# 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*)

Received: 5 December 2004 / Accepted: 29 March 2005 / Published online: 10 May 2005 © Springer-Verlag 2005

**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

Communicated by H. Nybom

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 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 genespecific. 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.

## 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 1related 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:  $PM \times FS$  (Maliepaard et al. 1998), Jonathan (JO)  $\times$  PM, and FS  $\times$  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 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 Tag (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 DraI, EcoRV, PvuII, and StuI 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 use	d for cloning of the	different $Mal \ d \ I$ genes				
Primer pairs	<sup>a</sup> Reference Sequence <sup>a</sup>	Primer name <sup>b</sup>	Primer sequence $(5'-3')$	$T_{\rm m}$ (°C/no. Proceeded) respectively.	duct cleotides)	Intron (nucleotides)	Isoallergen gene member
				Pfu Taq			
1	AY026910	Mla-For	ATCTCCAACACAATACTCCCAAC	58/25 60/2 800		168	Mal d 1.01
2	L42952 AF020542	Mald1b-For	AAAGUCACAAAULITUGAU CATCCTTGGTAGTTGCTTTC	52/30 54/2 723		171	Mal d 1.02
ŝ	Z48969	M1b-Kev M1a/b/c-For	TCATGGGTGTCTACACATTTG	50/35 50/2 439	; 483	0	Mal d 1.03B; Mal d 1ps1
4	Z72425–Z72428	Mlc-Rev HB20-For <sup>b</sup>	ATTAATTTAGTTGTAGGCATCC ATGGGTGTTTTCAATTACGAAACTGAG	54/30 56/2 480		0	Mal d 1.034; Mal d 1.08; Mal d 1.09;
S	AY026908 AY026909	HB22-Rev <sup>b</sup> M1.03-For1	TTAGTTGTAGGCATCGGAGTGTGCCAAGAGG ATGGGTGTTTTCACATACGAA	55/25 57/2 486		0	Mai d 1.07; Mai d 1.08; Mai d 1.09; Mai d 1.03; E. G
6	Z72426–Z72428	M1.03-Rev1 M1.04-s27	AAGAATTTAGTTGTAGGCATC GWYCACCTCTGTCATCCCTG	56/30 58/2 470		111	Mal d 1.04
7	Z72426–Z72428	M1.04-as496 M1.04-s62	TCTCAACATCACCTTAGCA 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	GAAGCTTGAAGAGACCGGAA CTCCTTCAATCATAGAGTAACTGTATGT	747			Mal d 1.04
6	Based on 6	M1.04UP-GSP2 M1.04DW-GSP1	CCAAAGGIAACTITCTIGATGATIGATICC ACATACAGTIACTCTATGATIGATGAAGGA	830			Mal d 1.04
10	Based on 7	M1.04DW-GSP2 M1.04UP-GSP1	CATCAAAACCACTAGCCACTACCATGC See above	747			Mal d 1.05
11	Based on 7	M1.05UP-GSP2 M1.04DW-GSP1	GAGAGAATATTGAGGGTTGTCTAGCTAA See above	477			Mal d 1.05
12 <sup>c</sup>	3262-PRI 3425-PRI	M1.05DW-GSP2 M1.06-For	ATTGACCGCATCTCCTGATGG CATGGGTGTCCTCACATACGAAAC	55/25 57/2 609-	-634	128–142	Mal d 1.06A- C
13	Based on 12	M1.06-Kev M1.07-fw M1.07 rev	LIAGLIGIAGGCALCAGGALIG ATGGGTGCTCCACATAGGAAACT TTAGGTATCAAGAATAGGAATAGACTAAAAGGATA	62/30 64/2 608		128	Mal d 1.06C
14 <sup>c</sup>	3076-PRI	M1.03-For2	ATCATGGGTGTTTTCACATAT	55/25 57/2 491		0	Mal d 1.034; Mal d 1.03B; Mal d 1.03F
15	Based on 14	M1.03-Rev2 M1.03-For2	CATATATTTAGTTGTAGGCGTCT See above	55/25 57/2 491		0	Mal d 1.034; Mal d 1.03B; Mal d 1.03B;
		M1.03-Rev3	CATATAATTTAGTTGTAGGCGT				JCO.I D IDM
<sup>a</sup> Primer pairs unpublished <sup>b</sup> HB20-for an <sup>c</sup> Annealing to	s $1-7$ were based on $M$ Mal d I sequences at 1 ad HB22-Rev are origi smperature	<i>al d I</i> sequences in Plant Research Inte nal primers of Hoff	the GenBank; 8–11 are gene-specific for genome walking trational ( <i>PRI</i> ), Wageningen mann-Sommergruber et al. (1997)	o upstream (UP)	and down	stream (DW);	12-15 were derived fron

Subfamily	Gene	Previous names <sup>a</sup>	New sequences (alleles)		Reference sequences <sup>b</sup>
			Prima	Fiesta	- (isoallergen and variants names)
	Mal d 1.01	Mal d la/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),
	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)	AJ41/551, AY4285/9 L42952 (Mal d 1.0201), AF020542 (Mal d 1.0206), AF074721, AF124824 (Mal d 1.0202), AF124824 (Mal d 1.0203), AF125825 (Mal d 1.0204), AF124826, AF124835 (Mal d 1.0204), AF124836, AF124834, AF124835 (Mal d 1.0207), AJ488060 AY026911 (Mal d 1.0207), AJ488060
Π	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)	(Mai d 1.0208), <b>A1428/8</b> Z72456 (Mai d 1.0401), Z72427 (Mai d 1.0400), Z72478 (Mai d 1.0403)
	Mal d 1.05		AY789245 (Mal d 1.0501)	AY789246 (Mal d 1.0501) AV780247 (Mal d 1.0502)	
III	Mal d 1.06A	Mal d 1e; Mal d 1f	AY789248 (Mal d 1.06A02.01)	AY789249 (Mal d 1.06A01.01a)	AY428580, AY428581
	Mal d 1.06B	Mal d 1g; Mal d 1h	AY 789250 (Mal d 1.06A01.01b) AY 789252 (Mal d 1.06B02) AV 780552 (Mal d 1.06B02)	AY789251 (Mal d 1.06B01)	AY428582, AY428583
	Mal d 1.06C		AY789255 (Mal d 1.06C02) AY789255 (Mal d 1.06C02)	AY789254 (Mal d 1.06C01)	
2	Mal d 1.07 Mal d 1.08	PR-10d; Mal d 1k; Mal d 1.03	AY789258 (Mai d 1.0702) AY789259 (Mai d 1.0702) AY789259 (Mai d 1.0801.01)	AY789257 (Mal d 1.0701) AY789260 (Mal d 1.0801.01)	AY026909 (Mal d 1.0303), <b>AY428586</b>
	Mal d 1.09		AY789262 (Mal d 1.0901)	(FO A CO F E TEAM CYCOGEXXA	1000 F F F X Y 2000 L
	Mal d 1.03A Mal d 1.03B Mal d 1.03C	Mal d 1.03 Mal d 1.03 Mal d 1 1	AY789264 (Mal d 1.03B01) AY789266 Mal d 1.03C01)	AY 789265 (Mal d 1.03B02) AY 789265 (Mal d 1.03B02)	Z/2425 (Mal d 1.0301) AY186248 (Mal d 1.0304) AY438587
	Mal d 1.03D Mal d 1.03E	PR-10b, Mal d 1.03 Mal d 1i	AY789267 (Mal d 1.03D02) AY789267 (Mal d 1.03D02) AY789270 (Mal d 1.03E01)	AY789268 (Mal d 1.03D01) AY789269 (Mal d 1.03E02)	AY026908 (Mal d 1.0302) AY428584
	Mal d 1.03F	Mal d 1j	AY789271 (Mal d 1.03F01) AY789272 (Mal d 1.03F01)	AY789273 (Mal d 1.03F02)	AY428585

