03G			AY789274 (Mal d 1.03G01)	
	Mal d 1ps1		AY789275		

[a] In the reference papers and Mal d 1 isoallergen list
[b] Accession numbers in **boldface** were released in the GenBank shortly before submission of this manuscript. Isoallergen and variant names in *parentheses* just behind a sequence were assigned by the Allergen Nomenclature Committee (see http://www.allergen.org)

pseudogene with a deletion of 44 nt in the middle. Of the 36 protein-coding sequences, 23 were present in PM and 16 in FS; three sequences were common to both cultivars (GenBank accessions AY789236–AY789275, Table 2). Apparently, PM is more heterozygous for *Mal d 1* than FS and is therefore more suitable to create markers for mapping. By comparison with the known reference cDNA sequences in the GenBank, the coding sequences and intron sizes (if present) could be deduced for the newly obtained *Mal d 1* genomic sequences. According to a phylogenetic tree of the 39 Mal d 1 protein-coding sequences, we classified the *Mal d 1* gene family into four subfamilies, which coincided with variation in intron size (Fig. 1). Subfamilies I–III contained members with a single intron, and the intron sizes are specific for each subfamily and gene, whereas subfamily IV included only intronless gene members.

By Clustal W alignment, all *Mal d 1* coding sequences in the database could be classified to these four subfamilies and members (Table 2). Subfamily I includes reference sequences that were formerly classified as Mal d 1a, Mal d 1b, Mal d 1c (Son et al. 1999), and Mal d 1 d (Beuning et al. 2004) isoforms or as PR-10a and PR-10c proteins (Ziadi et al. 2001). The new sequences of this subfamily were derived from the PCR primer pairs 1 and 2 (Table 1). Subfamily II includes representatives of the previously denoted isoallergen group Mal d 1.04 (Puehringer et al. 2003). The new sequences were derived from primer pairs 6 and 7 and subsequent genome walking in upstream and downstream directions (primer pairs 8–11, Table 1). Subfamily III contains sequences similar to Mal d 1e, Mal d 1f, Mal d 1g, and Mal d 1h (AY42580–AY42583, Beuning et al. 2004). Our sequences were obtained from primer pairs 12 and 13 based on two unpublished cDNA sequences (Tables 1, 2). Subfamily IV includes four reference sequences that were formerly classified as Mal d 1.03 (Puehringer et al. 2003) and four recently released sequences Mal d 1i, Mal d 1g, Mal d 1k, and Mal d 1l (AY428584–AY428587, Beuning et al. 2004). The new sequences were derived from primer pairs 3–5, 14, and 15 (Table 1), which not only represent all the reference sequences, but also include some new members.

Sequence-specific markers and mapping

To distinguish all 36 putatively functional sequences from PM and FS and a pseudogene sequence from PM, we created 34 SNAP markers and three SSR markers (Table 3). In addition, one SNAP marker (Mal d 1.03C02-GD) was developed according to a sequence from cv. Golden Delicious (GenBank accession AY822725). These markers were tested on the parental cultivars and on the mapping populations. As an example, Fig. 2 shows the specificity and segregation of the SNAP marker for AY78936 (Mal d 1.0105.01a). Most markers from PM were mapped in the population PM × FS, seven markers were homozygous in FS but
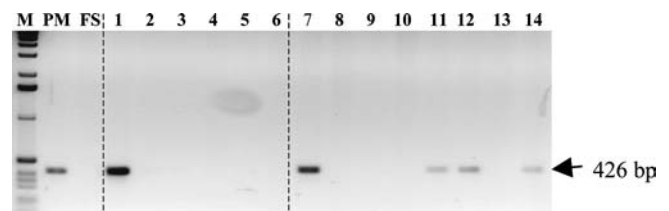
Fig. 2 Specificity and segregation of a Mal d 1.01 marker for sequence AY789236 (*from PM*). Lanes (*from left to right*): *M* Molecular 1-kb ladder; *1* positive cloned DNA of sequence AY789236; *2* and *3* cloned *Mal d 1.01* DNA of sequences AY789237 and AY789238, respectively; *4– 6* three *Mal d 1.02* cloned DNA; *7– 14* genomic DNA of eight descendants of the cross PM × FS

heterozygous in JO, which enabled their mapping in JO × PM (Table 3). The segregating markers allowed the mapping of 17 loci on three LGs: LG 6, LG13, and LG16 (Fig. 3). Seven genes were clustered around 0.6 cM on LG 13, including *Mal d 1.01* and six *Mal d 1.03* genes (*Mal d 1.03A–F*). *Mal d 1.01* and *Mal d 1.03F* were the anchor genes for the consensus *Mal d 1* gene cluster on LG 13 from JO and PM (Fig. 3). Nine *Mal d 1* genes were mapped in a cluster of 2 cM on LG 16, whereas one gene (*Mal d 1.05*) was mapped on LG 6. Within the *Mal d 1* gene cluster on LG 13, one recombination event was found between two subclusters, each of them having three or four genes located at identical map position (no recombination events were observed) (Fig. 3, LG13-cons). Similarly, two subclusters were found on LG16, one including *Mal d 1.02*, *Mal d 1.04*, *Mal d 1.07*, *Mal d 1.08*, and *Mal d 1ps1*, the other including three similar *Mal d 1.06* genes (Fig. 3, LG16-PM). The marker for *Mal d 1.03G* (AY789274) could not be mapped, because it did not segregate in any of the three mapping populations. Thus, this unmapped sequence represents a different gene rather than a different allele of the mapped genes. We fully clarified the allelic constitution of all seven intron-containing genes of PM and FS as of 3 of the 11 intronless genes.

