Available online at www.sciencedirect.com

ScienceDirect

Genomics 89 (2007) 567–579

GENOMICS

www.elsevier.com/locate/ygeno

Prokaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes

Filippo Passardi [a], Nenad Bakalovic [a], Felipe Karam Teixeira [b], Marcia Margis-Pinheiro [b,c], Claude Penel [a], Christophe Dunand [a,*]

[a] *Laboratory of Plant Physiology, University of Geneva, Quai Ernest-Ansermet 30, CH-1211 Geneva 4, Switzerland*
[b] *Department of Genetics, Institute of Biology, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil*
[c] *Department of Genetics, Federal University of Rio Grande do Sul, Rio Grande do Sul, Brazil*

Received 16 June 2006; accepted 18 January 2007
Available online 13 March 2007

Abstract

Members of the superfamily of plant, fungal, and bacterial peroxidases are known to be present in a wide variety of living organisms. Extensive searching within sequencing projects identified organisms containing sequences of this superfamily. Class I peroxidases, cytochrome *c* peroxidase (CcP), ascorbate peroxidase (APx), and catalase peroxidase (CP), are known to be present in bacteria, fungi, and plants, but have now been found in various protists. CcP sequences were detected in most mitochondria-possessing organisms except for green plants, which possess only ascorbate peroxidases. APx sequences had previously been observed only in green plants but were also found in chloroplastic protists, which acquired chloroplasts by secondary endosymbiosis. CP sequences that are known to be present in prokaryotes and in Ascomycetes were also detected in some Basidiomycetes and occasionally in some protists. Class II peroxidases are involved in lignin biodegradation and are found only in the Homobasidiomycetes. In fact class II peroxidases were identified in only three orders, although degenerate forms were found in different Pezizomycota orders. Class III peroxidases are specific for higher plants, and their evolution is thought to be related to the emergence of the land plants. We have found, however, that class III peroxidases are present in some green algae, which predate land colonization. The presence of peroxidases in all major phyla (except vertebrates) makes them powerful marker genes for understanding the early evolutionary events that led to the appearance of the ancestors of each eukaryotic group.

Keywords: Peroxidase superfamily; Gene transfer; Evolution; Organelle-mediated transmission

Oxygen is essential for most living organisms but the generation of dioxygen compounds termed reactive oxygen species (ROS) can be toxic. To counteract these molecules, aerobes possess a large panel of proteins, which include the heme-containing enzymes peroxidases. The term "peroxidases" is used to describe a very large group of enzymes that reduce peroxide and oxidize a wide variety of substrates, such as lignin subunits, lipid membranes, and some amino acid side chains. According to Karen Welinder [1,2], the "superfamily of plant, fungal and bacterial peroxidases" includes three classes of structurally related peroxidases. The common features of the peroxidases belonging to this superfamily are that they all possess a heme moiety (ferriprotoporphyrin IX), a conserved histidine as a proximal ligand, and conserved arginine and histidine residues as distal ligands. Through the common peroxidative cycle, heme-containing peroxidases use H_2O_2 as an acceptor and various substrates as donors, which will be converted into radicals. In this cycle, the iron in the peroxidase heme group is converted from Fe(III) to Fe(IV).

Class I peroxidases have been found in plants, fungi, and prokaryotes [3]. They are not glycosylated and do not have signal peptides, calcium ions, or disulfide bridges. Three EC numbers define this class: EC 1.11.1.5 (ferrocytochrome-*c*: hydrogen-peroxide oxidoreductase), EC 1.11.1.6 (hydrogen-

Abbreviations: APx, ascorbate peroxidase; CcP, cytochrome *c* peroxidase; CP, catalase peroxidase; LiP, lignin peroxidase; MnP, manganese peroxidase; VP, versatile peroxidase.

* Corresponding author. Fax: +41223793017.
E-mail address: christophe.dunand@bioveg.unige.ch (C. Dunand).

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2007.01.006

peroxide:hydrogen-peroxide oxidoreductase, or more commonly catalase peroxidase), and EC 1.11.1.11 (L-ascorbate: hydrogen-peroxide oxidoreductase).

Cytochrome c peroxidases (CcP) play a major role scavenging H_2O_2 generated during aerobic respiration from the oxidation of cytochrome *c* in the mitochondrial intermembrane space. Despite this crucial function, CcP are absent in yeast without impeding cell viability or respiration [4]. CcP sequences are detectable in both Viridiplantae and fungi [5], but not in bacteria that possess diheme CcP unrelated to the peroxidase superfamily.

Catalase peroxidases (CP) are present in prokaryotes and fungi and typically have a dual catalytic activity, acting both as a catalase and as a peroxidase. Evolutionarily, CP have the unique feature of containing two peroxidase-like domains, an N- and a C-terminal domain. They probably evolved by the fusion of two copies of a primordial peroxidase gene (which probably also gave rise to the ascorbate and cytochrome *c* peroxidases) [3]. Only the catalase peroxidase N-terminal domain contains heme and is fully functional, whereas the C-terminal domain has no known catalytic activity [6]. CP are found mainly in prokaryotes but they have also been shown in some members of the Ascomycetes, probably following lateral gene transfer [3]. Although they share the same EC number, catalase peroxidases and monofunctional catalases have no sequence similarity, which suggests a different evolutionary origin [7].

Ascorbate peroxidases (APx) are found in chloroplastic organisms and putatively in cyanobacteria [8]. They show a particularly strong specificity for the electron donor ascorbate. They play a key role in the elimination of toxic amounts of intracellular H_2O_2 [9]. In higher plants, APx are separated, according to their cellular localization, into cytosolic, peroxisomal, and chloroplastic [10]. Algae also possess a subgroup of APx localized in chloroplasts whose sequences are more similar to the chloroplastic APx subgroup than to the cytosolic one [3,11]. Finally, putative APx sequences are also present in Euglenozoa, described as either cytoplasmic or as an endoplasmic reticulum-bound form [12,13].

Class II peroxidases are secreted by a particular type of fungi, the Homobasidiomycetes, to degrade plant-derived lignin. This class of peroxidases consists mainly of the *lignin peroxidases* (LiP; EC 1.11.1.14) and the *manganese peroxidases* (MnP; EC 1.11.1.13). Class II peroxidases (unlike Class I) are glycosylated and contain calcium ions and disulfide bridges as well as a peptide signal directing the protein to the endoplasmic reticulum for secretion. LiP and MnP are distinct from the class I and III peroxidases in their ability to oxidize molecules with high redox potentials such as lignin. LiP and MnP are essential globally for the degradation of dead plant material and soil recycling. LiP have a wide variety of substrates and participate in aromatic ring-opening reactions and side-chain cleavages that lead to degradation of lignin [14,15]. Both LiP and MnP contain a ferric heme group and work through the same mechanism during the peroxidative cycle. However, they radically differ in their mode of action: whereas LiP attacks lignin bonds directly, MnP takes an electron from Mn(II), resulting in Mn(III) that diffuses from the enzyme surface and attacks lignin [14]. Another group of class II peroxidases, named *versatile peroxidases* (VP; no defined EC number exists) [16–18], encompasses multifunctional peroxidases sharing the catalytic properties of LiP and MnP.

Class III peroxidases or the secreted plant peroxidases (EC 1.11.1.7) are found only in plants, where they form large multigenic families [19,20]. Although their primary sequence differs in some points from the two other classes of peroxidases, their three-dimensional structures are very similar to those of class II, and they also possess calcium ions, disulfide bonds, and an N-terminal signal for secretion. Some class III peroxidases have an additional C-terminal extension that is suspected to direct the protein to the vacuole [21,22]. Class III peroxidases probably appeared in algae just before land colonization but no sequences exist from this time to confirm the hypothesis. Class III peroxidases possibly played a critical role during this major evolutionary event, either by allowing the formation of rigid plant structures or by adapting the organisms to a more oxygenated environment [19,23]. Class III peroxidases are additionally able to undertake a second cyclic reaction, called "hydroxylic," which is distinct from the peroxidative one [24,25]. During the hydroxylic cycle, peroxidases pass through a Fe(II) state and use mainly the superoxide anion ($O_2^{\bullet -}$) to generate hydroxyl radicals ($^{\bullet}OH$). Class III peroxidases, by using both these cycles, are known to participate in many different plant processes from germination to senescence, for example, auxin metabolism, cell wall elongation and stiffening, or protection against pathogens [19,26]. The precise in vivo role of a single peroxidase has, however, not been found yet, mainly because of the wide range of peroxidase substrates and the probable functional redundancy of these enzymes.

