



OPEN

The tree of life of polyamine oxidases

Daniele Salvi¹ & Paraskevi Tavladoraki²✉

Polyamine oxidases (PAOs) are characterized by a broad variability in catalytic properties and subcellular localization, and impact key cellular processes in diverse organisms. In the present study, a comprehensive phylogenetic analysis was performed to understand the evolution of PAOs across the three domains of life and particularly within eukaryotes. Phylogenetic trees show that PAO-like sequences of bacteria, archaea, and eukaryotes form three distinct clades, with the exception of a few procaryotes that probably acquired a PAO gene through horizontal transfer from a eukaryotic donor. Results strongly support a common origin for archaeal PAO-like proteins and eukaryotic PAOs, as well as a shared origin between PAOs and monoamine oxidases. Within eukaryotes, four main lineages were identified that likely originated from an ancestral eukaryotic PAO before the split of the main superphyla, followed by specific gene losses in each superphylum. Plant PAOs show the highest diversity within eukaryotes and belong to three distinct clades that underwent to multiple events of gene duplication and gene loss. Peptide deletion along the evolution of plant PAOs of Clade I accounted for further diversification of function and subcellular localization. This study provides a reference for future structure–function studies and emphasizes the importance of extending comparisons among PAO subfamilies across multiple eukaryotic superphyla.

Polyamines are small organic polycations with primary and secondary amino groups. They are widely present in all organisms and are involved in a variety of biological processes. The relative abundance of the different polyamines depends on species and varies in a tissue-specific way. The most common polyamines are the diamine putrescine (Put), triamine spermidine (Spd), and tetraamine spermine (Spm), although a large number of algal, fungal and bacterial species do not contain Spm^{1–3}. In some organisms a wider variety of polyamines has been observed, such as 1,3-diaminopropane (Dap), cadaverine, agmatine, thermospermine (T-Spm), norspermine (Nor-Spm), norspermidine, homospermidine, various long-chain and branched polyamines, as well as conjugated forms and acetylated derivatives of polyamines^{4–9}.

Polyamine biosynthetic pathway is an ancient and well conserved metabolic pathway. Most eukaryotes synthesize Put from ornithine through ornithine decarboxylase (ODC), whereas in plants and bacteria there is an additional pathway to Put which involves arginine decarboxylase (ADC). It has been suggested that eukaryotic ODC was inherited from an α -proteobacterial ODC progenitor and that the ADC pathway has been acquired by plants from the cyanobacterial precursor of the chloroplast¹⁰. Transfer of an aminopropyl group to Put by spermidine synthase (SPDS) results to the production of Spd, while spermine synthase (SPMS) and thermospermine synthase (TSPMS) incorporate a new aminopropyl group at the N^8 -(aminobutyl)- and N^1 -(aminopropyl)-end of Spd to synthesize Spm and T-Spm, respectively. Additional triamines, tetramines, as well as long-chain and branched-chain polyamines may be also formed by transfer of aminopropyl or aminobutyl groups to different polyamines^{8,11–15}. Phylogenetic studies have suggested that the SPDS genes of the various organisms derive from a common ancestor preceding the separation between prokaryotes and eukaryotes and that they have been the origin of SPMS and TSPMS activities through gene duplication and/or neofunctionalization¹⁶. Moreover, it has been hypothesized that plants acquired TSPMS early during evolution by horizontal gene transfer from archaea or bacteria^{16–18}.

Polyamine oxidases (PAOs) have an important role in polyamine metabolism and contribute to several physiological processes through regulation of polyamine levels and reaction products. PAOs are characterized by a broad variability in substrate specificity, catalytic mechanism and subcellular localization. They are FAD-dependent enzymes catalyzing the oxidation of the free, and/or acetylated form, of polyamines at the secondary amino groups^{19–21}. In mammals, the peroxisomal PAOs (PAOXs) preferentially oxidize N^1 -acetyl-Spm, N^1 -acetyl-Spd, and N^1, N^{12} -bisacetyl-Spm through an *exo*-mode to produce Spd, Put, and N^1 -acetyl-Spd, respectively, in addition to 3-acetamidopropanal and H_2O_2 ^{3,20,22}. Moreover, the mammalian spermine oxidases (SMOXs), which present cytosolic/nuclear localization, preferentially oxidize Spm to produce Spd, 3-aminopropanal and

¹Department of Health, Life and Environmental Sciences, University of L'Aquila, 67100 L'Aquila, Italy. ²Department of Science, University 'Roma Tre', 00146 Rome, Italy. ✉email: paraskevi.tavladoraki@uniroma3.it

H₂O₂^{19,23,24}. Unlike PAOXs and SMOXs, *Saccharomyces cerevisiae* PAO (FMS1) catalyzes the oxidation of both acetylated and non-acetylated polyamines^{25,26}.

In plants, the intracellular PAOs (e.g., the *Arabidopsis thaliana* AtPAO1 with a putative cytosolic localization and the three peroxisomal AtPAO2, AtPAO3, AtPAO4) preferentially oxidize the free form of Spd, Spm, T-Spm or Nor-Spm to produce 3-aminopropanal, H₂O₂ and Put or Spd^{27–32}. The cytosolic AtPAO5 (which has a higher activity as dehydrogenase than as oxidase) and its rice orthologue oxidize also *N*¹-acetyl-Spm^{9,32–34}. In contrast to the intracellular animal and plant PAOs, the extracellular *Zea mays* PAO (ZmPAO1, previously ZmPAO^{35,36}) and its orthologues in *Oryza sativa*, *Avena sativa* and *Hordeum vulgare* oxidize the carbon at the *endo*-side of the *N*⁴-nitrogen of the free forms of Spd and Spm with the production of Dap, H₂O₂ and an aminoaldehyde^{32,37}.

In some bacterial species, such as *Pseudomonas aeruginosa*, *Citrobacter freundii* and *Serratia marcescens*, Spd is oxidized by spermidine dehydrogenases with FAD and/or heme as prosthetic groups^{38–40}. Reaction products of these enzymes with Spd are Dap and 4-aminobutanol, indicating cleavage at the *endo*-side of the *N*⁴-nitrogen. The *P. aeruginosa* enzyme (SpdH) oxidizes also Spm through an *exo*-mode producing Spd and 3-aminopropanal^{38–41}. In several bacterial species, an FAD-dependent amine oxidase classified as putrescine oxidase (PuO) is also present. PuO additionally catalyses Spd oxidation, though less efficiently than Put oxidation^{42–45}. However, being active at the primary amino group, this enzyme cannot be considered as a PAO. Unlike the well characterized polyamine catabolic pathways in eukaryotes and bacteria, nothing is known about polyamine catabolism in archaea which possess distinct polyamine biosynthetic pathways and produce long-chain and branched polyamines^{15,46,47}.

While several studies have been performed on the evolutionary pathways of genes involved in polyamine biosynthesis^{10,11,16–18,48}, comparatively less is known about the evolutionary history of genes involved in polyamine catabolism. Most studies have focused on the genomic identification and biochemical characterization of PAO isoforms from single species and therefore they used phylogenetic methods mainly for delimiting clusters and subfamilies to which to assign PAO isoforms^{49–51}. Only a few studies considered a wide taxonomic representation of PAO sequences in large groups of organisms and enabled the elucidation of evolutionary relationships among PAO subfamilies and the processes underlying their functional and structural diversity. For example, the phylogenetic analyses of animal PAOs clarified that SMOX and PAOX subfamilies originate from a duplication event preceding the diversification of vertebrates and that subsequently SMOX and PAOX enzymes acquired differences in substrate specificity through divergent evolution and functional specialization^{52,53}. Likewise, a recent phylogenetic study on plant PAOs identified four main subfamilies and two main duplication events preceding angiosperm diversification⁵⁴. However, it is still unclear whether all plant PAO gene subfamilies originate from a common ancestral gene along the Viridiplantae lineage. The same applies for metazoan PAO gene subfamilies, whereas phylogenetic studies on fungal PAO subfamilies have not been performed at all. At a more general level, the phylogenetic origin of the extensive diversity of eukaryotic PAOs and their evolutionary relationships with the few putative PAOs recently identified in bacteria⁵⁵ are still unknown. In this study, a phylogenetic framework was developed to explore the relationships between eukaryotic PAOs and related proteins from bacteria and archaea, and to better understand the evolutionary root of eukaryotic PAO subfamilies, with a special focus on plants that show the highest PAO diversity. Analysis of gene structure and amino acid residues of the putative catalytic sites was also performed to better understand the evolutionary processes that led to functional diversification of PAO genes.

Results and discussion

Early origin of PAO-like proteins within the three domains of life. To investigate the evolution of PAOs, we assembled, through extensive iterative sequence similarities searches, a set of 428 sequences from bacteria, archaea, and different groups of eukaryotes including alveolates, amoebozoans, cryptista, excavates, fungi, green algae, haptista, land plants, metazoans, red algae, rhizarians, and stramenopiles (Supplementary Table S1). Searches targeting genomic data of Asgard archaea and eukaryote centrohelids, glaucophytes, and metamonads returned no hits (Supplementary Table S2).

Phylogenetic analyses including 300 PAO and PAO-like sequences of bacteria, archaea, and eukaryotes showed three main clades (Fig. 1): the 'Bacteria clade' including proteins from bacteria (Bootstrap support, BS = 100), the 'Archaea clade' (BS = 100) including 13 archaeal PAO-like proteins, one bacterial putrescine oxidase (*Rhodococcus erythropolis* PuO, coded as Bat-Re) and two eukaryotic monoamine oxidases (*Mus musculus* MAO-A and MAO-B, coded as Mm-MmMAO-A and Mm-MmMAO-B), and the 'Eukaryota clade' (BS = 91) including all eukaryotic PAOs plus two small clades of bacterial (N = 7) and archaeal (N = 3) proteins. Therefore, besides a few exceptions, each domain of life has specific PAO-like proteins that evolved from distinct ancestral proteins (Fig. 1).

