

Characterization of Profilin Polymorphism in Pollen with a Focus on Multifunctionality

Jose C. Jimenez-Lopez¹, Sonia Morales¹, Antonio J. Castro¹, Dieter Volkmann², María I. Rodríguez-García¹, Juan de D. Alché^{1*}

¹ Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, High Council for Scientific Research, Granada, Spain, ² Institute of Cellular and Molecular Botany, Department of Plant Cell Biology, University of Bonn, Bonn, Germany

Abstract

Profilin, a multigene family involved in actin dynamics, is a multiple partners-interacting protein, as regard of the presence of at least of three binding domains encompassing actin, phosphoinositide lipids, and poly-L-proline interacting patches. In addition, pollen profilins are important allergens in several species like *Olea europaea* L. (Ole e 2), *Betula pendula* (Bet v 2), *Phleum pratense* (Phl p 12), *Zea mays* (Zea m 12) and *Corylus avellana* (Cor a 2). In spite of the biological and clinical importance of these molecules, variability in pollen profilin sequences has been poorly pointed out until now. In this work, a relatively high number of pollen profilin sequences have been cloned, with the aim of carrying out an extensive characterization of their polymorphism among 24 olive cultivars and the above mentioned plant species. Our results indicate a high level of variability in the sequences analyzed. Quantitative intra-specific/variety polymorphism was higher in comparison to inter-specific/cultivars comparisons. Multi-optional posttranslational modifications, e.g. phosphorylation sites, physicochemical properties, and partners-interacting functional residues have been shown to be affected by profilin polymorphism. As a result of this variability, profilins yielded a clear taxonomic separation between the five plant species. Profilin family multifunctionality might be inferred by natural variation through profilin isoforms generated among olive germplasm, as a result of polymorphism. The high variability might result in both differential profilin properties and differences in the regulation of the interaction with natural partners, affecting the mechanisms underlying the transmission of signals throughout signaling pathways in response to different stress environments. Moreover, elucidating the effect of profilin polymorphism in adaptive responses like actin dynamics, and cellular behavior, represents an exciting research goal for the future.

Citation: Jimenez-Lopez JC, Morales S, Castro AJ, Volkmann D, Rodríguez-García MI, et al. (2012) Characterization of Profilin Polymorphism in Pollen with a Focus on Multifunctionality. PLOS ONE 7(2): e30878. doi:10.1371/journal.pone.0030878

Editor: Vladimir N. Uversky, University of South Florida College of Medicine, United States of America

Received: December 15, 2011; **Accepted:** December 28, 2011; **Published:** February 14, 2012

Copyright: © 2012 Jimenez-Lopez et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by the following European Regional Development Fund co-financed grants: MCINN BFU 2004-00601/BFI, BFU 2008-00629, BFU2011-22779, CICE (Junta de Andalucía) P2010-CVI15767, P2010-AGR6274 and P2011-CVI-7487, and by the coordinated project Spain/Germany MEC HA2004-0094. The funders had no role in the study, design or decision to publish.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: juandedios.alche@eez.csic.es

Introduction

Plant cytoskeleton plays a major role during cell events including division, expansion, morphogenesis and differentiation, and also in response to external stimuli like pathogen attack [1]. Actin, the central player in the cytoskeleton, has a paramount importance in the cytoskeleton structure. Its proper assembly and organization depends upon the expression of an appropriate and complex mixture of actin-binding proteins (ABPs) as well as signals mediated by various signaling pathways involving molecules like the Rho family of GTPases [2]. Plant cells respond to a wide range of internal or external stimuli by reorganizing their cytoplasm [3]. These modifications often correlate with changes in the actin filament network, being the ABPs at the crossroad between extracellular signals and rearrangements of the cytoskeleton. One of the best characterized examples of ABPs in plants is profilin [4], a large multigene family (Pfam accession number PF00235) [5], differentially expressed, with biochemical and functionally diverse isoforms [6]. They have been also found in lower eukaryotes [7], invertebrates [8], and vertebrates [9]. Viral profilins have been found, whose gene organization is homologous to mammalian profilins [10].

Plant profilins contain up to ten different genes in both mono- and dicotyledonous [11], divided in two classes differentially expressed: one is ubiquitously present, and constitutively expressed in all plant tissues, whereas the second class is restricted to the reproductive tissues [12]. The complexity of profilin expression and the number of isoforms in higher plants is correlated with the observation that the actin family is also more complex in plants than in other kingdoms [13].

Profilins display a molecular mass around 15 kDa. These proteins control actin polymerization in eukaryotic cells [14], promoting or inhibiting actin polymerization, depending on the profilin/G-actin ratio, ionic environment of the cell [15], and also depending of the interaction with other actin-binding proteins [16]. Profilins have been shown to be in a 1:1 stoichiometry rate in relation with total actin in pollen [17] and in tobacco suspension cells [18]. They are generally considered to be the main buffer of the actin monomer pool, suppressing filament spontaneous nucleation and providing a large population of subunits for formin-mediated polymerization [19].

Profilins have been revealed as key mediators of the membrane-cytoskeleton communication, acting at critical points of signaling

pathways initiated by events in the plasma membrane and transmitted by transduction cascades to promote cytoskeletal rearrangements [20]. This functionality arises from their binding capacity of interaction with phosphatidylinositides (PIP2), as well as with poly-L-proline-rich proteins [21].

Up to date, over 400 profilin sequences from 100 plant species are available at NCBI GenBank database [22]. Among these, about half have been isolated from pollen, mostly from allergenic plant species including the allergens *Ole e 2*, *Bet v 2*, *Cor a 2*, *Phl p 12* and *Zea m 12*. However, the presence of polymorphism has only been poorly pointed out in a low number of sequences in *Z. mays* [23], *T. aestivum* [24], *A. thaliana* [6], *N. tabacum* [25], *O. europaea* [26], *C. dactylon* [27], *P. pratense* [27], *Artemisia vulgaris* [28], and *P. Judaica* [29]. No data about interspecific comparisons, and cultivars sequence variability is available so far.

In the present study, we have cloned a representative number of profilin sequences from olive pollen and from other 4 worldwide distributed allergenic species (*Betula verrucosa*, *Corylus avellana*, *Phleum pratense* and *Zea mays*), to extensively characterize the polymorphism affecting profilin sequences. A comparative analysis of the intra- and inter-species/cultivars variability has been carried out, as well as an extensive bioinformatic analysis of how the polymorphism may affect these sequence motives considered of relevance for profilin functionality. These include sequences differentially affecting physicochemical properties, posttranslational modifications sites, i.e. phosphorylation, and also the interaction properties of ligand-binding partners.

Results

Analysis of profilin sequence polymorphism

RT-PCR amplification of total RNA with a set of degenerate primers resulted in 94 raw sequences from 24 olive cultivars, 10 from hazel, 8 from timothy-grass, 7 from maize and 2 identical sequences from birch. Each one of these sequences was individually analyzed by the nucleotide-nucleotide BLAST (blastn) program and ScanProsite software searching for specific profilin motif patterns. Non-redundant nucleotide sequences were deposited in the GenBank™/EMBL Database (Table 1).

The full-length cDNA of the profilin sequences ranged between 393 bp (*Olea europaea*, DQ663558 or *Zea mays*, DQ663565) to 405 bp (*Olea europaea* Y12429, Y124230, Y12425). Peculiarities of a multiple sequence alignment of the profilins are shown in the Figure 1A, which also includes several profilins previously characterized and deposited in the GenBank databases. Variability along nucleotide sequences were calculated and depicted in Figure 1B. A remarkable feature was the presence of deletions in frame with protein translation in all the species analyzed, ranging from 3 nucleotides in *Betula verrucosa* (sequences DQ650633 and M65179) and *Corylus avellana* (DQ663543–4, DQ663546 and DQ663548–52), to 12 nucleotides in olive (Picual DQ663558) or maize (DQ663565). Deletions of 9 nucleotides were the most common feature in the remaining sequences with deletions. Protein sequences included polypeptides of 130, 131, 133 and 134 amino acids length (Figure 2A). The region along the profilin sequences with the lower conservation levels corresponded to the first 5' half of the sequences, highly coinciding with the area of the main deletions (Figure 2B).

Frequent amino acid substitutions among the species were detected along the sequences. A total of 39 variable positions (variability index value, $viv > 3$) were found in olive profilin (Table 2). The table 2 shows the variable positions and variability index in the olive profilin sequences, and in the rest of species. For all species (with the exception of birch), the most highly-variable

residues were located in the N-terminal α -helix. The rest of variable positions were found all through the protein sequence. Different number of polymorphic positions was detected in *Bet v 2*, *Cor a 2*, *Phl p 12* and *Zea m 12* (5, 8, 3 and 4 positions respectively) (Table 2). Two key amino acids were affected by a high level of polymorphism in several species (positions 130 and 131 respectively, involved in the phosphoinositides lipids interaction), and the position 79, exclusively in olive tree (a position involved in actin interaction). Three key cysteines showed differential conservation as well, with the one in the position 106 displaying a high variability ($viv > 6$).

