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. Alché¹* 

¹Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, High Council for Scientific Research, Granada, Spain, ²Institute 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 up 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/varietal 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.

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, Phil p 12 and Zea m 12. However, the presence of polymorphism has only been poorly pointed out in a low number of sequences in *O. europaea* [23], *T. aestivum* [24], *A. thaliana* [6], *N. tabacum* [25], *O. europaea* [26], *C. donkey* [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 *O. europaea*), 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-binging 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<sup>TM</sup>/EMBL Database (Table 1).

The full-length cDNA of the profilin sequences ranged between 393 bp (*Olea europaea*, DQ663558 or *Oea mays*, DQ663565) to 405 bp (*Olea europaea* Y12429, Y124230, Y124245). 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, Phil 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 among 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. polypephalum* (P22271) 45.3–36.0%, *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 distribution 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.

| Olive Cultivar/Clone | Accession n° | Olive Cultivar/Clone | Accession n° | Species/Clone | Accession n° |
|----------------------|--------------|----------------------|--------------|--------------|--------------|
| Acebuche 1           | DQ138355     | Lucio 2              | DQ138363     | Betula pendula 1 | DQ650633    |
| Acebuche 2           | DQ138356     | Lucio 3              | DQ138365     | Corylus avellana 1 | DQ663543    |
| Acebuche 3           | DQ138357     | Lucio 4              | DQ138364     | Corylus avellana 2 | DQ663544    |
| Arbequina 1          | DQ138327     | Lucio 5              | DQ640908     | Corylus avellana 3 | DQ663545    |
| Arbequina 2          | DQ138328     | Manzanilla Sevilla 1 | DQ117911     | Corylus avellana 4 | DQ663546    |
| Arbequina 3          | DQ138329     | Manzanilla Sevilla 2 | DQ138324     | Corylus avellana 5 | DQ663547    |
| Arbequina 4          | DQ138330     | Manzanilla Sevilla 3 | DQ138325     | Corylus avellana 6 | DQ663548    |
| Bella de España 1    | DQ317563     | Manzanilla Sevilla 4 | DQ138326     | Corylus avellana 7 | DQ663549    |
| Bella de España 2    | DQ317564     | Morrut 1             | DQ317573     | Corylus avellana 8 | DQ663550    |
| Bella de España 3    | DQ640909     | Morrut 2             | DQ317574     | Corylus avellana 9 | DQ663551    |
| Bella de España 4    | DQ640910     | Morrut 3             | DQ317575     | Corylus avellana 10 | DQ663552   |
| Blanqueta 1          | DQ138335     | Morrut 4             | DQ317576     | Phleum pratense 1 | DQ663535    |
| Blanqueta 2          | DQ138336     | Picual 1             | DQ317580     | Phleum pratense 2 | DQ663536    |
| Blanqueta 3          | DQ138337     | Picual 2             | DQ317581     | Phleum pratense 3 | DQ663537    |
| Blanqueta 4          | DQ138338     | Picual 3             | DQ317582     | Phleum pratense 4 | DQ663538    |
| Cornicabra 1         | DQ138331     | Picual 4             | DQ640904     | Phleum pratense 5 | DQ663539    |
| Cornicabra 2         | DQ138332     | Picual 5             | DQ663553     | Phleum pratense 6 | DQ663540    |
| Cornicabra 3         | DQ138333     | Picual 6             | DQ663554     | Phleum pratense 7 | DQ663541    |
| Cornicabra 4         | DQ138334     | Picual 7             | DQ663555     | Phleum pratense 8 | DQ663542    |
| Empeltre 1           | DQ138342     | Picual 8             | DQ663556     | Zea mays 1       | DQ663559    |
| Empeltre 2           | DQ138343     | Picual 9             | DQ663557     | Zea mays 2       | DQ663560    |
| Farga 1              | DQ317565     | Picudo 1             | DQ117907     | Zea mays 3       | DQ663561    |
| Farga 2              | DQ317566     | Picudo 2             | DQ117908     | Zea mays 4       | DQ663562    |
| Farga 3              | DQ317567     | Picudo 3             | DQ117909     | Zea mays 5       | DQ663563    |
| Frantoio 1           | DQ317568     | Picudo 4             | DQ117910     | Zea mays 7       | DQ663565    |
| Frantoio 2           | DQ317569     | Sevillenca 1         | DQ317570     | Sevillenca 2     | DQ317571    |
| Galega 1             | DQ317570     | Sevillenca 2         | DQ317571     | Sevillenca 3     | DQ317572    |
| Hojiblanca 1         | DQ661979     | Sevillenca 3         | DQ317573     | Sourani 1        | DQ317574    |
| Hojiblanca 2         | DQ661980     | Sourani 2            | DQ317575     | Sourani 3        | DQ317576    |
| Hojiblanca 3         | DQ661981     | Sourani 4            | DQ317577     | Sourani 5        | DQ317578    |
| Leccino 1            | DQ138345     | Sourani 1            | DQ640905     | Sourani 6        | DQ317579    |
| Leccino 2            | DQ138346     | Verdial Huévar 1     | DQ117902     | Verdial Huévar 2 | DQ117903    |
| Leccino 3            | DQ138347     | Verdial Huévar 2     | DQ117903     | Verdial Huévar 3 | DQ117904    |
| Lechin de Granada 1  | DQ317571     | Verdial Huévar 3     | DQ117904     | Verdial Huévar 4 | DQ117905    |
| Lechin de Granada 2  | DQ317572     | Verdial Huévar 4     | DQ117905     | Verdial Huévar 5 | DQ117906    |
| Lechin de Sevilla 1  | DQ28766      | Verdial Vélez-Málaga 1 | DQ138358    |
| Lechin de Sevilla 2  | DQ61976      | Verdial Vélez-Málaga 2 | DQ138359    |
| Lechin de Sevilla 3  | DQ61977      | Verdial Vélez-Málaga 3 | DQ138360    |
| Lechin de Sevilla 4  | DQ61978      | Verdial Vélez-Málaga 4 | DQ138361    |
| Loaime 1             | DQ138339     | Villalonga 1         | DQ138351     | Villalonga 2     | DQ138352    |
| Loaime 2             | DQ138340     | Villalonga 2         | DQ138352     | Villalonga 3     | DQ138353    |
| Loaime 3             | DQ138341     | Villalonga 3         | DQ138353     | Villalonga 4     | DQ138354    |
| Loaime 4             | DQ640903     | Villalonga 4         | DQ138354     | 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
The average molecular weight was 14,334.40 ± 0.61 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 [17G(Q/L)hl(T/A)/A/S]22, [30GQdgSV25], [33GSvwAQ35], [64GMfvAG69], [67GLhlGG72], [93GGitSK98], 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 [17/19GQqlAA24], [30GSvwAQ35] and [66GLhlGG72], whereas 2 sites were detected in timothy-grass sequences with two possible motifs, [20GT vwlAQ35], and [43GMtvAT69]. In maize sequences, 1–2 myristoylation sites were detected, with the two possible motifs [30GA(t/a/v)wAQ35] and [63GLlGG69].

