

Systematic and Phylogenetic Analysis of the Ole e 1 Pollen Protein Family Members in Plants

José Carlos Jiménez-López,
María Isabel Rodríguez-García and Juan de Dios Alché
*Department of Biochemistry, Cell and Molecular Biology of Plants
Estación Experimental del Zaidín, CSIC, Granada,
Spain*

1. Introduction

Pollen allergens are specific substances able to cause IgE-mediated hypersensitivity (allergy) after contact with the immune system [D'Amato et al. 1998]. To date, about 50 plant species have been registered in the official allergen list of the International Union of Immunological Societies (IUIS) Allergen Nomenclature Subcommittee <http://www.allergen.org> as capable of inducing pollen allergy in atopic individuals [Mothes et al. 2004]. These plants are usually grouped as (1) trees (members of the orders: *Fagales*, *Pinales*, *Rosales*, *Arecales*, *Scrophulariales*, *Junglandales*, *Salicales*, and *Myrtales*), (2) grasses (members of the families: *Bambusoideae*, *Arundinoideae*, *Chloridoideae*, *Panicoideae*, and *Poideae*), and (3) weeds (components of families *Asteraceae*, *Chenopodiaceae* and *Urticaceae*) [Hauser et al. 2010].

Allergens are proteins with a broad range of molecular weights (~5 to 50 kDa), which exhibit different features of solubility and stability. More than 10 groups of pollen allergens have been reported. Among all groups of pollen allergens, Pollen Ole e I (Ole) domain-containing proteins are the major allergens, included like-members of the "pollen proteins of the Ole e 1 family" (Accession number: PF01190) within the Pfam protein families database [Finn et al. 2010].

Ole e 1 was the first allergen purified from *Olea europaea* L. [Lauzurica et al. 1998] and named as such according to the IUIS nomenclature [King et al. 1994]. This protein is considered the major olive pollen allergen on the basis of its high prevalence among atopic patients and the high proportion it represents within the total pollen protein content, in comparison with other olive pollen allergens. These include at present another 10 allergens already identified and classified like Ole e 2 to Ole e 11 [Rodríguez et al. 2002, Barral et al. 2004, Salamanca et al. 2010]. Ole e 1 consists of a single polypeptide chain of 145 amino acid residues with a MW of 18–22 kDa, displaying acidic pI and different forms of N-glycosylation [Villalba et al. 1990, Batanero et al. 1994]. Heterologous proteins with a relevant homology have been described in other members of the *Oleaceae* family, such as fraxinus, lilac, jasmine and privet. The polypeptides encoded by the *LAT52* gene from tomato and the *Zmc13* gene from maize pollens also exhibit a high similarity to Ole e 1 [Twell et al. 1989, Hanson et al. 1989]. These plant pollen proteins are structurally related but their biological function is not yet known; though they have been suggested to be

involved in important events of pollen physiology, such as hydration, germination and/or pollen tube growth, and other reproductive functions [Alché et al. 1999, 2004, Tang et al. 2000, Stratford et al. 2001].

Structurally, the Ole domain contains six conserved cysteines which may be involved in disulfide bonds, since no free sulfhydryl groups have been detected in the native protein [Villalba et al. 1993]. Olive Ole e 1 exhibits a high degree of microheterogeneity, mainly concentrated in the third of the molecule closer to the N- terminus. The Ole e I (Ole) domain defining the pollen proteins Ole e I family signature or consensus pattern sequences PS00925 [Sigrist et al. 2010], is characterized by the amino acid sequence [EQT]-G-x-V-Y-C-D-[TNP]-C-R, where “x” could be any residue.

There is a high diversity of proteins sharing the Ole domain among plant species. To date, eleven Ole domain-containing genes have been isolated and characterized from olive pollens [Rodríguez et al. 2002]. Ole-containing proteins include proline-rich proteins, proteins encoding extensin-like domains, phosphoglycerate mutase, tyrosine-rich hydroxyproline-rich glycoprotein, and hydroxyproline-rich glycoprotein. These Ole-containing proteins can exhibit: (1) the pollen Ole signature exclusively, e.g. the ALL1_OLEEU P19963 protein from *Olea europaea* L., (2) both the pollen Ole signature and the replication factor A protein 3 motive pattern (PF08661), e.g. the O49527 pollen-specific protein-like from *Arabidopsis thaliana* (842 residues), (3) both the pollen Ole domain and the phosphoglycerate mutase (PGAM) motif, e.g. the Q9SGZ6 protein from *Arabidopsis thaliana*., and finally (4) both the pollen Ole signature and the reverse transcriptase 2 (RVT2) motif, e.g. the A5AJL0 protein from *Vitis vinifera*.

Several efforts have been made to develop an understandable and reliable systematic classification of the diverse and increasing number of different allergen protein structures. As mentioned above, the classification system widely established for proteins that cause IgE-mediated atopic allergies in humans (allergens) was defined by Chapman et al. (2007). This system uses the first three letters of the genus; a space; the first letter of the species name; a space and an Arabic number. Despite this classification system, protein databases are full of allergen proteins lacking this systematic and comprehensive nomenclature. In other cases, many of the proteins described here have not been described as allergens, or their naming makes no reference to the Ole e 1 family that facilitates their identification. Otherwise, naming in databases is frequently given randomly, on the basis of chromosome location, addressing structural features and functional characterizations or simply using the name of the entire family. In this study, we used a combination of functional genomics and computational biology to name and classify the entire Ole e 1 family, as well as to characterize structurally and functionally the proteins of this superfamily. Our data indicate that the Ole e 1 protein family consists of at least 109 divergent families, which will likely expand as more genomic studies are undertaken, and fully sequenced plant genomes become available.

2. Material and methods

2.1 Database search for Ole e 1 family genes

Sequences of Ole e 1 and Ole e 1-like genes were retrieved from the US National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>), the Uniprot database (<http://www.uniprot.org/>), and the non-redundant expressed sequence tag (EST)

databases using BLASTX, BLASTN and BLAST (low complexity filter, Blosum62 substitution matrix) [Altschul et al. 1997]. Searches were conducted using previously characterized *Olea europaea* L. *Ole e 1* (GenBank Accession number P19963), *Solanum lycopersicum* LAT52 (GenBank Accession number P13447), *Zea mays* Zmc13 (GenBank Accession number B6T1A9), *Arabidopsis thaliana* pollen-specific protein-like (GenBank Accession number O49527), *Arabidopsis thaliana* PGAM containing domain protein (GenBank Accession number Q9SGZ6), and *Vitis vinifera* RVT2 containing domain protein (GenBank Accession number A5AJL0). Full-length amino acid sequences for Ole e 1 proteins were compiled and aligned using ClustalW [Thompson et al. 1994]. Genetic distances between pairs of amino acid sequences were calculated with Bioedit V7.0.5.3 [Hall 1999]. Consensus protein sequences were derived from these original alignment, and further analyzed for the presence of putative functional motifs using the PROSITE database [Sigrist et al. 2010], of biologically meaningful motif descriptors derived from multiple alignments and the ScanProsite program [de Castro et al. 2006], from the Expert Protein Analysis System (ExPASy) proteomics server of the Swiss Institute of Bioinformatics [Gasteiger et al. 2003]. Finally, the consensus protein sequences were submitted to BLASTP analysis to identify homologous proteins from other plant species.

2.2 Revised/unified nomenclature

In order to provide a revised and unified nomenclature for Ole e 1-like gene superfamily, we developed a sequence-based similarity approach to classify all the retrieved sequences using a previously developed gene nomenclature model [Kotchoni et al. 2010]. For this new nomenclature, Ole e 1 protein sequences that are more than 40% identical to previously identified Ole e 1 sequences compose a family, and sequences more than 60% identical within a family, compose a gene subfamily. Protein sequences that are less than 40% identical would describe a new Ole e 1 gene family. Taking olive protein Ole e 1_57A9 (previous name Ole e 1, major olive pollen allergen) as an example for the revised nomenclature (Table 1), Ole e 1 indicates the root; the digits (57) indicates a family and the first letter (A) a subfamily, while the final number (9) identifies an individual gene within a subfamily. The revised nomenclature is therefore composed of an assigned gene symbol (Ole e 1) (abbreviated gene name) for the whole gene superfamily. The gene symbol must be (i) unique and representative of the gene superfamily; (ii) contain only Latin letters and/or Arabic numerals, (iii) not contain punctuation, and (iv) without any reference to species. These newly developed criteria have been applied to database curators to generate the unified Ole e 1 gene families/classes regardless of the source of the cloned gene(s).

2.3 Sequence alignments and phylogenetic analyses

The retrieved Ole e 1 protein families were used to generate a phylogenetic tree using ClustalW [Thompson et al. 1994]. The alignment was created using the Gonnet protein weight matrix, multiple alignment gap opening/extension penalties of 10/0.5 and pairwise gap opening/extension penalties of 10/0.1. These alignments were adjusted using Bioedit V7.0.5.3 [Hall 1999]. Portions of sequences that could not be reliably aligned were eliminated. Phylogenetic tree was generated by the neighbourjoining method (NJ), and the branches were tested with 1,000 bootstrap replicates. The tree was visualized using Treedyn program [Chevenet et al. 2006].

2.4 Ole e 1 superfamily: Protein modeling and structural characterization

In order to study the structural and conformational variability between the Ole e 1 protein families, selected members of the Ole e 1 superfamily were modelled using SWISS-MODEL server, via the ExPASy web server [Gasteiger et al. 2003]. The initial modelled Ole e 1 structures were subjected to energy minimization with GROMOS96 force field energy [van Gunsteren et al. 1996] implemented in DeepView/Swiss-PDBViewer v3.7 [Guex and Peitsch 1997] to improve the van der Waals contacts and to correct the stereochemistry of the improved models. The quality of the models was assessed by checking the protein stereochemistry with PROCHECK [Laskowski et al. 1993] and the protein energy with ANOLEA [Melo et al. 1997, 1998]. Ramachandran plot statistics for the models were calculated to show the number of protein residues in the favoured regions.

3. Results

3.1 The Ole e 1 protein families: Revised and unified nomenclature

In order to provide a revised/international consensus and unified nomenclature for the Ole e 1 gene superfamily, we first retrieved all the Ole e 1 and Ole e 1-like gene sequences using PS00925 as the major molecular consensus defining the entire superfamily of Ole e 1 proteins. We next verified all annotated plant Ole e 1 open reading frames (ORFs) using Ole e 1 sequence domains. A complementary and comparative study was developed by using Uniprot database to validate the molecular function and previous denomination of each Ole e 1 protein. Our searches resulted in the identification of 571 sequences encoding Ole e 1 and Ole e 1 like proteins from a wide variety of plant species, with the diagnostic motif PS00925 (Table 1). According to the established criteria (see Material and Methods), these sequences integrated 109 Ole e 1 gene families which have been attributed to different functional categories including extensins and extensin-like proteins, proline-rich proteins, hydroxyproline-rich glycoproteins, tyrosine-rich/hydroxyproline-rich glycoproteins, hydrolases, phosphoglycerate mutases, arabinogalactan proteins, etc. (Table 1).