Genome characteristics of *Mal d 1* isoallergen genes

Sequence differences and similarities among *Mal d 1* genes can be examined for the coding, intron, and upstream regions. All 17 identified *Mal d 1* genes have a coding region of 480 nt, except for *Mal d 1.04* and *Mal d 1.05* in subfamily II with 483 nt, as in *Bet v 1*. Comparison of our new *Mal d 1* coding sequences revealed different levels of identity: 71–83% among the four subfamilies, 86–98.1% among genes within a subfamily, and 98.3–100% among alleles of a single gene. Sequence identity for the intronless genes on LG 13 was over 95%. Some characteristic polymorphic nucleotides for each subfamily or each gene could be found by alignment of all *Mal d 1* genomic sequences. For instance, counted from the ATG start codon, 174T, 239A, 240C, 241T,

Table 3 Description of sequence-specific primer pairs for *Mal d 1* isoallergen genes

Sequence and marker name	Primer sequence (5′–3′)[a]	Position[b] (gDNA)	PCR (nucleotides)	T_m[c]	Cycles[d]	Presence and segregation[e]	LG[f]
AY789236 Mal d 1.0105.01a	Fw-AGCTGAAATCCTTGAA CGAA Rv-CAATGTTTCCCTGGTG AGA	123–142 528–547	426	55/53	10/30	PM (PM × FS)	13
AY789237 Mal d 1.0105.03	Fw-TGAAGCACAGGATTGAC ACG Rv-CAATGTTTCCCTTGGTGT CG	371–390 528–547	177	57/55	10/30	PM (PM × FS)	13
AY789238 Mal d 1.0105.02	Fw-TGAAGCACAGGATTGAC GCA Rv-CCACACAACCTTCGACT CA	371–390 698–716	346	56	35	**FS**(JOxPM)	13
AY789239 Mal d1.0201.09	Fw-TACACCCTCGAGATTCCAC GAC Rv-TATGCGTCGGGGTGTCCCT AG	28–49 624–644	617	57/55	10/30	PM (PM × FS)	16
AY789240–41 Mal d 1.0201.01	Fw-TCACTTTTGGTGAAGGTC TG Rv-AGGCGTATGAGTAGTTT ATCC	170–189 402–421	252	57	35	PM/**FS**(JO × PM)	16
AY789242 Mal d 1.0404.01	Fw-TCCTTGAGGCGATAACC CT CGTATTGACTGCCTACAAAC	71–90 288–307	237	60/58	10/30	PM (PM × FS)	16
AY789243 Mal d 1.0404.02a	Fw-GGCATAAGTTCAAGTTTCTT CA Rv-TTAGGAATCAGAATGGGTG TG	1240–1261 1609–1629	390	57	35	PM (PM × FS)	16
AY789244 Mal d 1.0404.02b	Fw-TTGCCAGGTCTACTTTCA TC Rv-AACACCCTCCATGCCTTCG TT	49–68 727–747	699	61	35	**FS**	
AY789245–46 Mal d 1.0501.01	Fw-ATGGCGATAAACCTCATCC GA Rv-GCCTGCAAATGAATTAAAAC TA	77–96 284–305	229	57	35	**PM**/FS	16
AY789247 Mal d 1.0502.01	Fw-TCAAAGCACTGAAATCGTT Rv-TCAAGAAGCTTGAAGAGA GT	116–135 546–565	450	57	35	FS (PM × FS)	6
AY789248 Mal d 1.06A02.01	Fw-GAAGGTTAGTTTAATTTCCACA Rv-*GTTT*GCTGTATTCACTCCCTGTCA	181–202 321–340+4	164	56	35	PM (PM × FS)	16
AY789249 Mal d 1.06A01.01a	Fw-GAAGGTTAGTTTAATTTCCACA Rv-*GTTT*GCTGTATTCACTCCCTGTCA	181–202 309–328+4	152	56	35	**FS**	
AY789250 Mal d 1.06A01.01b	Fw-GAAGGTTAGTTTAATTTCCACA Rv-*GTTT*GCTGTATTCACTCCCTGTCA	181–202 312–330+4	154	56	35	PM (PM × FS)	16
AY789251 Mal d 1.06B01.01	Fw-GTA GAATGCCCTGTCCTTG TC Rv-TCTTTGCCAGCCTTAACATGC CT	57–78 544–567	511	57	35	**FS**	
AY789252 Mal d 1.06B02.01	Fw-CCACCATTTTCTCCATTAACTT CA Rv-GCCTTAAACATGCTCTTCCTTGA TT	198–221 534–557	360	62/60	5/35	PM (PM × FS)	16
AY789253 Mal d 1.06B03.01	Fw-GAAATTCACGAGGGAGATGG TA Rv-GCCTTAACATGCTCTTCCTTGAT TT	122–148 534–557	436	62/60	5/35	PM (PM ×FS)	16
AY789254 Mal d 1.06C01.01	Fw-GAAACTGAATACGCCT TCG Rv-GGCATCTCCTTCAATCAA GT	19–37 382–401	383	57	57	**FS**	
AY789255 Mal d 1.06C03.01	Fw-GAAACTGAATACGCCT TCG Rv-CCTTGACATGTTCTTCCT AA	19–37 513–531	513	55/53	5/35	PM (PM × FS)	16
AY789256 Mal d 1.06C02.01	Fw-GAAACTGAATACGCCT TCG Rv-CTTAAT GTCATAAGAGATCTT GTG	19–37 417–440	412	53/51	5/35	PM (PM × FS)	16
AY789257 Mal d 1.0701.01	Fw-GGTGTTTCACATACGAA CT Rv-CTCCCTCAATGACACTGTAT AG	3–23 247–268	266	60/58	5/35	**FS**(JO × PM)	16
AY789258 Mal d 1.0702.01	Fw-CAAGGAAAACTTTGTGTACC TG Rv-CAATCAACTTGAAGAGATGTG TA	228–249 420–442	215	56	35	PM (PM × FS)	16
AY789259–60 Mal d 1.0801.01	Fw-GACTGCTCTTGATGGTG TTG Rv-CAACATCACCCTTAGTGT CG	66–85 361–380	315	57/55	5/35	**PM**/FS	16
AY789261 Mal d 1.0801.02	Fw-GACTGCTCTTGATGGTG TTG Rv-TCAACATCACCCTTAGTGT CA	66–85 360–380	315	58/56	5/35	PM (PM × FS)	16
AY789262 Mal d 1.0901.01	Fw-TGCTAGGTTGTTCAATGCGAC Rv-GCTTGATGATCGAACCGTTGTT	46–68 325–346	299	65/64	5/35	PM (PM × FS)	16

Table 3 (Contd.)