Since 1992, numerous publications have reported independent studies of the different heme peroxidase classes or the situation in a single species, but the exhaustive phylogeny of the complete family has never been studied. The availability of an increasing number of genomic and EST sequencing projects gave us the opportunity to search for peroxidase-encoding sequences in all the major phyla of living organisms. For each class, using phylogenetic tree analysis, we have gone back into the evolutionary record and searched for the emergence of the diverse peroxidase families from putative ancestral sequences. We have also tried to correlate the presence or absence of particular peroxidases with specific roles and with particular species.

Our global phylogenetic analysis of the whole family revealed that APx sequences are absent from cyanobacteria and specific to chloroplastic organisms (green plants and protists) and that CcP sequences are detected in all mitochondria-containing organisms except animals, parasites, and plants. CP distribution within prokaryotes is very heterogeneous and can follow the normal inheritance (vertical transfer) or be dependent on horizontal gene transfer. The presence of CP sequences in particular fungal groups and occasionally in various unicellular organisms is probably related to isolated lateral gene transfers. Class II peroxidases, initially thought to be restricted to Homobasidiomycetes, have also been found in Ascomycetes. Analysis of new sequences from green algae has shown that the origin of class III peroxidase can be restricted and localized to algal species present just before the emergence of the Embryophytes.

Results and discussion

Class I peroxidases

The largest class of the superfamily, they are found in all living organisms except animals. Ascorbate and cytochrome *c* peroxidases are closely related (at least 50% similarity between distantly related organisms such as *Rhodophyta* (CmeAPx01) and Ascomycota species (YlCcP03)) and the 3D structures of APx and CcP are also very similar [27]. APx were found in chloroplast-containing organisms, whereas CcP were found in most mitochondria-containing organisms. Class I sequences were absent from Diplomonads, Parabasalids, Apicomplexa, animals, and Amoebozoa (Fig. 1). Except for the land plants, which contain only APx, the chloroplastic organisms contain both APx and CcP sequences (Table 1 and Fig. 1). There are exceptions to this observation (symbolized with question marks in Fig. 1 and Table 1), which can be explained by the small size of the libraries. As there is a high degree of similarity between APx and CcP sequences, it is probable that both sequences evolved from the same ancestral sequence. In many cases, hybrid sequences containing APx- and CcP-specific motifs (APx/CcP column in

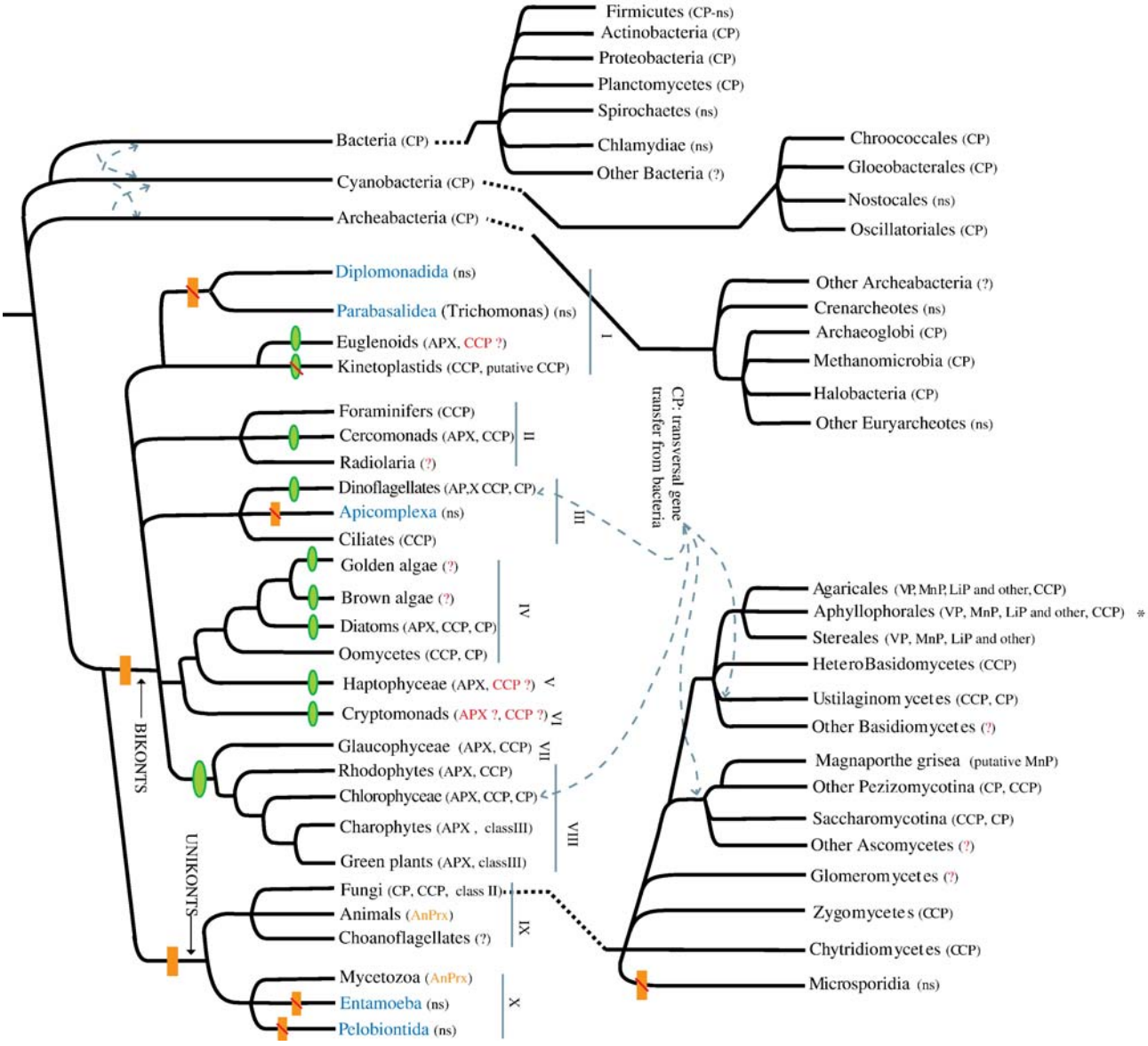

Fig. 1. Hypothetical phylogeny of the major phyla of living organisms. Endosymbiotic events that led to the appearance of mitochondria (orange rectangles) and chloroplasts (green ovals) are represented by bold lines. Acquisitions of chloroplast from a secondary endosymbiosis are represented with smaller icons. Losses of these organelles are represented with a red-slashed icon. Lateral gene transfers are represented with dashed lines. When at least one peroxidase-encoding gene was detected, the membership of the sequence is specified in parentheses for each phylum. APx, ascorbate peroxidase; AnPrx, animal peroxidase; CcP, cytochrome *c* peroxidase; CP, catalase peroxidase; LiP, lignin peroxidase; MnP, manganese peroxidase; VP, versatile peroxidase; "?", absence of sequence probably due to the low number of sequences available in the database; and ns, absence of sequence. When the presence of a particular class is expected but not detected, the class appears in red and is followed by "?". I, Excavata; II, Rhizaria; III, Alveolata; IV, Stramenopiles; V, Haptophyceae; VI, Cryptomonads; VII, Glaucophyceae; VIII, Viridiplantae; IX, Opisthokonts; X, Amoebozoa; *, Homobasidiomycetes. The phylum marked in blue stands for parasitic organisms. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
CcP and APx sequences compositions of chloroplastic organisms (other than green plants)