Concerning polymorphism at the family level, comparisons of the sequence identities between families showed that the variability between the *Oleaceae* and *Poaceae* families was 29.8 and 25.8% for nucleotide and amino acid sequences, respectively (they correspond to the most distant families), whereas *Oleaceae* and *Betulaceae* share higher identities, with variability percentages of 26.1 and 21.7% only (they are the most close-related families). The *Poaceae* and *Betulaceae* families showed 28.2 and 22.5% of variability, respectively (Table 3).

As regard to the polymorphism within the species analyzed, *Olea europaea* displayed the highest percentage of variability (28.2 and 24.6% for nucleotide and amino acid sequences respectively), while the lowest percentage of variability was found in *Phleum pratense* (20.1 and 20.7% for nucleotide and amino acid sequences, respectively). When polymorphism was analyzed between species, *Betula pendula* and *Phleum pratense* were the species with the lowest percentage of variability in nucleotide and amino acid sequences (1.3 and 3.0%, respectively). The most distant species in terms of sequence variability were *Olea europaea* and *Corylus avellana* with 21.6 and 19.8% of variability respectively (Table 3).

Additional pairwise alignments of more than 100 profilin sequences (data not shown) corresponding to a wide and representative number of plant species, showed a range of identity of 99.2–68.7% among plant reproductive profilin proteins, and 66.0–90.0% of identity among plant vegetative profilins. Furthermore, comparison of the plant reproductive profilins with profilins other than plant profilin proteins showed identity intervals: *P. polycephalum* (P22271) 45.8–36.0%, *D. discoideum* (P26199) 44.2–31.2%, *D. melanogaster* (P25843) 42.8–38.6%, *A. castellanii* (P68696), 42.7–34.3%, *S. cerevisiae* (P07274) 31.2–25.1%, *M. musculus* (P62962) 25.8–17.0%, *H. sapiens* (P07737) 25–17.7%; and virus (Q76ZN5) 12–6%.

A statistical analysis was performed to establish whether the differences in the variability at the level of both, nucleotide and amino acid sequence had statistical significance. The comparison of the variance for the distributions of variability showed significant differences for five species (F tests, $p < < 0.05$). To analyze the variability differences between species, a *post hoc* range analysis between each pair of species was performed, assuming not normal distributions (Shapiro-Wilk tests, $p < < 0.05$), and inequality of variances (Levene tests, $p < < 0.05$). The table 4 shows the range analysis with the differences between pair of species determined by Games-Howell tests. At the level of nucleotide sequences, clear polymorphism differences were found between almost all pairs of species. Furthermore, amino acid sequences polymorphism revealed differences between many pairs of species.

Physicochemical properties and posttranslational modifications affected by the polymorphism

Most profilin sequences showed a calculated molecular weight and a calculated isoelectric point within the range of the profilins described in the literature and the NCBI database sequences (Table S1), with the exception of two protozoa basic profilins [30].

Table 1. GenBank™/EMBL Database entries.

<i>Olea europaea</i> Cultivar/Clone	Accession n°	<i>Olea europaea</i> Cultivar/Clone	Accession n°	Species/Clone	Accession n°
Acebuche 1	DQ138355	Lucio 2	DQ138363	<i>Betula pendula</i> 1	DQ650633
Acebuche 2	DQ138356	Lucio 3	DQ138365	<i>Corylus avellana</i> 1	DQ663543
Acebuche 3	DQ138357	Lucio 4	DQ138364	<i>Corylus avellana</i> 2	DQ663544
Arbequina 1	DQ138327	Lucio 5	DQ640908	<i>Corylus avellana</i> 3	DQ663545
Arbequina 2	DQ138328	Manzanilla Sevilla 1	DQ117911	<i>Corylus avellana</i> 4	DQ663546
Arbequina 3	DQ138329	Manzanilla Sevilla 2	DQ138324	<i>Corylus avellana</i> 5	DQ663547
Arbequina 4	DQ138330	Manzanilla Sevilla 3	DQ138325	<i>Corylus avellana</i> 6	DQ663548
Bella de España 1	DQ317563	Manzanilla Sevilla 4	DQ138326	<i>Corylus avellana</i> 7	DQ663549
Bella de España 2	DQ317564	Morrut 1	DQ317573	<i>Corylus avellana</i> 8	DQ663550
Bella de España 3	DQ640909	Morrut 2	DQ317574	<i>Corylus avellana</i> 9	DQ663551
Bella de España 4	DQ640910	Morrut 3	DQ317575	<i>Corylus avellana</i> 10	DQ663552
Blanqueta 1	DQ138335	Morrut 4	DQ317576	<i>Phleum pratense</i> 1	DQ663535
Blanqueta 2	DQ138336	Picual 1	DQ317580	<i>Phleum pratense</i> 2	DQ663536
Blanqueta 3	DQ138337	Picual 2	DQ317581	<i>Phleum pratense</i> 3	DQ663537
Blanqueta 4	DQ138338	Picual 3	DQ317582	<i>Phleum pratense</i> 4	DQ663538
Cornicabra 1	DQ138331	Picual 4	DQ640904	<i>Phleum pratense</i> 5	DQ663539
Cornicabra 2	DQ138332	Picual 5	DQ663553	<i>Phleum pratense</i> 6	DQ663540
Cornicabra 3	DQ138333	Picual 6	DQ663554	<i>Phleum pratense</i> 7	DQ663541
Cornicabra 4	DQ138334	Picual 7	DQ663555	<i>Phleum pratense</i> 8	DQ663542
Empeltre 1	DQ138342	Picual 8	DQ663556	<i>Zea mays</i> 1	DQ663559
Empeltre 2	DQ138343	Picual 9	DQ663557	<i>Zea mays</i> 2	DQ663560
Empeltre 3	DQ138344	Picual 10	DQ663558	<i>Zea mays</i> 3	DQ663561
Farga 1	DQ317565	Picudo 1	DQ117907	<i>Zea mays</i> 4	DQ663562
Farga 2	DQ317566	Picudo 2	DQ117908	<i>Zea mays</i> 5	DQ663563
Farga 3	DQ317567	Picudo 3	DQ117909	<i>Zea mays</i> 6	DQ663564
Frantoio 1	DQ317568	Picudo 4	DQ117910	<i>Zea mays</i> 7	DQ663565
Frantoio 2	DQ317569	Sevillencia 1	DQ138348		
Galega 1	DQ317570	Sevillencia 2	DQ138349		
Hojiblanca 1	DQ061979	Sevillencia 3	DQ138350		
Hojiblanca 2	DQ061980	Sourani 1	DQ317577		
Hojiblanca 3	DQ061981	Sourani 2	DQ317578		
Hojiblanca 4	DQ061982	Sourani 3	DQ317579		
Leccino 1	DQ138345	Sourani 4	DQ640905		
Leccino 2	DQ138346	Verdial Huévar 1	DQ117902		
Leccino 3	DQ138347	Verdial Huévar 2	DQ117903		
Lechín de Granada 1	DQ317571	Verdial Huévar 3	DQ117904		
Lechín de Granada 2	DQ317572	Verdial Huévar 4	DQ117905		
Lechín de Granada 3	DQ640906	Verdial Huévar 5	DQ117906		
Lechín de Sevilla 1	DQ028766	Verdial Vélez-Málaga 1	DQ138358		
Lechín de Sevilla 2	DQ061976	Verdial Vélez-Málaga 2	DQ138359		
Lechín de Sevilla 3	DQ061977	Verdial Vélez-Málaga 3	DQ138360		
Lechín de Sevilla 4	DQ061978	Verdial Vélez-Málaga 4	DQ138361		
Loaime 1	DQ138339	Villalonga 1	DQ138351		
Loaime 2	DQ138340	Villalonga 2	DQ138352		
Loaime 3	DQ138341	Villalonga 3	DQ138353		
Loaime 4	DQ640903	Villalonga 4	DQ138354		
Lucio 1	DQ138362	Villalonga 5	DQ640907		

Accession numbers of the profilin cDNA sequences obtained after RT-PCR from pollen of five plant species: *Olea europaea*, *Betula pendula*, *Corylus avellana*, *Phleum pratense* and *Zea mays*.