Proteins of the Archaea clade are recovered as sister to the Eukaryota clade (BS = 100; Fig. 1) using either the midpoint or the MAD rooting methods, thus suggesting a shared evolutionary history between archaeal PAO-like proteins and eukaryotic PAOs. Moreover, Archaeal PAO-like proteins show a close phylogenetic relationship with the mammalian Mm-MmMAO-A and Mm-MmMAO-B⁵⁶ and the bacterial PuO⁴⁵, thus suggesting a common origin between PAOs and MAOs, likely from an oxidase protein carried by the common ancestor of archaea and eukaryotes⁵⁷. A common origin between PAOs and MAOs is further supported by their significant structural similarity^{35,58,59}. While biochemical information for archaeal PAO-like sequences is not available, their close evolutionary relationship with eukaryotic MAOs and bacterial PuO raises the question of whether the archaeal PAO-like proteins have catalytic properties more similar to MAOs and PuO than to PAOs.

Most of the bacterial PAOs are included in the well-supported Bacteria clade. This clade includes PAOs from beta- and gamma-proteobacterial species, as well as PAOs from some alpha- and epsilon-proteobacteria, acidobacteria, actinobacteria and Deinococcus-Thermus. These PAOs have a sequence identity ranging between 30

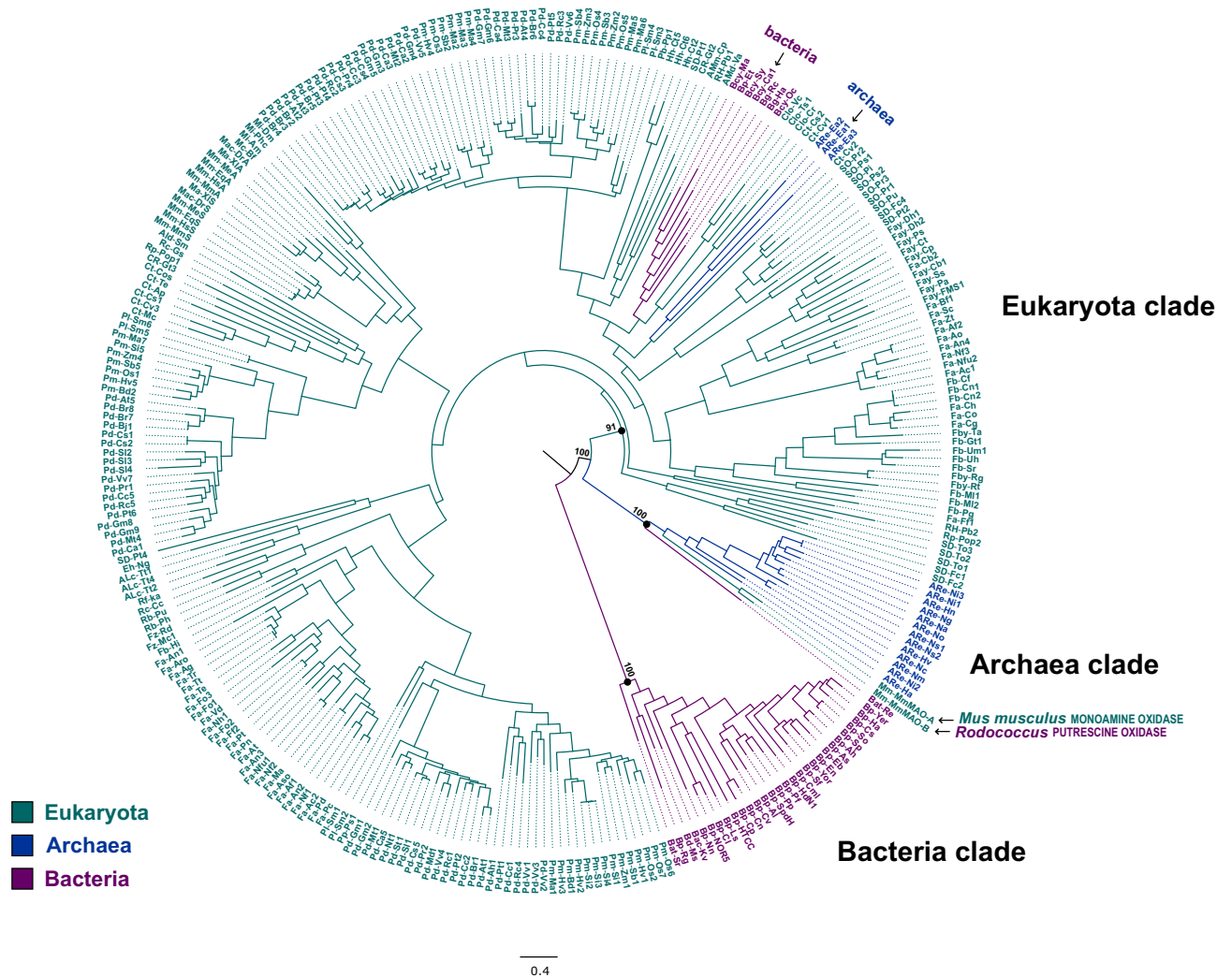


Figure 1. Maximum likelihood tree depicting the phylogenetic relationships between eukaryote PAOs and related PAO-like enzymes in bacteria and archaea; the tree is drawn to scale, with branch lengths measured in number of substitutions per site. Bootstrap support values (ultrafast bootstrap approximation) over 1000 replicates are reported in correspondence of the main nodes. The analysis involved 300 amino acidic sequences (see Supplementary Table S1). Sequences are coloured according to the three main domains of life: eukaryote = green, archaea = blue, and bacteria = red. The three main clades are named as follow: Eukaryota clade, Archaea clade, Bacteria clade; prokaryotic sequences clustering within the Eukaryota clade and non-archaeal sequences clustering within the Archaea clade are indicated by black arrows.

to 75% with *Pseudomonas aeruginosa* spermidine dehydrogenase (SpdH) which oxidizes both Spd and Spm⁴⁰. On the other hand, a few PAOs of cyanobacteria (Bcy-Oc, Bcy-Ma, Bcy-Sy, and Bcy-Ca1), chlorobacteria (Bg-Rc and Bg-Ha), and of the proteobacterium *Edwardsiella tarda* (Bp-Et) form a monophyletic clade nested within the Eukaryota clade (Fig. 1), and in particular within the PAO Clade IV (Fig. 2). This latter clade includes various eukaryotic lineages (amoebozoans, criptista, green algae, haptista, land plants, rhizaria, and stramenopiles), as well as a small clade of three archaeal proteins. This phylogenetic pattern suggests that the PAOs of these prokaryotes have been probably acquired through horizontal transfer from an eukaryotic lineage⁶⁰. However, further data on bacterial and archaeal PAO diversity, including detailed biochemical information, are required to corroborate this preliminary hypothesis.

Four main clades of eukaryotic PAOs. PAO sequences of eukaryotes exhibit significant diversity. Phylogenetic trees show that all eukaryotic PAOs evolved from a common ancestor (Fig. 1) and represent four main evolutionary lineages referred to as Clade I-IV (Fig. 2). The four main clades of eukaryotic PAOs and their sublineages are strongly supported by bootstrap approximation and/or SH-like approximate likelihood ratio test values ≥ 90 (Fig. 2). The midpoint rooting method recovered a close relationship between Clades I and II (plus a few sequences from stramenopiles, rhizarians and red algae), as well as between Clades III and IV (plus a small clade of stramenopiles); however, the placement of the root is not consistent between the midpoint and the MAD methods (result not shown).

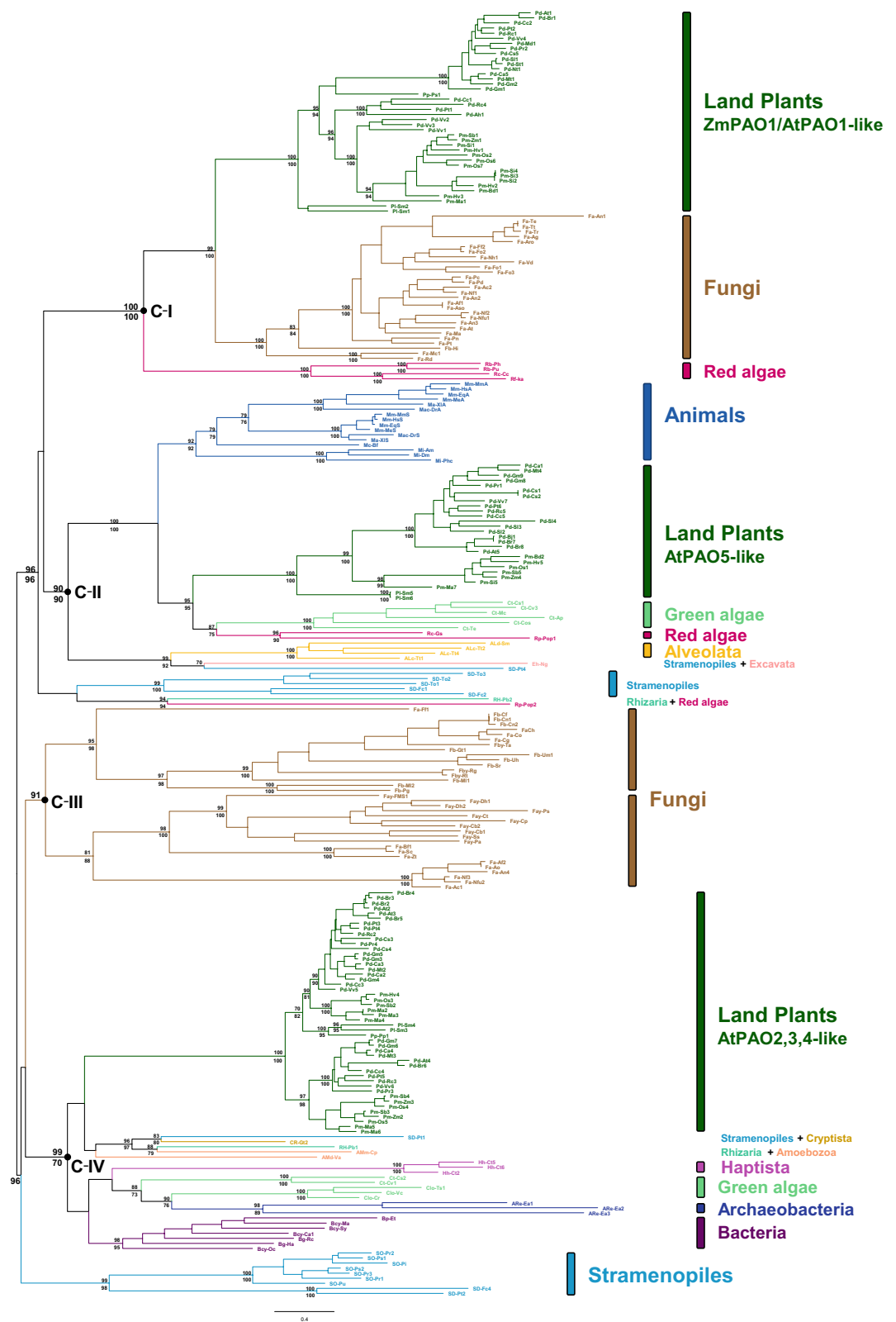


Figure 2. Phylogeny of eukaryote PAOs based on the 253 amino acid sequences of the Eukaryota clade of Fig. 1. The tree shown was estimated with the Maximum likelihood method under the WAG + I + G model of amino acid replacement. Nodal support (> 70) is reported along the main branches: above, ultrafast bootstrap approximation (1000 replicates); below, SH-like approximate likelihood ratio test (1000 replicates). Main clades and eukaryote groups are indicated.