Families	Species		APx non chloro	APx/CcP	Ccp
Rhodophyta (ii)	*Chodrus crispus*	?	?	CcriAPx01	?
				CcriAPx02	
	Cyanidioschyzon merolae	CmeAPx01		CmeAPx02	
	Galdieria partita	?	?	GpAPx	?
	Galdieria sulphuraria		GsuAPx01	GsuAPx02	GsuCcP01
				GsuAPx03	
				GsuAPx04	
	Porphyra yezoensis	?	?	PyAPx	?
Diatoms (iv)	*Thalassiosira pseudonana*	TpsAPx02	TpsAPx01		TpsCcP01
				TpsAPx03	TpsCcP02
				TpsAPx05	TpsCcP03
					TpsCcP04
	Phaeodactylum triconutum	PtrAPx02		PtrAPx-CcP	PtrCcP01
		PtrAPx01			PtrCcP02
					PtrCcP04
					PtrCcP04
Cercomonads (iv)	*Bigelowiella natans*	BnaAPx01	BnaAPx02		BnaCcP02
Dinoflagellates (iii)	*Heterocapsa triquetra*	HtrAPx01		?	?
		HtrAPx02			
	Karenia brevis	?		?	CcP ? (C0064570)
	Amphidinium carterae	?		AcaAPx-CcP	?
	Alexandrium tamarense	AtamAPx		?	AtamCCP
Euglenoids (iii)	*Euglena gracilis*	?	EgrAPx01		?
Haptophyceae (ii)	*Prymnesium parvum*	?	PparAPx01		?
	Isochrysis galbana	?		IgAPx-CcP	?
Glaucocystophyceae (ii)	*Glaucocystis nostochinearum*	?	GnoAPx01		GnoCcP01
					GnoCcP02
	Cyanophora paradoxa	?		?	?
Chlorophyceae (ii)	*Acetabularia acetabularum*	AacAPx01		AacAOPx02	?
		AacAPx03			
	Chlamydomonas reinhardtii	CreAPx02	CreAPx01	CreAPx-CcP	
	Chlorella vulgaris	CvAPx01		?	?
	Dunaliella salina	DsaAPx01		?	?
	Scenedesmus obliquus	SobAPx01		?	SobCcP01
Cryptomonadaceae (iv)		?		?	?

?: Absence of detected sequence probably due to the small size of the data database. The number in brackets after the family name corresponds to the number of membranes surrounding the chloroplast. APx/CcP column: sequences that possess both APx and CcP motifs.

Table 1) were classified with difficulty using BLAST comparisons and phylogenetic alignments. As shown in Fig. 2 (neighbor-joining (NJ) tree), three representative sequences from protists (complete sequences: EgrAPx01, TpsAPx01, and PparAPx01) form a well-supported group that appears at the origin of the APx branch. Maximum likelihood (ML) phylogenetic tree analysis supports the position of this group (data not shown). The relatively low number of CcP-encoding sequences in Basidiomycetes (*Coprinus, Phanerochaete, Ustilago*, and *Cryptococcus*) was most likely due to the very low number of sequences currently available (Fig. 1). Curiously, CcP were absent from the completed genome of *Encephalitozoon cuniculi* [28]. This intracellular parasite belongs to the Microsporidia group, which is considered the most basal group of fungi or the closest sister group to fungi [29,30]. The absence of CcP in Microsporidia is more suggestive of the second hypothesis (sister group) rather than their being a fungal progenitor.

Only one or two CcP-encoding sequences can be found per organism, whereas multiple isoforms of APx were detected in most of the *Viridiplantae* (e.g., 6 in *Arabidopsis* and 8 in *Oryza*). The localization of APx can be predicted to be chloroplastic, peroxisomal, or cytoplasmic, as is clearly seen in the class I phylogenetic tree (Fig. 2). APx-encoding sequences from unicellular algae, *Rhodophyta* (CcriAPx01, CmeAPx01, GpAPx, PyAPx) and *Chlamydomonas* (CreAPx01 and CspAPx01), were at the base of the chloroplastic APx subbranch (Figs. 2 and 4). *Trypanosoma* and *Leishmania* sequences (LmAPx-CcP and TcrAPx-CcP), described as APx proteins [13,31], appear at the origin of the APx/CcP branch (Figs. 2 and 4). These two hybrid sequences share similar motifs with both CcP and APx sequences from various organisms. New recognition profiles have been used to distinguish in silico APx clearly from CcP sequences (unpublished data), but in the case of hybrid sequences, the profiles recognize them as members of both categories.

The presence of catalase peroxidase sequences was determined in the completed genomes of Archaea, Bacteria, and Ascomycota and for the few protists and plants available (Figs. 1 and 4). In bacteria and archaebacteria, the nonhomogeneous distribution of CP does not always follow a strict phylogenetic descent (i.e., vertical gene transfer) but can also be due to lateral gene transfer. CP are also intron-free in eukaryotic organisms, and the similarities observed between all the sequences are over 70%. These findings confirm the idea previously proposed by Zamocky [3] that the presence of CP

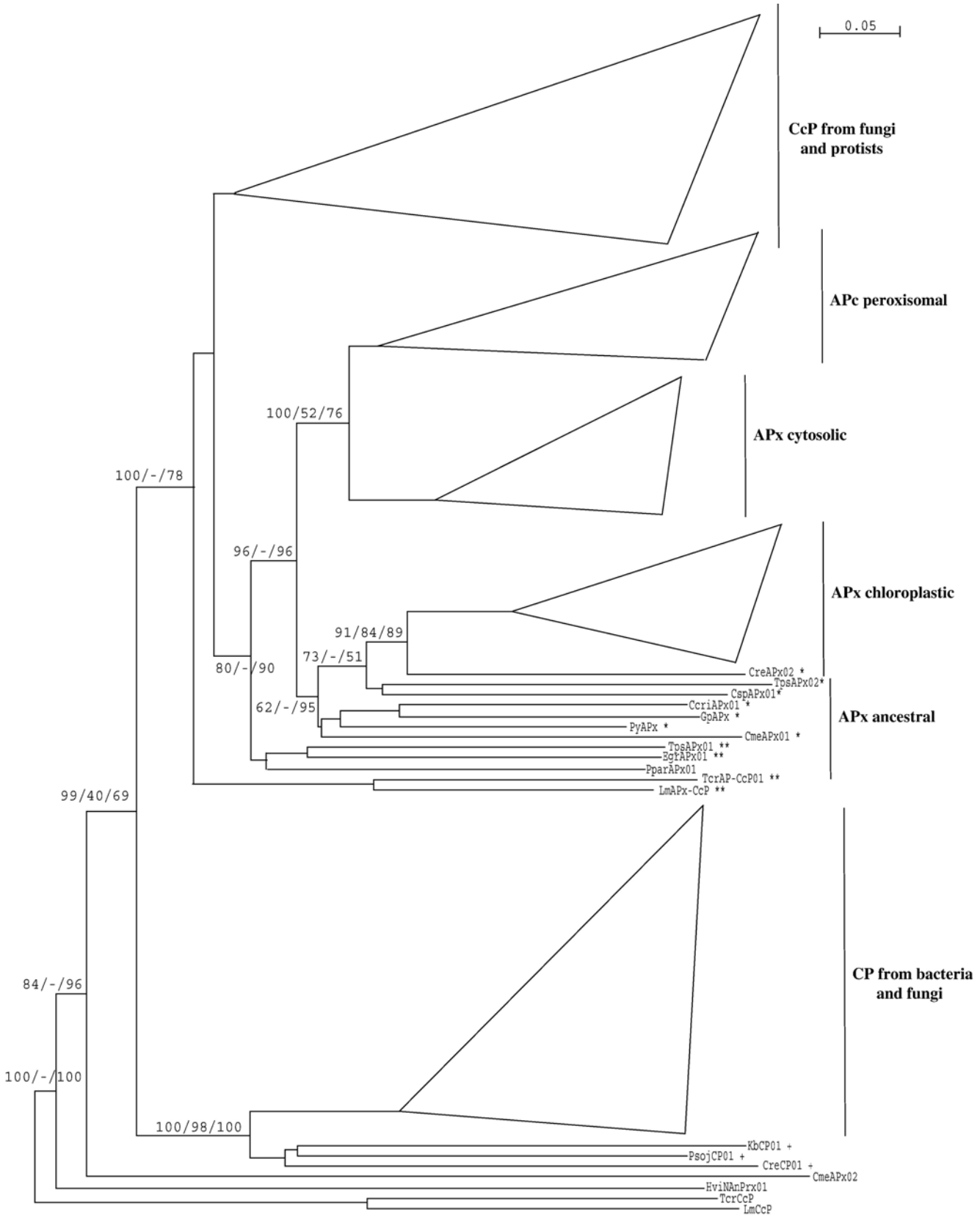

Fig. 2. Neighbor-joining tree of class I peroxidases based on predicted sequences. Only complete sequences were used to build the tree: 54 for APx, 22 for CcP, and 35 for CP. Numbers represent bootstrap values obtained from distance, MP, and ML methods. APx, ascorbate peroxidase; CcP, cytochrome *c* peroxidase; CP, catalase peroxidase; *, unicellular Viridiplantae; **, chloroplastic protists; +, CP sequences from other organisms than bacteria/fungi.

sequences in fungi is a consequence of bacterial gene transfer. When the N-terminal and C-terminal domains of CP are aligned separately with either CcP or APx, it can be seen that the N-terminal domain has preserved the critical hemoperoxidase residues responsible for catalytic activity [6], whereas they are absent from the C-terminal domain. Nevertheless, the C-terminal domain has many conserved residues, although its function is so far unknown. The N-terminal domain has also gained approximately 100 new residues that form large gaps when aligned to CcP and APx.