Sequence and marker name	Primer sequence (5′–3′)[a]	Position[b] (gDNA)	PCR (nucleotides)	T_m[c]	Cycles[d]	Presence and segregation[e]	LG[f]
AY789263 Mal d 1.03A01	Fw-TCTGAGTTCACCTCCGTCATT Rv-TTCCTTGATCTCAACGTCA GT	22–42 370–390	369	60	35	**FS**(JO × PM)	13
AY789264 Mal d 1.03B01	Fw-GTCTTCACATACGAATCCG CA Rv-AACCGCTGCCGGAA TCCACT	7–27 315–335	329	65/63	5/30	**PM**	
AY789265 Mal d 1.03B02	Fw-TCCTTGAAGGAGATGGCG AC Rv-TCAATCACACTGTACTTGT TA	131–150 243–263	133	57	35	**FS**(JO × PM)	13
AY789266 Mal d 1.03C01	Fw-CCTTATCCCCAAGATTGCAC AT Rv-AACATCTCCCTTGGTG CGA	87–108 360–378	292	60	35	**PM**	
AY822725 Mal d 1.03C02-GD	Fw-CACCGAGATCCTTGGAG TAA Rv-GCTGGTGTCTCTTGATG TTGc	123–142 335–354	232	57	40	JO (JO × PM)	13
AY789267 Mal d 1.03D01.01	Fw-AAGCAGTGAAGAGCGC AGAA Rv-TTGATGACGGAACCGCT TCT	110–129 325–344	235	57	35	PM (PM × FS)	13
AY789268 Mal d 1.03D02.01	Fw-GAGACTAAGTTGGTTGCTG ACG Rv-TTAGTTGTAGGCATCCTGA AG	304–325 460–480	177	58	35	PM/FS	
AY789269 Mal d 1.03E01.01	Fw-AGATCCTTGAAGGAGAT AGC Rv-TTAGTTGTAGGCATCCTGG CT	126–147 460–480	355	54	35	PM/FS	
AY789270 Mal d 1.03E02.01	Fw-CGAATCCGAGTTCACCT GTA Rv-TTAGTTGTAGGCATCCTGA AT	18–37 460–480	463	57	35	PM (PM × FS)	13
AY789271 Mal d 1.03F01.01	Fw-CTGCTAGGTTGTTCAAT CCA Rv-TGATGACGGAACCGCTGC TG	47–66 324–343	298	57/55	10/30	PM (PM × FS)	13
AY789272 Mal d 1.03F01.02	Fw-TGCTAGGTTGTTCAAT CCA Rv-CTTGATGACAGAACCGTT CCT	47–66 325–345	299	57/55	10/30	PM (PM × FS)	13
AY789273 Mal d 1.03F02.01	Fw-CTGCTAGGTTGTTCAAT CCA Rv-CTTGATGACAGAACCGTT CCT	47–66 325–345	399	55	35	**FS**(JO × PM)	13
AY789274 Mal d 1.03G01	Fw-AAGGACAACTTTGTGTACAAG AAT Rv-TCACCCTTGGTGTAGTAGT CA	229–252 354–374	146	57	35	**PM/FS**	?
AY789275 Mal d 1ps1	Fw-GTACAATGCCTTTATTCTTGA CT Rv-GAATATAGCCATCTCCTCCT	57–79 261–279	223	57	30	PM (PM × FS)	16

[a]Nucleotides in *boldface* and *underlined* indicate introduced mismatches

[b]Positions count from the ATG start codon of the gDNA sequence

[c,d]The PCR T_m and number of cycles, respectively. In the case of two values, a touchdown PCR was performed in two steps: the first number then refers to the T_m and number of cycles, respectively, of the first step

[e]*PM* Prima, *FS* Fiesta, *JO* Jonathan. *Boldface type* refers to homozygous presence of the sequence; segregating mapping population are shown in *brackets*

[f]Linkage group

and 248C nucleotides were exclusively present in *Mal d 1.01* and *Mal d 1.02* (data not shown). Gene-specific polymorphic sites combined with allele-specific polymorphisms in any position were frequently used to design sequence-specific primer pairs in this study.