Clade I includes three main subclades, one represented by plant PAOs that have high similarity (33–93%) to *A. thaliana* PAO isoform 1 (AtPAO1) and *Z. mays* ZmPAO1, another one consisting of fungal PAOs (mainly from ascomycetes) and a third one, sister to the other two, consisting of red algae PAOs. Clade II includes three subclades: one composed of animal PAOs (including vertebrate SMOXs and PAOXs) that is sister to an Archaeplastida sub-clade including land plant AtPAO5-like PAOs, green algae and red algae PAOs; a third subclade includes PAOs from alveolates, stramenopiles and excavata. Clade III consists exclusively of fungal PAOs grouped into two main subclades, both including yeast forms: one subclade is composed of ascomycete PAOs including *S. cerevisiae* FMS1²⁵ and the other is composed of basidiomycetes PAOs, including *Ustilago maydis* PAO (UmPAO; Fb-Um1 in Fig. 2)². Clade IV includes one subclade of plant PAOs with high sequence similarity to the AtPAO2, AtPAO3 and AtPAO4 (AtPAO2,3,4-like PAOs), as well as PAOs from green algae and various protists (amoebozoans, cryptista, rhizarians and stramenopiles) and a few prokaryotic PAOs that have been discussed above. In addition to these four main clades, putative PAOs of stramenopiles form two small clades, one including diatom proteins and the other one including both diatom and oomycete proteins (Fig. 2). PAO-like sequences of these two stramenopile clades have sequence identity of 21–31% to AtPAO1–AtPAO5, MmSMOX, MmPAOX and FMS1.

The phylogenetic distribution of PAOs of each main clade in multiple eukaryotic superphyla and the lack of monophyly of PAO isoforms of plants (Clades I, II, and IV), fungi (Clades I and III), green algae (Clades II and IV), red algae (Clades I and II) and stramenopiles (various clades), suggest a birth-and-death scenario, with the origin of the main lineages arising from the ancestral eukaryotic PAO before the split of the main superphyla followed by specific gene losses in each superphylum. According to this scenario PAO genes of Clade I would have been lost, for example, in animals and PAO genes of Clade II in fungi, whereas PAO of Clade III would have been retained only in fungi and those of Clade IV only in green plants and some protists. On the other hand, the low number of PAO-like sequences of protists available in the databases suggests caution with the interpretation of the absence of PAO genes, of one or more PAO lineages, in these groups.

The four unique groups of homologous PAO sequences identified within the broad phylogenetic framework used in this study provide a crucial reference for future structure–function studies and emphasize the importance of extending the comparisons among PAO subfamilies across multiple eukaryotic superphyla. This is particularly true, for example, for the plant PAO subfamilies ZmPAO1/AtPAO1-like and AtPAO5-like that show a closer relationship to either fungal or animal PAOs rather than to the plant AtPAO2,3,4-like PAOs subfamily. This finding is in agreement with previous studies showing that *Arabidopsis* AtPAO5 is more similar to animal PAOXs/SMOXs in terms of amino acid sequence (including amino acids of the catalytic site) and substrate specificity (specificity for *N*^l-acetyl-Spm), than to plant AtPAO1–4⁹. By analogy, the close phylogenetic relationships between plant ZmPAO1/AtPAO1-like proteins and ascomycete PAOs of Clade I provide directions for future comparative biochemical analyses and suggest that the available ZmPAO1 crystal structure may be a valuable resource for homology modelling and function prediction of related fungal proteins.

In the following subsections, we discuss in detail phylogenetic relationships between and within eukaryotic PAO clades in conjunction with gene structure, subcellular localization, substrate specificity and amino acid residues of the catalytic site.

Clade I: ZmPAO1/AtPAO1-like PAOs of plants and fungi. Plant PAOs of Clade I (Fig. 2) have high amino acid sequence identity to each other (from 40 to 55%). A phylogenetic analysis based on an extended dataset, including 81 ZmPAO1/AtPAO1-like PAOs from bryophytes, lycopodiophytes, pinophytes, angiosperms (eudicots and monocots) and their sister lineage *Amborella trichopoda* (Fig. 3) shows that plant PAOs of Clade I belong to two distinct groups (Fig. 3, box a). One includes ZmPAO1-like PAOs characterized by an extracellular localization (possessing a N-terminal signal peptide for secretion to the apoplast) and an *endo*-mode of substrate oxidation; whereas the second includes AtPAO1-like PAOs characterized by putative cytosolic localization (lacking any known targeting sequence to a specific subcellular compartment) and an *exo*-mode of substrate oxidation. ZmPAO1-like PAOs are widespread across Land Plants lineages (thought in eudicot angiosperms they are only present in a few species), whereas AtPAO1-like PAOs are exclusive of eudicot angiosperms (Fig. 3).

The phylogenetic tree indicates that bryophyte and lycopodiophyte ZmPAO1-like PAOs are sister to seed plant PAOs (BS = 100, node a, Fig. 3). Among the latter, two large clades are present, one including ZmPAO1-like isoforms of monocots, Vitales eudicots and the *A. trichopoda* Pa-Amt2 (Fig. 3; BS = 77, node c) and the other one being exclusively composed of AtPAO1-like isoforms of eudicots plus the *A. trichopoda* Pa-Amt1 (BS = 100, node f). Two smaller clades of extracellular PAOs were also recovered that have a close relationship with the AtPAO1-like clade (BS = 100, node b): one is composed of ZmPAO1-like isoforms found in some eudicots (BS = 98, node d), and the other of a third isoform of *A. trichopoda* (Pa-Amt3) and two PAOs from gymnosperms (BS = 98, node e). Overall, these results are consistent with an origin of the AtPAO1-like clade from a gene duplication event occurring on a ZmPAO1-like gene ancestor before angiosperm diversification⁵⁴. This duplication would have been followed by AtPAO1-like gene extinction in monocots and ZmPAO1-like gene extinction in several eudicots. However, the close relationship between the AtPAO1-like clade and the two newly discovered clades including ZmPAO1-like PAOs found in some eudicots, *Amborella* (Pa-Amt3) and two gymnosperms (BS = 93, node b) suggests an additional duplication event, followed by a gene loss in monocots, that preceded the origin of the AtPAO1-like clade (see the scheme in Fig. 3, box a). Indeed, albeit phylogenetic relationships among the AtPAO1-like clade and these two small ZmPAO1-like clades are not well resolved, the latter two have a closer affinity with the AtPAO1-like clade than with the other ZmPAO1-like clades (node c), strongly indicating that a single duplication event does not explain well the phylogenetic diversity of plant PAOs of Clade I. Moreover, the addition of representatives of gymnosperm PAOs compared to previous studies suggests that duplication