Class II peroxidases

After exhaustive data mining, we conclude that class II encoding sequences were found in only three saprophytic Homobasidiomycetes: Agaricales, Aphyllophorales, and Sterales. Four independent families of class II sequences can be detected in these orders: MnP, LiP, VP, and a fourth unnamed group (Fig. 3). The four families form well-defined clusters that are supported by phylogenetic analysis even when the sequences are closely related. The node separating MnP from the other main cluster (bootstrap "70/–/–") is not strongly supported by maximum parsimony (MP) and ML methods. However, on MP and ML trees, both the MnP and the "new ligninase" groups are well defined. Key residues related to the catalytic specificity [32] are well conserved within each family and new recognition profiles helped us to draw up hypotheses concerning the chronological evolution of the different families [20]. Several sequences from various fungi (TvMRP, GaMnP01, TceLiP01, ArMnP, and CcinCIP) that contain some of the major ligninase

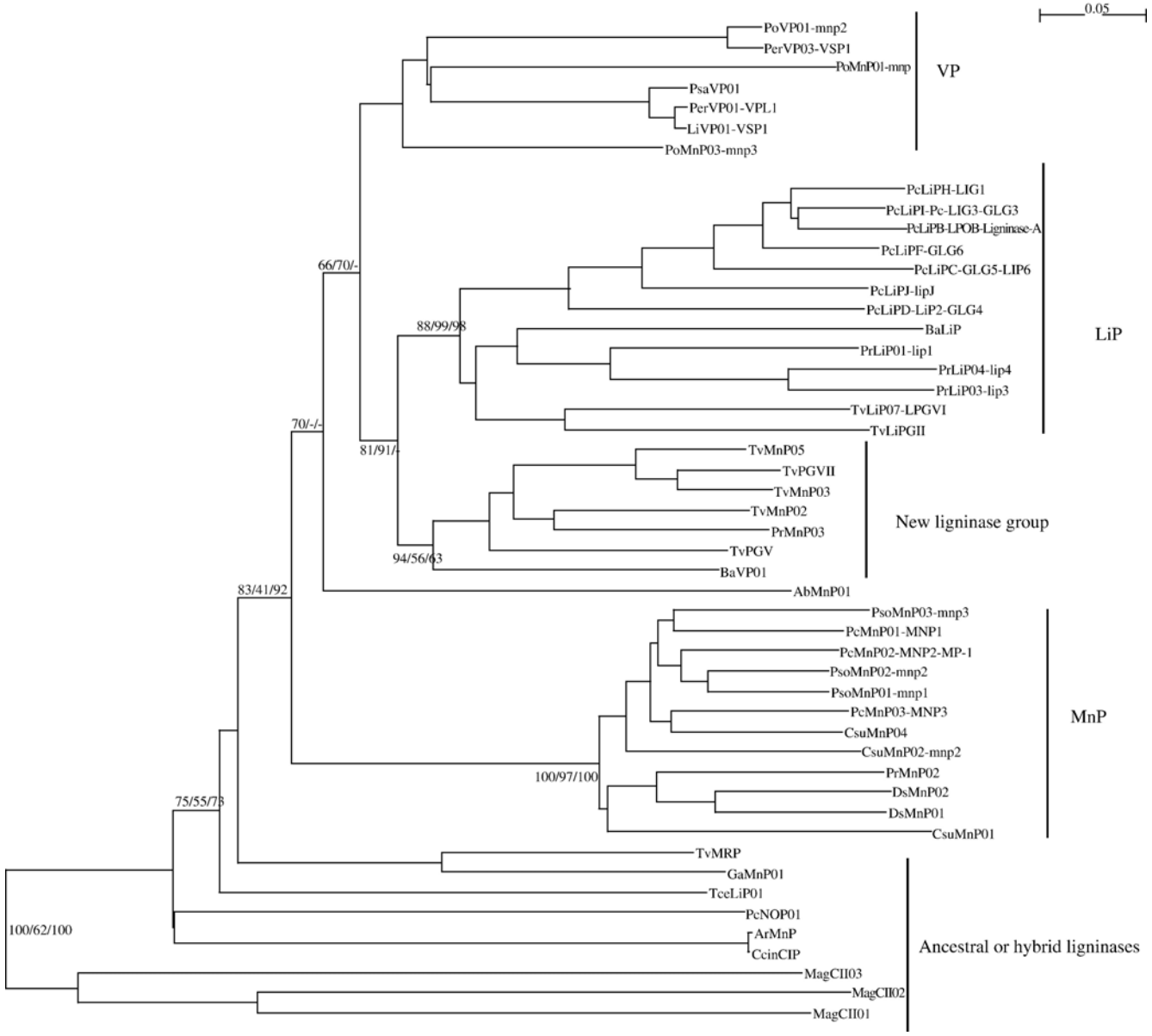

Fig. 3. Neighbor-joining tree of class II peroxidase based on predicted sequences. Only complete sequences were used to build the tree (50 sequences). Numbers represent the bootstrap values obtained from distance, MP, and ML methods. LiP, lignin peroxidase; MnP, manganese peroxidase; VP, versatile peroxidase.

motifs are not included in any of the four groups. These sequences all come at the base of the class II tree without forming a well-defined group (Figs. 3 and 4). The positions of these sequences are well supported by ML and MP analyses. Putative lignin peroxidase-encoding sequences could also be found in Ascomycetes such as *Magnaporthe grisea* (MagCII01, 02, 03) and in *Gibberella zeae* (GzCII01). These sequences contain motifs conserved between ligninases but show a low level of similarity over the whole sequence. Coevolution between ligninolytic activity and diversity of the plant cell wall composition has probably led the ligninolytic fungi to accumulate various copies of the same family and/or of different families (Table 2). Within a particular species, LiP and MnP genes are organized in clusters (e.g., PcLiP from *Phanerochaete chrysosporium* and TvMnP from *Trametes versicolor*), suggesting recent gene duplication events [33]. The repartition of the different families and number of copies within each organism may reflect the ligninolytic capacity and specificity of fungi (Table 2). The tree (Fig. 2) constructed in this work does not indicate which of the four different class II peroxidases might have been the progenitor. This question could be solved by sampling more genome data from early Homobasidiomycetes (cantharellales, phallales, hymenochaetales) [34,35].

Class III peroxidases

Analysis of the large number of sequences entered in PeroxiBase [20] further strengthens the hypothesis, already suggested in previous publications, that class III peroxidases first appeared with the emergence of land plants and subsequently the number of isoforms increased with the complexity of these organisms [19]. A partial sequence was recently found in a green alga belonging to Zygnemophycea, a basal branch of Streptophytes (*Closterium peracerosum–strigosum–littorale* complex, CpslPrx01) [36]. This sequence could narrow the origin of the appearance of class III peroxidases. Furthermore sequences from *Chlamydomona*s