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

<sup>a</sup>In the reference papers and Mal d 1 isoallergen list <sup>b</sup>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)

AY789274 (Mal d 1.03G01)

AY789275

Mal d 1.03G Mal d Ipsl

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 Mald1 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 li, 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  $\times$  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,

Sequence and marker name	Primer sequence (5'-3') <sup>a</sup>	Position <sup>b</sup> (gDNA)	PCR (nucleotides)	$T_{\rm m}^{\rm c}$	Cycles <sup>d</sup>	Presence and segregation <sup>e</sup>	LG <sup>f</sup>
AY789236 Mal d 1.0105.01a	Fw-AGCTGAAATCCTTGAA CGAA	123-142 528 547	426	55/53	10/30	PM (PM $\times$ FS)	13
AY789237 Mal d 1.0105.03	RV-CAALULLUCUCI 10010 AUA FwTGAGCACAGGATTGAC B CAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	371-390 371-390 528 547	177	57/55	10/30	PM (PM $\times$ FS)	13
AY789238 Mal d 1.0105.02	FW-CAALGI I JUCCUCI 100101 CU FW-CTAALGACGCACAGGATTGAC GCA BV-CTAACAACATTCGACT CA	371-390 371-390 608-716	346	56	35	FS(JOxPM)	13
AY789239 Mal d1.0201.09	Fw-TACACCTCTGAGATTCCAC GAC by TATCACCTCTGAGATTCCAC GAC	28-49 534 644	617	57/55	10/30	PM (PM × FS)	16
AY789240-41 Mal d 1.0201.01	KV-1ALGCOLCGGGGGGGC AG Fw-1CACTTTTGGTGAAGGCC TG Fw-1CACTTTTGGTGAAGGCC TG	024-044 170-189 402 421	252	57	35	$\rm PM/FS(JO  imes PM)$	16
AY789242 Mal d 1.0404.01	FW-TOCTCATGAGGCATAACC CT	71-90 71-90 288 307	237	60/58	10/30	PM (PM $\times$ FS)	16
AY789243 Mal d 1.0404.02a	EWIGITIOACIOCCIACAAAC FW-GGCATAGGTTCAAGTTTCTT CA B.: TTAGGAAATTCAAGTTTG TC	1240-1261 1240-1261	390	57	35	PM (PM $\times$ FS)	16
AY789244 Mal d 1.0404.02b	Fw-TIGCCAGGTCTACTTTCA IC	49–68 1007–1027	669	61	35	FS	
AY789245-46 Mal d 1.0501.01	FW-AACAUCIUCATUGULIUU II FW-ATGGCGATAACCTCATCC GA	77-96 77-96 284 205	229	57	35	PM/FS	
AY789247 Mal d 1.0502.01	Fw-TCC10CAAAAGCACTGAAATCGTT	116-135	450	57	35	FS (PM $\times$ FS)	9
AY789248 Mal d 1.06A02.01	Fw-forder for the former of th	240-202 181-202	164	56	35	PM (PM $\times$ FS)	16
AY789249 Mal d 1.06A01.01a	Fw-6111GC1G1A11CAC1CCC1G1CA	321-340+4 181-202	152	56	35	FS	
AY789250 Mal d 1.06A01.01b	Fw-6111GC1G1A11CAC1CCC1G1CA	309-328+4 181-202	154	56	35	PM (PM $\times$ FS)	16
AY789251 Mal d 1.06B01.01	Fw-6111GC1G1A11CAC1CUC1G1CA Fw-6111GATGCCTTGTCCTTG TC Fw-forcestococcttgTCCTTG TC	512-550+4 57-78 544 557	511	57	35	FS	
AY789252 Mal d 1.06B02.01	Fw-CCACCATTTTCTCCCATTAACTTCCC	198-221 224-50/ 524 557	360	62/60	5/35	$PM (PM \times FS)$	16
AY789253 Mal d 1.06B03.01	FW-GCC11AACA1GC1C11CC11GA11 FW-GCC11AACA1GG4GGGGGGAGA7GG TA FW-GCC1TTAACA1GCTCTTTCATTCA	122-148	436	62/60	5/35	PM (PM ×FS)	16
AY789254 Mal d 1.06C01.01	FW-OCCLIDAATACCICLICATION FW-GAAACTGAATACGCCT TCG BW-GGCATCTTTCAATCAACG	19-37 19-37 382 401	383	57	57	FS	
AY789255 Mal d 1.06C03.01	Fw-GOCATCLCULUCITCAACCCTTCG Du. COTTACACTCCTTCG	19-37	513	55/53	5/35	PM (PM × FS)	16
AY789256 Mal d 1.06C02.01	EW-CUITORCATALCULAA FW-CUITORCATACCCT TC EW-CUITORCATACCCT TC	19-37 19-37 117 440	412	53/51	5/35	PM (PM $\times$ FS)	16
AY789257 Mal d 1.0701.01	W-CLIAAL GICALAAGAGALCII GIG Fw-GGTGTTTCACATAGGAA D. CTCCCTCAAAGAAACTAT	41/-440 3-23 247 268	266	60/58	5/35	$FS(JO \times PM)$	16
AY789258 Mal d 1.0702.01	Fw-CACGUCAAAACTTTGGGGTACCTG Dw.CAAGGAAAACTTTGGGGTACCTG Dw.CAATCAAAAACTTTGAAAAAACTTTGTGTAC	228-249 228-249 430 443	215	56	35	PM (PM × FS)	16
AY789259–60 Mal d 1.0801.01	Fw-GACTGCTCTTGATGGTG TTG	720-772 66-85 361-380	315	57/55	5/35	PM/FS	
AY789261 Mal d 1.0801.02	FW-GACTGCTCTTGATGGTG TTG	001-300 66-85 360-380	315	58/56	5/35	PM (PM × FS)	16
AY789262 Mal d 1.0901.01	KV-ICAAUAIUAUUUI AUJUI UA FW-TGCTAGGTTGTTCAATGCGAC RV-GCTTGATGATCGAACCGTTGTT	500-580 46-68 325-346	299	65/64	5/35	PM (PM $\times$ FS)	16