doi:10.1371/journal.pone.0030878.t001

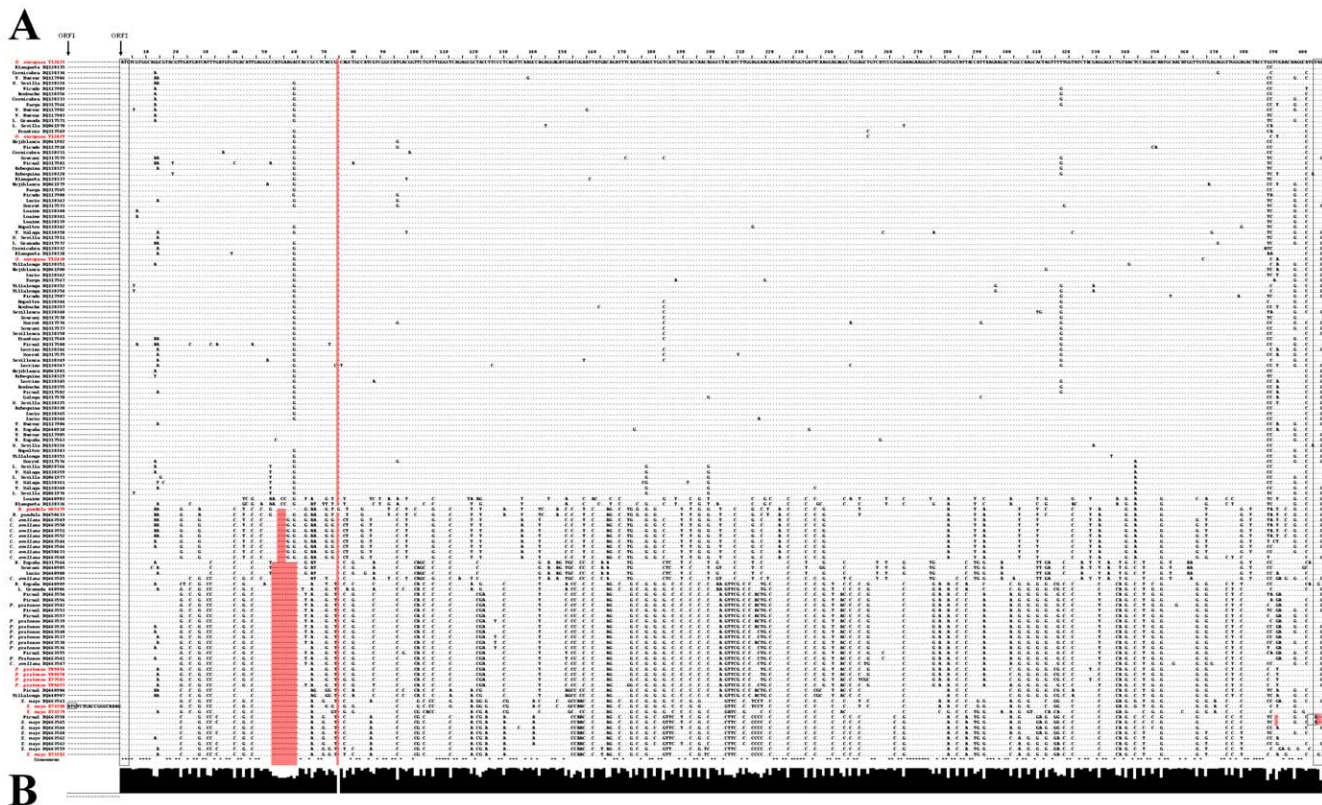


Figure 1. Multiple alignment of the nucleotide sequences of profilins. A) Two open reading frames (ORF) are indicated by vertical arrows. Numbering of the sequence positions begins at the ATG start codon from ORF2. Reference sequences of profilin from different species previously deposited in the GenBankTM/EMBL database were highlighted in red colour. Start and stop codons are indicated by colourless boxes. Significant deletions are depicted in red boxes. B) Sequence conservation index calculated with ClustalW software, 100% conservation corresponding to the maximal height of the bar at each position, and pointed out by an asterisk in the consensus.
doi:10.1371/journal.pone.0030878.g001

The average molecular weight was $14,334.40 \pm 144.01$ Da. The average of the calculated isoelectric points (5.14 ± 0.21) shows the acidic character of the profilin proteins (Table S1).

Most of the sequences exhibited hydrophilic character, as indicated by the negative average value (-0.15 ± 0.06) of the calculated Grand index (GRAVY) [31]. Profilin protein sequences could be considered as stable proteins, at the light of the average aliphatic index of 80.03 ± 3.62 [31], and the average stability index (27.95 ± 2.07) [32], where values lower than 40 are considered as a stable protein (Table S1).

Post-translational motifs implicated in protein function regulation were analyzed by matching all the sequences with the PROSITE database [33]. Table S2 shows the variability in the potential N-myristoylation motifs in the olive, with a variable number of post-translational sites ranging between 1 to 4, including different motifs such as [$^{17}G(Q/L)hl(T/A)(A/S)^{22}$], [$^{30}GQdgSV^{35}$], [$^{33}GSvwAQ^{38}$], [$^{64}GMfvAG^{69}$], [$^{67}GLHlGG^{72}$], [$^{93}GGitSK^{98}$], where the numbers represent the position of the amino acids in the profilin sequence. Profilin sequences from birch showed 2 or 3 sites with different motifs such as [$^{17/19}GQqlAA^{24}$], [$^{32}GSvwAQ^{37}$] and [$^{66}GLHlGG^{71}$], whereas 2 sites were detected in timothy-grass sequences with two possible motifs, [$^{30}GTvwAQ^{35}$], and [$^{64}GMfvAT^{69}$]. In maize sequences, 1–2 myristoylation sites were detected, with the two possible motifs [$^{30}GA(t/a/v)wAQ^{35}$] and [$^{64}GLHlGG^{69}$].

Larger homogeneity in the number of amidation motives was observed, with a unique predictive sequence [$^{84/87}rGKK^{90}$], present in all species. No glycosylation motifs were found in any of

the profilin sequences, with the exception of the olive sequence DQ138337 (cv. Blanqueta) [$^{51}NGTM^{54}$].

A variable number of multi-optional phosphorylation sites were found, involving serine, threonine and tyrosine residues [34]. Serine residues susceptible of phosphorylation ranged between 0 to 2, threonine between 1 to 4 and tyrosine between 1 to 3 (Table S3). Finally, the analysis of variability in the phosphorylation motifs for Mitogen-activated protein (MAP) kinase was performed for all profilin sequences [35]. These motifs have been described in a large number of MAP kinase-interacting proteins, including profilin [36]. The table S2 also shows several changes for many profilin sequences, i.e. ($^{79}Q \rightarrow A$), ($^{82}A \rightarrow P/S$), ($^{84}A \rightarrow V/R$), ($^{92}S \rightarrow A/T$), ($^{97}I \rightarrow V/S$), which may affect the kinase interaction motif (KIM domain) [$^{78}IQGEAGAVIRGKKGSGGITK^{98}$]. The major variability was found in the olive cultivar Bella de España, where all sequences were different. This kinase motif was also affected by the polymorphism at the level of species.

Clustering analysis

Phylogenetic analysis was performed in order to determine the relationships between profilin sequences in different species, and to infer the evolutive trends among the wide representation of the olive germplasm (Figure 3). When all the sequences were analyzed together, a clear separation between the five species was detected. Several exceptions for olive profilin, and two sequences for hazel were found, with sequences of these species located in a tree branch belonging to timothy-grass and maize (Figure 3).

Table 2. Summary of the polymorphic positions (viv>3) in the profilin sequences among the species analyzed in this work.

Sequence position	4	5	17	18	19	20	21	22	23	24	25	29	32	34	35	40	41	48	50	51	52
<i>Olea europaea</i>	3.06	5.47	3.10	4.62	3.55	4.16	3.13	3.59	-	5.71	3.38	3.55	3.13	3.35	-	3.23	3.42	3.06	3.55	3.46	3.20
<i>Betula pendula</i>	-	-	-	-	-	4.00	4.00	4.00	4.00	4.00	-	-	-	-	-	-	-	-	-	-	-
<i>Corylus avellana</i>	-	3.33	-	-	-	-	-	-	-	-	-	3.75	-	-	-	-	3.75	-	-	-	-
<i>Phleum pratense</i>	-	4.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Zea mays</i>	-	-	-	-	-	-	-	-	-	4.29	-	-	-	-	4.29	-	-	-	-	-	-

Sequence position	53	54	58	*62	70	72	73	*79**	*82	86	92	97	104	105	106	110	113	123	130***	131***	134
<i>Olea europaea</i>	3.06	4.26	3.50	3.13	3.23	5.51	-	3.13	3.91	3.10	3.42	4.26	3.10	3.10	6.64	3.60	3.46	3.06	5.45	3.20	4.36
<i>Betula pendula</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Corylus avellana</i>	-	-	-	-	-	-	3.75	-	3.75	-	-	-	-	-	-	-	3.75	-	10.00	6.00	-
<i>Phleum pratense</i>	-	-	-	-	-	7.14	-	-	-	-	-	-	-	-	-	-	-	-	-	3.75	-
<i>Zea mays</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.29	5.00

Polymorphic positions 62, 79 and 82 (asterisk), correspond to amino acids in the plant-specific binding pocket region. Position 79 (double-asterisk) is also one of the key amino acid for the profilin-actin interaction. Polymorphic positions 130 and 131 (triple-asterisk) correspond to important amino acids for the profilin-PIP interaction. doi:10.1371/journal.pone.0030878.t002

Table 3. Ranges of identity percentages calculated for both the nucleotide and the amino acid sequences of profilins in the species analysed.