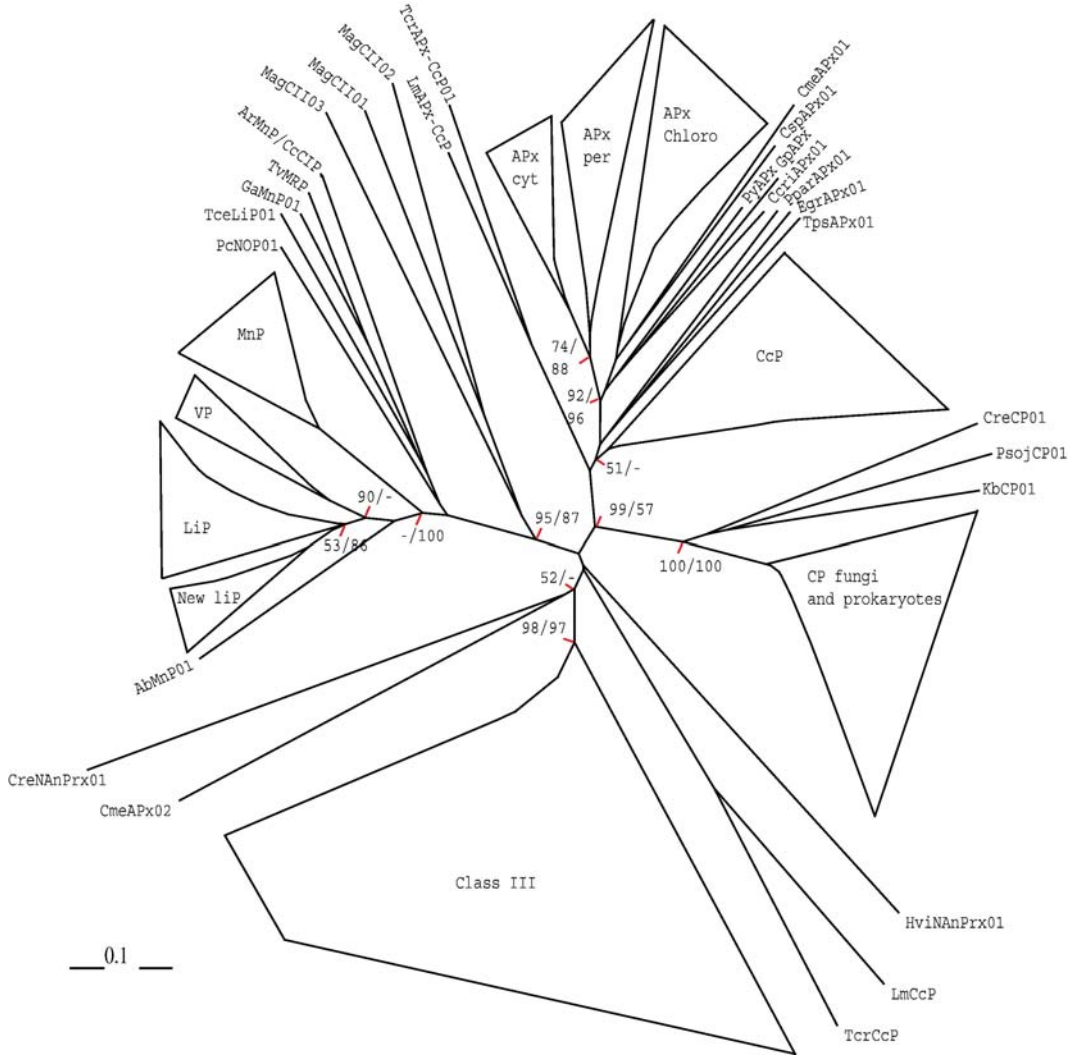

Fig. 4. Cluster tree of the nonanimal peroxidase superfamily. Complete sequences from each class (110 sequences for class I and 50 sequences for classes II and III) were used to build the phylogenetic tree. Schematic representation has been used for the cluster of sequences. Triangles were formed by linking the longer branch of each group to the shortest branch. Isolated sequences were excluded from the cluster. Numbers represent the bootstrap values obtained from distance and ML methods.

Table 2
Distribution of the different class II peroxidase isoforms in relation to the different lignolytic fungus families

	Aphyllophorales											Agaricales				
	Antrodia cinnamomea	Bjerkandera adusta	Bjerkandera sp. B33/3	Ceriporiopsis subvermispora	Coriolopsis gallica	Dichomitus squalens	Phanerochaete chrysosporium	Phanerochaete sordida	Phlebia radiata	Pycnoporus coccineus	Trametes cervina	Trametes versicolor	Laccaria	Lepista irina	Pleurotus eryngii	Pleurotus ostreatus	Pleurotus sapidus
MnP				5	1	2	4	3	1							1	
VP			3										1	1	3	3	1
LiP		1		2			10	2	3	1		3					
New ligninase	1	1							1			7					
Ancestral											1	1					

Numbers correspond to the number sequences detected. LiP: lignin peroxidase, MnP: manganese peroxidase, VP: versatile peroxidase.

(CreNAnPrx01), *Hydra viridis* (HviNAnPrx01), and *Rhodophyta* (CmeAPx02), which contain some class III-like conserved motifs but overall show low levels of identity, could be hybrid sequences between class III peroxidases and the ancestral sequence. These sequences appear at the origin of the class III branch on the NJ tree (Fig. 4). The basal position of these sequences is also supported by the bootstrap values obtained with the MP method (Fig. 4).

Some key organisms to follow the evolution of class II and III peroxidases

In *H. viridis* a putative peroxidase-encoding sequence has been detected (HviNAnPrx01). The presence of such a sequence in an animal is probably due to an ancestral symbiosis rather than the current one with *Chlorella,* which possesses an APx that is not particularly similar to HviNAnPrx01 [37]. This sequence is indeed also present in nonsymbiotic species such as *H. magnipapillata* as well as other symbionts such as *H. vulgaris*. HviNAnPrx01, HvuNAnPrx, and HmNAnPrx probably originated from a class I peroxidase (APx or CcP) horizontally transferred from an ancestral symbiont. The sequences have approximately the same percentage of identity with APx and class III peroxidase sequences. In addition, two cysteine residues characteristic of class III (and implicated in a short disulfide bridge near the distal histidine residue) are present in the *Hydra* sequences. The current sequence HviNAnPrx01 could be an intermediary form that is evolving toward a class III peroxidase. Similarly, we also found in *Chlamydomonas reinhardtii* two hybrid sequences named CreNAnPrx01 and 02. The predicted proteins could be fusion proteins with a middle region containing motifs similar to class III peroxidases and N- and C-terminal regions with similarity to a multicopper oxidase. The "peroxidase" domain cannot, however, be considered as a true class III peroxidase. Detailed sequence analysis and phylogenetic analysis show it to be a hybrid between a class I and a class III peroxidase (Fig. 4). This fusion protein could be functional because one transcript has been found at least for the multicopper oxidase region (Accession No. BI527303).

Trypanosoma and *Leishmania*, two chloroplast-free members of the Euglenozoa branch, contain a putative CcP-like sequence and a possible APx or CcP sequence (the motifs necessary to determine if the sequences are APx or CcP are not detectable). We have shown that the presence of APx strongly correlates with the presence of chloroplasts, which might suggest that chloroplasts have been lost from an ancestor of this organism [38]. An alternative hypothesis claims that these organisms have not evolved from a chloroplastic ancestor [39], and acquisition of APx probably occurred by lateral gene transfer that occurred during their parasitism of plants [40].

Among the fungi, *M. grisea* is a good candidate to study the evolution of peroxidases, as its genome contains an APx, two CcP, 2 CP, and three class II hybrid sequences (MagCII01, 02, and 03). These three sequences have similar degrees of identity with other ligninases and with APx and indeed appeared well separated from the class II branch (Figs. 3 and 4).

The "hybrid" peroxidases described above were very difficult to place in one of the three classes of the superfamily. They all share class I motifs, but also some class II or III motifs. They may thus be considered as extant remnants of intermediate forms between class I and class II/III peroxidases. Collecting more of such sequences from future databases will certainly help our understanding of the critical steps that allowed the conversion of class I peroxidases to the two other classes.

Fusion proteins

During our extensive search for peroxidases, we found two examples of sequences that are fusion proteins containing another domain as well as the peroxidase motifs: GzCP01 (*G. zeae*) and CreNAnPrx01/02 (*C. reinhardtii*). GzCP01 is a CP linked to a cytochrome P450 protein, and CreNAnPrx01/02 is a CP that also contains a multicopper oxidase motif. Only CreNAnPrx01 has been found to be expressed so far. *G. zeae* also has a separate CP gene (GzCP02) and several separate P450 genes (110 are listed in the P450 database: http://drnelson.utmem.edu/CytochromeP450.html). GzCP02 is located on the same chromosome as GzCP01, whereas the "separate" P450s are scattered on various chromosomes. Both cytochrome P450 and multicopper oxidases are involved in redox reactions [41,42], hence the proximity of peroxidases suggests a possible cooperation in specific redox mechanisms. To our knowledge, this is the first time that such a fusion event has been reported in the peroxidase superfamily.