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

177

sequence and marker name	Primer sequence $(5'-3')^a$	Position <sup>b</sup> (gDNA)	PCR (nucleotides)	$T_{\rm m}{}^{\rm c}$	Cycles <sup>d</sup>	Presence and segregation <sup>e</sup>	LG <sup>f</sup>
AY789263 Mal d 1.03A01	Fw-TCTGAGTTCACCTCCGTCATT	22-42 270-200	369	60	35	$FS(JO \times PM)$	13
AY789264 Mal d 1.03B01	Fw-ffCUTIGACUTCACUTCA GI	5.10-390 7-27	329	65/63	5/30	PM	
AY789265 Mal d 1.03B02	Rv-AACCGCTGCCGGAA TCCACT Fw-TCCTTGAAGGAGATGGCG <u>A</u> C	315-335 131-150	133	57	35	$FS(JO \times PM)$	13
AY789266 Mal d 1.03C01	Rv-TCAATCACACTGTACTTGT TA Fw-CCTTATCCCCAAGATTGCAC AT	243-263 87-108	292	60	35	PM	
AY822725 Mal d 1.03C02-GD	Rv-AACATCTCCCTTGGTG CGA  Fw-CACCGAGATCCTTGGAG TAA	360–378 123–142	232	57	40	$(MX \times PM)$	13
AY789267 Mal d 1.03D01.01	Rv-GCTGGTGCTCTTGATG TTGC Fw-AAGCAGTGAAGAGCGC AGAA	335–354 110–129	235	57	35	$PM (PM \times FS)$	13
AY789268 Mal d 1.03D02.01	Rv-11GA1GACGGAACCGCT1CT Fw-GAGACTAAGTTGGTTGCTG ACG	325-344 304-325	177	58	35	PM/FS	
AY789269 Mal d 1.03E01.01	Rv-TTAGTTGTAGGCATCCTGA AG Fw-AGATCCTTGAAGGAGAT <u>AG</u> C	460–480 126–147	355	54	35	PM/FS	
AY789270 Mal d 1.03E02.01	Rv-TTAGTTGTAGGCATCCTGG CT Fw-CGAATCCGAGTTCACCT GTA	460-480 18-37	463	57	35	$PM (PM \times FS)$	13
AY789271 Mal d 1.03F01.01	Rv-TTAGTTGTAGGCATCCTGA AT Fw-CTGCTAGGTTGTTCAAT CCA	460–480 47–66	298	57/55	10/30	$PM (PM \times FS)$	13
AY789272 Mal d 1.03F01.02	Rv-TGATGACGGAACCGCTGC TG Fw-TGCTAGGTTCATT CCA	324–343 47–66	299	57/55	10/30	$PM (PM \times FS)$	13
AY789273 Mal d 1.03F02.01	Fw-CI IGAI GACAGAACCGI I CCI Fw-CTGCTAGGTTGTTCAAT CCA	525545 47-66 225 - 545	399	55	35	$FS(JO \times PM)$	13
AY789274 Mal d 1.03G01	EW-CLIGATGACAGAGUCIT CCI EW-AAGGACATTTGTGGTACAAG AAT E-T-AAGGACAACTTGTAGAAG AAT	229-545 229-252 254 274	146	57	35	PM/FS	ċ
AY789275 Mal d 1ps1	Fw-GTACAATGCCTTTATTCTTGA CT Rv-GAATAGCATCTCCTCCT	57-79 57-79 261-279	223	57	30	PM (PM $\times$ FS)	16
Nucleotides in <i>boldface</i> and <i>und</i> Positions count from the ATG	<i>lerlined</i> indicate introduced mismatches start codon of the eDNA sequence						

<sup>c,d</sup>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 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* <sup>f</sup>Linkage group

178

Table 3 (Contd.)

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

Fig. 3 Map positions of 17 Mal d l 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

# 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 d1.02 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).



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).

GxGGxGxxK (P-loop)