Among the sequences retrieved, Ole e 1₄₈ is the most extensive family with 63 gene members encoding for different pollen-specific protein C13 homologues, followed by Ole e 1₅₇ family with 42 gene homologues encoding Ole e 1 (the olive major pollen allergen), Ole e 1₁₆ with 26 gene members encoding proline-rich proteins, and Ole e 1₅₂ with 22 members encoding LAT52 homologues (Table 1). The number of Ole e 1 genes greatly varied from one plant species to another. The genus *Oryza* included the highest number of Ole e 1 genes (143), followed by *Arabidopsis* with 95 genes (Table 1). At present, more than half of the catalogued Ole e 1 families encoded a single Ole e 1/Ole e 1-like gene, which was in most cases “uncharacterized” (Table 1).

The total number of genes in the Ole e 1 superfamily is expected to increase steadily with time, mainly due to the genomic sequencing of additional species like *Olea europaea* L. (http://www.gen-es.org/11_proyectos/PROYECTOS.CFM?pg=0106&n=1). Regardless of the plethora of Ole e 1 genes yet to be identified/characterized, their classification and relationship to the entire extended Ole e 1 gene superfamily will be easy owing to this nomenclature building block that catalogues newly identified/characterized Ole e 1 gene products only on the basis of sequence similarity to previously characterized Ole e 1 gene products.

Ole e 1 Family	Revised annotation	Previous annotation	GeneBank Accession number	Source
1	Ole e 1.1A1	At4g17215	Q6RAZ6	ARATH
1	Ole e 1.1A2	-	Q6LEVR	ARATH
1	Ole e 1.1A3	ARALYDRAFT_493155	D7MC15	ARALY
1	Ole e 1.1A4	40.00006	Q2A9B5	BRAOL
1	Ole e 1.1A5	31.00008	Q2A9F3	BRAOL
1	Ole e 1.1B1	ARALYDRAFT_403053	D7LEF7	ARALY
1	Ole e 1.1B2	At2g40113	Q58FY6	ARATH
1	Ole e 1.1B3	-	Q6LEB2	ARATH
1	Ole e 1.1B4	At5g47835	Q29PT1	ARATH
1	Ole e 1.1B5	ARALYDRAFT_330672	D7MP26	ARALY
2	Ole e 1.2A1	POPTDRRAFT_818926	B9HCD0	POPTR
2	Ole e 1.2A2	POPTDRRAFT_776772	B9H16	POPTR
2	Ole e 1.2B1	VIT_0006138001	D7SSK6	VITVI
2	Ole e 1.2C1	-	C0T3E9	SOYBN
2	Ole e 1.2D1	RCOM_0860870	B8SA60	RICCO
3	Ole e 1.3A1	Os1_33016	B9BG44	ORYSI
3	Ole e 1.3A2	Os10g0206500	Q109X3	ORYSJ
3	Ole e 1.3B1	OSJNB0014J14.3	Q7G7E7	ORYSJ
3	Ole e 1.3B2	QJ1004_F02.9	Q8RV11	ORYSJ
3	Ole e 1.3C1	OSJNB0014J14.29	Q8SEU0	ORYSJ
3	Ole e 1.3D1	Os10g0205000	Q109X1	ORYSJ
3	Ole e 1.3D2	Os1_33026	B9BG53	ORYSI
3	Ole e 1.3E1	SORBI DRAFT_01g013620	C5WR76	SORBI
3	Ole e 1.3F1	-	B4FE66	MAIZE
4	Ole e 1.4A1	SELMODRAFT_444021	D8SBK5	SEMLL
4	Ole e 1.4A2	SELMODRAFT_443385	D8SOX9	SEMLL
5	Ole e 1.5A1	ARALYDRAFT_481629	D7LGX1	ARALY
5	Ole e 1.5A2	At2g27385	Q6NLE8	ARATH
5	Ole e 1.5B1	-	C6SVU8	SOYBN
5	Ole e 1.5B2	-	C6T8F4	SOYBN
5	Ole e 1.5C1	-	C6GZD3	SOYBN
5	Ole e 1.5D1	POPTDRRAFT_821598	APIS57	POPTR
5	Ole e 1.5D2	RCOM_1281970	B9SCW4	RICCO
5	Ole e 1.5D3	VITISV_031997	ASBY12	VITVI
5	Ole e 1.5E1	At5g22430	Q9FMQ8	ARATH
5	Ole e 1.5E2	-	Q6LH14	ARATH
5	Ole e 1.5E3	ARALYDRAFT_351286	D7MOX5	ARALY
6	Ole e 1.6A1	-	B6TL01	MAIZE
6	Ole e 1.6A2	-	B4FQ66	MAIZE

Ole e 1 Family	Revised annotation	Previous annotation	GeneBank Accession number	Source
6	Ole e 1.6B1	-	B6TAH9	MAIZE
6	Ole e 1.6C1	ShB4g021840	C5XTZ8	SORBI
6	Ole e 1.6D1	B1136H02.23	Q6EPW8	ORYSJ
7	Ole e 1.7A1	SELMODRAFT_405039	D8QY68	SEMLL
7	Ole e 1.7A2	SELMODRAFT_414879	D8RTV5	SEMLL
8	Ole e 1.8A1	SELMODRAFT_449129	D8TAZ3	SEMLL
8	Ole e 1.8B1	SELMODRAFT_409805	D8RC11	SEMLL
8	Ole e 1.8C1	SELMODRAFT_448128	D8TAZ1	SEMLL
9	Ole e 1.9A1	At2g21140	Q8SKP5	ARATH
9	Ole e 1.9A2	Proline-rich protein 2	Q9M7P0	ARATH
9	Ole e 1.9A3	ARALYDRAFT_300523	D7LL03	ARALY
9	Ole e 1.9B1	Extensin-like protein	Q9M676	ARATH
9	Ole e 1.9B2	Proline-rich protein 4	Q9M7N8	ARATH
9	Ole e 1.9B3	At7g38770/09A11_50	Q9T855	ARATH
9	Ole e 1.9B4	ARALYDRAFT_430641	D7MFN2	ARALY
10	Ole e 1.10A1	Proline-rich protein	Q9M677	NICGL
10	Ole e 1.10B1	VIT_10024081001	D7U5A0	VITVI
10	Ole e 1.10B2	POPTDRRAFT_700888	B9HRA8	POPTR
10	Ole e 1.10C2	POPTDRRAFT_156915	B9H154	POPTR
10	Ole e 1.10D2	RCOM_0660490	B8STC5	RICCO
11	Ole e 1.11A1	Proline-rich protein	Q8Q666	SOLTU
12	Ole e 1.12A1	VITISV_029041	A5BQP2	VITVI
13	Ole e 1.13A1	VITISV_029838	A5BQP1	VITVI
13	Ole e 1.13A2	VITISV_029837	A5BQP0	VITVI
13	Ole e 1.13B1	VIT_00024076001	D7U587	VITVI
14	Ole e 1.14A1	proline-rich protein	Q93WF4	ORYSA
14	Ole e 1.14A2	Os10g0145100	C7J7T1	ORYSJ
14	Ole e 1.14A3	proline-rich protein	Q93WY5	ORYSA
14	Ole e 1.14A4	Os10g0145090	Q7XGT3	ORYSJ
14	Ole e 1.14A5	proline-rich protein	Q94H18	ORYSA
14	Ole e 1.14A6	Os10g0145290	Q7XGT1	ORYSJ
14	Ole e 1.14A7	OsJ_30733	A3C2J8	VITVI
14	Ole e 1.14A8	OsJ_12924	A2XKE9	ORYSI
14	Ole e 1.14A9	OsJ_12923	A2XKE7	ORYSI
14	Ole e 1.14A10	OsJ_12921	B8AP23	ORYSI
14	Ole e 1.14A11	OSJNB0031A07.6	Q84H17	ORYSJ
14	Ole e 1.14A12	OsJ_30737	A3C2K3	ORYSJ
14	Ole e 1.14A13	Os10g0145400	Q7XGT0	ORYSJ
14	Ole e 1.14A14	OsJ_30734	A3C2K0	ORYSJ

Table 1. The Ole e 1 protein superfamily: new and unified nomenclature. ARATH: *Arabidopsis thaliana*; ARALY: *Arabidopsis lyrata*; BETPN: *Betula pendula*; BRAOL: *Brassica oleracea*; BRARP: *Brassica rapa*; CAPAN: *Capsicum annuum*; CARAS: *Cardaminopsis arenosa*; CHE1: *Chenopodium album*; CROSA: *Crocus sativus*; DAUCA: *Daucus carota*; EUPPU: *Euphorbia pulcherrima*; FRAEX: *Fraxinus excelsior*; GOSBA: *Gossypium barbadense*; GOSHE: *Gossypium herbaceum*; GOSHI: *Gossypium hirsutum*; GOSKI: *Gossypoides kirkii*; HYAOR: *Hyacinthus orientalis*; LigVu: *Ligustrum vulgare*; LILLO: *Lilium longiflorum*; LOLPE: *Lolium perenne*; MAIZE: *Zea mays*; MEDTR: *Medicago truncatula*; NICAL: *Nicotiana alata*; NICGL: *Nicotiana glauca*; NicLa: *Vitis pseudoreticulata*; OleEu: *Olea europaea*; ORYSI: *Oryza sativa*; PETCR: *Petroselinum crispum*; PETHY: *Petunia hybrida*; PHAVU: *Phaseolus vulgaris*; PHEPR: *Phleum pratense*; PHYPA: *Physcomitrella patens*; PICSI: *Picea sitchensis*; PLALA: *Platanus lanceolata*; POPTR: *Populus trichocarpa*; RICCO: *Ricinus communis*; SALKA: *Salsola kali*; SAMNI: *Sambucus nigra*; SELML: *Selaginella moellendorffii*; SOLLI: *Solanum lycopersicum*; SOLTU: *Solanum tuberosum*; SORBI: *Sorgum bicolor*; SOYBN: *Glycine max*; TOBAC: *Nicotiana tabacum*; TRISU: *Trifolium subterraneum*; VITVI: *Vitis vinifera*; 9ROSI: *Cleome spinosa*; (-): uncharacterized.