	<i>Oleaceae Olea europaea</i>	<i>Corylus avellana</i>	<i>Betula pendula</i>	<i>Betulaceae (B. pendula+C. avellana)</i>	<i>Phleum pratense</i>	<i>Zea mays</i>	<i>Poaceae (P. pratense+Z. mays)</i>	
<i>Oleaceae Olea europaea</i>	71.8–100	72.3–98.4	73.6–81.9	72.3–98.4	74.4–97.7	70.4–92.9	70.4–100	Nucleotide
	75.4–100	76.3–96.1	74.4–86.5	74.4–96.1	76.1–100	74.2–100	74.2–100	Amino acid
<i>Corylus avellana</i>		75.7–100	74.3–95.5	74.3–100	75.0–98.4	70.2–86.1	70.2–98.4	Nucleotide
		77.4–100	75.1–96.9	75.1–100	77.8–96.9	76.6–84.7	76.6–96.9	Amino acid
<i>Betula pendula</i>			99.5–100	74.3–100	74.8–76.1	71.9–75.3	71.9–76.1	Nucleotide
			96.2–100	75.1–100	76.6–79.6	71.9–78.1	71.9–79.6	Amino acid
<i>Betulaceae (B. pendula+C. avellana)</i>				74.3–100	74.1–98.4	70.2–86.1	70.2–98.4	Nucleotide
				75.1–100	76.6–96.9	74.4–84.7	74.4–96.9	Amino acid
<i>P. pratense</i>					94.6–100	81.8–87.1	81.8–100	Nucleotide
					94.6–100	79.3–86.2	79.3–100	Amino acid
<i>Z. mays</i>						79.9–100	79.3–100	Nucleotide
						84.6–100	80.1–100	Amino acid
<i>Poaceae (P. pratense +Z. mays)</i>							79.9–100	Nucleotide
							79.3–100	Amino acid

It also includes the data corresponding to the accumulated percentage for the comparisons between families. When comparing species, nucleotide sequence identity varied from 70.2% (*C. avellana*×*Z. mays*) to 100%, whereas identities in the amino acid sequences ranged from 71.9% (*B. pendula*×*Z. mays*) to 100%. The higher range of intra- specific variability was that of the *Olea europaea* L. nucleotide sequences (28.2%), whereas the profilin nucleotide sequences of *Betula pendula* showed the shorter range of variability (0.5%). Among the families, the percentage of nucleotide identity varied from 70.2% (*Betulaceae*×*Poaceae*) to 100% in several cases, while the percentages applied to the amino acid sequences ranged from 74.2% (*Oleaceae*×*Poaceae*) also to 100%. Intra- family variability in the nucleotide sequences ranged from 28.2% (*Oleaceae*) to 20.1% (*Poaceae*). doi:10.1371/journal.pone.0030878.t003

Table 4. Statistical differences in polymorphism calculated among species for nucleotide and amino acid sequences.

Specie vs Specie	<i>Olea europaea</i>	<i>Betula pendula</i>	<i>Corylus avellana</i>	<i>Phleum pratense</i>	<i>Zea mays</i>	
<i>Olea europaea</i>		0.02 p = 0.71	0.16** p = 0.00	0.12** p = 0.00	0.03 p = 0.39	NUCLEOTIDE SEQUENCES
<i>Betula pendula</i>	0.83** p = 0.00		0.14** p = 0.00	0.11** p = 0.00	0.01 p = 0.97	
<i>Corylus avellana</i>	0.15** p = 0.00	0.67** p = 0.00		0.32** p = 0.00	0.13** p = 0.00	
<i>Phleum pratense</i>	1.42** p = 0.00	0.58** p = 0.00	0.09 p = 0.92		0.09** p = 0.00	
<i>Zea mays</i>	1.12** p = 0.00	0.28 p = 0.42	0.38** p = 0.01	0.30 p = 0.14		
	AMINO ACID SEQUENCES					

Comparisons in each box include the Game-Howell test and the level of significance ($p < 0.05$). Statistical significant differences were stood out by double asterisk. doi:10.1371/journal.pone.0030878.t004

Furthermore, the analysis of the individual species showed no clear relationships for olive cultivars, where the sequences were mixed and distributed around branches of the tree.

Discussion

Multiple profilin isoforms as a result of the sequence polymorphism reveal their multifunctionality

The presence of protein variants is one of the most widespread properties in higher plants, mainly as the result of the occurrence of multigene families [6]. Moreover, different mechanisms rising protein variability have been described such as single nucleotide polymorphisms (SNP) and post-translational modifications. The existence of profilin isoforms in pollen has been described in *Zea mays* [23], *Triticum aestivum* [24], *Nicotiana tabacum* [25], soybean [37] and *Olea europaea* [26]. In addition, microheterogeneities have been identified in different sources, including foods [38], grasses and tree pollen [26], [27]. Alternative splicing is also another mechanism to generate protein variability, i.e. profilin II in mouse [39], apple [38] and *Arabidopsis*, where sequence introns are implicated in the constitutive expression of profilin in most vegetative tissues, exhibiting functionally differential properties in reproductive tissues [40].

We have performed an extensive analysis of profilin sequence polymorphism, which includes the first report of such variability in hazel pollen profilin, as well as the largest number of olive pollen cultivars ever analyzed. Similarly to what occurs in other plant species, profilin variability in the 5 species analyzed might be risen by multiple gene copies [6]. In addition, sequence polymorphism described here would come up as new mechanism to generate variability, more concretely in plant species such as olive, with a wide germplasm (more than 2000 cultivars around the world) [41], suffering different stress or physiological conditions. Thus, sequence polymorphism is affecting protein functional and regulatory motifs, as reported here for Ser/Thr or Tyr phosphorylation, might be able to regulate profilin activity and those processes implicated by differentiated dynamics.

The complex expression patterns and functional diversity of profilin family members in other plant species has been correlated with the existence of specific cellular functions [12] and its complex regulation [21]. Our results have shown a wide variability generated through olive germplasm that would be enough to afford the functional requirements, in terms of actin dynamics and signal transduction, buffering different stress and physiological

conditions [29], and to provide an additional mechanism to protect cells from loss-of-functions with multiple forms of profilins, with overlapping and specific functions. Protein isoforms with unique functions could be required at different times or locations, while overlapping expression of different cytoskeletal proteins and functionally distinctive isoforms are required at the same time in the same tissue or cell [42]. Our result confirms that profilin isoforms exhibit different Mw and pI as result of the polymorphism. Birch and timothy grass pollen profilin have shown a rather acidic pI [43]. In addition, *Acanthamoeba* [44], *Dictyostelium* [45] and humans [46] have several profilin isoforms with differences in sequence, pI, function and expression pattern. Maize pollen profilins also have different tissue expression [23] and function [42].

Deletions seem to be an additional or alternative mechanism to generate profilin isoforms with differential physicochemical properties, e.g. the human PROF III and a virus profilin homologue [47]. We have shown in this work the presence of deletions in 47 sequences of profilin, which also exhibit differential properties (MW, pI, post-translational modification sites), supporting the existence of this mechanism to generate profilin variability.

Nevertheless, it is worthy to mention that many of the profilin sequences described here exhibited multi-optional post-translational sites generated by polymorphism, e.g. new or missed residues of serine or threonine [48], [34], or tyrosine implicated in phosphorylation and protein regulation by potential differential interactions with different partners such as PLP or PLP-proteins [49], [50].

Other evidence supporting the existing polymorphism as a mechanism to generate multiple forms of profilins in plants come up with its co-evolution with different forms of actin and other ABPs, more precisely profilin-interacting proteins. There are evidences of different actin isoforms as the result of multiple genes in *Arabidopsis* [51] and soybean [52], and strong evidences of functional differences among actin classes, e.g. isoelectric points of the plant actins vary over a relatively wide [53], [54].