**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%

												Г	1.0								
n-4		1 0101		Contraction of the local division of the loc	*		20	in No	*		40	-	*	I	60	DER LUZ	*		80		
Bet	v	1.1801	;	SVEN 121	STTOU	IPAARI	FK I	TLDG	DNL JP	ATAPI	ISSV	INTRO	IGGPGT	IKKIT	PD 65	PFK	ERVI	EVDHA	NICK	: 8	n
Bet	v	1.1301	÷	GVENTED	ATSV	APARI	FKSI	TVIDA	DNLIP	VAPE	NVSSA	INTEGN	GGRGT	IKRIT	FPEGS	FKM	RVD	EIDHA	NEK	: 8	0
Mal	d	1.0105	:	GVCIFIEN	FTSE	I P P S RI	FK	FVLDA	DNL IP	KIAP	AIKQA	EILEGN	IGGPGT	IKKIT	FGEGS	Q Y GYVK	RID	SIDEA	SYS	: 8	0
Mal	d	1.0201	:	GVYTERN	EYTSE	I P P <mark>P</mark> RI	FKA	FVLDA	DNL IP	KIAPQ	АІКНА	EILEGI	GGPGT	IKKIT	FGEGS	рх схлк	KID	SVDEA	NYS	: 8	0
Mal	d	1.0404	:	OVFTYDT	DFTSV	<b>AP</b> RI	FK I	ILDG	DNL IP	KIAPQ	AIKST	ILECT	GGVGT	TREAT	FGEGS	D X G Y V K	<b>OBAN</b>	ICID:D	NET	: 8	0
Mal	d	1.0501	÷	GVET MET	DESCAL	APS	PK 1	TLDG	DNL IP	KLAPU VTADO	AIRST		CEVET		FGEGS	UX GXV K	KVD	GIDEN	NP T	: 8 . 6	
Mal	d	1.06A01	÷	OVLITYPT	YASV	IPPARI	YN I	VIDA	DNLIP	KIAPO	VTV		GGVGT	UNIVS	FGEGS	ey syvik	KVE	GID D	NED	: 8	10
Mal	d	1.06A02	:	GVLTYPT	YASV	IPPARI	YN I	VIDA	DNLIP	KIAPO	AVKTV	LEGE	GGVGT	IKEVS	FGEGS	EYSYVK	KVE	GIDKD	NED	: 8	a
Mal	d	1.06B01	:	GVLTYPT	PYASI	IPPARI	YN I	VLDA	DNL IPI	KIAPO	AVKTV	EILEGU	GGVGT	IKRVS	FGEGS	EYSYVK	KVE	GIDKD	NEV	: 8	0
Mal	d	1.06B02	:	GVLTYET	BYASI	IPPARI	YN 1	VLDA	DNL IP	KIAPQ	AVKTV	EILEGI	GGVGT	IKRVS	FGEGS	EYNYVK	KVE	GIDKD	NEA	: 8	0
Mal	d	1.06B03	:	GVLOOT	PYASI	IPPARI	YN I	VIDA	DNLIP	KIAPU	AA TA	EILEGE	GSVGT	INTAR	FGEGS	EYNYAK	KVE	GID D	NEV	: 8	10
Mal	d	1.06002	1	SVLUVET	TASU	I PPARI	VN	VIA		KTAPU	N TV		ICGVCT	10005	FGEGS	SUSYUK	KVE	GTD D	DIDV		n a
Mal	d	1.06003	-	GVLTYPT	YASV	IPPARI	YN I	VIDA	DNLIP	KIAPO	AVETV	LEGE	GGVGT	IKEVS	FGEGS	статик	KVE	GIDED	NEV	: 8	0
Mal	d	1.0701	:	GVETYDE	FTSV	IPPARI	YN I	TVLDA	DNLIP	KIAP	AVKST	EILEGE	GGVGT	IKKIN	FGEGS	IYSYVK	KID	GVDED	NEV	: 8	a
Mal	d	1.0702	:	GVETYDE	PFTSV	IPPARI	YN I	FVLDA	DNLIP	KIAPO	AVKST	ELEGI	GGVGT	IKKIN	FGEGS	rysyvk	RID	GVDKE	NEV	: 8	0
Mal	d	1.0801	:	GVETYDS	DTTSV	IPPARI	FN.	FALDG	DELTIA	KLAP	AVKSI	EILEGI	GGVGT	VQUII	FGEGS	INGWIK	KRID	AID (D	NEA	: 8	0
Mal	d	1.0901	÷	GVF MDS	DSTSV	IPPARI	EN T	TALDG	DIKL HAI	KUAPU	A SA		CCACT	MRIS	FUPSS	Fri GYA IX	KRID	AIDE	RIEV	: 8	0
Mal	d	1.03801	:	OVETTOS	SFTSU	IPPARI	FN T	TUDA	DNI TPI	KTAPO	V SA		GGVGT	TREAT	FGEGS	TV SVVK	RTD	SVD E	Velse	: 8	10
Mal	d	1.03B02	:	OVETYDS	FTSV	IPPARI	FN I	VIDA	DNL IP	KIAPO	VISA		GGVGT	IKKIN	FGEGS	TYSYVK	RID	GVD	NEV	: 8	0
Mal	d	1.03C01	:	GVETYDS	FTSV	IPPARI	YN I	FVLDA	DNL IPI	KIAPQ	AVKST	TLEGI	GGVGT	IKKIN	FGEGS	<b>FYNYV</b> K	RID	GVDKD	NEV	: 8	0
Mal	d	1.03D01	:	GWFTYDS	PETSA	IPPARI	FNAR	FVLDA	DNLIP	KIAPQ	AVKSA	EILEGI	GGAGI	IKKIN	FGEGS	IYSYVK	RID	GAD KD	NFV	: 8	0
Mal	d	1.03D02	:	GVFTTDS	FTSV	IPPARI	FN I	VIDA	DNL IP	KIAPO	VSA	STILEGO	GGVGT	IKKIN	FGEGS	r Shuk	RID	GAD D	REV	: 8	0
Mal	a	1.03E01	1	GVE TOPS	OF TSL	TPPARI	EN I	EVILLA	DNL IP	KLAPU KTADO	VSA		GGVGT		FGEGS	TY SYVK	DTD	GVU D	DIEW	: c	E C
Mal	d	1.03F01	-	GVETYES	FISU	IPPARI	EN I	VIDA	DNLIP	KIAPO	V SA	ILEGE	GGVGT	IKRIN	FGEGS	TY SWVK	RID	GVD D	NEV	: 8	0
Mal	d	1.03F02	:	GVETYES	FTSV	IPPARI	ENA	TUDA	DNLIP	KIAPO	AVKSA	LLEGE	GGVGT	IKKIN	FGEGS	TYSYVK	RID	GVDKD	NEV	: 8	D
Mal	d	1.03601	:	GVE TYES	FISV	IPPARI	YN I	FVILDA	DNL IP	KIAPO	AVKST	LEGI	GGVGT	IKKIN	FGEGS	FYSYVK	RID	GLDKD	NEV	: 8	a
															THE REAL PROPERTY.						
														-				_			
						1	100				120				140						
Bet	v	1.0101	:	VNYSVIIE		1 DTL 3K	100 ISNE]	IKIVA	* TPDGG	SILKI	120 SNKYH	IKGDH	* V A 0	VKASK	140 MGET	LRAVS	* SMLL	AHSDA	YN :	159	
Bet Bet	* *	1.0101 1.1801 1.1301		VNYSVIE YSYSMIE VGYSTIE	* GGPIGI GGPIGI	J DTL X DTL X DTL X	100   SNB]   CNB]	IKI VA IKI VA IKI VA	* TPDGG TPDGG	STLKI STLKI	120 SNKYH SNKYH TSKYH	TKGDHE TKGDHE	* VKADQ MKADH	VKASK UKAIK UKAGK	140 MGET K EA	LRAVE		AHSDA AHSDA AHSDA	YN : YN : YN :	159 159	
Bet Bet Bet Mal	v v v d	1.0101 1.1801 1.1301 1.0105		YNYSVIE YSYSMIE YCYSIIE YSYTLIE	* GGPIGI GGALGI GGPLGI GDALTI	I DTLEKI DTLEKI DTLEKI DTLEKI	100 ISNE ICNE ISYE ISYE	IKIVA IKIVA IKIVA IKIVA	* TPDGG TPDGG APGGG C-GSG	STLKI STLKI STLKI	120 SNK <sub>YH</sub> SNK <sub>YH</sub> TS <mark>K</mark> YH ISHYH	TKGDHE TKGDHE TKGD <mark>IS</mark> TKGNIE	* V AEQ M AEH SLNDEE	VKASK NKAIK Ikagk Vkagk	140 MGET K EA KGAG	LLRAVE LLRAVE LFKAVE LFKLIE	* SVILL SVILL NVILV SVILB	AHSDA AHSDA AHSDA AHPNA DHPDA	YN : YN : YN : YN :	159 159 159	
Bet Bet Mal Mal	v v v v d d	1.0101 1.1801 1.1301 1.0105 1.0201		VNYSVIE YSYSMIE YCYSIIE YSYTLIE YAYTLIE	* GGPIGI GGALGI GGPLGI GDALTI GDALTI	J DTLEK DTLEK DTLEK DTLEK	100 ISNEJ ICNEJ ISYEJ ISYEJ	IKIVA IKIVA IKIVA IKIVA IKIVA	* TPDGG TPDGG APGGG C-GSG S-GSG	S ILKI S ILKI S ILKI S I I KS S I I KS	120 SNKYH SNKYH TSKYH TSKYH TSHYH	IKGDHI IKGDHI IKGD <mark>IS</mark> IKG <mark>NI</mark> I IKG <b>NI</b> I	* VKAEQ MKAEH INDEE IKDEH IKDEH	VKASK MKAIK Ikagk Vkagk Vkagk	140 Emget Ekaea Ekgag Ekahg Ekahg	LLRAVE LLRAVE LFKAVE LFKLIE LFKLIE	* SYLL SYLL NYLV SYLF SYLF	AHSDA AHSDA AHSDA AHPNA DHPDA GHPDA	YN : YN : YN : YN : YN :	159 159 159 158	
Bet Bet Mal Mal Mal	V V V d d d	1.0101 1.1801 1.1301 1.0105 1.0201 1.0404		YNYSVIE YSYSMIE YCYSIIE YSYTLIE YAYTLIE YSYSMIE	* GGPIGI GGALGI GGPLGI GDALTI GDALTI GDALTI	J DTLEKI DTLEKI DTLEKI DTLEKI DTLEKI DKLEKI	100 ISNEJ ICNEJ ISYEJ ISYEJ ISYEJ	IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA	* TPDGG TPDGG APGGG C-GSG S-GSG S-GSG	STLKI STLKI STLKI STLKS STLKS STLKS	120 SNK YH SNK YH TSK YH TSH YH TSH YH TSH YH	TKGDHE TKGDHE TKGDIS TKGNIE TKGDVE TKGDVE	* VKAEQ MKAEH IKEEH IKEEH IKEEH	VKASK IKAIK IKAGK VKAGK VKAGK	140 DMGET DKAEA DKAEG DKAHG DKAHG DKAHG	LRAVE LRAVE LFKAVE LFKLTE LFKLTE LFKLTE	* SYIL SYIL SYIL SYIL SYIL SYIL SYIL SYIL	AHSDA AHSDA AHSDA AHPNA DHPDA GHPDA ANPDA	УМ : УМ : УМ : УМ : УМ : УМ :	159 159 159 158 158	