14	Ole e 1_14A15	OSJNBa0031A07.9	QMH14	ORYSA
14	Ole e 1_14A16	Os10g149300	QJXG57	ORYSJ
14	Ole e 1_14A17	OsJ_32754	A32328	ORYSJ
14	Ole e 1_14B1	proline-rich protein	QAH15	ORYSA
14	Ole e 1_14C1	Os10g149800	Q7XG58	ORYSJ
14	Ole e 1_14D1	OsJ_30739	A3C2K6	ORYSJ
14	Ole e 1_14E1	OsJ_32753	A2Z527	ORYSJ
15	Ole e 1_15A1	OsJ_30735	A3C2K1	ORYSJ
16	Ole e 1_16A1	proline-rich protein	QAH12	ORYSA
16	Ole e 1_16A2	proline-rich protein	A6N1C2	ORYSJ
16	Ole e 1_16A3	Os10g150400	Q7XG54	ORYSJ
16	Ole e 1_16A4	OsJ_32757	A2Z531	ORYSJ
16	Ole e 1_16A5	proline-rich protein	QAH09	ORYSA
16	Ole e 1_16A6	Os10g150800	Q7XG52	ORYSJ
16	Ole e 1_16A7	proline-rich protein	Q8RVW6	ORYSA
16	Ole e 1_16A8	proline-rich protein	QAH10	ORYSA
16	Ole e 1_16A9	Os10g150700	Q7XG53	ORYSJ
16	Ole e 1_16A10	OsJ_32758	A2Z533	ORYSJ
16	Ole e 1_16A11	proline-rich protein	A8M2D9	ORYSJ
16	Ole e 1_16A12	proline-rich protein	QAH11	ORYSA
16	Ole e 1_16A13	Os10g150600	Q7GBX3	ORYSJ
16	Ole e 1_16A14	Proline-rich protein	A6N177	ORYSJ
16	Ole e 1_16A15	OsJ_32758	A2Z532	ORYSJ
16	Ole e 1_16A16	OsJ_32768	BRBF7	ORYSJ
16	Ole e 1_16A17	LOC_Os10g05990	Q3B16	ORYSJ
16	Ole e 1_16B1	Proline-rich protein	QAH13	ORYSA
16	Ole e 1_16B2	Os10g150300	Q7XG55	ORYSJ
16	Ole e 1_16B3	OsJ_32756	A2Z530	ORYSJ
16	Ole e 1_16C1	-	B6J44	MAIZE
16	Ole e 1_16C2	-	B6TBY5	MAIZE
16	Ole e 1_16C3	Proline-rich protein	B6TL3	MAIZE
16	Ole e 1_16C4	So10g062180	C5WP19	SORBI
16	Ole e 1_16D1	-	B6U45	MAIZE
16	Ole e 1_16E1	-	C0PAE2	MAIZE
17	Ole e 1_17A1	So10g053970	C5WP14	SORBI
17	Ole e 1_17A2	Proline-rich protein	Q9SBK4	MAIZE
18	Ole e 1_18A1	So10g026170	C5WP16	SORBI
19	Ole e 1_19A1	proline-rich protein	Q9GB6	ORYSA
19	Ole e 1_19A2	Os10g148100	Q7XG79	ORYSJ
19	Ole e 1_19A3	OsJ_30727	B8GJ22	ORYSJ
19	Ole e 1_19A4	OsJ_32749	A2Z524	ORYSJ
19	Ole e 1_19B1	-	C4JAB8	MAIZE

19	Ole e 1_19B2	-	B4FM91	MAIZE
19	Ole e 1_19B3	Proline-rich protein	So10g027290	MAIZE
19	Ole e 1_19B4	Proline-rich protein	C5WP20	SORBI
19	Ole e 1_19C1	OsJ_10729	A2XEH2	ORYSJ
19	Ole e 1_19D1	Os10g245200	Q10P65	ORYSJ
19	Ole e 1_19E1	So10g041230	C5WQ27	SORBI
19	Ole e 1_19E2	Proline-rich protein	B1T564	MAIZE
19	Ole e 1_19E3	Proline-rich protein	B4FQ36	MAIZE
20	Ole e 1_20A1	proline-rich protein	Q9AGB4	ORYSA
20	Ole e 1_20A2	Os10g0148700	Q7X076	ORYSJ
20	Ole e 1_20A3	OsJ_32750	B1BF44	ORYSJ
21	Ole e 1_21A1	OsJ_10730	A2XEH3	ORYSJ
21	Ole e 1_21A2	Os10g245300	Q10P64	ORYSJ
21	Ole e 1_21B1	So10g041220	C5WQ26	SORBI
21	Ole e 1_21B2	Proline-rich protein	B6TJK1	MAIZE
22	Ole e 1_22A1	OsJ_04802	A2WAZ6	ORYSJ
22	Ole e 1_22A2	Os10g089700	Q5N8V8	ORYSJ
22	Ole e 1_22A3	So10g042800	C5XF44	SORBI
23	Ole e 1_23A1	SELMODRAFT_448971	D9TAT4	SELM
23	Ole e 1_24A1	SELMODRAFT_449207	D9TDP2	SELM
25	Ole e 1_25A1	OH-proline-rich glycoprotein	Q9R8B5	PHAVU
26	Ole e 1_26A1	Tyr., OH-proline-rich glycoprotein	Q40793	PETCR
27	Ole e 1_27A1	SELMODRAFT_419487	D5S850	SELM
27	Ole e 1_27A2	SELMODRAFT_427121	D5SVK5	SELM
28	Ole e 1_28A1	SELMODRAFT_441903	D8RM00	SELM
28	Ole e 1_28A2	SELMODRAFT_430805	D8T809	SELM
28	Ole e 1_28B1	SELMODRAFT_438663	D8R0Q2	SELM
28	Ole e 1_28B2	SELMODRAFT_443338	D8TF68	SELM
29	Ole e 1_29A1	SELMODRAFT_438958	D8R0Q9	SELM
29	Ole e 1_29A2	SELMODRAFT_439721	D8R5C7	SELM
30	Ole e 1_30A1	SELMODRAFT_448085	D8T4Q2	SELM
31	Ole e 1_31A1	SELMODRAFT_413728	D8R013	SELM
31	Ole e 1_31A2	SELMODRAFT_416536	D8R2L3	SELM
32	Ole e 1_32A1	SELMODRAFT_419534	D8R2L1	SELM
32	Ole e 1_32A2	SELMODRAFT_413730	D8R015	SELM
33	Ole e 1_33A1	ARALYDRAFT_489472	D7RM10	ARALY
33	Ole e 1_33A2	At5g15780	Q1LFU8	ARATH
33	Ole e 1_33B1	-	Q1KUY6	PROSO
34	Ole e 1_34A1	Pollen Ole e 1 allergen/oxidant	Q2HSE9	MEDTR
34	Ole e 1_34B1	-	B7FK05	MEDTR
34	Ole e 1_34C1	RCOM_073050	B5SV7	RICCO
34	Ole e 1_34D1	VIT_0036543001	D7RMN1	VITV

35	Ole e 1_35A1	MirDRAFT_AC172742g21v1	A2Q629	MEDTR
36	Ole e 1_36A1	POPTDRRAFT_58421	B5N7M0	POPTR
37	Ole e 1_37A1	OsJ_33013	A2Z573	ORYSJ
37	Ole e 1_37A2	Os10g205700	Q9RV00	ORYSJ
37	Ole e 1_37A3	OsJ_26257	A2YM00	ORYSJ
37	Ole e 1_37B1	OsJ_31017	A3C3E5	ORYSJ
37	Ole e 1_37C1	So10g030310	C5YAM4	SORBI
37	Ole e 1_37C2	-	B4FWB4	MAIZE
38	Ole e 1_38A1	-	A5NKB0	PCIBI
38	Ole e 1_38A2	-	A5NZN8	PCIBI
38	Ole e 1_38A3	-	A5NYK9	PCIBI
39	Ole e 1_39A1	-	A5NZS6	PCIBI
39	Ole e 1_39A2	-	C0PTE8	PCIBI
39	Ole e 1_39A3	-	A5NTZ1	PCIBI
40	Ole e 1_40A1	ARALYDRAFT_564511	D7LC1	ARALY
40	Ole e 1_40A2	At3g26360	Q1L5C3	ARATH
40	Ole e 1_40A3	Pollen proteins Ole e 1 family	B6SLJ6	MAIZE
40	Ole e 1_40A4	ARALYDRAFT_355671	D7M1Z7	ARALY
40	Ole e 1_40A5	At5g41050	Q9FLM4	ARATH
40	Ole e 1_40B1	-	C6T792	SOYBN
40	Ole e 1_40B2	-	C6S215	SOYBN
40	Ole e 1_40B3	-	B7FIQ2	MEDTR
40	Ole e 1_40B4	POPTDRRAFT_1069266	B9GFK1	POPTR
40	Ole e 1_40B5	POPTDRRAFT_174592	B9GF09	POPTR
40	Ole e 1_40B6	POPTDRRAFT_566673	A9PAE7	POPTR
40	Ole e 1_40B7	RCOM_1417040	B9HJ38	RICCO
40	Ole e 1_40B8	VIT_00033047001	D7SVM0	VITV
41	Ole e 1_41A1	P1 clone: MOJ10	Q3L3D9	ARATH
42	Ole e 1_42A1	At5g13140	Q8F196	ARATH
42	Ole e 1_42A2	-	Q1LEU2	ARATH
42	Ole e 1_42A3	ARALYDRAFT_468130	D7M572	ARALY
43	Ole e 1_43A1	VIT5V_011139	ASAJL0	VITV
44	Ole e 1_44A1	POPTDRRAFT_341728	B9GMF5	POPTR
44	Ole e 1_44B1	Hydrolase	B9RM88	RICCO
44	Ole e 1_44C1	POPTDRRAFT_413678	B9QZC8	POPTR
45	Ole e 1_45A1	Ep60	A5Y588	EUPPU
46	Ole e 1_46A1	OsJ_31975	B6DDC0	ORYSJ
46	Ole e 1_46A2	Os10g050200	Q1J011	ORYSJ
46	Ole e 1_46A3	OsJ_29955	B9G4H1	ORYSJ
46	Ole e 1_46A4	So10g023420	C5X936	SORBI
46	Ole e 1_46A5	-	B4F591	MAIZE
47	Ole e 1_47A1	OsJ_38287	B9BPK9	ORYSJ