Plant genomes encode large vegetative and reproductive classes of actin and ABP gene families [51], with different ABP isoforms, e.g. profilin, with class-specific preferential interaction between the actin and profilin isoforms in plants [55]. Plant actin proteins have an unusually large number of non-conservative amino acid substitutions (6 to 10%) mapping to the surface of the molecule [51], which have a significant effect on protein-protein interactions; i.e. profilin and thymosin [56]. Fourteen additional

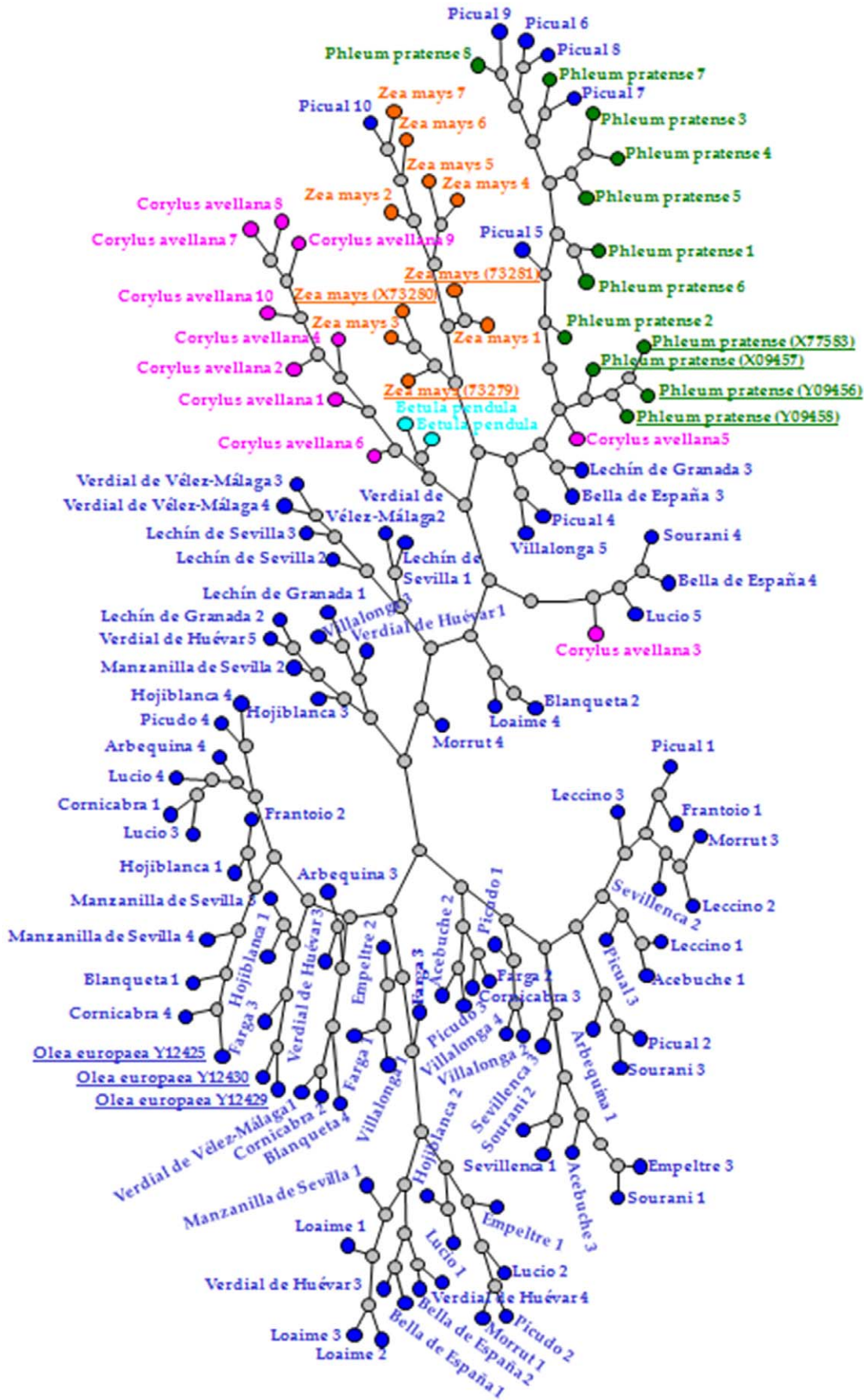


Figure 3. Phylogenetic analysis of profilin. Neighbor-Joining (NJ) method was used to perform a phylogenetic analysis of the deduced protein sequences of *O. europaea* L. (blue), *B. pendula* (magenta), *C. avellana* (pink), *P. pratense* (green), and *Z. mays* (orange) profilin. Reference sequences of profilin from different species previously deposited in the GenBankTM/EMBL database were highlighted by underline. doi:10.1371/journal.pone.0030878.g003

families of ABPs encoded by multiple genes are implicated in actin dynamics, where formin is the most prominent actin nucleator and profilin interacting protein family. This profilin-interacting protein is responsible for the assembly and elongation of unbranched actin filaments. Formin family is integrated at least by 21 genes in *Arabidopsis* [57], and shows a sequence identity range between 15 to 55% in plants [58]. This high variability would indirectly regulate the number of isoforms of interacting partners like profilin.

Finally, profilin variability is also reflected in their multiple subcellular localization, functions and regulation under different conditions. It is generally accepted that profilins have a cytoplasmic localization [59]. They are preferentially associated to plasma membrane domains, and differentially located within developing microspores [60]. Other possible locations are amoplasts [61], and generative and vegetative nuclei in pollen [62], since different profilin partners of nuclear localization have been found like PIP₂ [63] and actin [64]. Chloroplast is one of the last and newly characterized profilin locations, as a result of the interaction with the Chloroplast Unusual Positioning 1 (CHUP1) protein [65].

Profilin polymorphism as the result of post-translational modifications and changes in partners-binding regions is also an important factor affecting multifunctionality

Profilins have surface-exposed binding areas for actins [66], poly-L-proline (PLP) proteins [67], and phosphoinositide lipids [68], [69]. The interacting motives have been identified in many proteins, including plant profilins [70]. The affinity for the different ligands differs by orders of magnitude between species and even isoforms [71], [72]. Extreme examples include a minor splicing form in mouse that has been reported not to bind to G-actin [73], and Vaccinia virus profilin, which does not bind to PLP [74]. Differential properties as a result of sequence polymorphism in profilin isoforms would bring up a range of interacting affinities for ATP, actin and PLP [23], [75]. In this regard, sequence variability would constitute a mechanism able to increase the possibilities for plant responses to multiple stress and physiological conditions, which would be mediated by the transduction of external signals through actin dynamics [76].

The actin molecule exhibits a wide surface involved in profilin interaction (2,250 Å²) [66], and a large number of non-conservative amino acid substitutions have been exhibited in various plant actin molecules [51]. In addition, our results confirm that the N-terminal region of profilin involved in the interaction with actin is non-conservative, with a large number of variable residues located close to the actin-interacting area in the profilin surface. This variability in the actin-binding surface has been also reported in different species [70].

Furthermore, the so-called “plant specific binding pocket” of the profilins analyzed here, which is also implicated in the actin interaction, contains two highly variable residues (H62 and Q79). Our results indicate that the polymorphism surrounding the actin-interacting area, as well as the variability of the residues directly implicated in this interaction, would affect the binding properties of profilin isoforms. In addition, the variability affecting residues involved in phosphorylation within the actin interaction area might also play a regulatory role in the properties of this interaction with different profilins [77].

PLP-binding stretches play a major role for profilin interaction with proline-rich proteins [78]. The affinity of profilin interactions is quite variable [71], [72], and it is also regulated through phosphorylation [77], [50]. Profilin sequence polymorphism affecting phosphorylation sites, precisely tyrosines 6 and 128 in olive cultivars, would be a major mechanism to regulate the affinity of profilin-PLPs interactions, especially under different cellular processes and environmental stresses [49], [50]. In addition, a new domain located around tyrosine 75 has been also implicated in profilin-PLP and PI3K interactions. This interacting area, integrated by a characteristic plant pocket is defined by three regularly spaced aromatic residues highly conserved among profilin sequences, and followed by a pattern of three residues susceptible of phosphorylation (such as positions T66, Y75 and Y109). The analysis of polymorphism of this region in olive profilins has shown a variable number of combinations of phosphorylation sites among these three positions, which could be phosphorylated *in vivo* [49], probably because they are exposed in the protein surface and accessible to the solvent [70]. Phosphorylations within PLP domains might be a fundamental regulatory process, able to generate isoforms with differential interacting properties [42]. It was previously demonstrated that isoforms of profilins are generated by tyrosine phosphorylation in different tissues of *Phaseolus vulgaris*. Furthermore, these modifications would mimic physicochemical properties such as different maize profilin isoforms do it, raising the possibility that individual phosphorylated isoforms might display specific roles in different tissues, since tyrosine phosphorylation in poly-L-proline-binding regions inhibits binding to phosphoinositide 3-kinase in *Phaseolus vulgaris* [50].

The profilin region interacting with lipid phosphoinositides is divided in two areas of the protein surface. Changes in residues belonging to the actin-interacting area can either increase or decrease the affinity of profilin to PIP [79]. A second binding site for PIP₂ is overlapping with the poly-L-proline-binding site at the C-terminal region of the profilin molecule [39], [80]. This area allows the molecular interaction and a possible competition between PIP₂ and the PLP ligands at the C-terminal site. The profilin sequences analyzed here have shown a high variability in this area, concretely in the positions Leu130 and Glu131, which are directly implicated in the PIP interaction. This variability would make it possible to regulate the interaction affinity of PIP lipid in profilin isoforms [69], [71], [80], as well as the interaction with other lipid phosphoinositides (PI_(3,4)P₂ and PI_(3,4,5)P₂) [81]. These differential affinities to several PIP molecules would be a mechanism to regulate differential signal transduction, buffering different stress and physiological signals throughout actin rearrangements. Furthermore, changes in the PIP-PLP binding overlapping region would make a second competitive mechanism to control interactions with different poly-L-proline rich proteins.

Both overlapping regions together, actin-PIP-PLP might be a regulatory mechanism for a positive or negative interaction with different ligands, under different cellular environments, e.g. phosphorylation in serine 92 by protein kinase C zeta has been reported to increase the affinity for G-actin and PLP, while the interaction with PIP₂ remained unaltered [77]. This serine residue in highly variable in olive cultivars and timothy-grass sequences, which constitute another example to support that sequence

polymorphism is a mechanism generating profilin variability, regulating the differential properties of interaction with actin and PLP [71], [82].