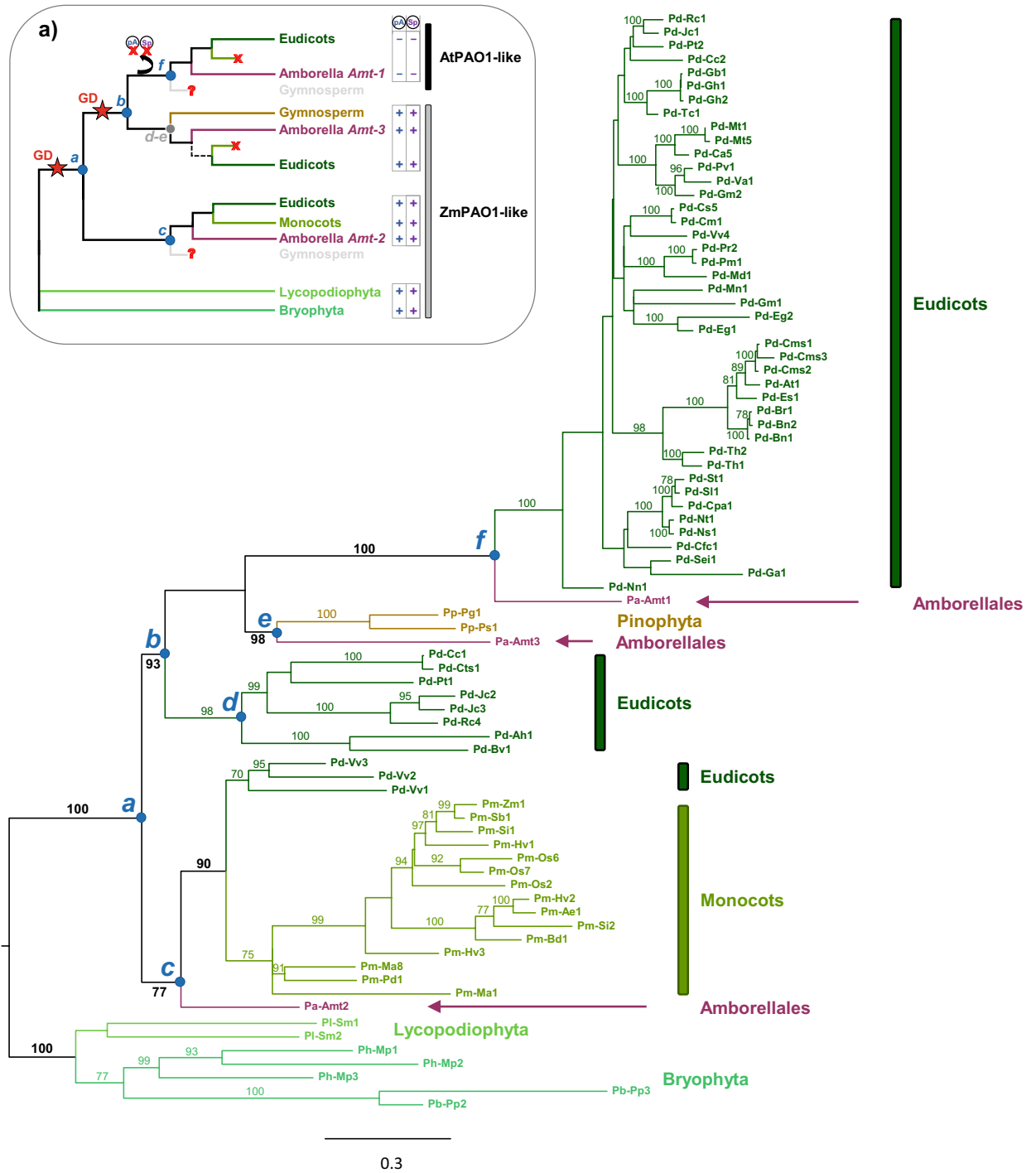


Figure 3. Maximum likelihood phylogeny of plant PAOs of the ZmPAO1/AtPAO1-like clade (see Fig. 2) based on 81 amino acid sequences. Main clades are labelled from *a* to *f*; bootstrap values of support (BS) are reported along the branches (BS > 70). Box (a) illustrates the evolutionary model for the AtPAO1-like proteins with two gene duplications (GD) in correspondence of node *a* and *b* (giving rise to the AtPAO1 subclade), and the loss of peptide A (pA) and of the signal peptide (sP) in correspondence of node *f*. Putative gene extinctions are indicated with 'x', whereas '?' denote missing information on the Gymnosperm group.

events within this clade might have been even older than previously thought, likely before the diversification of seed plants.

Comparative analysis of the available genomic and amino acid sequences showed three main differences between ZmPAO1-like and AtPAO1-like PAOs of Clade I. All ZmPAO1-like PAOs, including PAOs of the early divergent land plants, pinophytes and two isoforms of *Amborella* (Pa-Amt2 and Pa-Amt3), share a common structure with (i) 8 introns at highly conserved positions (plus an additional intron in the two *Selaginella moellendorffii*

PAOs and Pa-Amt3), (ii) a domain of 9 amino acids (peptide A, indicated as pA in Fig. 3, box a) close to Glu170 residue of ZmPAO1 catalytic site^{36,58} (aa174-aa182; numbering of mature ZmPAO1; Supplementary Figure S1), and (iii) the previously mentioned signal peptide for extracellular localization (indicated as Sp in Fig. 3, box a). In contrast, all AtPAO1-like PAOs, including *A. trichopoda* Pa-Amt1, show an additional intron at a position corresponding to the highly conserved amino acid residue Glu173 (numbering of mature ZmPAO1) (Supplementary Figure S1) and lack the peptide A and a signal peptide (given their intracellular localization). Therefore, these two peptides were lost after the gene duplication that gave rise to the AtPAO1-like clade (Fig. 3, box a).

The fungal PAO group of Clade I comprises mainly ascomycete PAOs with the single basidiomycete Fb-Hi PAO and the two zygomycete PAOs as sister to this clade (Fig. 2). Amino acid sequence identity between fungal and plant PAOs of Clade I is high (32–40%), including the amino acids of the catalytic site (see below Table 1). Most fungal PAOs (23 out of 29) have a predicted cleavable signal peptide suggesting extracellular localization; only three of them do not have typical features of the cleavable signal peptide in the N-terminal extension, whereas for the remaining three fungal PAOs only partial sequence data were available. Similarly to the extracellular ZmPAO1-like plant PAOs, the fungal PAOs possess a domain corresponding to the ZmPAO1 region aa174-aa182 (numbering of mature ZmPAO1), though with low sequence similarity. Analysis of the available genomic sequences revealed that the ascomycete PAOs share some common intron positions, which however are different from those of the AtPAO1-like and ZmPAO1-like plant PAO genes, as well from those of Fb-Hi PAO.

Within Clade I, sister to plant and fungal PAOs, there is a sub-clade of four red algae PAOs (Fig. 2), that have amino acid sequence identity of 21–30% with plant and fungal PAOs. Two of the four red algae PAOs have a putative signal peptide for extracellular localization, while for the other two only partial sequence were available thus preventing the identification of a signal peptide. Furthermore, similarly to the fungal PAOs, the four red algae PAOs possess a domain corresponding to the characteristic ZmPAO1 region aa174-aa182.

The identification of a putative signal peptide for extracellular localization in both the red algae and the fungal PAOs, as well as in most groups of plant PAOs of Clade I (except in the derived clade AtPAO1-like proteins), suggests that the ancestral PAO of this clade was extracellular and that it appeared early in the evolution of the eukaryotes.

Clade II: AtPAO5-like PAOs of plants and animal PAOs. Clade II includes the two reciprocally monophyletic sub-clades of animal PAOXs/SMOXs and of Archaeplastida AtPAO5-like PAOs (Fig. 2) with amino acid sequence identity among them in the range of 25–37%. Clade II also includes PAOs from two alveolates (*Tetrahymena thermophila* and *Symbiodinium microadriaticum*), a stramenopile (*Phaeodactylum tricoratum*) and an excavate heteroloboseans (*Naegleria gruberi*) which present amino acid sequence identity with plant and animal PAOs of Clade II in the range of 18–31%.

In agreement with the detailed study on animal PAOs by Polticelli et al.⁵², vertebrate PAOs consist of two sub-families, SMOXs and PAOXs, with different substrate specificity (free and acetylated form of Spm, respectively) and subcellular localization (cytosolic/nuclear and peroxisomal localization, respectively), probably derived from a duplication event followed by divergent evolution and functional specialization⁵². A recent study determined that the two PAO proteins of the cephalochordate amphioxus also show the same substrate specificity as vertebrate SMOXs and PAOXs, suggesting that gene duplication and functional specialization predates the diversification of chordates⁶¹.

As shown by a phylogenetic analysis based on an extended dataset (Fig. 4), phylogenetic relationships among plant PAO group of Clade II mirror the phylogenetic relationships among land plants and include three main sublineages, with the PAOs of *Marchantia polymorpha*, *S. moellendorffii* and *Selaginella lepidophylla* having a sister relationship to the clade formed by *Amborella* Pa-Amt6 and the groups of eudicot and monocot PAOs. Therefore, in contrast to the vertebrate PAOs, AtPAO5-like PAOs comprise a relatively homogeneous group of proteins that have orthologous relationships, except a few plant species that have multiple copies as a result of recent gene duplications. Plant AtPAO5-like enzymes have cytosolic localization (all lacking a targeting sequence to a specific subcellular compartment) and broad substrate specificity. Indeed, AtPAO5 and the AtPAO5-like enzymes of *O. sativa* and *S. lepidophylla* are able to oxidize the two substrates of the animal SMOXs/PAOXs (Spm and acetylated Spm) in addition to T-Spm and Nor-Spm^{9,34,62,63}. Furthermore, plant AtPAO5-like PAOs share a very simple gene structure with no intron⁹, with the exceptions of PAO genes in *Malus domestica* and *Tarenaya hassleriana* (Pd-Md6 and Pd-Th7, respectively) that have a single intron and in *S. moellendorffii* (Sl-Sm5 and Sl-Sm6) that have two introns. In contrast, animal PAO genes consist of 4 to 7 exons interspaced by 3 to 6 introns. Among animals, only *Trichoplax adhaerens* and *Nematostella vectensis*, representing early divergent animal phyla of Placozoa and echinoderms, have intron-less PAOs. These data indicate that animal and plant PAOs of Clade II experienced a very different evolutionary history from their common ancestor.

Clade III: a fungal-specific PAO clade. Clade III PAOs are exclusively from fungi and include one clade of ascomycete PAOs and another of PAOs found predominantly in basidiomycetes and a few ascomycetes (Fig. 2). Both clades include also yeast forms, such as FMS1 (Fay-FMS1) and UmPAO (Fb-Um1). Sequence identity among the PAOs of the ascomycete and basidiomycete PAO clades is relatively low, ranging from 19 to 28%. Moreover, available genomic sequences showed significant differences in gene structure between these two clades of fungal PAOs. In particular, ascomycete PAO genes possess from 2 to 4 introns at conserved positions, with the exception of yeast forms that have intron-less genes. Basidiomycete PAO genes have 1 to 3 introns at conserved positions, but different to the intron positions of ascomycete PAOs. Overall, the intron position of fungal PAOs of Clade III are different to those of PAOs of Clade I, II, and IV. Sequence analyses suggest intracellular localization of all fungal PAOs of Clade III in contrast to the fungal PAOs of Clade I which have extracellular localization.