Larger homogeneity in the number of amidation motives was observed, with a unique predictive sequence [84/87rGKK90], present in all species. No glycosylation motifs were found in any of the profilin sequences, with the exception of the olive sequence DQJ38337 (cv. Blanqueta) [51NGTM54].

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. [79Q→A], [82A→P/S], [84A→V/R], [95S→A/T], [97I→V/S], which may affect the kinase interaction motif (KIM domain) [78IQGEAGAVIRGKGGKITK98]. 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).
Figure 2. Multiple alignment of the deduced amino acid sequences of profilins. A) Key amino acid affecting protein folding and maintenance of the 3D structure, as well as Cys at positions 13, 106 and 118 are pointed by black and green arrows respectively at the top of the figure. Reference sequences of profilin from different species previously deposited in the GenBankTM/EMBL database were highlighted in red colour. Numbering of the sequence positions begins at the ATG start codon from ORF2. Significant deletions found in 1, 3 or 4 amino acids in the amino acid sequences are highlighted with red boxes. Amino acids integrating different ligand-binding regions are depicted with different coloured boxes: actin (orange), PLP (blue) and PIP (purple). Solvent filled plant-specific binding pockets are showed with green boxes. Alignment symbols: (.) same aa; (–) deletion. Consensus symbols: (*) same aa; (:) conservative change, (.) semi-conservative change, (±) non-conservative change. B) Sequence conservation index calculated by the Jalview program. 