Conclusion

The study presented here has revealed the possible functional and regulatory consequences of sequence polymorphism in pollen profilins. This polymorphism might represent a mechanism to generate multiple profilin isoforms among species, the germplasm of a particular species, their tissue or even their subcellular localization. These isoforms, exhibiting a wide range of physicochemical differences as well as differences in profilin-ligand binding properties, could have a direct influence in the cellular dynamics and the regulatory processes orchestrated by actin cytoskeleton, leading to more robust and a wide range of responses of cells to different physiological and stress conditions. In addition, the co-existence of different profilin variants in the same cell would allow more complex processes of signals integration through proteins (ABPs) interaction networks and cytoskeletal rearrangements.

Further research will unravel the effects of the polymorphism in different structures (folding) of the profilin isoforms, and its influence in the interaction with different ligands and in the cellular dynamics. Overall, the knowledge gained will help provide a comprehensive understanding on the stunning variety of functions of these small, ubiquitous proteins.

Materials and Methods

Olive (*Olea europaea* L.) pollen was individually collected during May and June from olive trees of 24 different cultivars, grown in different olive germplasm collections in Spain (CIFA “Alameda del Obispo”, Córdoba, CIFA “Venta del Llano”, Jaén, Olive Culture Museum, Baeza, Jaén, and Estación Experimental del Zaidín, CSIC, Granada). Pollen samples were collected in large paper bags by vigorously shaking the inflorescences, sequentially sieved through 150 and 50 µm mesh filters to eliminate debris and maintained at -80°C . Pollen from *Betula verrucosa* var. *Laciniata* and *Corylus avellana* var. *Avellana* was collected from well-identified trees at the Botanical Garden of the University of Bonn (Germany) using the same procedure described above. Commercially available pollen (Allergome, Sweden) was used in the case of *Phleum pratense* var. *Pratense* and *Zea mays* var. *Birko*.

RT-PCR, Cloning and sequencing of profilin transcripts

Total RNA was isolated from 100 mg pollen samples of each cultivar/species by using the RNeasy Plant Total RNA kit (Qiagen). cDNA synthesis was carried out by using Superscript II reverse transcriptase (Invitrogen) and a poly-dT adaptor as a primer, following manufacturer’s indications. PCR amplifications were carried out from 50–100 ng of the template cDNA, by using 0.2–0.5 µM of each one of the following degenerated primers [5'-AGAGAATTCCATATGTCGTGGCA(A/G)(A/G)CGTACGT-3'] (forward) and [5'-AGAAAGCTT(C/T)TACA(G/T)GCC(C/T)TGTTCA(C/G/T)(A/C/G)AGGTA-3'] (reverse), 1 µl (2.5 U) of the PfuUltra High-Fidelity DNA Polymerase (Stratagene), 250 µM each dNTP, final reaction buffer at 1X, and ultrapure water up to 50 µl of final reaction volume. PCR mixtures were subjected to the following conditions in a Biometra T-Gradient Thermocycler (Biotron, Germany): initial heating step at 95°C for 5 minutes, denaturation at 94°C for 30 sec, annealing at 56°C for 45 sec, and extension at 72°C for 1 min. A final extension step of 10 min at 72°C was included after 30 cycles. After analyzing the PCR products by agarose gel electrophoresis, bands (405 bp) were

excised and purified from gel with the Gel Purification Kit (Qiagen). Purified fragments were ligated into the pGEM-T easy Vector (Promega) and used to transform *Escherichia coli* DH $_{5\alpha}$ competent cells (Stratagene) according to the manufacturer’s instructions. Variable number (1–10) clones were sequenced.

Polymorphism analysis of profilin sequences

Both nucleotide and deduced amino acid sequences obtained in the current work were searched for identity by the nucleotide-nucleotide BLAST (blastn) and amino acid BLAST (Blastp) programs [83], respectively. A multiple sequence alignment and subsequent analysis was performed using ClustalW software [84], based on Blosom62 matrix (BLOck SUBstitution Matrix) [85], and viewed using the Jalview viewer 2.2 [86]. The Bioedit v 7.0.5.3 [87] software was used to calculate the sequence identity matrices.

In order to assess the variability present in nucleotide sequences, the alignment was used to calculate an entropy plot for each specie [88], [89] by measuring of the lack of “bits of information content” at each position in the alignment. For amino acid sequences, variability was calculated as the number of different residues occurring at each position of the alignment divided by the frequency of the most common one [90]. The numerical obtained regarding nucleotide and amino acid variability were calculated and summarized as intervals of identity percentages within families, species and cultivars throughout identity matrices.

Physicochemical properties and post-translational modification motifs

Physicochemical properties of the profilin sequences were analyzed by using the ExPASy Proteomics Server. The ProtParam tool [91] was implemented to calculate the MW/pI of the different profilins, as well as instability index, aliphatic index and grand average of hydropathicity (GRAVY).

Profilin consensus patterns were checked for each original sequence and further analysis were performed to highlight the presence of functional motifs by using the PROSITE database [92]. Biologically meaningful motifs, susceptible of posttranslational modifications were derived from multiple alignments and the ScanProsite program [93], from the Expert Protein Analysis System (ExPASy) proteomics server of the Swiss Institute of Bioinformatics [94], as well as phosphorylation motives were analyzed by using NETPhos v1.0 [95] and NETPhosK v1.0 [96].

Phylogenetic analysis of profilin sequences

Profilin proteins from 5 plant species were used to generate phylogenetic trees using ClustalW [84]. The alignment was created using the Blosom62 matrix (BLOck SUBstitution Matrix) [85], 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 [86]. Portions of sequences that could not be reliably aligned were eliminated. Phylogenetic tree was generated by the neighbor-joining method (NJ) [97], and the branches were tested with 1,000 bootstrap replicates. The tree was visualized using Treedyn program [98].

Statistical analysis of polymorphism

Statistical analysis was performed by using the SPSS v.18 statistical software package. A General comparison among multiple sample groups was performed throughout one-way analysis of variance (*One-way* ANOVA) on the basis of the Fisher-Snedecor distribution test ($\alpha = 0.05$ significance value) [99]. Normality and variances homogeneity of the data collection were checked by the Shapiro-Wilk test ($\alpha = 0.05$ significance value)

[100] and the Levene test ($\alpha=0.05$ significance value) [101], respectively, and *post hoc* range probes and pair of species comparisons were carried out with the parametric test of Games-Howell ($\alpha=0.05$ significance value) [102].

Supporting Information

Table S1 Physic-chemical properties deduced from the profilin sequences. Different physic-chemical parameters were calculated for the amino acid sequences of profilin from the five species studied. Parameters were: molecular weight (PM), isoelectric point (pI), extinction molar coefficient (C.E.M.) at 280 nm ($M^{-1} cm^{-1}$) (-SH-SH/-S-S-), instability index, GRAVY, and aliphatic index. (DOC)

Table S2 Analysis of the polymorphism affecting post-translational modification motifs in the profilin sequences. The most representative and important posttranslational modifications were examined for the profilin amino acid sequences. These motifs included: N-myristoylation, amidation, phosphorylation by MAP kinase, N-glycosylation and targeting signal for microsomal bodies. (DOC)