	Clade	E62	E170	Y298	K300	T402	F403	Y439
Clade I	Land plant AtPAO1-like (42)	A ₃₄ , Q ₁ , V ₇	E	Y	K	S	Y	Y
	Land plant ZmPAO1-like (36)	E ₂₅ , H ₁ , I ₁ , F ₃ , S ₁ , N ₂	E ₃₄ , T ₂	Y ₃₅ , F ₁	K	S ₂₃ , T ₁₄ , C ₁ , A ₁	F ₂₀ , Y ₁₆	Y
	Fungi (29)	Q ₁₃ , E ₁₁ , H ₄ , S ₁	E ₂₇ , G ₁ , D ₁	Y	K	S	F ₂₈ , Y ₁₁	F ₁₈ , Y ₁₁
	Red algae (4)	Q	E	Y	K	S ₃ , A ₁	Y ₃ , F ₁	Y ₃ , F ₁
Clade II	Land plant AtPAO5-like (35)	H	E ₁₂ , Q ₂₀ , H ₂ , N ₁	V	K	S	Y	T
	Animal SMOX-like (38)	H	E	T	K	S	Y	T
	Animal PAOX-like (24)	H	E ₂₃ , T ₁	N ₂₁ , S ₁ , C ₂	K	S	Y	T
	Animal PAOX/SMOX (12)	H	E ₁₁ , A ₁	V ₈ , A ₁ , I ₁ , T ₁ , L ₁	K	S ₁₀ , A ₂	Y	T
	Green Algae (6)	H	Q ₅ , T ₁	V	K	S	Y	T ₅ , C ₁
	Red Algae (2)	H	E	V	K	S	Y	T
	Alveolates, diatoms, Excavates (6)	H	ND	G ₄ , V ₁ , A ₁	K	S ₁ , T ₁ , A ₁ , N ₂ , G ₁	Y	T
Clade III	Fungi (36)	H	ND	L	K	S ₂₁ , A ₁₅	Y ₂₂ , T ₁₄	C ₁₅ , T ₆ , S ₁₃ , R ₁ , M ₁
Clade IV	Land plant AtPAO2,3,4-like (102)	H	E	E	K	S ₆₄ , C ₃₇ , A ₁	Y	S ₈₃ , T ₁₉
	Green algae (4)	H	E ₂ , Q ₂	Y ₂ , L ₂	K	S	Y	T
	Diatoms, amoebozoans cryptista, rhizarians (5)	H ₄ , V ₁	E	L	K	S ₄ , A ₁	Y ₄ , F ₁	T ₃ , S ₁ , C ₁
	Haptista (3)	H	E	V	K	A	Y	T
	Bacteria (7)	H ₆ , Q ₁	E	L	K	S ₆ , A ₁	Y	T
	Archaea (3)	H	A ₂ , S ₁	A	K ₂ , V ₁	A ₁ , G ₁ , S ₁	Y	T ₂ , A ₁
Archaea Clade	Archaea (9)	G	ND	V	K	G ₇ , C ₁ , A ₁	Y	F ₅ , Y ₃ , H ₁
	RePuO (1)	S	ND	V	K	A	Y	H
	Mm-MmMAO (2)	G	ND	V	K	C	Y	Y
Diatom Clade*	Diatoms (5)	Y	ND	M ₁ , V ₁ , S ₁ , F ₁ , L ₁	K	A ₃ , S ₁ , V ₁	Y ₄ , H ₁	T ₄ , N ₁
Stramenopile Clade	Oomycetes (7)	H	E	Y ₄ , C ₃	K	A ₄	Y ₄	A ₅
	Diatoms (2)	H	E	Y	K	S ₁ , A ₁	Y	S

Table 1. Amino acid residues of the catalytic site of the various PAOs. Amino acid numbering refers to ZmPAO1 mature protein³⁵. Numbers in parentheses indicate the number of PAOs analyzed. Subscript numbers indicate the number of PAOs sequence with each amino acid residue at the corresponding position. ND: not determined, due to the presence of gaps and regions of low sequence homology. *SD-To1,2,3 and SD-Fc1,2.

Clade IV: AtPAO2,3,4-like PAOs from plants, green algae and photosynthetic bacteria. Land plant PAOs of Clade IV have high sequence identity (55–93%) to AtPAO2, AtPAO3 and AtPAO4. Furthermore, within the Clade IV (Fig. 2), PAO-like sequences from amoebozoans, cryptista, green algae, haptista, rhizarians, stramenopiles (diatoms; SD-Pt1), and prokaryotes (archaea and bacteria) have sequence identity to AtPAO2, AtPAO3 and AtPAO4 in the range of 25–38%.

Phylogenetic analysis based on an extended dataset (Fig. 5) showed that plant AtPAO2,3,4-like PAOs are widespread across main lineages of land plants including bryophytes, lycophytes, gymnosperms and angiosperms. The phylogenetic tree further shows that angiosperm isoforms of *Amborella*, monocots and eudicots belong to two sister clades (BS = 100, node *a*; Fig. 5), one including AtPAO2,3-like PAOs (BS = 75, node *c*), and the other including AtPAO4-like PAOs (BS = 100, node *b*). Therefore, in keeping with previous studies⁵⁴, AtPAO2,3-like and AtPAO4-like PAOs arose through a gene duplication before the origin of angiosperms (Fig. 5, box *a*). Furthermore, within the AtPAO4-like clade, all Poales monocots have two PAO copies clustered into two sister clades, suggesting that an additional duplication event took place in the AtPAO4-like PAO of this lineage of monocots (Fig. 5, box *a*).

AtPAO2,3,4-like PAO genes of land plants have highly conserved intron positions and upstream untranslated open reading frames⁶⁴ (uORFs). Of note, the *AtPAO2* and *AtPAO3* uORFs are more conserved with each other than with that of *AtPAO4*⁶⁴, which is consistent with phylogenetic results. All plant PAO proteins of Clade IV have a type I peroxisomal targeting signal (PTS1) at the carboxyl terminal and indeed peroxisomal localization has been shown for *AtPAO2*, *AtPAO3*, *AtPAO4*^{28,29}, as well as for *OsPAO3*, *OsPAO4*, and *OsPAO5*³¹. Interestingly, AtPAO4-like PAOs of one of the two monocot groups bear a non-canonical, but functional, peroxisomal targeting signal at the carboxyl terminal (CRT)³¹. Despite shared subcellular localization AtPAO4-like PAOs and AtPAO2,3-like PAOs differ in catalytic properties. Indeed, while AtPAO2 is equally active with either Spm or Spd, AtPAO3 has greater activity with Spd and AtPAO4 with Spm³⁰. In a similar way, while *OsPAO3* of the AtPAO2,3 group is active mainly with Spd, *OsPAO4* and *OsPAO5* of the AtPAO4 group is mainly active with Spm and T-Spm³¹. Whether these differences reflect distinct physiological roles is still unknown.

Within Clade IV, the AtPAO2,3,4-like clade of land plant PAOs is closely related to the PAOs of green algae (chlorophytes and trebouxiophyceae). The PAOs of green algae *V. carteri* and *C. reinhardtii* have a similar gene structure to each other sharing some intron positions, but different from those of the PAO genes of plants and of the other green algae (trebouxiophyceae) for which genomic sequences are available (e.g., Ct-Cv1 of *Chlorella variabilis*). Furthermore, *C. reinhardtii* PAO, but not the *V. carteri* PAO, has a PTS1 signal for peroxisomal

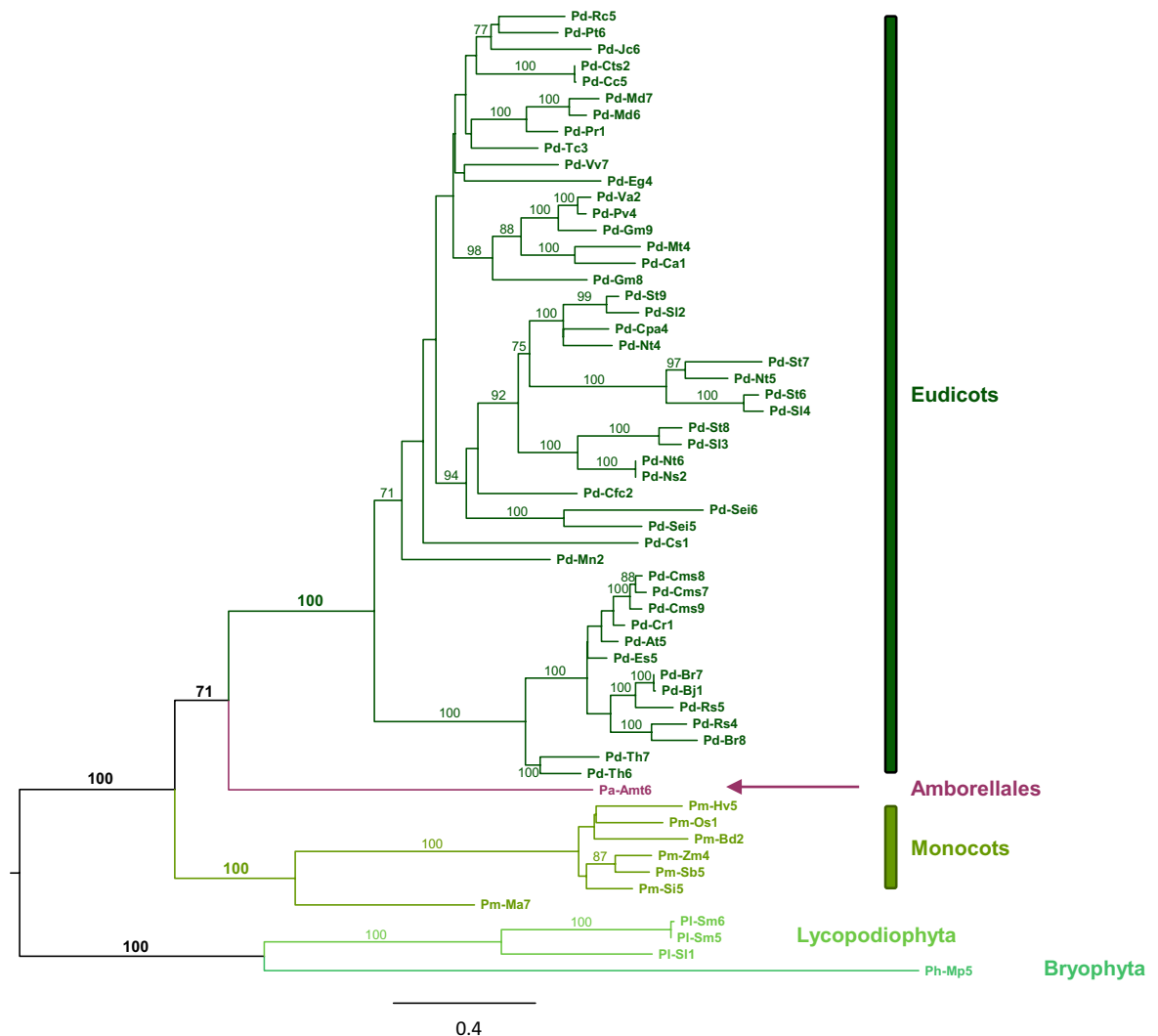


Figure 4. Maximum likelihood phylogeny of plant PAOs of the AtPAO5-like clade (see Fig. 2) based on 59 amino acid sequences. Bootstrap values of support (BS) are reported along the branches (BS > 70).

localization, though peroxisomal localization has still to be demonstrated. The PAO gene of *C. reinhardtii* is also an uORF which however does not exhibit similarity to the uORF of the AtPAO2,3,4 plant PAO genes. The close relationships of PAOs of many protists (amoebozoans, cryptista, haptista, rhizarians, and stramenopiles) with land plant PAOs of the AtPAO2,3,4-like clade allows identification of their common ancestors along the early diversification of eukaryotes, thus much earlier than during the diversification of streptophytes as suggested in previous studies⁵⁴.