As shown in Fig. 1, the *Mal d 1* gene family can be split into two categories, with or without intron. For the intron-containing genes, the intron size was specific for each subfamily and gene. *Mal d 1.01* has an intron of 168 nt and *Mal d 1.02* of 171 nt, which are the same as previously reported (Ziadi et al. 2001). *Mal d 1.04* has an intron of 111 nt (Hoffmann-Sommergruber et al. 1997), and its new closely related *Mal d 1.05* gene has an intron of 119 nt. Although three genes in subfamily III are highly similar in their coding regions, they have three different intron sizes—*Mal d 1.06A* has an intron of variable size (130–142 nt) because of variation in a

simple sequence (CA) repeat, *Mal d 1.06B* of 153 nt, and *Mal d 1.06C* of 128 nt. Their introns started always at position 184 and had the same 5′ splicing site of AG/GT, whereas their 3′ splicing sites were different. The *Mal d 1.02* had two 3′ splicing patterns, AG/GC and AG/GT. The *Mal d 1.01* of subfamily I and all genes of subfamily II (*Mal d 1.04* and *Mal d 1.05*) had AG/GC. All genes of subfamily III (*Mal d 1.06A*, *Mal d 1.06B*, and *Mal d 1.06C*) had AG/GG. Another feature in the intron was the putative branchpoint sequence for the processing of pre-mRNA within 40 nt before the 3′ splicing site: CTAAC was present in *Mal d 1.02* and *Mal d 1.06A*; CTAAT in *Mal d 1.01*, *Mal d 1.04*, *Mal d 1.05*, and *Mal d 1.06B*; and CTAGT in *Mal d 1.06C*. Only the first two putative branchpoint sequences fit to the consensus mRNA sequence CURAY for efficient intron splicing (Simpson et al. 2002).


Fig. 3 Map positions of 17 *Mal d 1* genes on linkage group (*LG*) 6, 13, and 16 of the cultivars Jonathan, PM, and FS. The linkage groups 13 and 16 are homoeologous, which is reflected by the presence of common restriction fragment length polymorphism (RFLP) markers (MC001 and MC041) and *Mal d 1* loci at similar mutual distances. The order of *Mal d 1* genes that are located at identical map position, is arbitrary. Reference RFLP and SSR markers flanking the *Mal d 1* clusters were included in the simplified consensus maps
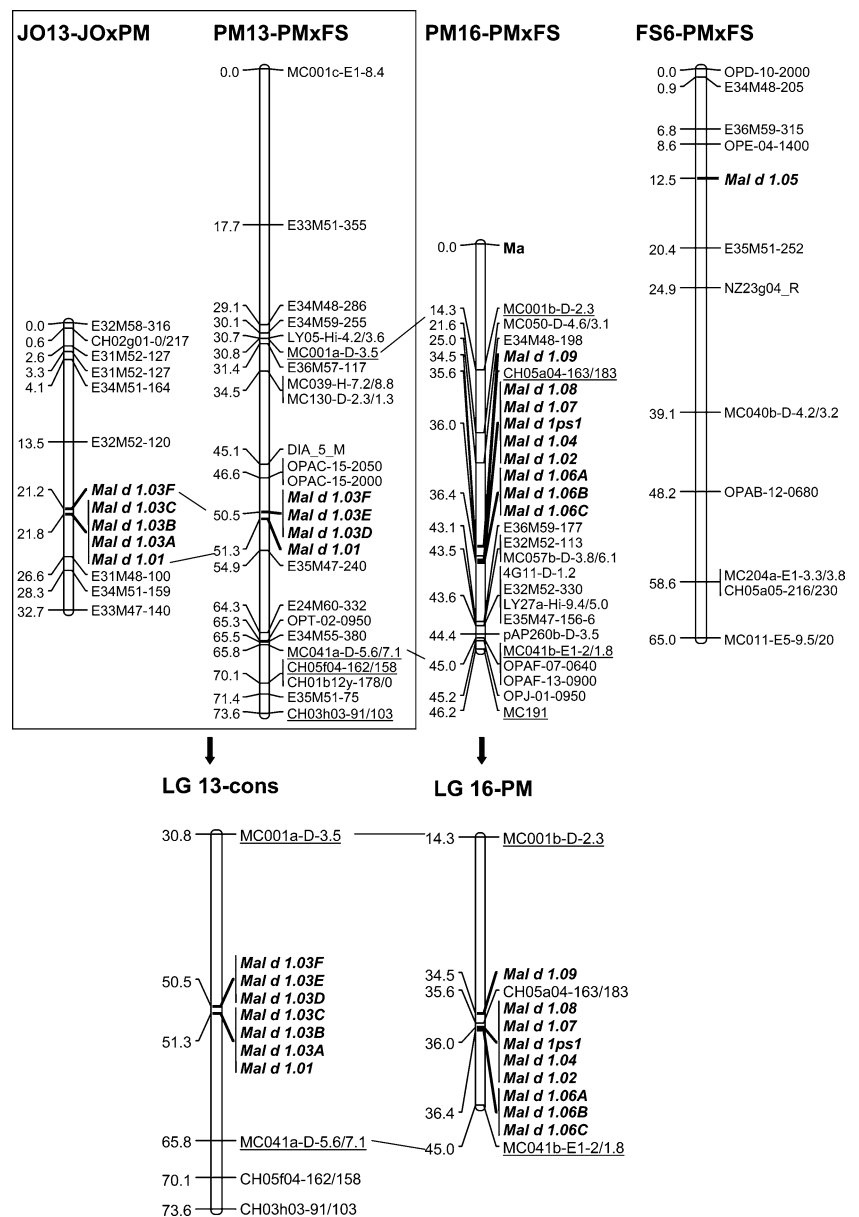
With regard to variation in the upstream region, we examined the sequence of 60 nt before the start codon ATG of eight genes. Intron-containing genes (*Mal d 1.01*, *Mal d 1.02*, *Mal d 1.04*, and *Mal d 1.06A*) showed a conserved TCATC sequence directly preceding to ATG, whereas the intronless genes (*Mal d 1.03A*, *Mal d 1.03D*, *Mal d 1.03E*, and *Mal d 1.03F*) had the GAGAATC sequence. The five *Bet v 1* sequences (accession X77599, X77601, X15877, X81972, and X82028) that have upstream parts also showed a conserved CATC (Swoboda et al. 1995). By scanning our sequences with the online PlantCARE database (Lescot et al. 2002) for putative binding sites, 594 nt upstream region of our three se-

quences of *Mal d 1.04* contained the same transcription factors as those found previously in four genomic sequences corresponding to *Mal d 1.01*, *Mal d 1.02*, *Mal d 1.03D*, and *Mal d 1.07* (AY026910, AY026911, AY026908 and AY026909), such as the Box-W motif (the type member of the PR-10 family), ERE-element, TCA-element, AuxRR-core, and ERRE-motif (involved in fungal elicitor, ethylene, salicylic acid production, auxin, and biotic elicitor responsive elements, respectively) (Ziadi et al. 2001). *Mal d 1.04* had a second TATA box 44 nt upstream to the start codon apart from a common TATAAAT box at around −100 nt found in four *PR-10* genes (Ziadi et al. 2001).