Organisms without peroxidases

Peroxidases are absent from a large variety of bacteria. They do not seem to be essential to their survival, since in many bacterial genera, some species possess CP and others do not. It was even observed that, within one species, some strains contain CP sequences and others do not [43]. When peroxidases are absent, other proteins must be able to cope with ROS production, such as superoxide dismutase, catalase, or chloroperoxidase, which were frequently found in bacteria lacking CP sequences.

None of the currently known amitochondriate eukaryotic organisms (Apicomplexa, Diplomonadida, Parabasalidea, and Microsporidia, as well as some Amoebozoa) showed any peroxidase sequences after extensive BLAST searches. Although these organisms possess a "reduced" form of mitochondria, called either mitosomes or hydrogenosomes [44], it is thought that they once possessed true mitochondria [45–47].

Finally, among nonparasitic and mitochondriate organisms, only the Mycetozoa, animals, and choanoflagellates lack peroxidases belonging to the superfamily (Fig. 1). Both animals and Mycetozoa have another set of peroxidases that belong to the animal peroxidase superfamily (myeloperoxidase, lactoperoxidase, eosinophil peroxidase, and thyroid peroxidase: Prosite Pattern PS50292), which most likely functionally replace the nonanimal superfamily peroxidases. For the choanoflagellates, we did not find even these animal peroxidases, most probably because of the limited amount of data so far available. Curiously, there is no similarity between the two kinds of peroxidases; thus we cannot determine whether animal peroxidases are phylogenetically related to nonanimal peroxidases. The only exceptions were the *Hydra* sequences, which are similar to class I peroxidases, although they have clearly been derived from an ancestral symbiosis with green algae and are not remnants of a hypothetical transition form between nonanimal and animal peroxidases.

Conclusion

Origins of the superfamily of nonanimal peroxidases

Following our study on the peroxidase superfamily, we can conclude that APx is present only in chloroplastic organisms and that CcP can be detected in organisms containing mitochondria except for the green plants, animals, and Amoeba (Table 1 and Fig. 1). We propose that, hundreds of millions years ago, ancestral bacteria possessed a nonduplicated peroxidase sequence from which APx, CcP, and CP originated (Fig. 5). The ancestral (nonfused) peroxidase was probably present in the bacteria that gave rise to the mitochondrion and the chloroplast. Eukaryotes acquired the CcP during a first endosymbiosis event (mitochondrion bacterial precursor) and APx during a second endosymbiosis event with the bacterial ancestor of the chloroplast [48–50]. Regarding CP, they are the result of a duplication and subsequent fusion of the ancestral peroxidase in bacteria. The existence of this ancestral sequence cannot be confirmed: extant bacteria possess only the catalase peroxidase fusion forms.

An alternative hypothesis [7] proposes that current catalase peroxidases are the ancestral class I form. A CP would have been laterally transferred to the nucleus of a eukaryotic ancestor. APx and CcP derived from the insertion of an intron between the N- and the C-terminal parts of the CP. The C-terminal part, probably not functional, would have then been eliminated, leaving the N-terminal part to evolve eventually into APx and CcP. If this hypothesis were true, then the associations mitochondrion-CcP and chloroplast-APx would be more difficult to explain. Moreover, if an intron were inserted into a CP sequence in an ancestral eukaryotic organism, at least traces of the C-terminal part should be found in some eukaryotes, which is not the case according to our studies.

Our hypothesis appears hence as the most parsimonious. To distinguish definitely between these two divergent hypotheses, it would be necessary to date the fusion event and determine whether it occurred before or after organellar acquisition by eukaryotes.

Current theories state that eukaryotes can be split into two groups, termed Unikonts and Bikonts [50]. The first group contains animals, fungi, and Amoeba, while the second group consists of plants, algae, and other protists. The observation that class I is absent from Unikonts (except fungi) suggests the following hypothesis. A Unikont ancestor acquired an earlier form of mitochondrion lacking CcP. A more evolved form of mitochondrion containing a CcP sequence was later transferred to a Bikont ancestor. The timeline of this transfer is supported

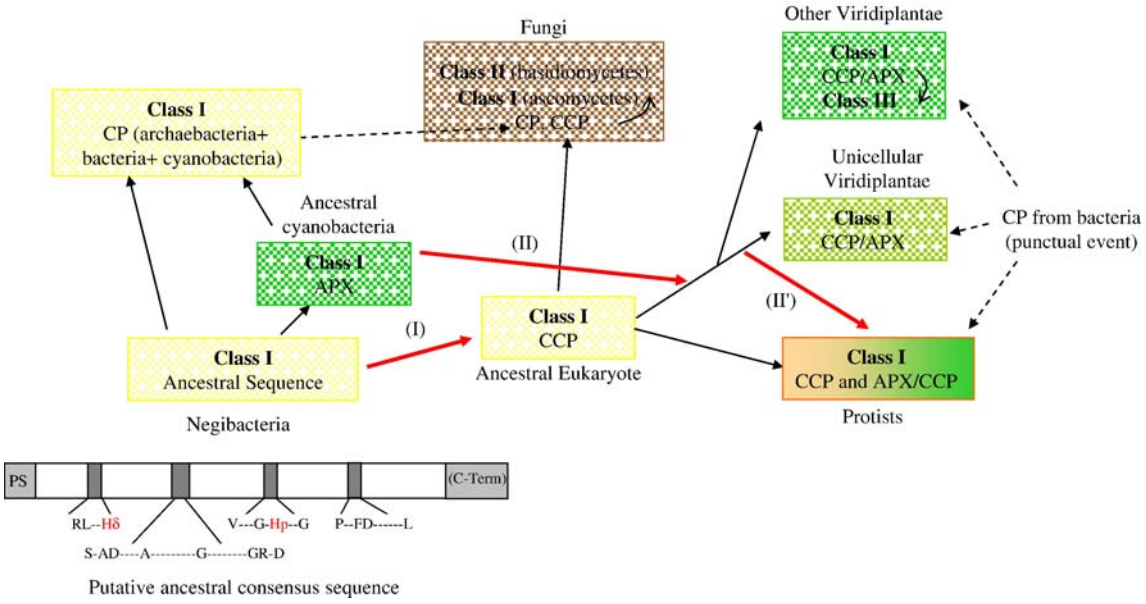

Fig. 5. Model for the emergence and the evolution of the three peroxidase classes from one ancestral sequence. Arrows represent the putative evolution, bold arrows the major endosymbiosis events (I and II for primary and secondary endosymbiosis, respectively), and dashed arrows the putative CP sequence transfer.

by the fact that Unikonts contain separate dihydrofolate reductases (DHFR) and thymidylate synthase (TS) genes, whereas Bikonts exclusively possess a fused DHFR–TS gene, thus indicating a posterior origin for Bikonts [51]. Fungi would have acquired CcP by a subsequent lateral gene transfer, but after the emergence of the Microsporidia branch, which does not contain a CcP sequence. Although there is no support yet for two different mitochondrial acquisition steps, it seems to us that this hypothesis is the most probable, to be consistent with the current view that mitochondria were acquired by all existing eukaryotes [52]. In land plants, CcP probably evolved into class III peroxidases. We indeed observed that all land plants, which encode class III peroxidases, do not have any CcP genes. Similarly, the charophycean alga *Cl. peracerosum* encodes a class III peroxidase, and no CcP gene was found. CcP and class III peroxidase sequences are, however, quite divergent: in addition to differences in their primary sequence, CcP lack the four disulfide bridges and the two calcium ions of class III peroxidases. In this respect, the "hybrid" peroxidase found in *Chlamydomonas* (CreNAnPrx01) may represent an intermediate form between CcP and class III peroxidases.