Active site analysis of PAOs. The amino acid residues Glu62, Glu170, Tyr298, Lys300, Phe403 and Tyr439 of ZmPAO1 (Clade I) are key amino acids of the catalytic site as shown by resolution of the crystal structure^{58,65}, molecular modelling studies and site-directed mutagenesis^{36,66}. Glu62 and Glu170 are located close to the cofactor FAD and residue Tyr298 is found in close proximity to Lys300, the ϵ -amino group of which is H-bonded through a water molecule with the N5 atom of FAD, an atom which participates in the catalytic mechanism⁵⁸. Moreover, Phe403 and Tyr439 flank the catalytic tunnel on opposite sides and form a kind of 'aromatic sandwich'⁵⁸. Sequence alignments of PAOs of Clades I to IV, as well as of Archaea, Diatom and Stramenopile clades showed that Lys300 residue of ZmPAO1 is strictly conserved in all PAOs (Table 1). This residue is also present in RePuO and Mm-MAOs (Table 1). Furthermore, except for some fungal PAOs, all PAOs analysed (and also RePuO, Mm-MmMAO-A and Mm-MmMAO-A) have an aromatic amino acid (either Phe or Tyr) at position Phe403 of ZmPAO1 (Table 1). Tyr439 is highly conserved in Clade I (including fungal PAOs) and is also present in the Archaea Clade, while it is substituted mostly by Thr or Ser residues in the PAOs of the other clades, with the exception of the oomycete proteins in which it is substituted by Ala. Tyr439 is also conserved in Mm-MmMAO-A and Mm-MmMAO-B and substituted by His in RePuO (Table 1). These observations suggest that the 'aromatic sandwich' Phe403/Tyr439 is a particular characteristic of Clade I PAOs, as well as of the archaea PAO/MAO-like sequences. Glu62 is highly conserved in the ZmPAO1-like PAOs of monocots, *Amborella* (Pa-

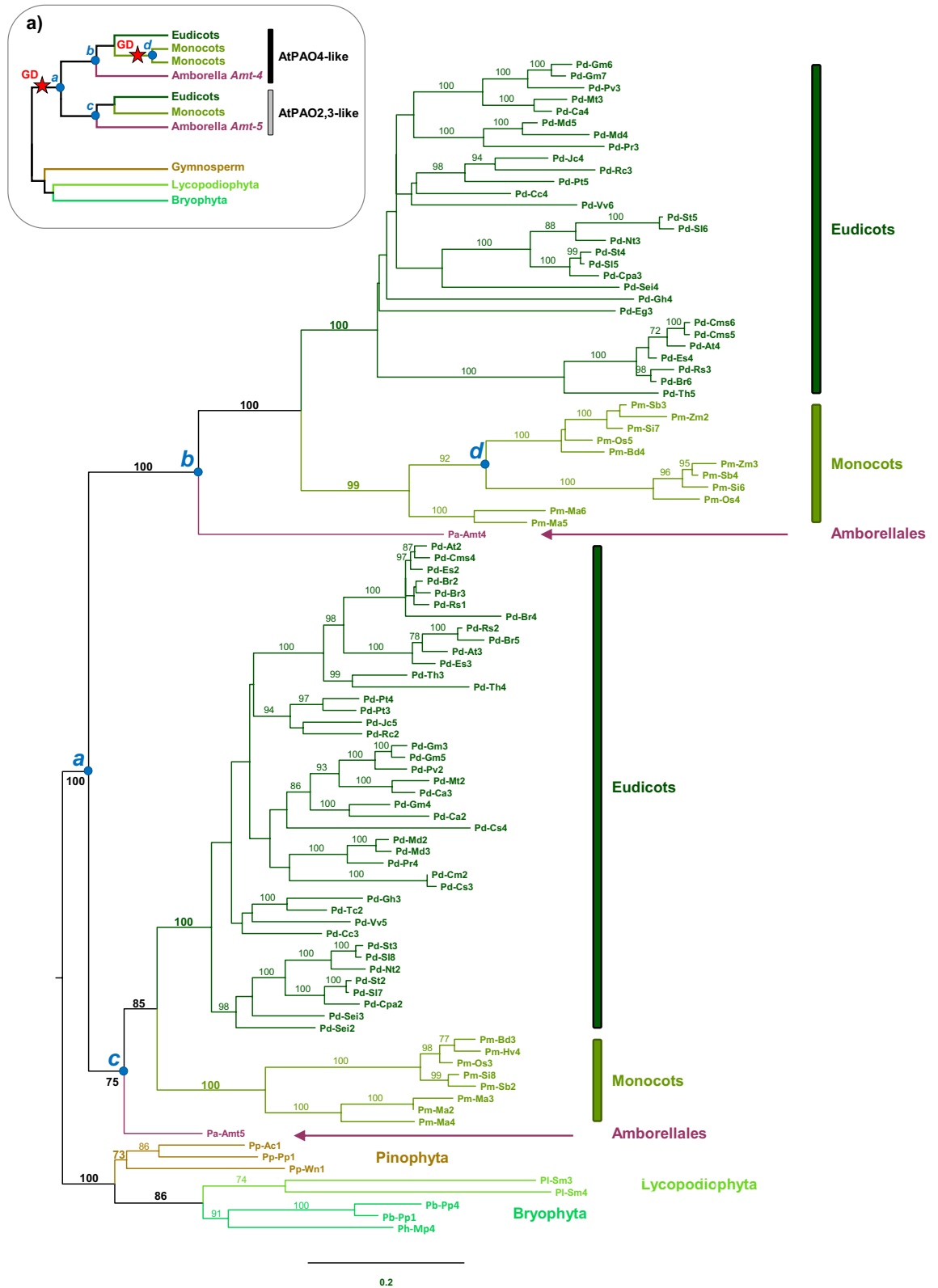


Figure 5. Maximum likelihood phylogeny of plant PAOs of the AtPAO2,3,4-like clade (see Fig. 2) based on 100 amino acidic sequences. Main clades are labelled from *a* to *c*; bootstrap values of support (BS) are reported along the branches (BS > 70). Box (**a**) illustrates the evolutionary model for the AtPAO2,3,4-like proteins with two gene duplications (GD) in correspondence of node *a* (giving rise to the AtPAO2,3 and AtPAO4 subclades) and *d*.

Amt2 and Pa-Amt3), and the dicot *Vitis vinifera* (Pd-Vv1, Pd-Vv2, Pd-Vv3), but highly varies (Asn, Ile; Phe, His) in the extracellular PAOs present in some other dicots (node *d*, Fig. 3). It is also present in several fungal PAOs of Clade I, while it is substituted by an Ala residue in all AtPAO1-like PAOs and by a His residue in most of the other PAOs clades (Clades II, III, and IV, and stramenopile clades). Glu170 is also well conserved in PAO-like sequences of the various clades (including the bacterial and archaeal proteins of Clade IV, and of those of the stramenopile clades), with the exception of the land plant AtPAO5-like PAOs in which it is substituted by Gln. Furthermore, in some PAO-like sequences ZmPAO1 Glu170 residue is not well-defined (Table 1, ND) due to the presence of gaps and/or regions of low sequence homology. Ser402, which in the Mm-MmMAO-A and Mm MmMAO-B is substituted by a Cys residue involved in covalent binding to the isoalloxazine ring of the FAD, is also highly conserved (Table 1). Only the AtPAO4-like PAOs of Clade IV have a Cys residues at this position which, however, is not involved in covalent binding of the FAD³⁰. Unlike the other residues of the ZmPAO1 catalytic site, Tyr298 highly varies across the PAOs. In particular, while an aromatic residue is present at this position in the PAOs of Clade I, it is substituted by a Thr residue in vertebrate SMOXs, by a Val residue in the plant and algal PAOs of Clade II and some invertebrate PAOs (such as insect PAOs and the two PAOs from the amphioxus *Branchiostoma floridae*), and by a Asn residue in all vertebrate PAOXs⁶⁷. Further studies are necessary to understand whether these variations in amino acid residues of the PAO catalytic sites correlate to variations in substrate specificity.

Conclusions

The tree of life of polyamine oxidases suggests a common origin for archaeal PAO-like proteins and eukaryotic PAOs, which probably also involved the evolution of monoamine oxidases. Within eukaryotes, four main clades of PAOs were identified, likely originated from an ancestral eukaryotic PAO before the split of the main supergroups and followed by specific gene losses in each supergroup. As a result, while some eukaryotes present a high diversity of PAO isoforms belonging to multiple clades (e.g. land plants and stramenopiles), some others have PAOs belonging to one (animals) or two clades (e.g. fungi and green algae). Within each of these clades, phylogenetic patterns revealed that PAOs have undergone several diversification events. Evolution of Clade I and Clade IV is shaped by multiple gene duplications. Conversely, only a few gene duplication events occurred within Clade II and Clade III. Clade I PAOs have additionally experienced peptide deletion leading to functional changes and diversification in subcellular localization. The latter has been a pervasive process along eukaryotic PAO evolution, most organisms having PAOs in two or three different subcellular compartments (extracellular space, cytosol and peroxisomes), which suggests different physiological roles. The large variety of PAOs analysed in the present study may facilitate structure–function studies.