Fig. 4 Amino acid sequences of the Mal d 1 isoallergens and variants from cultivars PM and FS together with those of Bet v 1.0101, Bet v 1.1301, and Bet v 1.1801. *Dashes* at position 108 indicate gaps. Four levels of identity are shown by *shade*: (1) *black*, 100%; (2) *gray*, 80–99%; (3) *light gray*, 60–79%; (4) *white* <60%
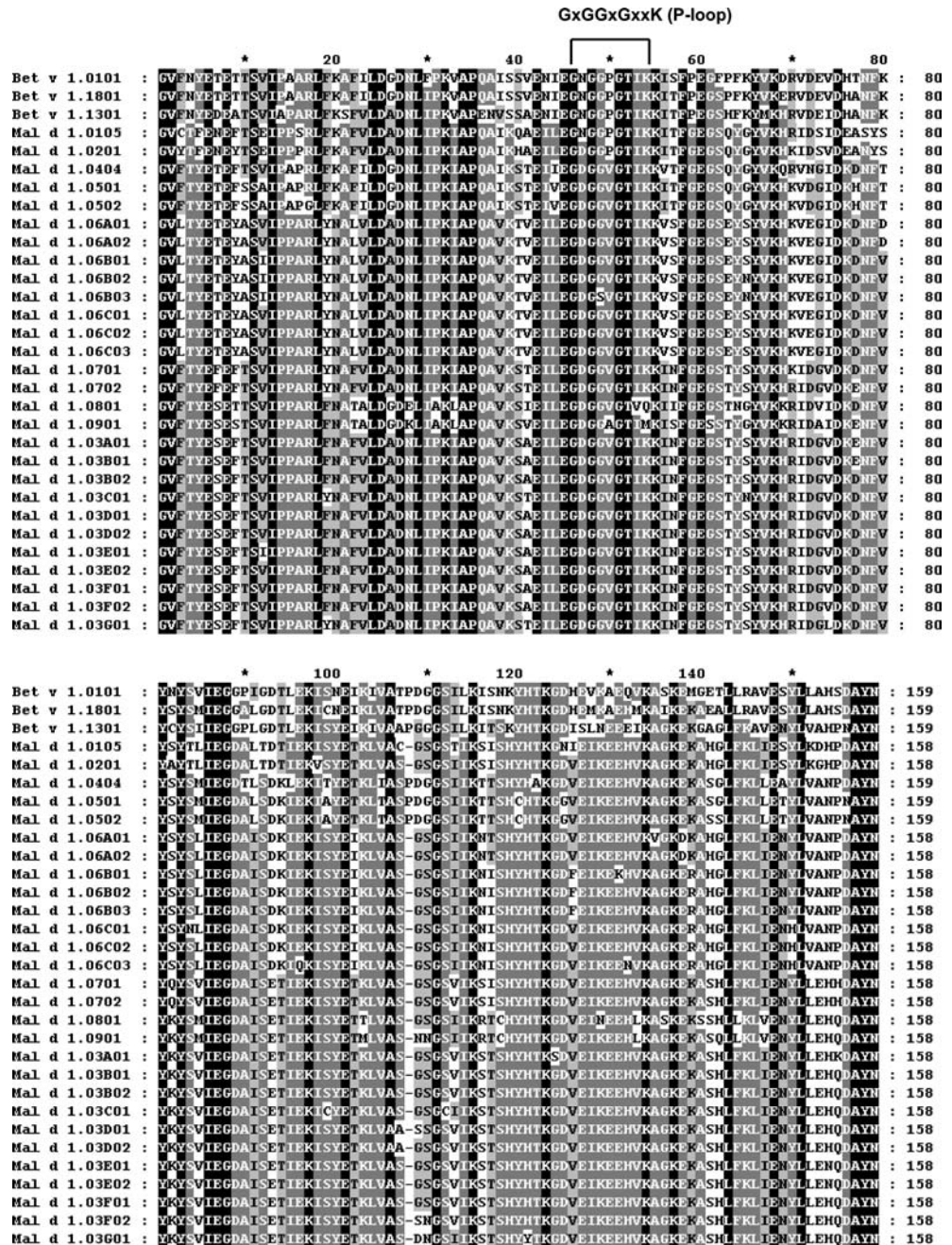
Deduced amino acid sequences

The 39 genomic sequences from PM and FS represented 28 different potential proteins (Fig. 4). Apparently, various silent mutations occurred owing to which allelic variation at the genomic level did not result in variation at the protein level. For instance, three allelic gDNA (AY789236–AY789238) of *Mal d 1.01* encoded the same amino acid sequence (Mal d 1.0105, Table 2). Comparison between protein sequences showed that their identity varied from 65% to 81% among subfamilies, 82% to 100% among genes within a subfamily, and 97.5% to 100% among alleles within one gene. Mal d 1 proteins in subfamily III were 95.6–96.9% identical, so we named them as one isoallergen group, Mal d 1.06. Within subfamily IV, *Mal d 1* genes located on LG 13 were more than 95% identical in amino acid sequence and were therefore classified as isoallergen Mal d 1.03, whereas three genes on LG 16 shared less than 95% identity to Mal d 1.03 and thus were assigned as different isoallergens (Mal d 1.07–09). As a special case, Mal d 1.03F01 (encoded by AY789271 and AY789272 from PM) and Mal d 1.03B02 (encoded by AY789265 from FS) had an identical amino acid sequence although these DNA sequences showed 3.5% dissimilarity. The unmapped gene was intronless and coded for a protein that is more closely related to Mal d 1.03 isoallergens than to Mal d 1.07, Mal d 1.08 and Mal d 1.09, and was subsequently denoted as *Mal d 1.03G*. The predicted molecular weights of our Mal d 1 isoallergens were in a narrow range of 17.3–17.6 kDa, whereas their calculated isoelectric point varied from 5.1 to 6.2.