Bet Bet Mal Mal Mal	v v v d d d d	1.0101 1.1801 1.1301 1.0105 1.0201 1.0404 1.0501		VNVSVIE VSVSMIE VGVSIIE VSVTLIE VAVTLIE VSVSMIE VSVSMIE	* GGPIGI GGALGI GDALTI GDALTI GDALSI GDALSI	DTLEK DTLEK DTLEK DTLEK DTLEK DTLEK DKLEK	100 ISNE ICNE ISYE ISYE ISYE ITYE IAYE	TKIVA TKIVA TKIVA TKIVA TKIVA TKITA	* TPDGG TPDGG APGGG C-GSG S-GSG S-DGG S-PDGG S-PDGG	STLKI STLKI STLKI STLKS STLKS STLKI	120 SNK УН SNK УН TSK YH ISH YH ISH YH TSH YH TSH CH	TKGDH IKGDH IKGDIS IKGNI IKGDV IKGDV IKGDV	★ MKAEQ MKAEH IKEEH IKEEH IKEEH IKEEH	VKASK IKAGK VKAGK VKAGK VKAGK VKAGK	140 MGET KAEA KGAG KAHG KAHG KASG	LRAVE LRAVE FKAVE FKLIE FKLIE FKLIE	× SYTL SYTL SYTL SYTL SYTL SYTL SYTL SYTL	AHSDA AHSDA AHSDA DHPDA GHPDA GHPDA ANPDA ANPDA	YN : YN : YN : YN : YN : YN : YN :	159 159 159 158 158 159	
Bet Bet Mal Mal Mal Mal	v v v d d d d d	1.0101 1.1801 1.0105 1.0201 1.0404 1.0501 1.0502 1.0502			* GGP I GI GGALGI GDALTI GDALTI GDALSI GDALSI GDALSI GDALSI	J DTLEK DTLEK DTLEK DTLEK DTLEK DKLEK DKLEK DKLEK	100 ISNE ISNE ISYE ISYE ISYE ITYE IAYE IAYE	TKIVA TKIVA TKIVA TKIVA TKIVA TKITA TKITA	* TPDGG TPDCG C-GSG S-GSG SPDGG SPDGG S-DSG	SILKI SILKI SILKI SILKI SIIKI SIIKI SIIKI SIIKI	120 SNK VH SNK VH T SK VH I SHVH I SHVH T SH CH T SH CH	IKGDH IKGDH IKGDIS IKGNI IKGNI IKGNI IKGNI IKGNI IKGNI	* V:ACQ H:ACH IKEEH IKEEH IKEEH IKEEH	VKASK IKAGK VKAGK VKAGK VKAGK VKAGK	140 PMGET DKAEA DKGAG DKAHG DKAHG DKAHG DKASG DKAHG	LRAVE LRAVE FRAVE FRITE FRITE FRITE FRITE	* SYTL SYTL SYTL SYTL SYTL SYTL SYTL SYTL	AHSDA AHSDA AHPNA DHPDA GHPDA ANPDA ANPNA ANPNA	YN : YN : YN : YN : YN : YN : YN : YN :	159 159 158 158 158 159 159	
Bet Bet Mal Mal Mal Mal Mal Mal	v v v d d d d d d	1.0101 1.1801 1.1301 1.0105 1.0201 1.0404 1.0501 1.0502 1.06A01 1.06A02		YNYSYIE YSYSHIE YSYSHIE YSYTLIE YSYSHIE YSYSHIE YSYSHIE YSYSLIE YSYSLIE YSYSLIE	* GGPIG GGALG GCPLG GDALT GDALS GDALS GDALS GDALS GDALS GDALS	DTLEK DTLEK DTLEK DTLEK DTLEK DTLEK DTLEK DKLEK DKTEK DKTEK	100 ISNE ISNE ISYE ISYE ISYE IAYE IAYE ISYE	TKLVA IKLVA IKLVA IKLVA IKLVA IKLVA IKLVA	* TPDGG TPDGG C-GSG S-GSG SPDGG S-DSG S-DSG S-GSG S-GSG	SILKI SILKI SILKI SIIKS SIIKI SIIKI SIIKI SIIKI SIIKI	120 SNK2H SNK2H TSK2H ISH2H ISH2H TSH2H TSH2H TSH2H TSH2H TSH2H TSH2H	IKGDHI IKGDHI IKGD <mark>IS</mark> IKG <mark>NI</mark> IKGDVI IKG <b>C</b> VI IKG <b>C</b> VI IKG <b>DV</b> I IKG <b>DV</b> I	★ YKAEQ HKAEH IKEEH IKEEH IKEEH IKEEH IKEEH IKEEH	VKASK IKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK	140 PMGET PKAEA PKARG PKARG PKARG PKASG PKASS DKARG DKARG	LLRAVE LLRAVE LFKAVE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE	× SYTL SYTL SYTL SYTL SYTL SYTL SYTL TYTU TYTU NYTU	AHSDA AHSDA AHPNA DHPDA GHPDA GHPDA ANPDA ANPNA ANPNA ANPDA ANPDA	YN : YN : YN : YN : YN : YN : YN : YN :	159 159 158 158 159 159 159 159	
Bet Bet Mal Mal Mal Mal Mal Mal Mal	v v d d d d d d	1.0101 1.1801 1.1301 1.0105 1.0201 1.0404 1.0501 1.0502 1.06A01 1.06B01		YNYSYIE YSYSHIE YSYTLIE YSYTLIE YSYSHIE YSYSHIE YSYSLIE YSYSLIE YSYSLIE YSYSLIE	* GGPIG GGALG GGPLG GDALT GDALS GDALS GDALS GDAIS GDAIS	DTLEK DTLEK DTLEK DTLEK DTIEK DKLEK DKLEK DKLEK DKLEK DKLEK	100 ISNE ICNE ISYE ISYE ISYE IAYE IAYE ISYE ISYE	IKIVA IKIVA IKIVA IKIVA IKINA IKINA IKIVA IKIVA	* TPDG G TPDG G C - GS G S - GS G S - GS G S - GS S	S 11 KI S 11 KI S 11 KI S 11 KS S 11 KS S 11 KI S 11 KI S 11 KN S 11 KN S 11 KN	120 SNK TSK TSK TSK TSK TSK TSK TSK TSK TSK TS	IKGDH IKGDH IKGDIS IKGNI IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV	★ M:AE0 M:AEH IKEEH IKEEH IKEEH IKEEH IKEEH IKEEH IKEEH IKEEH	VKASK IKAIK IKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK	140 MGET KGAG KGAG KHG KGC KSG KSG KSG KSS DKHG RG KG	LLRAVE LFAVE FKITE FKITE FKITE FKITE FKITE FKITE FKITE	× SYIII SYIII SYIII SYIII SYIII SYIII SYIII TYIII TYIII TYIII NYIII	AHSDA AHSDA AHPNA DHPDA GHPDA GHPDA GHPDA ANPDA ANPNA ANPNA ANPDA ANPDA ANPDA	YN : YN : YN : YN : YN : YN : YN : YN :	159 159 158 158 159 159 159 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal	v v d d d d d d d d	1.0101 1.1801 1.0105 1.0201 1.0404 1.0501 1.0502 1.06A02 1.06A02 1.06B01 1.06B01		YNYSYIE YSYSHIE YSYTLIE YSYTLIE YSYSHIE YSYSHIE YSYSLIE YSYSLIE YSYSLIE YSYSLIE YSYSLIE	* GCPIG GCALG GCPLG GDALTI GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS	DTLEK DTLEK DTLEK DTLEK DTTEK DKLEK DKLEK DKTEK DKTEK DKTEK DKTEK	100 ISNE ISNE ISYE ISYE ITYE IAYE ISYE ISYE ISYE ISYE	IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA	* TPD G TPD G G C - G S G S - G S - G S S - G S S - G S S - G S S - G S S - G S S - G	SILKI SILKI SILKI SIIKS SIIKI SIIKI SIIKI SIIKN SIIKN SIIKN	120 SNKVH SNKVH TSKVH ISHVH TSHVH TSHVH TSHVH TSHVH TSHVH ISHVH	IKGDH IKGDH IKGDIS IKGDY IKGDY IKGDY IKGDY IKGDY IKGDY IKGDY IKGDY IKGDY	VEACO MEACH INCOM INCOM INCOM INCOM INCOM INCOM INCOM INCOM INCOM	VKASK TKAIK UKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK	140 MGET KARA KGAG KAGG KAGG KAGG KAGG KAG KAG RAG RAG RAG	LIRAVE LIRAVE LFKAVE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE	* SYTL SYTL SYTL SYTL SYTL SYTL SYTL SYTL	AHSDA AHSDA AHPNA DHPDA (GHPDA (GHPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA	YN : YN : YN : YN : YN : YN : YN : YN :	159 159 158 158 159 159 159 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.0105 1.0201 1.0404 1.0501 1.0502 1.06A01 1.06A02 1.06B02 1.06B02		VNYSVIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE VSYSUIE	GCPICI GCALCI GCPLCI GDALTI GDALSI GDALSI GDALSI GDALSI GDATSI GDATSI GDATSI GDATSI	DTLEK DTLEK DTLEK DTTEK DTTEK DKLEK DKLEK DKTEK DKTEK DKTEK DKTEK	100 ISNE ISNE ISYE ISYE IAYE IAYE ISYE ISYE ISYE ISYE	IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA	* TPDG G TPDG G C-GS G S C-GS G S C-GS G S C-GS G S C-GS G S C-GS G S C-GS G S C-GS G S C-GS G S C-GS G C-GS	SILKI SILKI SILKI SIIKS SIIKI SIIKI SIIKI SIIKI SIIKI SIIKI SIIKI	120 SNK2H SNK2H I SKH I SKH I SKH T SKCH T SKCH T SKCH I SKH I SKH I SKH I SKH	IKGDH IKGDH IKGDI IKGDV IKGV IKGV IKGV IKGV IKGV IKGV IKGV IKG	TKEEN IKEEN IKEEN IKEEN IKEEN IKEEN IKEEN IKEEN IKEEN IKEEN	VKASK IKAIK IKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK	140 MGET KAEA KAG KAG KAG KASG KASG KASG DKAG DKAG DKAG DRAG DRAG	LIRAVE LFKAVE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE LFKLTE	× SYTL SYTL SYTL SYTL SYTL SYTL SYTL SYTL	AHSDA AHSDA AHPDA COHPDA COHPDA COHPDA COHPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA	YN : YN : YN : YN : YN : YN : YN : YN :	159 159 158 158 159 159 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0105 1.0201 1.0501 1.0502 1.06A02 1.06B01 1.06B03 1.06C01 1.06C01		YNYSYIE YSYSIE YSYTIE YSYTIE YSYTIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE	GPTC GCALC CCPLC GDALT GDALTS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS	DTLEK DTLEK DTLEK DTTEK DTTEK DKLEK DKLEK DKTEK DKTEK DKTEK DKTEK DKTEK	100 ISNE ISNE ISYE ISYE ISYE ISYE ISYE ISYE ISYE ISY	IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA	* TPDG G TPDG G C-GSG C-GSG SPDG SPDG SPDG SPDG G SPDG G SPGG G S-GSG		120 SNK TH SNK TH TSK TH TSK TH TSH TSH T TSH TSH T TSH TH TSH TH TSH TH TSH TH TSH TH TSH TH TSH TH TSH TH TSH TH TSH TH	IKGDH IKGDH IKGDIS IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV	A A A A A A A A A A A A A A	VKASK IKAIK IKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK	140 MGET KGAG KAGG KAGG KAGG KAG KAG KAG RAG RAG RAG RAG	LIRAVE FRAVE FRAVE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE	* SYII SYII SYIF SYIF SYIF SYIF SYIF NYIV NYIV NYIV NYIV NYIV NYIV	AHSDA AHSDA AHPNA CHPDA CHPDA CHPDA CHPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA	XN : XN : XN : XN : XN : XN : XN : XN :	159 159 158 158 159 159 159 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.0105 1.0201 1.0501 1.0501 1.06001 1.06602 1.06602 1.06603 1.06602		VNYSVIE VSYSVIE VSYVIE VSYVIE VSYSVIE VSYSVIE VSYSLE VSYSLE VSYSLE VSYSLE VSYSLE VSYSLE VSYSLE VSYSLE VSYSLE	* GCPIC GCALCI GCALCI GDALTI GDALSI GDALSI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI	DTLEK DTLEK DTLEK DTLEK DTTEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK	100 ISNE ISNE ISNE ISNE ISNE ISNE ISNE ISNE	IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA IRIVA	* TPDG G TPDG G C - G S - - - - - - - - - - - - -		120 SNK TH SNK TH TSK TH TSK TH TSH CH TSH CH TSH CH TSH TSH TH TSH TH	IKGDH IKGDH IKGDH IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV	VAC MARACI MARACI INCOMI	VKASK IKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK	140 MGET KGAG KGAG KGAG KGAG KGG KGG KGG KGG KG KG KG KG KG KG KG K	LIRAVE FRAVE FRAVE FRITE FRITE FRITE FRITE FRITE FRITE FRITE FRITE FRITE FRITE	* SYIL SYIL SYIL SYIL SYIL SYIL SYIL SYIL	AHSDA AHSDA AHPNA DHPDA GHPDA GHPDA GHPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA	XM       :         XM <td< td=""><td>159 159 158 159 159 159 158 158 158 158 158 158</td><td></td></td<>	159 159 158 159 159 159 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0105 1.0201 1.0501 1.0502 1.06601 1.06602 1.06603 1.06C01 1.06C03		YNYSYIE YCYSIE YCYSIE YSYIIE YSYIE YSYNE YSYNE YSYSIE YSYSIE YSYIE YSYIE YSYIE YSYIE YSYIE YSYIE YSYIE	* GCPIC GCALCI GCALCI GDALTI GDALS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS	DTLEK DTLEK DTLEK DTLEK DTLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK	100 ISTE ISTE ISTE ISTE ISTE ISTE ISTE ISTE	IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA	* TPDG G TPDG G C C C C C C C C C C C C C		120 SMK 2H SMK 2H T SK 2H T SK 2H T SH CH T SH CH T SH CH T SH 2H T SH	IKGDH IKGDH IKGDH IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV	* M A C M A C H I KEEH I KEEH	VKASK IKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK VKAGK	140 MGET K EAC K GAG K GAG K GAG K GG K GG K GG R G R HG R HG R HG R HG R HG R HG R	LRAVE FFAVE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE	× SVII SVII SVII SVII SVII SVII TVIV TVIV	AHSDA AHSDA AHPNA DHPDA CHPDA CHPDA CHPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA	XM 2 XM 2 XM 2 XM 2 XM 2 XM 2 XM 2 XM 2	159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0105 1.0201 1.0502 1.06A01 1.06B02 1.06B03 1.06C02 1.06C03 1.0701		YNYS YIE YSYSHIE YSYSHIE YSYTIIE YSYTIIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYTIE YSYTIE YSYTIE YSYTIE YSYTIE YSYTIE YSYTIE	* GGPIG GGALG GCPLG GDALTI GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS	DTLEK DTLEK DTLEK DTLEK DTLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK	LOO I SITE I CHE I SITE I SITE	IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA IKIVA	* TPDG G TPDG G C C C C C C C C C C C C C		120 SMK 241 T SK 241	KGDH IKGDH IKGDI IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV IKGDV		VRASK IKAIK VRAGK VRAGK VRAGK VRAGK VRAGK VRAGK VRAGK VRAGK VRAGK VRAGK VRAGK	140 MGET KAGAG KAGAG KAGAG KAGG KAGG KAGG KAGG	LRAVE FFAVE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE FKLTE	STIL STIL STIL STIL STIL STIL STIL STIL	AHSDA AHSDA AHPNA CHPDA CGHPA CGHPA ANPA ANPA ANPA ANPA ANPA ANPA ANPA A	XM : 1 XM	159 159 159 159 159 159 159 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0105 1.0201 1.0601 1.06602 1.06603 1.06603 1.06603 1.06603 1.06603 1.0701 1.0702 1.0801		YNYSYN E YSYSM E YSYSM E YSYTH E YSYTH E YSYSM E YSYSH	* GGPIG GGALG GCPLG GDALTI GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS GDATS	DTLEK DTLEK DTLEK DTLEK DTLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK DKLEK ETLEK	LUU ISTE ISTE ISTE ISTE ISTE ISTE ISTE ISTE	TKI VA TKI VA TKI VA TKI VA TKI TA TKI TA TKI VA TKI VA TKI VA TKI VA TKI VA TKI VA	* TPDG G TPDG G G G G G G G G G G G G G	5 TLKI 5 TLKI	120 SNK TSK2H TSK2H TSK2H TSH2 TSH2 TSH2 TSH2 TSH2 TSH2 TSH2 TS	KGDH KGDH KGDY KGDY KGDY KGDY KGDY KGDY KGDY KGDY		UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK	140 MGET KGAG KAGAG KAGAG KAGAG KAG KAG KAG RAG RAG RAG RAG RAG RAG RAG RAG RAG R	LRAVE FRAVE FRAVE FRI TE FRI TE	* SYILL SYILL NYIV SYIP AYIV SYIP AYIV YYV NYIV NYIV NYIV NYIV NYILL NYILL	AHSDA AHSDA AHPNA CHPDA CGHPA CGHPA ANPA ANPA ANPA ANPA ANPA ANPA ANPA A	XM       :         XM <td< td=""><td>159 159 159 159 159 159 159 158 158 158 158 158 158 158 158</td><td></td></td<>	159 159 159 159 159 159 159 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0201 1.0201 1.0501 1.0502 1.06801 1.06802 1.06803 1.06603 1.06603 1.06003 1.0701 1.0701 1.0701 1.0801 2.0801		YNYS VIE YSYSHIE	* GCPIC GCALC GCALC GDALT GDALS GDALS GDALS GDAIS GDAIS GDAIS GDAIS GDAIS GDAIS GDAIS GDAIS GDAIS GDAIS GDAIS GDAIS	DTLEK DTLEK DTLEK DTLEK DTLEK DTLEK DKEK DKEK DKEK DKEK DKEK DKEK DKEK D	100 ISTE ISTE ISTE ISTE ISTE ISTE ISTE ISTE		* TPDGG TPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SCGG SCGG SCGG SCGG SCGG SCGG SCGG S		120 SNK TSK2H TSK2H ISH2H TSH2H TSH2H TSH2H ISH2	IKGDIG IKGDIG IKGDIG IKGDIG IKGDVE IKGDVE IKGDVE IKGDVE IKGDVE IKGDVE IKGDVE IKGDVE IKGDVE IKGDVE IKGDVE		UKASK IKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK	140 MGET KAGG KAGG KAGG KAGG KAGG KAGG RAG RAG RAG RAG RAG RAG RAG RAG RAG	LIRAVE LIRAVE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE FKALTE	* SVIL SVIL SVIL SVIL SVIL SVIL SVIL SVIL	AHSDA AHSDA AHSDA CHPDA CGHPDA CGHPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA ANPDA CHUDA EHQDA	XM         :	159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.0105 1.0201 1.0501 1.0501 1.0502 1.06801 1.06801 1.06602 1.06803 1.06603 1.06603 1.0701 1.0702 1.0801 1.0801 1.0801		YNYS VIE YSYSME YSYSME YSYSME YSYSME YSYSME YSYSME YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE YSYSLE	GOPICI GCALCI GCALTI GDALTI GDALTI GDALSI GDALSI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI GDAISI	DTL:K DTL:K DTL:K DTL:K DTL:K DK:K DK:K DK:K DK:K DK:K DK:K DK:K D	100 1 S 10 1 C 10 1 S 1		* TPDCG TPDCG CCSCC CCSC CCSC CCSC CCSC CCSC CCSC CCSC CCSC			I KOD I I KOD I KOD I I KOD I I KOD I KOD I I KOD I I KOD I KOD I I KOD I KOD I I KOD I KOD I I KOD I KOD I I KOD I I KOD I KOD I I KOD I KOD I I KOD I KOD I I KOD I I KOD I KOD I I KOD I I KOD I KOD I I KOD I I KOD I I KOD I KOD I I KOD		UKASK IKASK IKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK	140 MGET KAG KAG KAG KAG KAG KAG KAG RAG RAG RAG RAG RAG RAG RAG RAG RAG R	LIRAVE LIRAVE FFAUE FFAUE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE FKLIE	* SYIII NYIY SYIIF SYIF SYIF SYIF SYIF SYIF NYIF NYIF NYIF NYIF NYIF NYIF	AHS DA AHS DA AHS DA ANDE DA GHE DA CGHE DA ANE DA ANE DA ANE DA ANE DA ANE DA ANE DA ANE DA ANE DA ANE DA EHC DA EHC DA EHC DA	AM 2	159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Nal Nal Nal Nal Nal Nal Nal Nal Nal Nal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.0105 1.0201 1.0501 1.0501 1.06801 1.06802 1.06801 1.06603 1.06603 1.06603 1.0701 1.0702 1.0801 1.03801 1.03801		YNYSYIE YCYSIIE YCYSIIE YSYIIE YSYNE YSYNE YSYNE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YGYSYE YGYSYE YGYSYE YKYSIE YKYSYE	GPEGLG GGALGI GDALTI GDALTI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI	J TIEK OTIEK DTIEK	100 ISTE ISTE ISTE ISTE ISTE ISTE ISTE ISTE		* TPDCG TPDCG CCSC C	11 K1 11	120 SHK VH SKK VH SKK VH ISK V	IKCD I IKCD I I IKCD I I IKCD I I I I I I I I I I I I I I I I I I I	VK20 INC2 INC2 INC2 INC2 INC2 INC2 INC2 INC2	UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK	140 HGET KGAG KGAG KGAG KGAG KGAG KGAG KGAG RHG RHG RHG RHG RHG RHG RHG RHG RHG RH	L RAVE FRAVE FRA F FRA F F FRA F F F F F F F F F F F F F F F F F F F	* SYIII NYIY SYIII SYIII SYIII SYIII SYIII NYIY NYIY	AHS JA AHS JA DHD JA GHU JA GHU JA GHU JA JANU JA ANU JANU JA ANU	XXN : 2 XXN :	159 159 158 159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Nal Nal Nal Nal Nal Nal Nal Nal Nal Nal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.0201 1.0201 1.0501 1.0501 1.06801 1.06802 1.06603 1.06603 1.0601 1.0702 1.0801 1.03801 1.03801 1.03802		YNYSYIE YCYSIE YCYSIE YSYTEE YSYTEE YSYSEE	GOPICI GALGI GALGI GDALTI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI GDALSI	J TIEK DITEK	100 ISTE ISTE ISTE ISTE ISTE ISTE ISTE ISTE		* TPDGG TPDGG C=65G SPDGG SPDGG SPDGG SPDGG SPDGG S=65G S=65G S=65G S=65G S=65G S=65G S=65G S=65G S=65G S=65G S=65G S=65G S=65G		120 SINK OFF SINK OFF TS NOT TS NOT	IKOD M IKOD I IKO I IKO I IKO V IKO	VRACO HAASH IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM	UKASK IKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK UKASK	140 HGET KGAG K	L RAVE PRAVE FRA DE FRA DE	SYTE SYTE SYTE SYTE SYTE SYTE SYTE SYTE	AHS DA AHS DA AHE MA CHE DA CHE DA CHE DA ANE DA AN	XM 2	159 159 158 159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0105 1.0201 1.0502 1.06802 1.06602 1.06603 1.06703 1.0701 1.0701 1.0702 1.0801 1.03801 1.03801 1.03801		YNYSYIE YSYSHIE YSYSHIE YSYTHIE YSYTHIE YSYSHIE	G GP I GI G GALGI G GALGI G DALSI G GALSI G GALSI	J LEK J LEK	100 ISTR ISTR ISTR ISTR ISTR ISTR ISTR ISTR		* TPDGG AP 6G6 SPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SPGG SP		120 SIRK OH SIRK OH ISNY	KCD I KCD I KCD I KCD I KCD I KCD V KCD V	VIACON NACH IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM IKEEM	UKASK WASK WASK WASK WASK WASK WASK WASK	140 HEET KEAG K HG K SS K SG K SG K SG R HG R HG R HG R HG R HG R HG K SH K SH K SH K SH K SH K SH K SH K SH	L RAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE PRAVE	SYTE SYTE SYTE SYTE SYTE SYTE SYTE SYTE	ANS DA ANG DA AN	XNN 2 XNN 2 XNNN 2 XNN 2	159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0105 1.0404 1.0501 1.06801 1.06802 1.06803 1.06603 1.06603 1.0701 1.0701 1.03801 1.03802 1.		