doi:10.1371/journal.pone.0030878.g002
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 | 20 | 21 | 22 | 23 | 24 | 25 | 29 | 32 | 34 | 35 | 40 | 41 | 48 | 50 | 51 | 52 |
|-------------------|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Olea europaea     | 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 |
| Betula pendula    | - | - | - | - | 4.00 | 4.00 | 4.00 | 4.00 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Corylus avellana  | 3.33 | - | - | - | - | - | - | - | - | - | 3.75 | - | - | - | - | - | - | - | - | - | - | - |
| Phleum pratense   | 4.00 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Zea mays          | - | - | - | - | - | - | - | - | - | - | 4.29 | - | - | - | - | - | 4.29 | - | - | - | - | - |

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.

| Family | Olea europaea | Corylus avellana | Betula pendula | Betulaceae (B. pendula + C. avellana) | Phleum pratense | Zea mays | Poaceae (P. pratense + Z. mays) |
|--------|---------------|-----------------|---------------|--------------------------------------|----------------|---------|---------------------------------|
| Oleaceae | 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** |
| Olea europaea | 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** |
| Corylus avellana | 75.7–100 | 74.3–95.5 | 74.3–100 | 74.3–95.5 | 75.0–98.4 | 70.2–86.1 | 70.2–98.4 | **Nucleotide** |
| 77.4–100 | 75.1–96.9 | 75.1–100 | 77.5–96.9 | 77.5–96.9 | 76.6–84.7 | 76.6–96.9 | 76.6–96.9 | **Amino acid** |
| Betula pendula | 99.5–100 | 74.3–100 | 74.8–76.1 | 74.3–100 | 71.9–75.3 | 71.9–76.1 | 71.9–79.6 | **Amino acid** |
| 96.2–100 | 75.1–100 | 76.6–79.6 | 71.9–78.1 | 71.9–79.6 | 71.9–79.6 | 71.9–79.6 | 71.9–79.6 | **Amino acid** |
| Betulaceae (B. pendula + C. avellana) | 74.3–100 | 74.1–98.4 | 70.2–86.1 | 70.2–98.4 | 70.2–98.4 | 70.2–98.4 | 70.2–98.4 | **Nucleotide** |
| 75.1–100 | 76.6–96.9 | 74.4–84.7 | 74.4–96.9 | 74.4–96.9 | 74.4–96.9 | 74.4–96.9 | 74.4–96.9 | **Amino acid** |
| P. pratense | 94.6–100 | 81.8–87.1 | 81.8–100 | 81.8–100 | 79.3–100 | 79.3–100 | 79.3–100 | **Nucleotide** |
| 94.6–100 | 79.3–86.2 | 79.3–100 | 79.3–100 | 79.3–100 | 79.3–100 | 79.3–100 | 79.3–100 | **Amino acid** |
| Z. mays | 79.9–100 | 84.6–100 | 80.1–100 | 80.1–100 | 80.1–100 | 80.1–100 | 80.1–100 | **Amino acid** |
| 84.6–100 | 80.1–100 | 80.1–100 | 80.1–100 | 80.1–100 | 80.1–100 | 80.1–100 | 80.1–100 | **Amino acid** |
| Poaceae (P. pratense + Z. mays) | 79.9–100 | 79.9–100 | 79.9–100 | 79.9–100 | 79.9–100 | 79.9–100 | 79.9–100 | **Nucleotide** |
| 79.3–100 | 79.3–100 | 79.3–100 | 79.3–100 | 79.3–100 | 79.3–100 | 79.3–100 | 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 | Olea europaea | Betula pendula | Corylus avellana | Phleum pratense | Zea mays |
|------------------|--------------|----------------|------------------|----------------|----------|
| Olea europaea    | 0.02         | 0.16**         | 0.12**           | 0.03           | NUCLEOTIDE SEQUENCES |
|                  | p = 0.71     | p = 0.00       | p = 0.00         | p = 0.39       |          |
| Betula pendula   | 0.83**       | 0.14**         | 0.11**           | 0.01           | p = 0.97 |
|                  | p = 0.00     | p = 0.00       | p = 0.00         | p = 0.00       |          |
| Corylus avellana | 0.15**       | 0.67**         | 0.32**           | 0.13**         | p = 0.00 |
|                  | p = 0.00     | p = 0.00       | p = 0.00         | p = 0.00       |          |
| Phleum pratense  | 1.42**       | 0.58**         | 0.09             | 0.09**         |          |
|                  | p = 0.00     | p = 0.00       | p = 0.92         | p = 0.00       |          |
| Zea mays         | 1.12**       | 0.38**         | 0.30             | 0.03           |          |
|                  | p = 0.00     | p = 0.42       | p = 0.01         | p = 0.14       |          |