References

- Volkman D, Baluška F (1999) The actin cytoskeleton in plants: From transport networks to signaling networks. *Microsc Res Tech* 47: 135–154.
- Staiger CJ (2000) Signaling to the actin cytoskeleton in plants. *Annu Rev Plant Physiol Plant Molec Biol* 51: 257–288.
- Vantard M, Blanchoin L (2002) Actin polymerization processes in plant cells. *Curr Opin Plant Biol* 5: 502–506.
- Valenta R, Ferreira F, Grote M, Swoboda I, Vrtala S, et al. (1993) Identification of profilin as an actin-binding protein in higher plants. *J Biol Chem* 268: 22777–22781.
- Bateman A, Coin L, Durbin R, Finn RD, Hollich V, et al. (2004) The Pfam protein families database. *Nucleic Acids Res* 32: D138–D141.
- Huang SR, McDowell JM, Weise MJ, Meagher RB (1996) The Arabidopsis profilin gene family. Evidence for an ancient split between constitutive and pollen-specific profilin genes. *Plant Physiol* 111: 115–126.
- Wilke DE, Otto JJ (2003) Profilin functions in cytokinesis, nuclear positioning, and stomatogenesis in *Tetrahymena thermophila*. *J Eukaryot Microbiol* 50: 252–262.
- Polet D, Lambrechts A, Ono K, Mah A, Peelman F, et al. (2006) *Caenorhabditis elegans* expresses three functional profilins in a tissue-specific manner. *Cell Motil Cytoskeleton* 63: 14–28.
- Witke W, Sutherland JD, Sharpe A, Arai M, Kwiatkowski DJ (2001) Profilin I is essential for cell survival and cell division in early mouse development. *Proc Natl Acad Sci USA* 98: 3832–3836.
- Blasco R, Cole NB, Moss B (1991) Sequence analysis, expression, and deletion of a vaccinia virus gene encoding a homolog of profilin, a eukaryotic actin-binding protein. *J Virol* 65: 4598–4608.
- Ren HY, Xiang Y (2007) The function of actin-binding proteins in pollen tube growth. *Protoplasma* 230: 171–182.
- Kandasamy MK, McKinney EC, Meagher RB (2002) Plant profilin isoforms are distinctly regulated in vegetative and reproductive tissues. *Cell Motil Cytoskeleton* 52: 22–32.
- McDowell JM, Huang S, McKinney EC, An YQ, Meagher RB (1996) Structure and evolution of the actin gene family in *Arabidopsis thaliana*. *Genetics* 142: 587–602.
- Karlsson R, Lindberg U (2007) Profilin, an essential control element for actin polymerization. In *In Actin Monomer-binding Proteins*. P. Lappalainen, ed. Landes Bioscience, Georgetown.
- Schlüter K, Jockusch BM, Rothkegel M (1997) Profilins as regulators of actin dynamics. *Biochim Biophys Acta* 27: 97–109.
- Staiger CJ, Gibbon BC, Kovar DR, Zonia LE (1997) Profilin and actin-depolymerizing factor: modulators of actin organization in plants. *Trends Plant Sci* 2: 275–281.
- Snowman BN, Kovar DR, Shevchenko G, Franklin-Tong VE, Staiger CJ (2002) Signal-mediated depolymerization of actin in pollen during the self-incompatibility response. *Plant Cell* 14: 2613–2626.
- Wang H-Y, Yu Y, Chen Z-L, Xia G-X (2005) Functional characterization of *Gossypium hirsutum* profilin 1 gene (GhPFN1) in tobacco suspension cells. Characterization of in vivo functions of a cotton profilin gene. *Planta* 222(4): 594–603.
- Staiger CJ, Blanchoin L (2006) Actin dynamics: old friends with new stories. *Curr Opin Plant Biol* 9: 554–562.
- Balaska F, Volkman D (2002) Actin-driven polar growth of plant cells. *Trends Cell Biol* 12: 14.
- Kovar DR, Drobak BK, Collings DA, Staiger CJ (2001) The characterization of ligand-specific maize profilin mutants. *Biochem J* 358: 49–57.
- Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 35: D61–D65.
- Staiger CJ, Goodbody KC, Hussey PJ, Valenta R, Drobak BK, et al. (1993) The profilin multigene family of maize: Differential expression of three isoforms. *Plant J* 4: 631–641.
- Rihs HP, Rozynek P, May-Taube K, Welticke B, Baur X (1994) Polymerase chain reaction based cDNA cloning of wheat profilin: a potential plant allergen. *Int Arch Allergy Immunol* 105: 190–194.
- Mittermann I, Swoboda I, Pierson E, Eller N, Kraft D, et al. (1995) Molecular cloning and characterization of profilin from tobacco (*Nicotiana tabacum*): increased profilin expression during pollen maturation. *Plant Mol Biol* 27: 137–146.
- Asturias JA, Arilla MC, Gómez Bayón N, Martínez J, Martínez A, et al. (1997b) Cloning and expression of the panallergen profilin and the major allergen (Ole e 1) from olive tree pollen. *J Allergy Clin Immunol* 100: 365.
- Gao ZS, van de Weg WE, Schaart JG, van Arkel G, Breiteneder H, et al. (2005) Genomic characterization and linkage mapping of the apple allergen genes Mal d 2 (thaumatin-like protein) and Mal d 4 (profilin). *Theor Appl Genet* 111(6): 1087–1097.
- Asturias JA, Arilla MC, Gómez Bayón N, Martínez J, Martínez A, et al. (1997a) Cloning and high level expression of *Cynodon dactylon* (Bermuda grass) pollen profilin (Cyn d 12) in *Escherichia coli*: purification and characterization of the allergen. *Clin Exp Allergy* 27: 1307.
- Wopfner N, Willeroidee M, Hebenstreit d, Van Ree R, Aalbers M, et al. (2002) Molecular and immunological characterization of profilin from mugwort pollen. *Biol Chem* 383(11): 1779–1789.
- Asturias JA, Ibarrola I, Eserverri J, Arilla MC, Gonzalez-Rioja R, et al. (2004) PCR-based cloning and immunological characterization of *Parietaria judaica* pollen profilin. *J Investig Allergol Clin Immunol* 14: 43–48.
- Binder M, Ortner S, Erben H, Scheiner O, Wiedermann G, et al. (1995) The basic isoform of profilin in pathogenic *Entamoeba histolytica* cDNA cloning, heterologous expression, and actin-binding properties. *Eur J Biochem* 233: 976–981.
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157(1): 105–32.
- Guruprasad K, Reddy BVP, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Prot Eng* 4: 155–164.
- Sigrist CJA, Cerutti L, Hulo N, Gattiker A, Falquet L, et al. (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform* 3: 265–274.
- Luan S (2003) Protein phosphatases in plants. *Annu Rev Plant Biol* 54: 63–92.

Table S3 The polymorphism of putative phosphorylation residues. Serine, threonine and tyrosine residues of the profilin sequences were analyzed. Multiple combinations of residues for susceptible phosphorylation were found among profilin sequences. (DOC)

Acknowledgments

JCJL and SM thank Spanish CSIC for their respective I3P-CSIC grants. The authors also want to thank Concepción Martínez for her excellent technical assistance. The authors want to deeply thank the sources of pollen collections in Spain: 1) World olive germplasm bank, CIFA “Alameda del Obispo” (Junta de Andalucía), Córdoba; 2) Estación Experimental del Zaidín (CSIC), Granada; 3) Olive Culture Museum, Baeza, Jaén; 4) CIFA “Venta del Llano”, Jaén; and in Germany: The Botanic Garden (Bonn).

Author Contributions

Conceived and designed the experiments: JDA MIR. Performed the experiments: JCJ SM. Analyzed the data: JCJ SM AJC JDA DV. Contributed reagents/materials/analysis tools: DV. Wrote the paper: JCJ JDA MIR.