YNYS VIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YSYSHIE YKYSHIE YKYSHIE YKYSHIE YKYSYIE YKYSYIE YKYSYIE YKYSYIE YKYSYIE YKYSYIE	GOPICI COALCI COALT COALT COALT COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI COALSI	1 TIEK DIEK DIEK DIEK DIEK DIEK DIEK DIEK D	100 ISTR ISTR ISTR ISTR ISTR ISTR ISTR ISTR		* TPDCG TPDCG C= 65 C= 65 C SPDCG SPDCG SPDCG C= 65 C S= 75 C S= 75 C S S= 75 C S S S= 75 C S S S S S S S S S S S S S S S S S S		120 SINK YIH SINK YIH ISHYM	I KGD I ( KGD I ( KGD I ( KGD Y ) KGD Y ( KGD Y )	VIAC INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE INFEE	UKASK UKASK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK UKAGK	140 HGET KEAG K HG K SC K SC K SC K SC K SC R HG R HG R HG R HG R HG R HG R HG K SH K SH K SH K SH K SH K SH K SH K SH	L RAVE RAVE FKL 10 FKL 10	* SYDI SYDI SYDI SYDI SYDI SYDI SYDI SYDI	ANS A ANG ANG ANG ANG ANG ANG ANG ANG ANG ANG	XM : 2 XM	159 159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0201 1.0404 1.0501 1.0502 1.06801 1.06801 1.06803 1.06603 1.06003 1.0701 1.0701 1.03001 1.03801 1.03802 1.03801 1.0		YNYS VIE YSYME	GOPICI GOALGI GOALTI GOALTI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI GOALSI	1 1 LEK 1 TEK 1 TE	100 ISTE STE STE STE STE STE ISTE ISTE ISTE		+ TPD66 - - - - - - - - - - - - -		120 SINK YII SIN YII SIN YII SINYII	KCD I KCD I K KCD I K K	VRACO INFER	UKASK IKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK	140 He F Ka G Ka G Ka G Ka G Ka G Ka G Ka G Ka G	U RAVE FRAVE FRAVE FRA D FRA D F FRA D F FRA D F F F F F F F F F F F F F F F F F F F	* SYDI SYDI SYDI SYDI SYDI SYDI SYDI SYDI	AHS DA AHS DA AHS DA CHE DA CH	AM 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	159 159 159 159 159 159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * * * * * * * * * * * * * * * * * *	1.0101 1.1801 1.1301 1.0105 1.0201 1.0502 1.06A02 1.06A02 1.06B01 1.06B03 1.06C01 1.06C03 1.06C03 1.06C03 1.0701 1.0701 1.03B01 1.		YNYSYIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YSYSIE YKYSYE YKYSYE YKYSYE YKYSYE YKYSYE	GPFICI GPFICI GPFICI GPALTI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI GPALSI	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	100 ISTE STE STE STE STE STE ISTE ISTE ISTE		HPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SPDGG SSGS SSGS		120 SINK YII SIN YII ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI ISIYYI TSIYYI TSIYYI TSIYYI TSIYYI TSIYYI TSIYYI TSIYYI	KCD I KCD I KCD I KCD I KCD Y KCD Y	VRACO NRACIN INCER INCER INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN INCEN	UKASK IKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK UKACK	140 MEET KAAG KAAG KAAG KAG KAG KAG KAG KAG RAG RAG RAG RAG RAG RAG RAG RAG RAG R	U RAVE PRAVE PRAVE PRAVE PRU	* SVIII SVIR SVIR AVIV TVIV TVIV NVIV NVIV NVIV NVIV NVIV N	AHS DA AHS DA AHE PA AHE PA CHE DA GHE DA CANE DA ANE PA ANE DA ANE DA A	XM 2	159 159 159 159 159 159 159 158 158 158 158 158 158 158 158 158 158	
Bet Bet Mal Mal Mal Mal Mal Mal Mal Mal Mal Mal	* * * & & & & & & & & & & & & & & & & &	1.0101 1.1801 1.0105 1.0201 1.0501 1.0501 1.06A02 1.06B01 1.06B01 1.06B03 1.06C01 1.06C02 1.06C01 1.0701 1.0702 1.0801 1.03B01 1.03B01 1.03B01 1.03B01 1.03B01 1.03E01 1.03E02 1.03F01 1.0		YNYSYIE YCYSIIE YCYSIIE YSYIE YSYNE YSYNE YSYNE YSYNE YSYSE	GPF (G) GPLG GPLG GPLG GDALTI GDALTI GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS GDALS	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	100 1 S 10 1 S 1		+ TPDGG - C-65G - SPDGG SPDGG - SPDGG - SPDGG - C-65G - C-55G - C-55G		120 SINK OFF SINK OFF	KCD I KCD I K KCD I KCD I KD I KD I KD I KD I KD I KD I KD I K		UKASK TKACK UKACK	140 He EX Ka G Ka G Ka G Ka G Ka G Ka G Ka G R	L RAVE PRAVE	* SVIII SVIR SVIR SVIR AVIV TYIV NYIV NYIV NYIV NYIV NYIV NYIV NYIV N	AHS DA AHS DA AHE PA AHE PA CHE DA GHE DA ANE DA AN	XM : 2 XM	159 158 158 158 158 158 158 158 158 158 158	

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 In (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 Im 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 In 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 v1 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.

Acknowledgements This research was carried out with financial support from the Commission of the European Communities for the projects SAFE QLK1-CT-2000-01394 and exploited unpublished results from FAIR5 CT97-3898 (DARE: Durable Apple Resistance in Europe). We thank Bernard Koller and Cesare Gessler for their SSR data to update the linkage maps of the cross  $PM \times FS$ . We also thank professors Evert Jacobsen and Richard Visser for critical reading of this manuscript. Z.S. Gao is grateful to the International Agricultural Centre, Wageningen, and the ICSC-World Laboratory, Lausanne for providing fellowships to conduct this research.

# 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 nontemplated 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 B 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 miniexon. 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