47	Ole e 1_47A2	Os10g0472800	Q2QR52	ORYSJ
47	Ole e 1_47B1	-	B4FY58	MAIZE
47	Ole e 1_47B2	-	B6TZW5	MAIZE
47	Ole e 1_47B3	So10g15012840	C6JRP6	SORBI
47	Ole e 1_47C1	-	C0PAV0	MAIZE
48	Ole e 1_48A1	Pollen ole e 1 allergen	D7M779	ARALY
48	Ole e 1_48A2	AT5G0130-like protein	B7JUS9	CARAS
48	Ole e 1_48A3	At5g10130	Q1LX15	ARATH
48	Ole e 1_48A4	Pollen-specific protein C13	C3JAS6	ARALY
48	Ole e 1_48A5	B9A08_10	Q4ABQ7	BRARP
48	Ole e 1_48B1	-	B9HFN2	POPTR
48	Ole e 1_48B2	-	B9H515	POPTR
48	Ole e 1_48B3	Pollen-specific protein C13	B9TH40	RICCO
48	Ole e 1_48C1	-	D7R8H5	ARALY
48	Ole e 1_48C2	Allergen-like protein BRSn20	Q9S2Y5	ARATH
48	Ole e 1_48C3	Pollen specific protein	Q4Z077	ARATH
48	Ole e 1_48D1	-	A9PS56	POPTR
48	Ole e 1_48D2	-	A9PS60	POPTR
48	Ole e 1_48D3	-	A9PCW2	POPTR
48	Ole e 1_48D4	-	A9PHV1	POPTR
48	Ole e 1_48D5	Pollen-specific protein C13	B9RJG6	RICCO
48	Ole e 1_48D6	Pollen-specific protein	Q2I307	NICLA
48	Ole e 1_48D7	-	A5BPK1	VITV
48	Ole e 1_48D8	scaffold_357 assembly12x	D7TRB4	VITV
48	Ole e 1_48D9	-	E2LMG1	CRPE
48	Ole e 1_48D10	Allergen-like protein BRSn20	Q9SE84	SAMNI
48	Ole e 1_48E1	Pollen-specific protein	Q8IPK7	GOSBA
48	Ole e 1_48E2	Pollen-specific protein	Q8IPK9	GOSBA
48	Ole e 1_48E3	Pollen-specific protein	Q8IPK5	GOSBA
48	Ole e 1_48E5	Pollen-specific protein	Q8IPK5	GOSBA
48	Ole e 1_48F1	-	C5SVN7	SOYBN
48	Ole e 1_48F2	-	C5SVR9	SOYBN
48	Ole e 1_48F3	-	B7FN25	MEDTR
48	Ole e 1_48G1	LLP-B3 protein	Q6QZ57	LILLO
48	Ole e 1_48H1	H2020H04 protein	Q8L134	ORYSA
48	Ole e 1_48H2	-	A2XSZ2	ORYSJ
48	Ole e 1_48H3	Os10g0338700 protein	Q7XKG2	ORYSJ
48	Ole e 1_48H4	Pollen-specific gene protein	Q40715	ORYSA
48	Ole e 1_48H5	So10g014740	C5YF43	SORBI
48	Ole e 1_48H6	Pollen-specific protein C13	P30350	MAIZE
48	Ole e 1_48H7	Pollen-specific protein C13	B6T1A5	MAIZE

Table 1. (continued). The Ole e 1 protein superfamily: new and unified nomenclature.

4E	Ole e 1_48H8	-	B4FKG2	MAIZE
4E	Ole e 1_48H9	Pollen-specific protein C13	B6T728	MAIZE
4E	Ole e 1_48I1	Putative pollen specific prot.C13	Q8RU50	ORYSJ
4E	Ole e 1_48J2	Os10g0371000 protein	Q0Y39	ORYSJ
4E	Ole e 1_48J3	-	A2Z6J6	ORYSJ
4E	Ole e 1_48J4	Pollen-specific protein C13	B6S340	MAIZE
4E	Ole e 1_48J5	Pollen-specific protein C13	B6T594	MAIZE
4E	Ole e 1_48J6	Sb0012s014530	C6JRR2	SORBI
4E	Ole e 1_48J7	Pollen-specific protein	Q67C4	HYAOR
4E	Ole e 1_48J1	Major pollen allergen Lol p 11	Q7M1X5	LOLPR
4E	Ole e 1_48J2	Pollen allergen Phi p 11	Q8H6L7	PHLPR
4E	Ole e 1_48J3	Sb03g091020	C5XK96	SORBI
4E	Ole e 1_48J4	Pollen allergen Phi p 11	B6T728	MAIZE
4E	Ole e 1_48J5	-	A2YF17	ORYSJ
4E	Ole e 1_48J6	Os06g055800 protein	Q5Z710	ORYSJ
4E	Ole e 1_48J1	Sb02g07280	C5YU02	SORBI
4E	Ole e 1_48J2	-	B4FC11	MAIZE
4E	Ole e 1_48J3	Pollen allergen Phi p 11	B6T595	MAIZE
4E	Ole e 1_48J1	-	B6B6U8	ORYSJ
4E	Ole e 1_48J2	-	A3C1R5	ORYSJ
4E	Ole e 1_48J3	Putative Pollen specific protein C13	Q8S0Z4	ORYSJ
4E	Ole e 1_48J4	Os09g0572800 protein	Q0ZF00	ORYSJ
4E	Ole e 1_48J5	Os07g0590500 protein	Q6ZLH6	ORYSJ
4E	Ole e 1_48J6	-	A3BLQ7	ORYSJ
4E	Ole e 1_48J7	-	A2YH83	ORYSJ
4E	Ole e 1_48J8	Sb02g0912930	C5XK96	SORBI
4E	Ole e 1_48J1	-	C5T353	SOYBN
50	Ole e 1_50A1	Pollen ole e 1 allergen	D7K6V2	ARALY
50	Ole e 1_50A2	Pollen ole e 1 allergen	D7K6V8	ARALY
50	Ole e 1_50A3	P-glycerate mutase 1 like prot.	Q8LD45	ARATH
50	Ole e 1_50A4	-	Q8H7B9	ARATH
50	Ole e 1_50A5	Pollen specific protein	Q42043	ARATH
50	Ole e 1_51A1	F2B1K1.76 protein	Q9G0Z6	ARATH
50	Ole e 1_52A1	A1g29140	Q8IWD6	ARATH
50	Ole e 1_52A2	F2B2A.16 protein	Q8I1A4	ARATH
50	Ole e 1_52B1	A1g45580	Q8B0L9	ARATH
50	Ole e 1_52B2	Ole e 1-like protein	Q8I1P9	ARATH
50	Ole e 1_52B3	-	D7M5G6	ARATH
50	Ole e 1_52B4	A1g18595	Q8NMJ2	ARATH
50	Ole e 1_52B5	Pollen ole e 1 allergen	D7MCQ3	ARALY
50	Ole e 1_52B6	Ole e 1-like protein	Q9F4J8	ARATH

52	Ole e 1_52C1	Ole e 1-like protein	Q49813	BETHP
52	Ole e 1_52D1	Pollen allergen Che e 1	B5S8A5	RICCO
52	Ole e 1_52E1	PN40024	D7JL11	VITI
52	Ole e 1_52F1	-	C6TL27	SOYBN
52	Ole e 1_52F2	-	B7FGN2	MEDTR
52	Ole e 1_52G1	Pollen allergen Che e 1	Q8LGR0	CHE1
52	Ole e 1_52G2	Pollen allergen Cro e 1	Q29W25	CROSA
52	Ole e 1_52H1	Sal k 4	E2D0Z8	SALKA
52	Ole e 1_52I1	-	B8N635	POPTR
52	Ole e 1_52J2	-	B8P8Z0	POPTR
52	Ole e 1_52J3	-	B8I1V1	POPTR
52	Ole e 1_52J1	Anther-specific prot. LAT52	B5S8K9	RICCO
52	Ole e 1_52K1	A81	D7R0W3	GOSHI
52	Ole e 1_52L1	Anther-specific prot. LAT52	P13447	SOLLC
52	Ole e 1_53A1	-	D7KDG8	ARALY
53	Ole e 1_54A1	Pollen-specific protein - like	Q49527	ARATH
53	Ole e 1_55A1	Putative Ole e 1-like protein	A3FA46	NICLA
56	Ole e 1_56A1	Major pollen allergen Phi 1	P82242	PLALA
57	Ole e 1_57A1	Allergen Fra e 1.0101	Q7XAV4	FRAEX
57	Ole e 1_57A2	Fra e 1.0102 major allergen	Q5XEL6	FRAEX
57	Ole e 1_57A3	Major pollen allergen Lig v 1	Q8Z015	LIGVU
57	Ole e 1_57A4	Ole e 1 olive pollen allergen	X76397	OleEu
57	Ole e 1_57A5	Ole e 1 olive pollen allergen	AF532765	OleEu
57	Ole e 1_57A6	Ole e 1 olive pollen allergen	AF532766	OleEu
57	Ole e 1_57A7	Ole e 1 olive pollen allergen	AF532767	OleEu
57	Ole e 1_57A8	Ole e 1 olive pollen allergen	X76396	OleEu
57	Ole e 1_57A5	Ole e 1 olive pollen allergen	P19963	OleEu
57	Ole e 1_57A10	Ole e 1 olive pollen allergen	Ole e 1 Educat	OleEu
57	Ole e 1_57A11	Ole e 1 olive pollen allergen	X76395	OleEu
57	Ole e 1_57A12	Ole e 1 olive pollen allergen	Y12426	OleEu
57	Ole e 1_57A13	Ole e 1 olive pollen allergen	Y12427	OleEu
57	Ole e 1_57A14	Ole e 1 olive pollen allergen	AF500908	OleEu
57	Ole e 1_57A15	Ole e 1 olive pollen allergen	AF515277	OleEu
57	Ole e 1_57A16	Ole e 1 olive pollen allergen	AF515278	OleEu
57	Ole e 1_57A17	Ole e 1 olive pollen allergen	AF515280	OleEu
57	Ole e 1_57A18	Ole e 1 olive pollen allergen	AF515279	OleEu
57	Ole e 1_57A19	Ole e 1 olive pollen allergen	AF515281	OleEu
57	Ole e 1_57A20	Ole e 1 olive pollen allergen	AF532755	OleEu
57	Ole e 1_57A21	Ole e 1 olive pollen allergen	AF532756	OleEu
57	Ole e 1_57A22	Ole e 1 olive pollen allergen	AF532757	OleEu
57	Ole e 1_57A23	Ole e 1 olive pollen allergen	AF532768	OleEu