Discussion

Our results on genomic cloning and linkage mapping revealed that the *Mal d 1* gene family consists of at least 18 members. Except for *Mal d 1.05* on LG 6 and an unmapped gene, all these genes were located in two clusters on the two homoeologous LG 13 and 16. This study forms the basis for a better understanding of the genetics of *Mal d 1* and enables further research on the occurrence of allelic diversity among cultivars in relation to allergenicity and biological functions.

Mal d 1 gene family organization and evolutionary origin

We identified 18 *Mal d 1* genes, of which 16 are organized in a duplicated cluster located between two common random amplified polymorphic DNA markers (MC001 and MC041, Fig. 3) of LG 13 and 16 at similar mutual distances. We thus confirmed that LG 13 and LG 16 are homoeologous LG, which fits with the known duplicate nature of the apple genome (Maliepaard et al. 1998). The position of the *Mal d 1*

genes both in the phylogenetic tree (Fig. 1) and on the linkage maps (Fig. 3) gives some clues to the relationships among members in view of gene duplication. Being present on both the homoeologous LG 13 and LG 16, genes of subfamilies I and IV reflect the amphidiploid origin of the apple genome. Moreover, the intronless genes of subfamily IV are located closely above the intron-containing gene of subfamily I on both LG. Genes of subfamilies II and III are only present on LG 16 but not on LG 13, indicating that these two LG evolved differently. The presence of *Mal d 1.05* on LG 6 is unexpected, but its mapping is based on two reliable SNAP markers for the specific allele of FS. Two additional *Mal d 1*-related genes, *Mal d 1m* (AY428588) and *Mal d 1n* (AY428589) described by Beuning et al. (2004), were not covered, because these sequences were not available at the time of our study. A DNA phylogenetic tree analysis indicated that *Mal d 1m* belongs to subfamily II, but is distinct from both *Mal d 1.04* and *Mal d 1.05* in sequence identity (data not shown) and in the size of its coding region (486 nt instead of 483 nt). *Mal d 1n* does not meet the threshold level of 67% sequence identify for assignment of different proteins to the same allergen (King et al. 1995), making its *Mal d 1* membership questionable.

Comparisons between Mal d 1 and Bet v 1 amino acid sequences

Amino acid sequence comparison of 28 Mal d 1 isoallergens and variants from cultivars PM and FS with all known isoallergens of Bet v 1 revealed the highest identity with Bet v 1.13 (accession X77601) and Bet v 1.18 (accession Z724231) in ranges of 58–64% and 59–62%, respectively. Figure 4 shows the alignment of these 28 amino acid sequences together with Bet v 1 isoallergen 01, 13, and 18. Subfamily II (Mal d 1.04 and Mal d 1.05) is the only one that has the same protein size as Bet v 1 (159 amino acids) and shows the highest identity (59–67.7%) with Bet v 1, whereas all other Mal d 1 proteins consist of 158 amino acids. All these Mal d 1 and Bet v 1 sequences shared 44 conserved amino acids. The Mal d 1 amino acids in segment 108–113 are identical to Bet v 1.01 and Bet v 1.18 including the residue S112, which is essential for IgE binding and cross-reactivity (Son et al. 1999). The conserved sequence motif GXGGXGXXK or P loop (Spangfort et al. 1997) at amino acid residues 47–52 was observed in all the Mal d 1 isoallergens except for Mal d 1.06B03, Mal d 1.08, and Mal d 1.09 (Fig. 4). Ferreira et al. (2000) suggested that amino acid positions 10, 30, 57, 112, 113, and 125 are important for IgE-binding of Bet v 1 and Mal d 1. Sequence alignment of our deduced Mal d 1 amino acids showed some substitutions in five of these six positions (Fig.4). Recombinant by expressed proteins of these isoallergens and variants identified in this study can be used to assess their allergenicity.

Common features of *Bet v 1* homologues

Our results are in line with the four common features of *Bet v 1* (*PR-10*) homologue genes. First, *Mal d 1* genes have either a 480- or 483-nt open reading frame (ORF) with one or two exons, which is similar to *Bet v 1* and its homologues (ORF of 465–489 nt, Hoffmann-Sommergruber et al. 1997; Liu and Ekramoddoullah 2004). Second, all seven *Mal d 1* intron-containing genes have the same 5′ splicing site at DNA position 184 (amino acid position 62) as most other *Bet v 1* homologues (Hoffmann-Sommergruber et al. 1997). Third, *Mal d 1* comprise a multigene family, and most members are clustered in the genome. At least three *Bet v 1* genes (clone Sc1-3, also called Ypr-10 a–c) are within a genome segment of 14 kb (Hoffmann-Sommergruber et al. 2004). In soybean, the Bet v 1 homologue SAM22 allergen consists of ten genes present in a tandem array (Crowell et al. 1992; Kleine-Tebbe et al. 2002). Fourth, *Bet v 1*-related genes are expressed in different tissues and under biotic or abiotic stresses (Hoffmann-Sommergruber and Radauer 2004). Several expression studies on apple *Mal d 1* genes also showed some differences for individual genes. *Mal d 1.02* (GenBank accession L42952, AF020542) is fruit ripening-related (Atkinson et al. 1996) and stress and pathogen inducible (Puehringer et al. 2000). PR-10c (*Mal d 1.01*), and PR-10a (*Mal d 1.02*) will express much more strongly than PR-10d (*Mal d 1.03C*) and PR-10b (*Mal d 1.03D*) upon a salicylic acid analogue induction (Ziadi et al. 2001), whereas *Mal d 1.06A*, *Mal d 1.03E*, and *Mal d 1.03F* are auxin responsive (Kodde and Van der Geest, unpublished). A recent comprehensive expression study on *Mal d 1*-related genes in apple demonstrated that 8 out of 12 genes were expressed in tree-ripened fruit, and most of these were also expressed in leaves in response to a challenge with *Venturia inaequalis*—a fungus causing apple scab (Beuning et al. 2004).