AMINO ACID SEQUENCES

Comparisons in each box include the Game-Howell test and the level of significance (p < 0.05). Statistical significative 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 protein variability have been described such as single nucleotide of multigene families [6]. Moreover, different mechanisms rising properties in higher plants, mainly as the result of the occurrence of 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 pl as result of the polymorphism. Birch and timothy grass pollen profilin have shown a rather acidic pl [43]. In addition, *Acanthamoeba* [44], *Dictyostelium* [45] and humans [46] have several profilin isoforms with differences in sequence, pl, 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 isovariants 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, pl, 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 acts 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 isovariants, e.g. profilin, with class-specific preferential interaction between the actin and profilin isovariants 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
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 cytoplasmatic localization [59]. They are preferentially associated to plasma membrane domains, and differentially located within developing microspores [60]. Other possible locations are amiplasts [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 iso variants 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 iso variants. 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 iso variants with differential interacting properties [42]. It was previously demonstrated that iso forms of profilins are generated by tyrosine phosphorylization in different tissues of Phaseolus vulgaris. Furthermore, these modifications would mimic physicochemical properties such as different maize profilin iso variants do it, raising the possibility that individual phosphorylated iso forms 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 PIPI [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 PIP2 and profilin iso variants [69], [71], [80], as well as the interaction with other lipid phosphoinositides [P1,3,4,5 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
Pollen Profilin Polymorphism and Multifunctionality

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 isovariants among species, the germplasm of a particular species, their tissue or even their subcellular localization. These isovariants, exhibiting a wide range of physiochemical 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'-AGAGAATTTCCATATGCTGGCA(A/G)/A/G/GTACGT-3'] (forward) and [5'-AGAAAGCTT/C/T/TACA/G/TGCCC/C/T/TTGTCA/G/T/AGGTA-3'] (reverse), 1 μl (2.5 U) of the PhiUltra 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 DH5α 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 Blosum62 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 species [88], [99] 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 was performed to highlight the presence of functional motifs by using the PROSITE database [92]. Biologically meaningful motifs, susceptible of postranslational 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 Blosum62 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 three 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 (α = 0.05 significance value) [99]. Normality and variances homogeneity of the data collection were checked by the Shapiro-Wilk test (α = 0.05 significance value)

[References]

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 Blosum62 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 species [88], [99] 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.

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 Blosum62 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 species [88], [99] 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 was performed to highlight the presence of functional motifs by using the PROSITE database [92]. Biologically meaningful motifs, susceptible of postranslational 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 Blosum62 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 three 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 (α = 0.05 significance value) [99]. Normality and variances homogeneity of the data collection were checked by the Shapiro-Wilk test (α = 0.05 significance value)
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  **Physico-chemical properties deduced from the profilin sequences.** Different physico-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) ), instability index, GRAVY, and aliphatic index.