Methods

Protein sequence homology search and retrieval. The amino acid sequence of PAOs were retrieved by sequence similarity searches using BLASTP⁶⁸ (NCBI, Uniprot and Phytozome databases) and TBLASTN (NCBI TSA database). As query sequences the amino acid sequence of the following PAOs, for which enzymatic activity had been previously verified, were used: *Arabidopsis thaliana* AtPAO1, AtPAO2, AtPAO3, AtPAO4, AtPAO5^{9,27,30}, *Zea mays* ZmPAO1³⁵, *Pseudomonas aeruginosa* SpdH⁴⁰, *Saccharomyces cerevisiae* FMS1²⁵, *Ustilago maydis* UmPAO², *Mus musculus* SMOX²⁴ and *M. musculus* PAOX²³. Following an initial search on the entire databases, several protist lineages were not represented in the dataset. To further assess the presence of PAO-like proteins in these lineages, we repeated the search using the same query sequences and specifically targeting, for each eukaryotic super-group, those species for which genomic resources were available (Supplementary Table S1). Among retrieved sequences, we selected those having a sequence identity with the query sequence $\geq 20\%$, a coverage $\geq 60\%$ and an E-value $\leq 1e^{-6}$. Selected sequences were further validated based on sequence length (selecting sequences in the range of 400–650 amino acids), annotation of protein function, and the presence of particular domains. FAD-dependent PAO-like sequences with SWIRM domains, which are involved in histone oxidative demethylation⁶⁹ rather than in polyamine metabolism, were excluded.

Sequence Analysis. Subcellular localization was inferred based on amino acid sequences using PSORT and SignalP. Genomic exon–intron structure comparison was performed by means of alignment between genomic and cDNA sequences. Amino acid residues of catalytic site were retrieved by multiple sequence alignments performed using Clustal Omega 1.2.1⁷⁰ and based on ZmPAO1 and Fms1 crystal structure^{26,58,65,66}.

Phylogenetic analysis. Multiple amino acid sequence alignments were performed using Clustal Omega 1.2.1. On large data sets, Clustal Omega outperforms other packages in terms of execution time and quality⁷⁰. Multiple sequence alignments were not trimmed. Phylogenetic analyses of the amino acid sequences were performed using the Maximum Likelihood (ML) method on five distinct datasets for a total of 428 PAO-like sequences (see Supplementary Table S1). Multiple sequence alignments and phylogenetic trees are provided in Supplementary Data. The first dataset included 300 sequences of Bacteria (37), Archaea (16), and Eukaryotes (247). Subsequently, ML analyses were performed on the monophyletic group including all eukaryotic PAOs and a few prokaryotic PAOs (253 sequences) based on a new alignment. Additionally, to increase the taxonomic representation, we built, through additional similarity searches, three extended datasets of land plant PAOs for each of the AtPAO1-like (Clade I), AtPAO5-like (Clade II), and AtPAO2,3,4-like (Clade IV) clades, and we made new alignments for each of these clades. Phylogenetic trees were rooted using the midpoint method, which is a valuable method when a proper outgroup is not available or difficult to identify⁷¹. Additionally, we tested the root position using the Minimal Ancestor Deviation (MAD) method, that has been shown to outperform

existing methods⁷². Details on numbers of taxa, sites and informative sites are reported for each alignment in Supplementary Table S3.

For each dataset the best-fit model of amino acid replacement was selected by ModelTest-NG 0.1.5⁷³, using an optimize Maximum-Likelihood topology and branch lengths for each model (*-t ml*) and the Akaike Information Criterion (AIC). The WAG model⁷⁴ with gamma distributed rates across site (+G) and a proportion of invariant sites (+I) was selected for the eukaryote dataset, and the JTT model⁷⁵ with gamma distributed rates and a proportion of invariant sites (+I) was selected for each plant PAO dataset. ML tree searches were performed with IQ-tree⁷⁶ (for dataset larger than 200 sequences) and PhyML 3.0⁷⁷ (for dataset smaller than 100 sequences) using the best-fit model and 100 random starting trees. Node support for the resulting phylogenetic tree was evaluated by 1000 bootstrap replicates in IQ-tree (using both the ultrafast bootstrap approximation and the SH-like approximate likelihood ratio test) and by 100 bootstrap replicates in PhyML. Phylogenetic analyses were carried out on the T-REX webserver⁷⁸ and the CIPRES Science Gateway 3.3⁷⁹ (at <https://www.phylo.org/>).

Code availability

GenBank accession numbers of all sequences used in this study are reported in Supplementary Table S1.

Received: 6 February 2020; Accepted: 29 September 2020

Published online: 20 October 2020

References

- Hamana, K. & Matsuzaki, S. Widespread occurrence of norspermidine and norspermine in eukaryotic algae. *J. Biochem.* **91**, 1321–1328 (1982).
- Valdés-Santiago, L., Guzmán-de-Peña, D. & Ruiz-Herrera, J. Life without putrescine: disruption of the gene-encoding polyamine oxidase in *Ustilago maydis* odc mutants. *FEMS Yeast Res.* **10**, 928–940 (2010).
- Tavladoraki, P. *et al.* Polyamine catabolism: target for antiproliferative therapies in animals and stress tolerance strategies in plants. *Amino Acids* **42**, 411–426 (2012).
- Hamana, K., Niitsu, M., Samejima, K. & Matsuzaki, S. Linear and branched pentaamines, hexaamines and heptaamines in seeds of *Vicia sativa*. *Phytochemistry* **30**, 3319–3322 (1991).
- Casero, R. A. Jr. & Pegg, A. E. Spermidine/spermine N¹-acetyltransferase—the turning point in polyamine metabolism. *FASEB J.* **7**, 653–661 (1993).
- Grienenberger, E. *et al.* A BAHF acyltransferase is expressed in the tapetum of Arabidopsis anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. *Plant J.* **58**, 246–259 (2009).
- Morimoto, N. *et al.* Dual biosynthesis pathway for longer-chain polyamines in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *J. Bacteriol.* **192**, 4991–5001 (2010).
- Michael, A. J. Molecular machines encoded by bacterially-derived multi-domain gene fusions that potentially synthesize, N-methylate and transfer long chain polyamines in diatoms. *FEBS Lett.* **585**, 2627–2634 (2011).
- Ahou, A. *et al.* A plant spermine oxidase/dehydrogenase regulated by the proteasome and polyamines. *J. Exp. Bot.* **65**, 1585–1603 (2014).
- Illingworth, C. *et al.* The diverse bacterial origins of the Arabidopsis polyamine biosynthetic pathway. *FEBS Lett.* **549**, 26–30 (2003).
- Ober, D. & Hartmann, T. Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. *Proc. Natl. Acad. Sci. USA* **96**, 14777–14782 (1999).
- Ohnuma, M. *et al.* N¹-aminopropylagmatine, a new polyamine produced as a key intermediate in polyamine biosynthesis of an extreme thermophile, *Thermus thermophilus*. *J. Biol. Chem.* **280**, 30073–30082 (2005).
- Cacciapuoti, G. *et al.* The first agmatine/cadaverine aminopropyl transferase: biochemical and structural characterization of an enzyme involved in polyamine biosynthesis in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **189**, 6057–6067 (2007).
- Knott, J. M. Biosynthesis of long-chain polyamines by crenarchaeal polyamine synthases from *Hyperthermus butylicus* and *Pyrobaculum aerophilum*. *FEBS Lett.* **583**, 3519–3524 (2009).
- Okada, K. *et al.* Identification of a novel aminopropyltransferase involved in the synthesis of branched-chain polyamines in hyperthermophiles. *J. Bacteriol.* **196**, 1866–1876 (2014).
- Minguet, E. G. *et al.* Evolutionary diversification in polyamine biosynthesis. *Mol. Biol. Evol.* **25**, 2119–2128 (2008).
- Fuell, C., Elliott, K. A., Hanfrey, C. C., Franceschetti, M. & Michael, A. J. Polyamine biosynthetic diversity in plants and algae. *Plant Physiol. Biochem.* **48**, 513–520 (2010).
- Solé-Gil, A., Hernández-García, J., López-Gresa, M. P., Blázquez, M. A. & Agustí, J. Conservation of thermospermine synthase activity in vascular and non-vascular plants. *Front. Plant Sci.* **10**, 663. <https://doi.org/10.3389/fpls.2019.00663> (2019).
- Wang, Y. *et al.* Cloning and characterization of a human polyamine oxidase that is inducible by polyamine analogue exposure. *Cancer Res.* **61**, 5370–5373 (2001).
- Wu, T., Yankovskaya, V. & McIntire, W. S. Cloning, sequencing, and heterologous expression of the murine peroxisomal flavoprotein, N¹-acetylated polyamine oxidase. *J. Biol. Chem.* **278**, 20514–20525 (2003).
- Cona, A., Rea, G., Angelini, R., Federico, R. & Tavladoraki, P. Functions of amine oxidases in plant development and defence. *Trends Plant Sci.* **11**, 80–88 (2006).
- Vujcic, S., Liang, P., Diegelman, P., Kramer, D. L. & Porter, C. W. Genomic identification and biochemical characterization of the mammalian polyamine oxidase involved in polyamine back-conversion. *Biochem. J.* **370**, 19–28 (2003).
- Vujcic, S., Diegelman, P., Bacchi, C. J., Kramer, D. L. & Porter, C. W. Identification and characterization of a novel flavin-containing spermine oxidase of mammalian cell origin. *Biochem. J.* **367**, 665–675 (2002).
- Cervelli, M., Polticelli, F., Federico, R. & Mariottini, P. Heterologous expression and characterization of mouse spermine oxidase. *J. Biol. Chem.* **278**, 5271–5276 (2003).
- Landry, J. & Sternglanz, R. Yeast Fms1 is a FAD-utilizing polyamine oxidase. *Biochem. Biophys. Res. Commun.* **303**, 771–776 (2003).
- Huang, Q., Liu, Q. & Hao, Q. Crystal structures of Fms1 and its complex with spermine reveal substrate specificity. *J. Mol. Biol.* **348**, 951–959 (2005).
- Tavladoraki, P. *et al.* Heterologous expression and biochemical characterization of a polyamine oxidase from Arabidopsis involved in polyamine back conversion. *Plant Physiol.* **141**, 1519–1532 (2006).
- Moschou, P. N. *et al.* Bridging the gap between plant and mammalian polyamine catabolism: a novel peroxisomal polyamine oxidase responsible for a full back-conversion pathway in Arabidopsis. *Plant Physiol.* **147**, 1845–1857 (2008).
- Kamada-Nobusada, T., Hayashi, M., Fukazawa, M., Sakakibara, H. & Nishimura, M. A putative peroxisomal polyamine oxidase, AtPAO4, is involved in polyamine catabolism in *Arabidopsis thaliana*. *Plant Cell Physiol.* **49**, 1272–1282 (2008).
- Fincato, P. *et al.* Functional diversity inside the Arabidopsis polyamine oxidase gene family. *J. Exp. Bot.* **62**, 1155–1168 (2011).