Application in defining different *Mal d 1* genes and allelic variations

Our mapping results clearly demonstrated that alleles of the same locus have over 98.3% sequence identity at the DNA level. A diploid cultivar allows at the most two different alleles for a single locus. These facts can be used to judge whether different *Mal d 1* sequences reflect allelic variation of a single gene or originated from different genes. As shown in Table 2, four very recent sequences identified by Beuning et al. (2004) in subfamily III are classified to two genes with two allelic sequences for each independent gene. If more than two almost identical sequences are identified from a single diploid cultivar, then some of these may contain PCR errors or belong to duplicated genes.

Breeding aspects

Once the alleles for high and low allergenicity have been identified and their linkage phase in the parental cultivars is known, genotyping within a breeding program becomes useful. In this respect, it is worthwhile to investigate the allelic diversity of each *Mal d 1* gene for the founders of a breeding program. We have indications that especially LG 16 is involved in allergenicity (Z.S. Gao et al., unpublished results). LG 16 also contains genes for taste of apple fruit, including the *Ma* gene for acidity (Maliepaard et al. 1998; King et al. 2000) and a major QTL for juiciness and crispness (King et al. 2000). However, these genes are at a 35-cM distance from the *Mal d 1* cluster (Fig.3). This distance is large enough to allow frequent recombination, which opens ways to breeding of new cultivars for good taste and low allergenicity.

In conclusion, this research provides fundamental knowledge of the genomic sequences and linkage map position of the *Mal d 1* gene family. Further investigation on allelic diversity of different *Mal d 1* genes and their protein expression in various cultivars would be one more step toward determining the genetic causes for the difference in allergenicity among apple cultivars.

References

Atkinson RG, Perry J, Matsui T, Ross GS, Macrae EA (1996) A stress-, pathogenesis-, and allergen-related cDNA in apple fruit is also ripening-related. NZ J Crop Hortic Sci 24:103–107

Beuning LL, Bowen JH, Persson HA, Barraclough D, Bulley S, MacRae EA (2004) Characterisation of Mal d 1-related genes in *Malus*. Plant Mol Biol 55:369–388

Breiteneder H, Ebner C (2000) Molecular and biochemical classification of plant-derived food allergens. J Allergy Clin Immunol 106:27–36

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20:1004–1010

Crowell DN, John ME, Russell D, Amasino RM (1992) Characterization of a stress-induced, developmentally regulated gene family from soybean. Plant Mol Biol 18:459–466

Drenkard E, Richter BG, Rozen S, Stutius LM, Angell NA, Mindrinos M, Cho RJ, Oefner PJ, Davis RW, Ausubel FM (2000) A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in *Arabidopsis*. Plant Physiol 124:1483–1492

Ebner C, Birkner T, Valenta R, Rumpold H, Breitenbach M, Scheiner O, Kraft D, Sakamoto T, Hayashi Y, Yamada M, Torii S, Urisu A, Vieths S, Schoning B, Petersen A, Janek K,

Aulepp H, Rudeschko O, Fahlbusch B, Henzgen M, Schlenvoigt G, Herrmann D, Jager L (1991) Common epitopes of birch pollen and apples: studies by Western and Northern blot. J Allergy Clin Immunol 88:588–594

Ferreira F, Hebenstreit D, Kramer B, Himly M, Breiteneder H, Scheiner O, Britza P, Ebner C, Breitenbach M (2000) Amino acid positions involved in the formation of IgE-binding epitopes of Api g 1 and Mal d 1 allergens. J Allergy Clin Immunol 105:S137

Gao ZS, Van de Weg WE, Schaart JG, Van der Meer IM, Kodde LP, Laimer M, Breiteneder H, Hoffmann-Sommergruber K, Gilissen LJWJ (2005) Linkage map positions and allelic diversity of two Mal d 3 (non-specific lipid transfer protein) genes in the cultivated apple (Malus domestica). Theor Appl Genet 110:479–491

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998) Simple sequence repeats for the genetic analysis of apple. Theor Appl Genet 96:1069–1076

Hoffmann-Sommergruber K, Radauer C (2004) Bet v 1-homologous allergens. In: Mills ENC, Shewry PR (eds) Plant food allergens. Blackwell, Oxford, pp 125–140

Hoffmann-Sommergruber K, Vanek-Krebitz M, Radauer C, Wen J, Ferreira F, Scheiner O, Breiteneder H (1997) Genomic characterization of members of the Bet v 1 family: Genes coding for allergens and pathogenesis-related proteins share intron positions. Gene 197:91–100

Holm J, Baerentzen G, Gajhede M, Ipsen H, Larsen JN, Lowenstein H, Wissenbach M, Spangfort MD (2001) Molecular basis of allergic cross-reactivity between group 1 major allergens from birch and apple. J Chromatogr 756:307–313