**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.

**References**

1. Volkman D, Baluska F (1999) The actin cytoskeleton in plants: From transport networks to signaling networks. Microsc Res Tech 57: 135–154.
2. Staiger CJ (2000) Signaling to the actin cytoskeleton in plants. Annu Rev Plant Physiol Plant Mol Biol 51: 257–288.
3. Vantard M, Blanchon I (2002) Actin polymerization processes in plant cells. Curr Opin Plant Biol 5: 562–566.
4. Valenta R, Ferreira F, Guse M, Sloboda I, Vrtaľa S, et al. (1993) Identification of profilin as an actin-binding protein in higher plants. J Biol Chem 268: 22777–22781.
5. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, et al. (2004) The Pfam protein families database. Nucleic Acids Res 32: D138–D141.
6. 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.
7. Wilke DE, Otto JJ (2003) Profilin functions in cytokinesis, nuclear positioning, and stomagenesis in Tetrahymena thermophila. J Eukaryot Microbiol 50: 252–262.
8. 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: 14–21.
9. Wilke W, Sutherland JD, Sharp R, 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: 3032–3036.
10. Blasco R, Cole NR, 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: 4508–4608.
11. Ren HY, Xiang Y (2007) The function of actin-binding proteins in pollen tube growth. Protoplasma 230: 171–182.
12. Kandasamy MK, Kimney EC, Meagher RB (2002) Plant profilin isoforms are distinctly regulated in vegetative and reproductive tissues. Cell Motil Cytoskeleton 52: 22–32.
13. McDowell JM, Huang S, Kimney EC, An YQ, Meagher RB (1996) Structure and evolution of the actin gene family in Arabidopsis thaliana. Genetics 142: 587–602.
14. Karlsson R, Lindberg U (2007) Profilin, an essential control element for actin polymerization. In: In Actin Monomer-Binding Proteins F. Lappalainen, ed. Landes Bioscience, Georgetown.
15. Schluter K, Jechslehner BM, Rokhsigiel M (1997) Profilins as regulators of actin dynamics. Biochim Biophys Acta 27: 97–109.
16. Staiger CJ, Gibbon BC, Kovar DR, Zonia LE (1997) Profilin and actin-depolymerizing factor: modulators of actin organization in plants. Trends Plant Sci 2: 273–281.
17. Snowman BN, Kovar DR, Shevechenko G, Franklin-Tong VE, Staiger CJ (2002) Signal-mediated depolymerization of actin in pollen during the self-incompatibility response. Plant Cell 14: 2613–2626.
18. Wang HY, Yu Y, Chen ZL, Xia GX (2000) 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.
19. Staiger CJ, Blanchon L (2006) Actin dynamics: old friends with new stories. Curr Opin Plant Biol 9: 554–562.
20. Baharke F, Volkman D (2002) Actin-driven polar growth of plant cells. Trends Cell Biol 12: 14.
21. Kocár M, Drobak BK, Collings DA, Staiger CJ (2001) The characterization of ligand-specific maize profilin mutants. Biochim Biophys Acta 150: 49–57.
22. Pruitt KD, Tanosova T, Maggott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35: D61–D65.
23. 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.
24. Rihs HP, Roseneck P, May-Taube K, Wetcbe B, Baur X (1994) Polymerase chain reaction based cDNA cloning of wheat profilin: a potential plant allergen. Int Arch Allergy Immunol 105: 190–194.
25. Mittermann I, Sloboda 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.
26. Asturias JA, Arilla MC, Gómez Bayón N, Martínez J, Martínez A, et al. (1997a) Cloning and expression of the panallergen profilin and the major allergen (Ole e 1) from olive tree pollen. J Allergy Clin Immunol 100: 165.
27. 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 (thiamatin-like protein) and Mal d 4 profilin. Theor Appl Genet 111(6): 1087–1097.
28. Asturias JA, Arilla MC, Gómez Bayón N, Martínez J, Martínez A, et al. (1997a) Cloning and high level expression of Cydonia oblonga (Bermuda grass) pollen profilin (Cyn d 12) in Escherichia coli. Purification and characterization of the allergen. Clin Exp Allergy 27: 1307.
29. Wopfner N, Willerode M, Hebenstreit D, Van Kees R, Aalbers M, et al. (2002) Molecular and immunological characterization of profilin from mugwort pollen. Biol Chem 383(11): 1779–1789.
30. Asturias JA, Barrola I, Eversheer J, Arilla MC, Gonzalez-Ripsa R, et al. (2004) PCR-based cloning and immunological characterization of Parietaria judaica pollen profilin. J Investig Allergol Clin Immunol 14: 43–48.
31. Binder M, Ottner S, Ebner H, Scheiner O, Wiedermann G, et al. (1995) The basic isoform of profilin in pathogenic Entamoeba histolytica cDNA cloning, heterodispergous expression, and actin-binding properties. Eur J Biochem 233: 976–981.
32. Kite J, Doolittle RF (1982) A simple method for displaying the hydrophilic character of a protein. J Mol Biol 157(1): 105–132.
33. Guruprasad K, Reddy BVP, Pandit MV (1998) Correlation between stability of a protein and its disopyridine composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Prot Eng 1: 155–164.
34. Sigrist CJA, Cerutti L, Gattiker A, Hulo N, F אחרת O, et al. (2002) PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform 3: 265–274.
35. Luan S (2003) Protein phosphatases in plants. Annu Rev Plant Biol 54: 63–92.