57	Ole e 1_57A24	Ole e 1 olive pollen allergen	AF532753	OleEu
57	Ole e 1_57A25	Ole e 1 olive pollen allergen	AF532754	OleEu
57	Ole e 1_57A26	Ole e 1 olive pollen allergen	AY137487	OleEu
57	Ole e 1_57A26	Ole e 1 olive pollen allergen	AY137488	OleEu
57	Ole e 1_57A29	Ole e 1 olive pollen allergen	AY137469	OleEu
57	Ole e 1_57A30	Ole e 1 olive pollen allergen	S75766	OleEu
57	Ole e 1_57A31	Ole e 1 olive pollen allergen	Y13426	OleEu
57	Ole e 1_57A32	Ole e 1 olive pollen allergen	AF532758	OleEu
57	Ole e 1_57A33	Ole e 1 olive pollen allergen	AF532761	OleEu
57	Ole e 1_57A34	Ole e 1 olive pollen allergen	AF532762	OleEu
57	Ole e 1_57A35	Ole e 1 olive pollen allergen	AF532759	OleEu
57	Ole e 1_57A36	Ole e 1 olive pollen allergen	AF532764	OleEu
57	Ole e 1_57A37	Ole e 1 olive pollen allergen	AF532763	OleEu
57	Ole e 1_57A38	Ole e 1 olive pollen allergen	AY150888	OleEu
57	Ole e 1_57A39	Ole e 1 olive pollen allergen	AY150881	OleEu
57	Ole e 1_57A40	Allergen Fra e 1	Q6U740	FraEx
57	Ole e 1_57A41	Ole e 1 olive pollen allergen	X76541	OleEu
57	Ole e 1_57A42	Ole e 1 olive pollen allergen	X76540	OleEu
57	Ole e 1_57A43	Ole e 1 olive pollen allergen	X76539	OleEu
58	Ole e 1_58A1	-	B7JNF5	MEDTR
58	Ole e 1_58A2	-	B7JRF3	MEDTR
58	Ole e 1_58B1	-	C6YFJ3	SOYBN
58	Ole e 1_59A1	Extensin-like protein	A9WMR0	PCSI
58	Ole e 1_59A2	-	A9HPL2	PCSI
58	Ole e 1_59A3	Extensin-like protein	E0ZE62	PCSI
58	Ole e 1_59A4	Extensin-like protein	E0ZE90	PCSI
58	Ole e 1_59A5	Extensin-like protein	E0ZE78	PCSI
58	Ole e 1_60A1	AT1g27100/T7N9_16	Q04EJ3	ARATH
58	Ole e 1_60A2	-	Q1LJ22	ARATH
58	Ole e 1_60A3	Pollen ole e 1 allergen	D7MUX1	ARALY
58	Ole e 1_60A4	-	C6SVQ8	SOYBN
58	Ole e 1_60A5	-	C6T474	SOYBN
58	Ole e 1_60A6	-	A9P8A0	POPTR
58	Ole e 1_60A7	-	A9PFL1	POPTR
58	Ole e 1_60A8	-	B9SAF6	RICCO
58	Ole e 1_60A9	PN40024	D7T895	VITI
58	Ole e 1_60A10	-	A2X417	ORYSJ
58	Ole e 1_60A11	Os02g057800 protein	Q6Z411	ORYSJ
58	Ole e 1_60A12	Sb07g069530	C5YJW7	SORBI
58	Ole e 1_60B1	-	B6T3A4	MAIZE
58	Ole e 1_60B1	-	BLR17	PCSI
58	Ole e 1_61A1	-	Q8RWG3	ARATH

61	Ole e 1_61A2	A1g216630	Q0SLF4	ARATH
61	Ole e 1_61A3	-	D7L7M3	ARALY
61	Ole e 1_61B1	-	B9N158	POPTR
61	Ole e 1_61B2	-	A9PG40	POPTR
61	Ole e 1_61B3	-	B9SQJ5	RICCO
61	Ole e 1_61B4	-	D7U593	VITI
62	Ole e 1_62A1	PN40024	A2WZ98	ORYSJ
62	Ole e 1_62A2	-	A3A255	ORYSJ
62	Ole e 1_62B1	-	B6TF27	MAIZE
62	Ole e 1_62B2	-	B4FZU6	MAIZE
62	Ole e 1_62B3	-	B6TR02	MAIZE
63	Ole e 1_63A1	-	B6UFQ0	MAIZE
64	Ole e 1_64A1	-	A9RQ15	PHYPA
64	Ole e 1_64B1	-	A9SHJ0	PHYPA
65	Ole e 1_65A1	-	D8TAV3	SEML
65	Ole e 1_65A2	-	D8TDP3	SEML
66	Ole e 1_66A1	-	D7MY20	ARALY
66	Ole e 1_66A2	proline-rich glycoprotein	O64586	ARATH
66	Ole e 1_66A3	-	D7LH37	ARALY
67	Ole e 1_67A1	-	D7LGE2	ARALY
67	Ole e 1_67A2	A1g233790	P33013	ARATH
67	Ole e 1_67B1	A1g26290	Q9FAZ4	ARATH
67	Ole e 1_67B2	proline-rich protein	Q9WPA7	ARATH
67	Ole e 1_67B3	-	D7RCU8	ARALY
67	Ole e 1_67C1	HyPRP1	Q9IWH3	GOSHI
67	Ole e 1_67D1	Arabinogalactan protein	C6YU77	GOSHI
68	Ole e 1_68A1	-	C6TL02	SOYBN
68	Ole e 1_68A2	proline-rich protein	Q41122	PHYVU
68	Ole e 1_68B1	-	B7F159	MEDTR
68	Ole e 1_68C1	-	A9PAW5	POPTR
68	Ole e 1_68C2	-	A9PA42	POPTR
68	Ole e 1_68D1	-	B9N307	POPTR
68	Ole e 1_68D2	-	B9H7T3	POPTR
68	Ole e 1_68D3	-	A9P845	POPTR
68	Ole e 1_68E1	Structural constituent of cell wall	B9RBC9	RICCO
68	Ole e 1_68F1	PN40024	D7TGB4	VITI
68	Ole e 1_69A1	Arabinogalactan protein	Q9FSW6	DAUCA
68	Ole e 1_69A2	hybrid proline-rich protein PRP1	Q9XES6	TRISU
70	Ole e 1_70A1	Proline-rich protein	Q9T8R4	NICAL
70	Ole e 1_70A2	-	COLLS3	PETHY
70	Ole e 1_70A3	Proline-rich protein 1	Q6QNA3	CAPAN

Table 1. (continued). The Ole e 1 protein superfamily: new and unified nomenclature.

71	Ole e 1_71A1	Pistil extensin-like protein	Q40385	NICAL
71	Ole e 1_71A2	Pistil-specific extensin-like prot.	Q03211	PEXLP
71	Ole e 1_71B1	Pistil extensin-like protein	Q40549	TOBAC
71	Ole e 1_71B1	Pistil extensin-like protein	Q40552	TOBAC
72	Ole e 1_72A1	120 kDa stipe glycoprotein	Q49986	NICAL
72	Ole e 1_72A2	120 kDa pistil extensin-like prot.	Q49728	NiCa
72	Ole e 1_72A3	120 kDa pistil extensin-like prot.	Q49729	NiCa
72	Ole e 1_72A4	120 kDa pistil extensin-like prot.	Q49727	NiCa
72	Ole e 1_72A5	120 kDa pistil extensin-like prot.	Q49732	NiCa
72	Ole e 1_72A6	120 kDa pistil extensin-like prot.	Q49733	NiCa
72	Ole e 1_72A7	120 kDa pistil extensin-like prot.	Q49734	TOBAC
73	Ole e 1_73A1	120 kDa pistil extensin-like prot.	Q49730	NiCa
74	Ole e 1_74A1	Pollen ole e 1 allergen	D7M476	ARALY
74	Ole e 1_74A2	Aldg02270	Q81417	ARATH
75	Ole e 1_75A1	-	C87917	SOYBN
75	Ole e 1_75A2	Drought resistance protein	E0A235	SOYBN
75	Ole e 1_75A3	-	C87425	SOYBN
75	Ole e 1_75B1	-	B9G9P3	POPTR
75	Ole e 1_75B2	-	B9G9P2	POPTR
75	Ole e 1_75C1	-	B8M440	POPTR
75	Ole e 1_75C2	-	B9P957	POPTR
75	Ole e 1_75D1	-	B9GSD7	POPTR
76	Ole e 1_76A1	-	B9SAV3	RICCO
76	Ole e 1_76B1	Structural constituent cell wall	B9SAV4	RICCO
77	Ole e 1_77A1	-	B9GSD1	POPTR
77	Ole e 1_77A2	Structural constituent cell wall	B9SAV3	RICCO
77	Ole e 1_77B1	PN40024	D7U2C5	VITV1
77	Ole e 1_77B2	-	A5B127	VITV1
77	Ole e 1_77C1	PN40024	D7U2C3	VITV1
77	Ole e 1_77C2	-	A5B125	VITV1
77	Ole e 1_77D1	Pollen ole e 1 allergen	D7LGP2	ARALY
77	Ole e 1_77D2	At2g47540	Q22257	ARATH
78	Ole e 1_78A1	-	A5B126	VITV1
78	Ole e 1_78A2	PN40024	D7U2C4	VITV1
78	Ole e 1_78A1	-	D7KQ21	ARALY
78	Ole e 1_78A2	-	D7KQ24	ARALY
78	Ole e 1_79A3	Proline-rich protein 1	Q9F235	ARATH
78	Ole e 1_79A4	Proline-rich protein 1	Q9M7P1	ARATH
78	Ole e 1_79A5	Proline-rich protein	Q9LZJ7	ARATH
78	Ole e 1_79A6	Proline-rich protein 3	Q9M7N3	ARATH
78	Ole e 1_79A7	-	D7L786	ARALY
80	Ole e 1_80A1	-	D7LGP0	ARALY