Present-day CcP and APx sequences originated from an ancestral peroxidase form and have evolved into more compartmentalized functions. However, CcP duplications within an organism are the exception rather than the rule, and their significance is not clearly known. For the APx, however, the two well-defined plant cytoplasmic and chloroplastic APx have the same origin and have resulted from duplication events in the basal Viridiplantae. In the case of Euglenozoa, chloroplasts may have been gained not by an early ancestor of the whole group, but only by some members (Euglenids) in later endosymbiotic events [39,49]. The presence of peroxisomal APx isoforms only in land plants reinforces the idea of more recent duplications in the higher plants. These isoforms probably originated from cytosolic isoforms that gained an exon encoding the membrane-bound domain and the targeting peptide [10]. The detection of alternative splicing in chloroplast genes of certain eudicot lineages (Cucurbitaceae, Fabaceae, Caryophyllales, and some Solanaceae) reinforces the concept of an active evolution of the class I family in green plants. It should be noted, however, that some eudicots (Brassicaceae and Solanaceae) have distinct genes encoding stromal and thylakoid-bound isoforms. Repartition of APx into cytoplasmic, peroxisomal, and chloroplastic is probably due to subfunctionalization or, less likely, to neofunctionalization, as their only known function is detoxification of the cell. We cannot discard, though, the gain of new functions (neofunctionalization) as a consequence of the repartition of APx in different cellular locations. Subfunctionalization can lead to very subtle variations of APx localization, such as the separation observed between stromal and thylakoid-bound chloroplastic APx in *Arabidopsis thaliana* (respectively AtAPx05 and AtAPx06) or in *Lycopersicon esculentum* (respectively LeAPx 07 and LeAPx08), and gave rise to organisms bearing up to eight APx in their genome (*Glycine max*).

Our extensive search of peroxidase sequences in all living organisms did not detect any peroxidase sequence in several distinct groups, from animals to amitochondriate parasites and some bacteria. If peroxidases played such an important role in detoxification and, later in evolution, in many other functions, why did so many organisms lose (or not acquire) this gene? What other gene did they find to cope with the loss of peroxidases? Did the parasitic organisms use the host machinery to deal with ROS production? It is certain that peroxidases are not the only proteins able to protect organisms against oxidative damage. Other peroxidases such as superoxide dismutase (EC 1.15.1.1), catalase (EC 1.11.1.6), glutathione peroxidase (EC 1.11.1.9), chloroperoxidase (EC 1.11.1.10), or

peroxiredoxin (EC 1.11.1.15) were found in bacteria lacking CP [43]. Some of these enzymes are also present in amitochondriate and/or parasitic eukaryotes lacking nonanimal superfamily peroxidases. Parasitic organisms probably also deal with elimination of ROS by using their host enzymes. Finally, animals and Mycetozoa possess peroxidases belonging to the "animal superfamily."

Interestingly, a study on yeast (*Saccharomyces cerevisiae*) showed that disruption of its unique CcP gene was not essential for cell viability or respiration, though it rendered the yeast more sensitive to H_2O_2 [4]. Although CP and CcP may not be essential to viability of, respectively, bacterial organisms or yeast, peroxidases certainly play a major role in plants. The ancestral class I sequence probably played a crucial role in the bacteria that gave rise to chloroplasts and mitochondria. Class I peroxidases have been retained during evolution, implying that they still bring an advantage to the organisms that possess them.

Evolution of class I peroxidases to classes II and III

Specific duplication events have occurred for APx and class III peroxidases in Viridiplantae and for class II in fungi. The conservation of duplicated genes could be related to neofunctionalization or subfunctionalization specific to a particular phylum or branch. In fungi, class II peroxidases are present only in one saprophytic fungal family and are organized in clusters of highly conserved and duplicated sequences. Class II peroxidases could have evolved from existing class I genes with the appearance and subsequent diversification of lignin polymers in plant cell walls. The ancestor of class II peroxidases is most probably a cytochrome *c* peroxidase that, in response to the very high redox potentials of lignin polymers, modified key amino acid residues to enable the enzyme to oxidize these new targets. The evolutionary mechanism of class III peroxidases from class I is less clear, but may be related to the emergence of land plants and thus could be a consequence of the dramatic modifications of growth conditions (new organs, structures, atmospheric conditions) [53]. Interestingly, whenever an organism possesses class III peroxidases, it lacks any CcP-encoding sequences. Therefore, it is possible that class I peroxidases, which do not have disulfide bridges, are at the origin of class II and III peroxidases, which in contrast both form four disulfide bridges. In this respect, sequences from the Ascomycetes (MagCII01, 02, and 03), *Chlamydomonas* (CreNAnPrx01 and 02), and *H. viridis* (HviNAnPrx01) may represent intermediary forms between CcP and class II/III peroxidases (Fig. 4). Quantification of the cysteine residues in these sequences shows that they could form a maximum of three disulfide bridges. As class III peroxidases possess four disulfide bridges, and class I have none, this observation is in favor of our theory for the origin of class III peroxidases. The discovery of two hybrid class I-class III peroxidases in *C. reinhardtii* and their phylogenetic position at the origin of the class III peroxidase branch suggest that reasons other than land colonization may have led to the appearance of class III peroxidases. The number of class III peroxidase-encoding genes has dramatically increased in land plants during their evolution. The most recent (and thus highly evolved) plants often contain more than 100 peroxidase genes [20]. Several reasons may explain this "explosion" in numbers of class III peroxidases: organ diversification, climatic changes, colonization of new biotopes, the constant appearance of new pathogens, and human impact on cultivated plants [53].

The need for a revision of the peroxidase superfamily

In 1992, Karen Welinder described, using all the information available at that time, the "plant, fungal and bacterial peroxidase superfamily." Subsequently, extensive genome sequencing of very diverse organisms has allowed the discovery of peroxidases from Welinder's superfamily in new classes of organisms. We therefore have now chosen to update her peroxidase classification and propose the new name "superfamily of nonanimal peroxidases." This denomination does not predict which organisms will have peroxidase-encoding genes, but it reflects the fact that peroxidases can be present in almost any living organism.

Peroxidase sequences found in protistean organisms have evolved uniquely, and their current form does not allow their precise classification, even after enzymatic activity experiments. In the coming years, the amount of data on protistean organisms will increase, and many more sequences of this kind should appear that will help us to draw up a more precise evolution of the nonanimal peroxidase superfamily. It is probable that with the constant discovery of "hybrid" forms of peroxidases, new classes will need to be defined. However, as evolution becomes more understood, the superfamily of nonanimal peroxidases remains an exciting and challenging domain for research in gene evolution.

Experimental procedures

Retrieval of the peroxidase sequences

A portion of the protein sequences used for phylogenetic study was obtained from the UniProt database (http://www.expasy.uniprot.org/). Then, to find peroxidase sequences in a large variety of organisms, class I and III sequences from *Arabidopsis* and class I and II sequences from *Phanerochaete* were used as input sequences in tBLASTn searches within different databases. Peroxidases were searched for in the NCBI Web site (www.ncbi.nih.gov/BLAST) and in several specialized databases such as the PEP EST database (www.moss.leeds.ac.uk), the PlantGDB database (www.plantgdb.org/), and the DOE Joint Genome Institute Web site (genome.jgi-psf.org/). Some catalase peroxidases were also found through the J. Craig Venter Institute BLAST server (https://research.venterinstitute.org/blast/). Any nonannotated sequences were analyzed for the presence of the gene with different programs such as FGenesh (http://www.softberry.com/berry.phtml) and GenScan (http://genes.mit.edu/GENSCAN.html). The corresponding coding sequence was translated with the "translate" tool on Expasy (http://us.expasy.org/tools/dna.html) and checked for presence of specific motifs. Sequences of protein used in this study can be found at the following link: http://peroxidase.isb-sib.ch [20].

Comprehensive phylogenetic analysis of peroxidase sequences

Among the high number of available peroxidase sequences in the databases, for each class complete sequences representative of each phylum were used to

build the global tree containing all three classes. For each independent tree, a large number of sequences were aligned. Peroxidase protein sequences were aligned using ClustalW [54]. Only a part of the sequence containing highly conserved residues was used for the alignment. For this reason, signal peptides and any C-terminal extensions were removed. For CP sequences, only the N-termini were used for alignments. Alignments were further inspected and visually adjusted and realigned with ClustalX. The distance tree was constructed with the "neighbor" option of the PHYLIP 3.6a3 package [55] under the JTT substitution frequency matrix, and 1000 bootstrap replicates were carried out. Maximum parsimony trees were built with the PHYLO_WIN software, using 100 bootstrap replicates [56]. Maximum likelihood trees were inferred with the PHYML algorithm, under the JTT substitution frequency matrix, by using the BIONJ starting tree [57]. Njplot software was used to visualize phylogenetic trees and BioEdit software to obtain the different consensus sequences.