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

**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 their pollen collections in Spain: 1) World olive germplasm bank, CIFA “Alameda del Obispo” (Junta de Andalucia), 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. Lammonqui A, Giuliana C, Valerana R, Mittermann I, Hebele-Bors E, et al. (2004) MAP kinase phosphorylation of plant profilin. Biochem Biophys Res Commun 324: 382–386.

37. Sharrocks AD, Yang SH, Galanin A (2000) Docking domains and substrate specificity determination for MAP kinases. Trends Biochem Sci 25: 448–453.

38. Rota-H Pi, Chen Z, Roett F, Petersen A, Royczek P, et al. (1999) IgE binding of the recombinant allergen soybean profilin (6gly m 3) is mediated by conformational epitopes. J Aller Clin Immun 104(6): 1293–1301.

39. Lambrecht A, Vercellelli 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: 418–424.

40. Jeong Y-M, Mun J-H, Lee I, Woo JG, Hong CB, et al. (2006) Distinct roles of profilin-I and profilin-II shows end to end homology with vertebrate profilins and a small family of profilin genes. Cell Motil Cytoskel 56: 285–294.

41. Haugwitz M, Noegel AA, Rieger D, Lottspeich F, Schleicher M (1991) Dictyostelium discoidium contains two profilin isoforms that differ in subcellular localization and function. J Cell Biol 100: 481–499.

42. Honore B, Madsen P, Andersen AH, Leffers H, et al. (2007) A novel profilin variant in human brain. Brain Res 1131: 151–157.

43. Giehl K, Valenta R, Rothkegel M, Ronsiek M, Mannherz HG, et al. (1994) Phosphorylated profilin-I and profilin-II shows end to end homology with vertebrate profilins and a small family of profilin genes. Cell Motil Cytoskel 29: 169–172.

44. Butler-Cole C, Wagner MJ, Da Silva M, Brown GD, Burke RD, et al. (2007) An electron microscopy of high-pressure frozen Ledebouria socialis Roth pollen grains and tubes of allergen. J Allergy Clin Immunol 98(1): 189–198.

45. Honore B, Madsen P, Andersen AH, Leffers H (1993) Cloning and expression of profilin-1 from human brain. J Neurochem 59: 176.

46. Meagher RB, McLean BG (1990) Diversity of plant actins. Cell Motil Cytoskel 17: 276–290.

47. Kovar DR, Drobak BK, Staiger CJ (1998) Maize profilin isoforms are functionally distinct. Plant Cell 12: 583–598.

48. Field K, Valenta R, Rothkegel M, Mannherz HG, et al. (1994) Interactions of profilin with mammalian actin. Eur J Biochem 269: 681–691.

49. Pollard TD, Rimm DL (1991) Analysis of cDNA clones for Actinophila profilin-I and profilin-II shows end to end homology with vertebrate profilins and a small family of profilin genes. Cell Motil Cytoskel 20: 169–172.

50. Haugwitz M, Noegel AA, Rieger D, Lottspeich F, Schleicher M (1991) Dictyostelium discoidium contains two profilin isoforms that differ in subcellular localization and function. J Cell Biol 100: 481–499.

51. Butler-Cole C, Wagner MJ, Da Silva M, Brown GD, Burke RD, et al. (2007) An electron microscopy of high-pressure frozen Ledebouria socialis Roth pollen grains and tubes of allergen. J Allergy Clin Immunol 98(1): 189–198.

52. Haugwitz M, Noegel AA, Rieger D, Lottspeich F, Schleicher M (1991) Dictyostelium discoidium contains two profilin isoforms that differ in subcellular localization and function. J Cell Biol 100: 481–499.
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 x 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.