31. Ono, Y. *et al.* Constitutively and highly expressed *Oryza sativa* polyamine oxidases localize in peroxisomes and catalyze polyamine back conversion. *Amino Acids* **42**, 867–876 (2012).
32. Tavladoraki, P., Cona, A. & Angelini, R. Copper-containing amine oxidases and FAD-dependent polyamine oxidases are key players in plant tissue differentiation and organ development. *Front. Plant Sci.* **7**, 824. <https://doi.org/10.3389/fpls.2016.00824> (2016).
33. Kim, D. W. *et al.* Polyamine oxidase 5 regulates *Arabidopsis thaliana* growth through a thermospermine oxidase activity. *Plant Physiol.* **165**, 1575–1590 (2014).
34. Liu, T., Kim, D. W., Niitsu, M., Berberich, T. & Kusano, T. *Oryza sativa* polyamine oxidase 1 back-converts tetraamines, spermine and thermospermine, to spermidine. *Plant Cell Rep.* **33**, 143–151 (2014).
35. Tavladoraki, P. *et al.* Maize polyamine oxidase: primary structure from protein and cDNA sequencing. *FEBS Lett.* **426**, 62–66 (1998).
36. Polticelli, F. *et al.* Lys300 plays a major role in the catalytic mechanism of maize polyamine oxidase. *Biochemistry* **44**, 16108–16120 (2005).
37. Liu, T. *et al.* Polyamine oxidase 7 is a terminal catabolism-type enzyme in *Oryza sativa* and is specifically expressed in anthers. *Plant Cell Physiol.* **55**, 1110–1122 (2014).
38. Tabor, C. W. & Kellogg, P. D. Identification of flavin dinucleotide and heme in a homogeneous spermidine dehydrogenase from *Serratia marcescens*. *J. Biol. Chem.* **245**, 5424–5433 (1970).
39. Hisano, T. *et al.* Characterization of membrane-bound spermidine dehydrogenase of *Citrobacter freundii*. *Biosci. Biotechnol. Biochem.* **56**, 1916–1920 (1992).
40. Dasu, V. V., Nakada, Y., Ohnishi-Kameyama, M., Kimura, K. & Itoh, Y. Characterization and a role of *Pseudomonas aeruginosa* spermidine dehydrogenase in polyamine catabolism. *Microbiology* **152**, 2265–2272 (2006).
41. Okada, M., Kawashima, S. & Imahori, K. Substrate binding characteristics of the active site of spermidine dehydrogenase from *Serratia marcescens*. *J. Biochem.* **85**, 1235–1243 (1979).
42. van Hellemond, E. W., van Dijk, M., Heuts, D. P., Janssen, D. B. & Fraaije, M. W. Discovery and characterization of a putrescine oxidase from *Rhodococcus erythropolis* NCIMB 11540. *Appl. Microbiol. Biotechnol.* **78**, 455–463 (2008).
43. Kopacz, M. M., Rovida, S., van Duijn, E., Fraaije, M. W. & Mattevi, A. Structure-based redesign of cofactor binding in putrescine oxidase. *Biochemistry* **50**, 4209–4217 (2011).
44. Foster, A., Barnes, N., Speight, R. & Keane, M. A. Genomic organisation, activity and distribution analysis of the microbial putrescine oxidase degradation pathway. *Syst. Appl. Microbiol.* **36**, 457–466 (2013).
45. Kopacz, M. M., Heuts, D. P. & Fraaije, M. W. Kinetic mechanism of putrescine oxidase from *Rhodococcus erythropolis*. *FEBS J.* **281**, 4384–4393 (2014).
46. Michael, A. J. Polyamines in eukaryotes, bacteria, and archaea. *J. Biol. Chem.* **291**, 14896–14903 (2016).
47. Michael, A. J. Polyamine function in archaea and bacteria. *J. Biol. Chem.* **293**, 18693–18701 (2018).
48. Shaw, F. L. *et al.* Evolution and multifarious horizontal transfer of an alternative biosynthetic pathway for the alternative polyamine sym-homospermidine. *J. Biol. Chem.* **285**, 14711–14723 (2010).
49. Kitashiba, H., Honda, C. & Moriguchi, T. Identification of polyamine oxidase genes from apple and expression analysis during fruit development and cell growth. *Plant Biotechnol.* **23**, 425–429 (2006).
50. Takahashi, Y. *et al.* Characterization of five polyamine oxidase isoforms in *Arabidopsis thaliana*. *Plant Cell Rep.* **29**, 955–965 (2010).
51. Wang, W. & Liu, J. Genome-wide identification and expression analysis of the polyamine oxidase gene family in sweet orange (*Citrus sinensis*). *Gene* **555**, 421–429 (2015).
52. Polticelli, F. *et al.* Molecular evolution of the polyamine oxidase gene family in Metazoa. *BMC Evol. Biol.* **12**, 90. <https://doi.org/10.1186/1471-2148-12-90> (2012).
53. Cervelli, M., Salvi, D., Polticelli, F., Amendola, R. & Mariottini, P. Structure–function relationships in the evolutionary framework of spermine oxidase. *J. Mol. Evol.* **76**, 365–370 (2013).
54. Bordenave, C. D., Granados Mendoza, C., Jiménez Bremont, J. F., Gárriz, A. & Rodríguez, A. A. Defining novel plant polyamine oxidase subfamilies through molecular modeling and sequence analysis. *BMC Evol. Biol.* **19**, 28. <https://doi.org/10.1186/s12862-019-1361-z> (2019).
55. Samasil, K., Lopes de Carvalho, L., Mäenpää, P., Salminen, T. A. & Incharoensakdi, A. Biochemical characterization and homology modelling of polyamine oxidase from cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol. Biochem.* **119**, 159–169 (2017).
56. Edmondson, D. E., Binda, C., Wang, J., Upadhyay, A. K. & Mattevi, A. Molecular and mechanistic properties of the membrane-bound mitochondrial monoamine oxidases. *Biochemistry* **48**, 4220–4230 (2009).
57. Gribaldo, S., Poole, A. M., Daubin, V., Forterre, P. & Brochier-Armanet, C. The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse?. *Nat. Rev. Microbiol.* **8**, 743–752 (2010).
58. Binda, C. *et al.* A 30-angstrom-long U-shaped catalytic tunnel in the crystal structure of polyamine oxidase. *Structure* **7**, 265–276 (1999).
59. Binda, C., Mattevi, A. & Edmondson, D. E. Structure-function relationships in flavoenzyme-dependent amine oxidations: a comparison of polyamine oxidase and monoamine oxidase. *J. Biol. Chem.* **277**, 23973–23976 (2002).
60. Ochman, H., Lawrence, J. G. & Groisman, E. A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304 (2000).
61. Wang, H., Liu, B., Li, H. & Zhang, S. Identification and biochemical characterization of polyamine oxidases in amphioxus: Implications for emergence of vertebrate-specific spermine and acetyl polyamine oxidases. *Gene* **575**, 429–437 (2016).
62. Liu, T., Wook Kim, D., Niitsu, M., Berberich, T. & Kusano, T. Polyamine oxidase 1 from rice (*Oryza sativa*) is a functional ortholog of Arabidopsis Polyamine oxidase 5. *Plant Signal. Behav.* **9**, 29773 (2014).
63. Sagor, G. H. *et al.* The polyamine oxidase from lycophyte *Selaginella lepidophylla* (SelPAO5), unlike that of angiosperms, back-converts thermospermine to norspermidine. *FEBS Lett.* **589**, 3071–3078 (2015).
64. Guerrero-González, M. L., Rodríguez-Kessler, M. & Jiménez-Bremont, J. F. uORF, a regulatory mechanism of the Arabidopsis polyamine oxidase 2. *Mol. Biol. Rep.* **41**, 2427–2443 (2014).
65. Binda, C., Angelini, R., Federico, R., Ascenzi, P. & Mattevi, A. Structural bases for inhibitor binding and catalysis in polyamine oxidase. *Biochemistry* **40**, 2766–2776 (2001).
66. Fiorillo, A. *et al.* The structure of maize polyamine oxidase K300M mutant in complex with the natural substrates provides a snapshot of the catalytic mechanism of polyamine oxidation. *FEBS J.* **278**, 809–821 (2011).
67. Sjögren, T. *et al.* The structure of murine N¹-acetylspermine oxidase reveals molecular details of vertebrate polyamine catabolism. *Biochemistry* **56**, 458–467 (2017).
68. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
69. Spedaletti, V. *et al.* Characterization of a lysine-specific histone demethylase from *Arabidopsis thaliana*. *Biochemistry* **47**, 4936–4947 (2008).
70. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539. <https://doi.org/10.1038/msb.2011.75> (2011).
71. Hess, P. N. & De Moraes Russo, C. A. An empirical test of the midpoint rooting method. *Biol. J. Linn. Soc.* **92**, 669–674 (2007).
72. Tria, F., Landan, G. & Dagan, T. Phylogenetic rooting using minimal ancestor deviation. *Nat. Ecol. Evol.* **1**, 0193 (2017).

73. Darriba, D. et al. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. bioRxiv 612903; <https://doi.org/10.1101/612903> (2019).
74. Whelan, S. & Goldman, N. A. General empirical model of protein evolution derived from multiple protein families using a Maximum-Likelihood approach. *Mol. Biol. Evol.* **18**, 691–699 (2001).
75. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 (1992).
76. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–327 (2015).
77. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
78. Boc, A., Diallo, A. B. & Makarenkov, V. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res.* **40**(W1), W573–W579 (2012).
79. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. In: SC10 Workshop on Gateway Computing Environments (GCE10) (2010).

Acknowledgements

We thank Anthony Michael (UT Southwestern) for useful suggestions. This work was supported by the Italian Ministry of Education, University and Research (Grant to Department of Science, University ‘Roma Tre’ ‘Dipartimenti di Eccellenza’, ARTICOLO 1, COMMI 314–337. LEGGE 423 232/2016; PRIN 2017– CUP F84I19000730005), and University ‘Roma Tre’.

Author contributions

D.S. and P.T. jointly conceived the study, performed the analyses, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-74708-3>.

Correspondence and requests for materials should be addressed to P.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020