80	Ole e 1_80A2	-	Q22258	ARATH
81	Ole e 1_81A1	-	D8TC68	SEML
81	Ole e 1_81A2	-	D8TF48	SEML
81	Ole e 1_81B1	-	D8TC67	SEML
82	Ole e 1_82A1	-	D8TCF0	SEML
83	Ole e 1_83A1	Pollen ole e 1 allergen	D7L7X8	ARALY
83	Ole e 1_83A2	At5g05500	Q8FFG5	ARATH
83	Ole e 1_83A1	-	B9HHU1	POPTR
83	Ole e 1_83B2	-	B9HSK7	POPTR
83	Ole e 1_83B3	-	B9SGR0	RICCO
83	Ole e 1_83B4	PN40024	D7T4L7	VITV1
83	Ole e 1_83B5	-	A5C9V2	VITV1
84	Ole e 1_84A1	-	A2WL03	ORYSJ
84	Ole e 1_84A2	-	Q8L1N2	ORYSJ
84	Ole e 1_84A3	-	A2WL01	ORYSJ
84	Ole e 1_84A4	-	B9ETU7	ORYSJ
84	Ole e 1_84A5	B1189A09.32	Q5VR32	ORYSJ
84	Ole e 1_84A6	-	A2WL00	ORYSJ
84	Ole e 1_84A7	-	A2WL05	ORYSJ
84	Ole e 1_84A8	-	A2WL09	ORYSJ
84	Ole e 1_84A9	-	Q8LJM3	ORYSJ
84	Ole e 1_84A10	-	A2WL07	ORYSJ
84	Ole e 1_84A11	-	Q8LJM3	ORYSJ
84	Ole e 1_84A12	-	A2WL04	ORYSJ
84	Ole e 1_84A13	-	Q8LJN3	ORYSJ
84	Ole e 1_84A14	-	A2ZPK8	ORYSJ
84	Ole e 1_84A15	B1189A03.34	Q5VR30	ORYSJ
84	Ole e 1_84B1	-	A2WL02	ORYSJ
84	Ole e 1_84C1	-	A2WL12	ORYSJ
84	Ole e 1_84C2	-	Q8LJM5	ORYSJ
84	Ole e 1_84C3	-	B9ADE3	ORYSJ
84	Ole e 1_84C4	-	A2ZPL3	ORYSJ
84	Ole e 1_84C5	B1189A03.42	Q5VR18	ORYSJ
84	Ole e 1_84C6	-	A2WL11	ORYSJ
84	Ole e 1_84C7	-	Q8LJM6	ORYSJ
84	Ole e 1_84D1	B1189A09.38	Q5VR19	ORYSJ
84	Ole e 1_84E1	-	B9ET08	ORYSJ
84	Ole e 1_84F1	-	A2ZPL6	ORYSJ
84	Ole e 1_84F2	B1189A09.43	Q5VR17	ORYSJ
84	Ole e 1_84G1	-	A2WL13	ORYSJ
84	Ole e 1_84G2	Sb03g005060	C5XP43	SORBI
85	Ole e 1_85A1	-	D8SHH8	SEML

85	Ole e 1_85B1	-	D8T5S3	SEML
86	Ole e 1_86A1	-	D8SHH9	SEML
87	Ole e 1_87A1	-	D8T5S4	SEML
88	Ole e 1_88A1	-	D8R4E3	SEML
88	Ole e 1_88A2	-	D8RK87	SEML
89	Ole e 1_89A1	Pollen ole e 1 allergen	D7L7E4	ARALY
89	Ole e 1_89A2	-	Q3EBA2	ARATH
89	Ole e 1_89B1	-	B9M7K3	POPTR
89	Ole e 1_89B2	-	B9S0B7	RICCO
89	Ole e 1_89C1	-	C873U0	SOYBN
90	Ole e 1_90A1	PN40024	D7T0V3	VITV1
90	Ole e 1_90A2	PN40024	D7T0X1	VITV1
91	Ole e 1_91A1	-	A2Z9X7	ORYSJ
91	Ole e 1_91A2	Ox10g0546100	Q5AV31	ORYSJ
91	Ole e 1_91B1	Sb01g030980	C5WTH1	SORBI
92	Ole e 1_92A1	-	D7LH66	ARALY
92	Ole e 1_92A2	At2g41400	Q6DBF8	ARATH
92	Ole e 1_92A3	RAF122-03-M12	Q6T2J7	ARATH
92	Ole e 1_92B1	At2g41400	Q6ZVC5	ARATH
92	Ole e 1_92B1	At2g41390	Q6ZVC4	ARATH
93	Ole e 1_93A1	At5g05020	Q9FF72	ARATH
94	Ole e 1_94A1	-	D7KW90	ARALY
95	Ole e 1_95A1	At3g16660	Q9GY16	ARATH
95	Ole e 1_95A2	MGL6	Q6LUR8	ARATH
95	Ole e 1_95A3	-	D7L657	ARALY
95	Ole e 1_95B1	-	D7L658	ARALY
95	Ole e 1_95B2	AT3g16670-MGL6.12	Q6LUR6	ARATH
95	Ole e 1_95C1	-	B9HXT5	POPTR
95	Ole e 1_95C2	PN40024	D7SYD8	VITV1
95	Ole e 1_95C3	Phytoplanin	B9RT72	RICCO
95	Ole e 1_95D1	-	C6TK68	SOYBN
95	Ole e 1_95E1	Phytoplanin	Q5E5S9	PHYLL
95	Ole e 1_95F1	Phytoplanin	Q1PCF2	TOBAC
96	Ole e 1_96A1	-	B9APE1	ORYSJ
96	Ole e 1_96A2	-	B9F8E1	ORYSJ
96	Ole e 1_96B1	-	A2YPV4	ORYSJ
96	Ole e 1_96B2	Ox07g0874400 protein	Q6ZDW8	ORYSJ
96	Ole e 1_96B3	Sb02g042730	C5X19	SORBI
96	Ole e 1_96B4	-	B9FF51	MAIZE
96	Ole e 1_96B5	Sb01g035810	C5X0G0	SORBI
97	Ole e 1_97A1	-	A2YPV5	ORYSJ
97	Ole e 1_97A2	Ox07g0614500	Q6ZDW7	ORYSJ

97	Ole e 1_97A3	Sb02g042740	C5X5J0	SORBI
98	Ole e 1_98A1	-	A2XGJ7	ORYSJ
98	Ole e 1_98A2	-	Q10L N4	ORYSJ
98	Ole e 1_98B1	-	A2XGJ8	ORYSJ
98	Ole e 1_98B2	Ox03g0342100	Q10L N2	ORYSJ
98	Ole e 1_98C1	Sb01g035830	C5X0Q2	SORBI
99	Ole e 1_99A1	-	B9UHT3	MAIZE
100	Ole e 1_100A1	-	A2JWU2	ORYSJ
100	Ole e 1_100A2	Ox01g0725900	Q6S158	ORYSJ
100	Ole e 1_100A3	-	A2ZXF1	ORYSJ
100	Ole e 1_100A4	Sb03g033350	C5X1F5	SORBI
100	Ole e 1_100B1	Sb03g033360	C5X1F6	SORBI
100	Ole e 1_100C1	QJ1131. E09.17	Q7K5K3	ORYSJ
100	Ole e 1_100C2	-	B9FH38	ORYSJ
100	Ole e 1_100C3	-	A2Y6T6	ORYSJ
100	Ole e 1_100C4	Ox05g0531400 protein	Q6DGH6	ORYSJ
100	Ole e 1_100D1	Sb03g026510	C5YU6	SORBI
100	Ole e 1_100D2	Arabidopsis thaliana protein	B6SLV3	MAIZE
101	Ole e 1_101A1	-	A2Y6T4	ORYSJ
101	Ole e 1_101A2	Ox05g0531200 protein	Q7K5K5	ORYSJ
101	Ole e 1_101B1	Pistil-specific extensin-like protein	B6UHM8	MAIZE
101	Ole e 1_101B2	-	B4FGH8	MAIZE
101	Ole e 1_101B3	Sb08g028580	C5Z1H9	SORBI
102	Ole e 1_102A1	-	A2WUJ3	ORYSJ
102	Ole e 1_102A2	Ox01g0726100 protein	Q6S154	ORYSJ
102	Ole e 1_102B1	Sb03g033370	C5X1F7	SORBI
102	Ole e 1_102B2	Pistil-specific extensin-like prot.	B6UHE3	MAIZE
103	Ole e 1_103A1	-	D8RZL9	SEML
104	Ole e 1_104A1	-	B4FMQ8	MAIZE
105	Ole e 1_105A1	-	D8RCI2	SEML
106	Ole e 1_106A1	-	D8T9B5	SEML
107	Ole e 1_107A1	-	A9SC13	PHYPA
108	Ole e 1_108A1	-	D9Q167	SEML
109	Ole e 1_109A1	-	D8S5V6	SEML

Table 1. (continued). The Ole e 1 protein superfamily: new and unified nomenclature.

The crystallographic structural coordinates of relatively few proteins of the Ole e 1 family have been deposited in the Protein Database (PDB) up to date. To our knowledge, detailed comparative studies of the structural and conformational features of members of the Ole e 1

extended protein families have not been performed in higher plants. Using computational modelling analysis, we have determined and modelled the molecular-structural features of selected members of the Ole e 1 extended families. A first overview of the generated models (Figure 2) indicated a relatively high level of similitude.

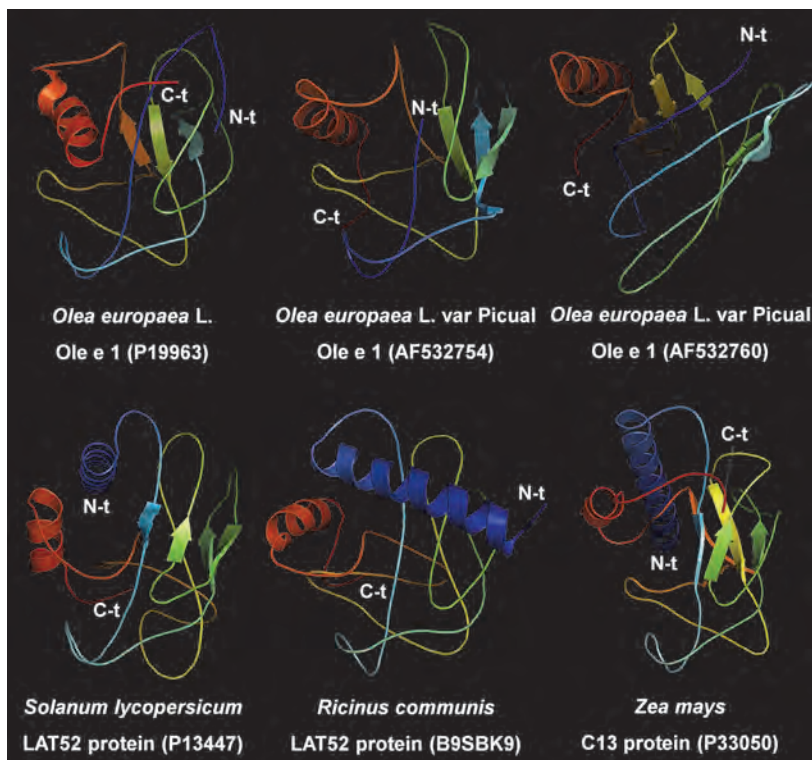


Fig. 2. Three-dimensional structure analysis of selected members of Ole e 1 family proteins. The model proteins are depicted as cartoon diagrams. The secondary elements of the crystallographic structures are rainbow coloured, with N-terminus in blue, and C-terminus in red.