Hsieh LS, Moos M, Lin Y (1995) Characterization of apple 18 and 31 kDa allergens by microsequencing and evaluation of their content during storage and ripening. J Allergy Clin Immunol 96:960–970

King TP, Hoffman D, Lowenstein H, Marsh David G, Platts Mills Thomas AE, Thomas W (1995) Allergen nomenclature. J Allergy Clin Immunol 96:5–14

King GJ, Maliepaard C, Lynn JR, Alston FH, Durel CE, Evans KM, Griffon B, Laurens F, Manganaris AG, Schrevens E, Tartarini S, Verhaegh J (2000) Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (Malus pumila Mill.). Theor Appl Genet 100:1074–1084

Kleine-Tebbe J, Wangorsch A, Vogel L, Crowell DN, Haustein UF, Vieths S (2002) Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1-related PR-10 protein in soybean, SAM22. J Allergy Clin Immunol 110:797–804

Lescot M, Dehais P, Moreau Y, De Moor B, Rouze P, Rombauts S (2002) PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30:325–327

Liu JJ, Ekramoddoullah AKM (2004) Characterization, expression and evolution of two novel subfamilies of Pinus monticola cDNAs encoding pathogenesis-related (PR)-10 proteins. Tree Physiol 24:1377–1385

Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, Van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, Den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink-Van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (Malus pumila Mill.) using multi-allelic markers. Theor Appl Genet 97:60–73

Ooijen J van, Voorrips R (2001) JoinmapTM Version 3.0: Powerful genetic linkage mapping made easy. Plant Research International, Wageningen, The Netherlands

Pauli G, Oster JP, Deviller P, Heiss S, Bessot JC, Susani M, Ferreira F, Kraft D, Valenta R (1996) Allergens, IgE, mediators, inflammatory mechanisms: Skin testing with recombinant allergens rBet v 1 and birch profilin, rBet v 2: Diagnostic value for birch pollen and associated allergies. J Allergy Clin Immunol 97:1100–1109

Puehringer H, Moll D, Hoffmann Sommergruber K, Watillon B, Katinger H, Laimer da Camara Machado M (2000) The promoter of an apple Ypr10 gene, encoding the major allergen Mal d 1, is stress- and pathogen-inducible. Plant Sci 152:35–50

Puehringer H, Zinoecker I, Marzban G, Katinger H, Laimer M (2003) MdAP, a novel protein in apple, is associated with the major allergen Mal d 1. Gene 321:173–183

Roche P, Alston FH, Maliepaard C, Evans KM, Vrielink R, Dunemann F, Markussen T, Tartarini S, Brown LM, Ryder C, King GJ (1997) RFLP and RAPD markers linked to the rosy leaf curling aphid resistance gene (Sd1) in apple. Theor Appl Genet 94:3–4

Schoning B, Ziegler WH, Vieths S, Baltes W (1996) Apple allergy: the cDNA sequence of the major allergen of apple, determined by performing PCR with a primer based on the N-terminal amino acid sequence, is highly homologous to the sequence of the major birch pollen allergen. J Sci Food Agr 71:475–482

Simpson CG, Thow G, Clark GP, Jennings SN, Watters JA, Brown JWS (2002) Mutational analysis of a plant branchpoint and polypyrimidine tract required for constitutive splicing of a mini-exon. RNA 8:47–56

Son DY, Scheurer S, Hoffmann A, Haustein D, Vieths S (1999) Pollen-related food allergy: Cloning and immunological analysis of isoforms and mutants of Mal d 1, the major apple allergen, and Bet v 1, the major birch pollen allergen. Eur J Nutr 38:201–215

Spangfort MD, Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen JN, Joost Van Neerven RJ, Schou C, Lowenstein H (1997) Three-dimensional structure and epitopes of Bet v 1. Int Arc Allergy Immunol 113:243–245

Swoboda I, Jilek A, Ferreira F, Engel E, Hoffmann-Sommergruber K, Scheiner O, Kraft D, Breiteneder H, Pittenauer E, Schmid E, Vicente O, Heberle BE, Ahorn H, Breitenbach M (1995) Isoforms of Bet v 1, the major birch pollen allergen, analyzed by liquid chromatography, mass spectrometry, and cDNA cloning. J Biol Chem 270:2607–2613

Vanek-Krebitz M, Hoffmann Sommergruber K, Laimer Da Camara Machado M, Susani M, Ebner C, Kraft D, Scheiner O, Breiteneder H (1995) Cloning and sequencing of Mal d 1, the major allergen from apple (Malus domestica), and its immunological relationship to Bet v 1, the major birch pollen allergen. Biochem Biophys Res Commun 214: 538–551

Vieths S, Jankiewicz A, Schoning B, Aulepp H (1994) Apple allergy: the IgE-binding potency of apple strains is related to the occurrence of the 18-kDa allergen. Allergy 49:262–271

Vieths S, Janek K, Aulepp H, Petersen A (1995) Isolation and characterization of the 18-kDa major apple allergen and comparison with the major birch pollen allergen (Bet v I). Allergy 50: 421–430

Voorrips R (2001) MapChart version 2.0: Windows software for the graphical presentation of linkage maps and QTLs. Plant Research International, Wageningen

Wen J, Vanek-Krebitz M, Hoffmann-Sommergruber K, Scheiner O, Breiteneder H (1997) The potential of Betv1 homologues, a nuclear multigene family, as phylogenetic markers in flowering plants. Mol Phylogenet Evol 8:317–333

Ziadi S, Poupard P, Brisset Marie N, Paulin Jean P, Simoneau P (2001) Characterization in apple leaves of two subclasses of PR-10 transcripts inducible by acibenzolar-S-methyl, a functional analogue of salicylic acid. Physiol Mol Plant Pathol 59:33–43