36. Limmongkon A, Giuliani C, Valenta R, Mittermann I, Heberle-Bors E, et al. (2004) MAP kinase phosphorylation of plant profilin. *Biochem Biophys Res Commun* 324: 382–386.
37. Sharrocks AD, Yang SH, Galanis A (2000) Docking domains and substrate-specificity determination for MAP kinases. *Trends Biochem Sci* 25: 448–453.
38. Rihs H-P, Chen Z, Rueff F, Petersen A, Rozynek P, et al. (1999) IgE binding of the recombinant allergen soybean profilin (rGly m 3) is mediated by conformational epitopes. *J Allergy Clin Immunol* 104(6): 1293–1301.
39. Lambrechts A, Verschelde JL, Jonckheere V, Goethals M, Vandekerckhove J, et al. (1997) The mammalian profilin isoforms display complementary affinities for PIP2 and proline-rich sequences. *EMBO J* 16: 484–494.
40. Jeong Y-M, Mun J-H, Lee I, Woo JC, Hong CB, et al. (2006) Distinct Roles of the First Introns on the Expression of Arabidopsis Profilin Gene Family Members. *Plant Physiol* 140: 196–209.
41. Rallo L, Barranco D, Caballero JM, Del Rio C, Martin A, et al. (2005) Variedades de olivo en España. Junta de Andalucía, M.A.P.A. and Ediciones Mundi-Prensa (Eds), Madrid.
42. Kovar DR, Drobak BK, Staiger CJ (2000) Maize profilin isoforms are functionally distinct. *Plant Cell* 12: 583–598.
43. Giehl K, Valenta R, Rothkegel M, Ronsiek M, Mannherz HG, et al. (1994) Interaction of plant profilin with mammalian actin. *Eur J Biochem* 226: 681–689.
44. Pollard TD, Rimm DL (1991) Analysis of cDNA clones for *Acanthamoeba* profilin-I and profilin-II shows end to end homology with vertebrate profilins and a small family of profilin genes. *Cell Motil Cytoskel* 20: 169–177.
45. Haugwitz M, Noegel AA, Rieger D, Lottspeich F, Schleicher M (1991) Dictyostelium discoideum contains two profilin isoforms that differ in structure and function. *J Cell Sci* 100: 481–489.
46. Honoré B, Madsen P, Andersen AH, Leffers H (1993) Cloning and expression of a novel human profilin variant, profilin II. *FEBS Lett* 330: 151–155.
47. Butler-Cole C, Wagner MJ, Da Silva M, Brown GD, Burke RD, et al. (2007) An ectromelia virus profilin homolog interacts with cellular tropomyosin and viral A-type inclusion protein. *Virology Journal* 4: 76.
48. Vemuri B, Singh SS (2001) Protein kinase C isozyme-specific phosphorylation of profilin. *Cell Signal* 13(6): 433–439.
49. Guillen G, Valdes-Lopez V, Noguez R, Olivares J, Rodriguez-Zapata LC, et al. (1999) Profilin in *Phaseolus vulgaris* is encoded by two genes (only one expressed in root nodules) but multiple isoforms are generated in vivo by phosphorylation on tyrosine residues. *Plant J* 19: 497–508.
50. Aparicio-Fabre R, Guillen G, Estrada G, Olivares-Grajales J, Gurrola G, et al. (2006) Profilin tyrosine phosphorylation in poly-L-proline-binding regions inhibits binding to phosphoinositide 3-kinase in *Phaseolus vulgaris*. *Plant J* 47(4): 491–500.
51. McDowell JM, Huang S, McKinney EC, An YQ, Meagher RB (1996) Structure and evolution of the actin gene family in Arabidopsis thaliana. *Genetics* 142: 587–602.
52. McLean BG, Eubanks S, Meagher RB (1990a) Tissue-specific expression of divergent actins in soybean root. *Plant Cell* 2: 335–344.
53. McLean BG, Huang S, McKinney EC, Meagher RB (1990b) Plants contain highly divergent actin isoforms. *Cell Motil Cytoskel* 17: 276–290.
54. Meagher RB, McLean BG (1990) Diversity of plant actins. *Cell Motil Cytoskeleton* 16: 164–166.
55. Kandasamy MK, Burgos-Rivera B, McKinney EC, Ruzicka DR, Meagher RB (2007) Class-Specific Interaction of Profilin and ADF Isovariants with Actin in the Regulation of Plant Development. *Plant Cell* 19: 3111–3126.
56. Larsson H, Lindberg U (1988) The effect of divalent cations on the interaction between calf spleen profilin and different actins. *Biochim Biophys Acta* 953: 95–105.
57. Deeks MJ, Hussey PJ, Davies B (2002) Formins: intermediates in signal-transduction cascades that affect cytoskeletal reorganization. *Trends Plant Sci* 7: 1360–1385.
58. Chalkia D, Nikolaidis N, Makalowski W, Klein J, Nei M (2008) Origins and evolution of the formin multigene family that is involved in the formation of actin filaments. *Mol Biol Evol* 25: 2717–2733.
59. Vidali L, Hepler PK (1997) Characterization and localization of profilin in pollen grains and tubes of *Lilium longiflorum*. *Cell Motil Cytoskeleton* 36: 323–338.
60. von Witsch M, Baluska F, Staiger CJ, Volkman D (1998) Profilin is associated with the plasma membrane in microspores and pollen. *Eur J Cell Biol* 77: 303–312.
61. Fischer S, Grote M, Fahlbusch B, Muller WD, Kraft D, et al. (1996) Characterization of Phl p 4, a major timothy grass (*Phleum pratense*) pollen allergen. *J Allergy Clin Immunol* 98(1): 189–198.
62. Hess MW, Valenta R (1997) Profilin revealed in pollen nuclei: immunoelectron microscopy of high-pressure frozen *Ledebouria socialis* Roth (Hyacinthaceae). *Sex Plant Reprod* 10: 283–287.
63. Mazzotti G, Zini N, Rizzi E, Rizzoli R, Galanzi A, et al. (1995) Immunocytochemical detection of phosphatidylinositol 4,5-bisphosphate localization sites within the nucleus. *J Histochem Cytochem* 43: 181–191.
64. Jockusch BM, Becker M, Hindennach I, Jockusch H (1974) Slime mold actin. Homology to vertebrate actin and presence in the nucleus. *Exp Cell Res* 89: 241–246.
65. von Braun S, Schleiff E (2008) The chloroplast outer membrane protein CHUP1 interacts with actin and profilin. *Planta* 227: 1151–1159.
66. Schutt CE, Myslik JC, Rozycki MD, Goonesekere NC, Lindberg U (1993) The structure of crystalline profilin-beta-actin. *Nature* 365: 810–816.
67. Mahoney NM, Jamney PA, Almo SC (1997) Structure of the profilin-poly-L-proline complex involved in morphogenesis and cytoskeletal regulation. *Nat Struct Biol* 4: 953–960.
68. Lassing I, Lindberg U (1985) Specific interaction between phosphatidylinositol 4, 5-bisphosphate and profilactin. *Nature* 314: 472–474.
69. Fedorov AA, Magnus KA, Graupe MH, Lattman EE, Pollard TD, et al. (1994) X-ray structures of isoforms of the actin-binding protein profilin that differ in their affinity for phosphatidylinositol phosphates. *Proc Natl Acad Sci USA* 91: 8636–8640.
70. Thorn KS, Christensen HEM, Shigeta R, Huddler D, Shalaby L, et al. (1997) The crystal structure of a major allergen from plants. *Structure* 5: 19–32.
71. Schlüter K, Jockusch BM, Rothkegel M (1997) Profilins as regulators of actin dynamics. *Biochim Biophys Acta* 27: 97–109.
72. Schlüter K, Schleicher M, Jockusch BM (1998) Effects of single amino acid substitutions in the actin-binding site on the biological activity of bovine profilin I. *J Cell Sci* 111: 3261–3273.
73. Di Nardo A, Gareus R, Kwiatkowski D, Witke W (2000) Alternative splicing of the mouse profilin II gene generates functionally different profilin isoforms. *J Cell Sci* 113: 3795–3803.
74. Machesky LM, Cole NB, Moss B, Pollard TD (1994b) Vaccinia virus expresses a novel profilin with a higher affinity for polyphosphoinositides than actin. *Biochemistry* 33: 10815–10824.
75. Gibbon BC, Zonia LE, Kovar DR, Hussey PJ, Staiger CJ (1998) Pollen profilin function depends on interaction with proline-rich motifs. *Plant Cell* 10: 981–993.
76. Jockusch BM, Murk K, Rothkegel M (2007) The profile of profilins. *Rev Physiol Biochem Pharmacol* 159: 131–149.
77. Sathish K, Padma B, Munugalavada V, Bhargavi V, Radhika KV, et al. (2004) Phosphorylation of profilin regulates its interaction with actin and poly (L-proline). *Cell Signal* 16: 589–596.
78. Kovar DR, Drobak BK, Collings DA, Staiger CJ (2001) The characterization of ligand-specific maize profilin mutants. *Biochem J* 358: 49–57.
79. Skare P, Karlsson R (2002) Evidence for two interaction regions for phosphatidylinositol (4,5)-bisphosphate on mammalian profilin I. *FEBS Lett* 522: 119–124.
80. Lambrechts A, Jonckheere V, Dewitte D, Vandekerckhove J, Ampe C (2002) Mutational analysis of human profilin I reveals a second PIP2 binding site neighbouring the poly(L-proline)binding site. *BMC Biochem* 3: 12.
81. Lu PJ, Shieh WR, Rhee SG, Yin HL, Chen CS (1996) Lipid products of phosphoinositide 3-kinase bind human profilin with high affinity. *Biochemistry* 35: 14027–14034.
82. Sohn RH, Goldschmidt-Clermont PJ (1994) Profilin: At the crossroads of signal transduction and the actin cytoskeleton. *BioEssays* 16: 465–472.
83. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–402.
84. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673–4680.
85. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89(22): 10915–10919.
86. Clamp M, Cuff J, Searle SM, Barton GJ (2004) The Jalview Java alignment editor. *Bioinformatics* 20: 426–427.
87. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids* 41: 95–98.
88. Pierce JR (1980) An introduction to information theory: symbols, signals and noise, 2nd ed Dover Publications, Inc., New York, NY.
89. Schneider TD, Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 18: 6097–6100.
90. Radauer C, Willerroider M, Fuchs H, Hoffmann-Sommergruber K, Thalhamer J, et al. (2006) Cross-reactive and species-specific immunoglobulin E epitopes of plant profilins: an experimental and structure-based analysis. *Clin Exp Allergy* 36: 920–929.
91. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, et al. (2005) Protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook*. In: *The Proteomics Protocols Handbook*. pp 571–607.
92. Sigrist CJ, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, et al. (2010) PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* 38: D161–166.
93. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, et al. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 34: 362–365.
94. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, et al. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31: 3784–3788.
95. Blom N, Gammeltoft S, Brunak S (1999) Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* 294(5): 1351–1362.

96. Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4(6): 1633–49.
97. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4): 406–425.
98. Chevenet F, Brun C, Banuls AL, Jacq B, Christen R (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 7: 439.
99. Mehta CR, Patel NR (1983) A network algorithm for performing Fisher's exact test in $r \times c$ contingency tables. *J Am Stat Assoc* 78: 427–434.
100. Shapiro SS, Wilk MB (1965) An analysis of variance for normality (complete samples). *Biometrika* 52: 591–611.
101. Levene H (1960) Robust tests for equality of variances. In: Ingram Olkin, Harold Hotelling, et alia Stanford University Press. pp 278–292.
102. Games PA, Howell JF (1976) Pairwise multiple comparison procedures with unequal n 's and/or variances: A Monte Carlo study. *J Educ Stat* 1: 113–125.