However, a more detailed analysis allowed identifying certain differences in the generated models, particularly consisting in 2D structural features. These differences can be distinguished even between very close proteins like P19963, AF532754 and AF532760 (Ole e 1_57A9, Ole e 1_57A25 and Ole e 1_57A23 with the new nomenclature), corresponding to the olive pollen major allergen cloned from different varietal sources or even to different clones of the same cultivar (Figure 2). The differences become higher when models of the same protein obtained from different plant species are compared. This is the case of P13447 and B9SBK9 (Ole e 1_52L1 and Ole e 1_52J1), which correspond to the LAT52 gene product in tomato and *Ricinus*, respectively (Figure 2). Divergences are even more obvious between the models indicated above and that of a P33050 (Ole e 1_48H6), a different member of the Ole e 1 superfamily corresponding to a pollen protein from maize (C13 protein) (Figure 2).

4. Discussion

Research as regard to the proteins of the Ole e 1 family has been carried out steadily since its definition. At present, many genes from the allergen Ole e 1 family of proteins have been characterized, and data are available concerning the sequence, structure, expression and biological function (e.g. extensin-like proteins constituting part of the cell wall). However, and as depicted in this chapter, the precise identification of more than half members of this family remains uncompleted. Up to now, Ole e 1 and Ole e 1-like genes are deposited into the databases, many of them with repetitive or arbitrary naming system by authors. This nomenclature includes a variety of generic names, such as Ole e 1 major olive pollen allergen, putative Ole e 1-like protein, anther-specific Ole e 1-like protein, and others depending of the protein location in the chromosome, e.g. At3g26960, Os09g0508200, or simply giving a random name e.g. P1 clone: MOJ10. For those members of the Ole e 1 family which have been recognized like allergens, a more sustainable and precise nomenclature has been built, by following the recommendations of the International Union of Immunological Societies (IUIS) (<http://www.allergen.org/>). However, these allergenic proteins only represent a part of the members of the Ole e 1 family, and this nomenclature still does not display the relationships among these proteins. In several cases, it is still common for researchers to use different names for the same allergen. Allergen biochemistry is now entering a new time of structural biology and proteomics that will require sophisticated tools for data processing and bioinformatics, and might require further definition of the nomenclature. Increasingly, the wealth of structural information is enabling the biologic function of allergens to be established and the assignment of allergen function to diverse protein families. Therefore, the arbitrary nomenclature currently in use is not sustainable for adequate comparative mega-functional genomics studies, especially as the number of Ole e 1 genes has increased steadily and will continue with this upward trend with the completion of the sequencing projects corresponding to more plant genomes.

The implementation of modifications in the nomenclature as proposed here may assist further developments of allergy understanding and new clinical approaches. As an example, nomenclature and structural biology have been proposed to play a crucial role in defining allergens for research studies and for the development of new clinical products [Chapman et al. 2007]. Sequence comparisons and assignments to protein families provide a molecular basis for clinical cross-reactions between food, pollen, and latex allergens that give rise to oral allergy syndromes [Wagner et al. 2002, Scheiner et al. 2004, van Ree 2004]. For food and pollen allergens, intrinsic protein structure probably plays an important role in determining allergenicity by conferring, for example, heat stability or resistance to digestion in the digestive tract, e.g. storage proteins from seed/nuts or legumes [Orruño and Morgan 2011]. Interestingly, analysis of databases, e.g. pFAM shows that there are currently more than 120 molecular architectures that are responsible for eliciting IgE responses. It will be important to link nomenclature with classification of allergens into protein families and subfamilies to provide complete definition of allergens and their structure-functional relationships as part of a comprehensive bioinformatics database. The practical consequences of this approach are seen most clearly with genetically modified foods, in which sequence comparisons can be used for safety assessment of genetically modified organisms [Goodman and Tetteh 2011].

The success of our new and unified nomenclature lies in its simplicity, with genetic basis and structural-functional characterizations of the proteins, regardless of the species origin,

with the possibility to further nomenclature expansion, to include as-yet-unidentified protein allergens from different sources or species: mites, insects, pollens, molds and foods. It might be also possible to include in the system engineered protein molecules, such as hypoallergens, or others being described as non-protein allergens. Allergens entered into the nomenclature could be used to develop allergen-specific diagnostics and to formulate recombinant allergen vaccines that will benefit patients [Chapman et al. 2000, Ferreira et al. 2004, Jutel et al. 2005, Sastre 2010].

The proposed system may also assist to clarify the importance of allergen polymorphism. Allergens often display numerous variants. These are proteins with typically greater than 90% sequence identity, but with enough differences in their amino acid sequences to make worth individual structural and or functional characterization and identification. This polymorphism has been deeply analyzed in mites, as their allergens present an extensive number of isoforms: 23 for Der p 1 and 13 for Der p 2 [Smith et al. 2001, Smith et al. 2001]. Furthermore, these polymorphisms might affect T-cell responses or alter antibody-binding sites. These differences can be structurally characterized to distinguish isoforms in a well-defined nomenclature system, by mean of structural-functional differentiation, helping to design allergen formulations for immunotherapy [Jutel et al. 2005, Piboonpocanun et al. 2006]. In the case of pollen allergens, Ole e 1 from olive pollen is a clear example of extreme polymorphism, both in its peptide and in its carbohydrate moieties, as demonstrated by peptide mapping and N-glycopeptide analysis [Castro et al. 2010]. Olive cultivar origin is a major cause of polymorphism for Ole e 1 pollen allergen [Hamman-Khalifa et al. 2008, Castro et al. 2010]. The olive tree has an extremely wide germplasm, with over 1200 varieties cultivated over the world [Bartolini et al. 1994]. Therefore, the number of Ole e 1 isoforms yet to be characterized in olive pollen is expected to be enormous. A similar situation is also likely to occur in many other plant species.

Overall, our developed unified nomenclature system is helpful in a quick functional prediction of any newly cloned Ole e 1 gene(s), because from the nomenclature point of view, the newly sequenced gene(s) will always be characterized/named with sequence similarity with previously characterized Ole e 1 genes/proteins, as well as a protein structure-functional characterization and comparison. The changes that have been introduced reflect into which extended family or subfamily a certain Ole e 1 protein belongs. Accordingly, the new nomenclature will have no significant impact on already published data with old/arbitrary naming system. However, we urge scientists working on Ole e 1's to adopt this new and easy nomenclature system. In this regard, we have made an effort to preserve the user friendly linkage between the old and the new designations, which we hope will help researchers to adapt the new names. As the revised nomenclature should facilitate communication and understanding within the community interested in Ole e 1 allergen proteins, we advocate that this new naming system be used in all future studies.

The classification model used here has been developed under the basis of a previously designed gene nomenclature model for male fertility restorer (RF) proteins in higher plants [Kotchoni et al. 2010]. The increasing numbers of RF genes described in the literature represented an ongoing challenge in their clear identification and logical classification which was solved using the proposed nomenclature. Undoubtedly, similar approaches could be applied to numerous protein families involving relevant levels of nomenclature heterogeneity, many of them registered in specialized databases like pFam. In the case of allergens, other numerous protein families like profilins (Ole e 2 in the case of olive pollen)

prolamins, cupins, Bet v 1-related proteins etc., which are currently included in the AllFam database [Radauer et al. 2008] (<http://www.meduniwien.ac.at/allergens/allfam/>) could benefit of the use of similar approaches.

5. Conclusion

We propose for first time a unified naming system for Ole e 1-like genes and pseudogenes across all plant species, which accommodates the numerous sequences already deposited in several databases, offering the needed flexibility to incorporate additional Ole e 1-like proteins as they become available. Additionally, we provide an analysis of the phylogenetic relationships displayed by the members of the Ole e 1-like family and use computational protein modelling to determine structural features of selected members of this family. These data are of particular relevance for the understanding of their biological activity and allergenic cross-reactivity.

6. Acknowledgment

Support of the Spanish Ministry of Science and Innovation (ERDF-cofinanced project BFU2008-00629) and Andalusian Regional Government (ERDF-cofinanced Proyectos de Excelencia CVI5767 and AGR6274) is gratefully acknowledged.

7. References

- Alché, J.D.; Castro, A.J.; Olmedilla, A.; Fernández, M.C.; Rodríguez, R.; Villalba, M. and Rodríguez-García, M.I. (1999). The major olive pollen allergen (Ole e 1) shows both gametophytic and sporophytic expression during anther development, and its synthesis and storage takes place in the RER. *Journal of Cell Science*, Vol.112, pp.2501-2509
- Alché, J.D.; M'rani-Alaoui, M.; Castro, A.J. and Rodríguez-García, M.I. (2004). Ole e 1, the major allergen from olive (*Olea europaea* L.) pollen, increases its expression and is released to the culture medium during in vitro germination. *Plant Cell Physiology*, Vol.45, pp.1149-1157
- Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, Vol.215, No.3, pp.403-410
- Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, Vol.25, No.17, pp.3389-402
- Barral, P.; Batanero, E.; Palomares, O.; Quiralte, J.; Villalba, M. and Rodríguez, R. (2004). A major allergen from pollen defines a novel family of plant proteins and shows intra- and interspecies cross-reactivity. *Journal of Immunology*, Vol.172, pp.3644-3651
- Bartolini, G.; Prevost, G. and Messeri, C. (1994). Olive tree germplasm: descriptor lists of cultivated varieties in the world. *Acta Horticulturae*, Vol.365, pp.116-118
- Batanero, E.; Villalba, M. and Rodríguez, R. (1994). Glycosylation site of the major allergen from olive tree. Allergenic implications of the carbohydrate moiety. *Molecular Immunology*, Vol.31, pp.31-37

- Castro, A.J.; Bednarczyk, A.; Schaeffer-Reiss, C.; Rodríguez-García, M.I.; Van Dorsselaer, A.; Alché, J.D. (2010). Screening of Ole e 1 polymorphism among olive cultivars by peptide mapping and N-glycopeptide analysis. *Proteomics*, Vol. 10, No 5, pp.953-962
- Chapman, M.D.; Pomés, A.; Breiteneder, H. and Ferreira, F. (2007). Nomenclature and structural biology of allergens. *Journal of Allergy and Clinical Immunology*, Vol.119, No.2, pp.414-420
- Chapman, M.D.; Smith, A.M.; Vailes, L.D.; Arruda, K.; Dhanaraj, V. and Pomes, A. (2000). Recombinant allergens for diagnosis and therapy of allergic diseases. *The Journal of Allergy and Clinical Immunology*, Vol.106, pp.409-418
- Chevenet, F.; Brun, C.; Banuls, A.L.; Jacq, B. and Christen, R. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics*, Vol.7, pp.439
- D'Amato, G.; Spiekma, F.T.; Liccardi, G.; Jager, S.; Russo, M.; Kontou-Fili, K.; Nikkels, H.; Wuthrich, B. and Bonini, S. (1998). Pollen-related allergy in Europe. *Allergy*, Vol.53, pp.67-78
- de Castro, E.; Sigrist, C.J.A.; Gattiker, A.; Bulliard, V.; Langendijk-Genevaux, P.S.; Gasteiger, E.; Bairoch, A. and Hulo, H. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Research*, Vol.34, pp.362-365
- Ferreira, F.; Wallner, M. and Thalhamer, J. (2004). Customized antigens for desensitizing allergic patients. *Advances in Immunology*, Vol.84, pp.79-129
- Finn, R.D.; Mistry, J.; Tate, J.; Coghill, P.; Heger, A.; Pollington, J.E.; Gavin, O.L. Gunasekaran, P.; Ceric, G. Forslund, K.; Holm, L.; Sonnhammer, E.L.; Eddy, S.R. and Bateman, A. (2010). The Pfam protein families database. *Nucleic Acids Research*, Database Issue 38, pp.D211-222
- Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel R.D. and Bairoch A. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, Vol.31, pp.3784-3788
- Goodman, R.E. and Tetteh, A.O. (2011). Suggested Improvements for the Allergenicity Assessment of Genetically Modified Plants Used in Foods. *Current Allergy and Asthma Reports*, doi: 10.1007/s11882-011-0195-6
- Guex, N. and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, Vol.18, No.15, pp.2714-2723
- Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, Vol.41, pp.95-98
- Hamman-Khalifa, A.M.; Castro A.J.; Jimenez-Lopez, J.C.; Rodríguez-García, M.I. and Alché, J.D. (2008). Olive cultivar origin is a major cause of polymorphism for Ole e 1 pollen allergen. *BMC Plant Biology*, Vol.8, 10
- Hanson, D.D.; Hamilton, D.S.; Travis, J.L.; Bashe, D.M. and Mascarenhas, J.P. (1998). Characterization of a pollen-specific cDNA clone from *Zea mays* and its expression. *Plant Cell*, Vol.1, pp.173-179
- Hauser, M.; Roulias, A.; Ferreira, F. & Egger, M. (2010). Panallergens and their impact on the allergic patient. *Allergy, Asthma & Clinical Immunology*, Vol.6, pp.1-

- Jutel, M.; Jaeger, L.; Suck, R.; Meyer, H.; Fiebig, H. and Cromwell, O. (2005). Allergenspecific immunotherapy with recombinant grass pollen allergens. *The Journal of Allergy and Clinical Immunology*, Vol.116, pp.608-613
- Laskowski, R.A.; MacArthur, M.W.; Moss, D.S. and Thornton, J.M. (1993). PROCHECK: A program to check the stereo-chemical quality of protein structures. *Journal of Applied Crystallography*, Vol.26, pp.283-291
- Lauzurica, P.; Gurbindo, C.; Maruri, N.; Galocha, B.; Diaz, R.; Gonzalez, J.; García, R. and Lahoz, C. (1988). Olive (*Olea europea*) pollen allergens—I. Immunochemical characterization by immunoblotting, CRIE and immunodetection by a monoclonal antibody. *Molecular Immunology*, Vol.25, pp.329-335
- King, T.P.; Hoffman, D.; Lowenstein, H.; Marsh, D.G.; Platts-Mills, T.A. and Thomas, W. (1994). Allergen nomenclature. WHO/IUIS Allergen Nomenclature Subcommittee. *International Archives of Allergy and Immunology*, Vol. 105, pp. 224-233
- Kotchoni, S.O.; Jimenez-Lopez, J.C.; Gachomo, W.E. and Seufferheld, M.J. (2010). A new and unified nomenclature for male fertility restorer (RF) proteins in higher plants. *PLoS ONE*, Vol.5, No.12, pp.e15906
- Melo, F. and Feytmans, E. (1997). Novel knowledge-based mean force potential at atomic level. *Journal of Molecular Biology*, Vol.267, No.1, pp.207-222
- Melo, F. and Feytmans, E. (1998). Assessing protein structures with a non-local atomic interaction energy. *Journal of Molecular Biology*, Vol.277, No.5, pp.1141-1152
- Mothes, N.; Horak, F. & Valenta, R. (2004). Transition from a botanical to a molecular classification in tree pollen allergy: implications for diagnosis and therapy. *International Archives of Allergy and Immunology*, Vol.135, pp.357-373
- Orruño, E. and Morgan, M.R.A. (2011). Resistance of purified seed storage proteins from sesame (*Sesamum indicum* L.) to proteolytic digestive enzymes. *Food Chemistry*, in press
- Piboonpocanun S, Malinual N, Jirapongsananuruk J, Vichyanond P, Thomas WR. (2006). Genetic polymorphisms of major house dust mite allergens. *Clinical & Experimental Allergy*, Vol.36, pp.510-516
- Radauer, C.; Bublin, M.; Wagner, S.; Mari, A. and Breiteneder, H. (2008). Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *Journal of Allergy and Clinical Immunology*, Vol.121, pp.847-852
- Rodríguez, R.; Villalba, M.; Batanero, E.; González, E.M.; Monsalve, R.I.; Huecas, S.; Tejera, M.L. and Ledesma, A. (2002). Allergenic diversity of the olive pollen. *Allergy*, Vol.57, pp.6-16
- Rodríguez, R.; Villalba, M.; Monsalve, R.I.; Batanero, E.; González, E.M.; Monsalve, R.I.; Huecas, S.; Tejera, M.L. and Ledesma, A. (2002). Allergenic diversity of the olive pollen. *Allergy*, Vol.57, pp.6-15
- Salamanca, G.; Rodríguez, R. Quiralte, J.; Moreno, C.; Pascual, C.Y.; Barber, D. and Villalba, M. (2010). Pectin methylesterases of pollen tissue, a major allergen in olive tree. *FEBS Journal*, Vol.277, No.13, pp.2729-2739
- Sastre, J. (2010). Molecular diagnosis in allergy. *Clinical & Experimental Allergy*, Vol.40, No.10, pp.1442-1460
- Scheiner, O.; Aberer, W.; Ebner, C.; Ferreira, F.; Hoffmann-Sommergruber, K.; Hsieh, L.S.; Kraft, D.; Sowka, S.; Vanek-Krebitz, M. and Breiteneder, H. (1997). Cross-reacting

- allergens in tree pollen and pollen-related food allergy: implications for diagnosis of specific IgE. *International Archives of Allergy and Immunology*, Vol.113, pp.105-108
- Shultz, J.L.; Kurunam, D.; Shopinski, K.; Iqbal, M.J.; Kazi, S.; Zobrist, K.; Bashir, R.; Yaegashi, S.; Lavu, N.; Afzal, A.J.; Yesudas, C.R.; Kassem, M.A.; Wu, C.; Zhang, H.B.; Town, C.D.; Meksem, K. and Lightfoot, D.A. (2006). The Soybean Genome Database (SoyGD): a browser for display of duplicated, polyploid, regions and sequence tagged sites on the integrated physical and genetic maps of *Glycine max*. *Nucleic Acids Research*, Vol.34(suppl 1), pp.D758-D765
- Sigrist, C.J.A.; Cerutti, L.; de Castro, E.; Langendijk-Genevaux, P.S.; Bulliard, V.; Bairoch, A. and Hulo, N. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Research*, Vol.38 (Database issue), pp.161-166
- Smith, A.S.; Benjamin, D.C.; Hozic, N.; Derewenda, U.; Smith, W.A.; Thomas, W.R.; Gafvelin, G.; van Hage-Hamsten, M. and Chapman, M.D. (2001). The molecular basis of antigenic cross-reactivity between the group 2 mite allergens. *The Journal of Allergy and Clinical Immunology*, Vol.107, pp.977-984
- Smith, W.A.; Hales, B.J.; Jarnicki, A.G. and Thomas W.R. (2001). Allergens of wild house dust mites: environmental Der p 1 and Der p 2 sequence polymorphisms. *The Journal of Allergy and Clinical Immunology*, Vol.107, pp.985-992
- Stratford, S.; Barne, W.; Hohorst, D.L.; Sagert, J.G.; Cotter, R.; Golubiewski, A.; Showalter, A.M.; McCormick, S. and Bedinger, P. (2001). A leucine-rich repeat region is conserved in pollen extensin-like (Pex) proteins in monocots and dicots. *Plant Molecular Biology*, Vol.46, pp.43-56
- Tang, B.; Banerjee, B.; Greenberger, P.A.; Fink, J.N.; Kelly, K.J. and Kurup, V.P. (2000). Antibody binding of deletion mutants of Asp f 2, the major *Aspergillus fumigatus* allergen. *Biochemical and Biophysical Research Communications*, Vol.270, pp.1128-1135
- Thompson, J.D.; Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, Vol.22, pp.4673-4680
- Twell, D.; Wing, R.; Yamaguchi, J. and McCormick, S. (1989). Isolation and expression of an anther-specific gene from tomato. *Molecular and General Genetics*, Vol.217, pp.240-245
- van Gunsteren, W.F.; Billeter, S.R.; Eising, A.A.; Hünenberger, P.H.; Krüger, P.; Mark, A.E.; Scott, W.R.P. and Tironi, I.G. (1996). Biomolecular Simulations: The GROMOS96 Manual and User Guide. Zürich, VdF Hochschulverlag ETHZ
- van Ree R. (2004). Clinical importance of cross-reactivity in food allergy. *Current Opinion in Allergy & Clinical Immunology*. Vol.4, pp.235-240
- Villalba, M.; Batanero, E.; Lopez-Otin, C.; Sanchez, L.M.; Monsalve, R.I.; Gonzalez de la Pena, M.A.; Lahoz, C. and Rodriguez, R. (1993). The amino acid sequence of Ole e I, the major allergen from olive tree (*Olea europaea*) pollen. *European Journal of Biochemistry*, Vol.216, pp.863-869
- Villalba, M.; López-Otín, C.; Martín-Orozco, E.; Monsalve, R.I.; Palomino, P.; Lahoz, C. and Rodríguez, R. (1990). Isolation of three allergenic fractions of the major allergen from *Olea europaea* pollen and N-terminal amino acid sequence. *Biochemical and Biophysical Research Communications*, Vol.172, pp.523-528
- Wagner, S. and Breiteneder, H. (2002). The latex-fruit syndrome. *Biochemical Society Transactions*, Vol.6, pp.935-940