Acknowledgments

We thank William Deakin and Jan Pawlowski for their critical reading and constructive suggestions and Laurent Falquet for the use of his unpublished profiles. The financial support of the Swiss National Science Foundation (Grant 31-068003.02) to C.P. and C.D. is gratefully acknowledged; N.B. is paid by the Office Cantonal de l'Emploi de Genève.

References

[1] K.G. Welinder, Superfamily of plant, fungal and bacterial peroxidases, Curr. Opin. Struct. Biol. 2 (1992) 388–393.
[2] H.B. Dunford, Heme peroxidase nomenclature, Plant Peroxidase Newsl. 13 (1999) 65–71.
[3] M. Zamocky, Phylogenetic relationships in class I of the superfamily of bacterial, fungal, and plant peroxidases, Eur. J. Biochem. 271 (2004) 3297–3309.
[4] M. Kwon, S. Chong, S. Han, K. Kim, Oxidative stresses elevate the expression of cytochrome c peroxidase in *Saccharomyces cerevisiae*, Biochim. Biophys. Acta 1623 (2003) 1–5.
[5] M. Zamocky, C. Dunand, Divergent evolutionary lines of fungal cytochrome c peroxidases belonging to the superfamily of bacterial, fungal and plant heme peroxidases, FEBS Lett. 580 (2006) 6655–6664.
[6] Y. Yamada, T. Fujiwara, T. Sato, N. Igarashi, N. Tanaka, The 2.0 A crystal structure of catalase-peroxidase from *Haloarcula marismortui*, Nat. Struct. Biol. 9 (2002) 691–695.
[7] M.G. Klotz, P.C. Loewen, The molecular evolution of catalatic hydroperoxidases: evidence for multiple lateral transfer of genes between prokaryota and from bacteria into eukaryota, Mol. Biol. Evol. 20 (2003) 1098–1112.
[8] C. Miyake, F. Michihata, K. Asada, Scavenging of hydrogen peroxide in prokaryotic and eukaryotic algae: acquisition of ascorbate peroxidase during the evolution of cyanobacteria, Plant Cell Physiol. 32 (1991) 33–43.
[9] K. Asada, The water–water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons, Annu. Rev. Plant Physiol. Plant Mol. Biol. 50 (1999) 601–639.
[10] F.K. Teixeira, L. Menezes-Benavente, R. Margis, M. Margis-Pinheiro, Analysis of the molecular evolutionary history of the ascorbate peroxidase gene family: inferences from the rice genome, J. Mol. Evol. 59 (2004) 761–770.
[11] S. Kitajima, et al., Stable form of ascorbate peroxidase from the red alga Galdieria partita similar to both chloroplastic and cytosolic isoforms of higher plants, Biosci. Biotechnol. Biochem. 66 (2002) 2367–2375.
[12] T. Ishikawa, T. Takeda, H. Kohno, S. Shigeoka, Molecular characterization of Euglena ascorbate peroxidase using monoclonal antibody, Biochim. Biophys. Acta 1290 (1996) 69–75.
[13] S.R. Wilkinson, S.O. Obado, I.L. Mauricio, J.M. Kelly, Trypanosoma cruzi expresses a plant-like ascorbate-dependent hemoperoxidase localized to the endoplasmic reticulum, Proc. Natl. Acad. Sci. USA 99 (2002) 13453–13458.
[14] M.H. Gold, M. Alic, Molecular biology of the lignin-degrading basidiomycete *Phanerochaete chrysosporium*, Microbiol. Rev. 57 (1993) 605–622.
[15] A. Conesa, P.J. Punt, C.A. van den Hondel, Fungal peroxidases: molecular aspects and applications, J. Biotechnol. 93 (2002) 143–158.
[16] M.J. Martinez, F.J. Ruiz-Duenas, F. Guillen, A.T. Martinez, Purification and catalytic properties of two manganese peroxidase isoenzymes from Pleurotus eryngii, Eur. J. Biochem. 237 (1996) 424–432.
[17] M.T. Moreira, et al., Role of organic acids in the manganese-independent biobleaching system of Bjerkandera sp. strain BOS55, Appl. Environ. Microbiol. 64 (1998) 2409–2417.
[18] F.J. Ruiz-Duenas, S. Camarero, M. Perez-Boada, M.J. Martinez, A.T. Martinez, A new versatile peroxidase from Pleurotus, Biochem. Soc. Trans. 29 (Pt 2) (2001) 116–122.
[19] F. Passardi, D. Longet, C. Penel, C. Dunand, The class III peroxidase multigenic family in rice and its evolution in land plants, Phytochemistry 65 (2004) 1879–1893.
[20] N. Bakalovic, et al., PeroxiBase: a class III plant peroxidase database, Phytochemistry 67 (2006) 534–539.
[21] J.M. Neuhaus, Protein targeting to the plant vacuole, Plant Physiol. Biochem. 34 (1996) 217–221.
[22] T. Matsui, H. Nakayama, K. Yoshida, A. Shinmyo, Vesicular transport route of horseradish C1a peroxidase is regulated by N- and C-terminal propeptides in tobacco cells, Appl. Microbiol. Biotechnol. 62 (2003) 517–522.
[23] L. Duroux, K.G. Welinder, The peroxidase gene family in plants: a phylogenetic overview, J. Mol. Evol. 57 (2003) 397–407.
[24] F. Passardi, C. Penel, C. Dunand, Performing the paradoxical: how plant peroxidases modify the cell wall, Trends Plant Sci. 9 (2004) 440–534.
[25] A. Liszkay, B. Kenk, P. Schopfer, Evidence for the involvement of cell wall peroxidase in the generation of hydroxyl radicals mediating extension growth, Planta 217 (2003) 658–667.
[26] S. Hiraga, K. Sasaki, H. Ito, Y. Ohashi, H. Matsui, A large family of class III plant peroxidases, Plant Cell Physiol. 42 (2001) 462–468.
[27] W.R. Patterson, T.L. Poulos, Crystal structure of recombinant pea cytosolic ascorbate peroxidase, Biochemistry 34 (1995) 4331–4341.
[28] M.D. Katinka, et al., Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi, Nature 414 (2001) 450–453.
[29] T.D. Bruns, et al., Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences, Mol. Phylogenet. Evol. 1 (1992) 231–241.
[30] Y. Tanabe, M.M. Watanabe, J. Sugiyama, Evolutionary relationships among basal fungi (Chytridiomycota and Zygomycota): insights from molecular phylogenetics, J. Gen. Appl. Microbiol. 51 (2005) 267–276.
[31] A.C. Ivens, et al., The genome of the kinetoplastid parasite, Leishmania major, Science 309 (2005) 436–442.
[32] A.T. Martinez, Molecular biology and structure function of lignin-degrading heme peroxidase, Enzyme Microb. Technol. 30 (2002) 425–444.
[33] P. Stewart, D. Cullen, Organization and differential regulation of a cluster of lignin peroxidase genes of Phanerochaete chrysosporium, J. Bacteriol. 181 (1999) 3427–3432.
[34] M. Binder, D.S. Hibbett, Higher-level phylogenetic relationships of Homobasidiomycetes (mushroom-forming fungi) inferred from four rDNA regions, Mol. Phylogenet. Evol. 22 (2002) 76–90.
[35] J.M. Moncalvo, et al., One hundred and seventeen clades of euagarics, Mol. Phylogenet. Evol. 23 (2002) 357–400.
[36] R.M. McCourt, C.F. Delwiche, K.G. Karol, Charophyte algae and land plant origins, Trends Ecol. Evol. 19 (2004) 661–666.
[37] M. Habetha, T.C. Bosch, Symbiotic Hydra express a plant-like peroxidase gene during oogenesis, J. Exp. Biol. 208 (Pt 11) (2005) 2157–2165.
[38] V. Hannaert, et al., Plant-like traits associated with metabolism of Trypanosoma parasites, Proc. Natl. Acad. Sci. USA 100 (2003) 1067–1071.
[39] B.S. Leander, Did trypanosomatid parasites have photosynthetic ancestors? Trends Microbiol. 12 (2004) 251–258.
[40] R.F. Waller, M.J. McConville, G.I. McFadden, More plastids in human parasites? Trends Parasitol. 20 (2004) 54–57.

[41] D.W. Nebert, F.J. Gonzalez, P450 genes: structure, evolution, and regulation, Annu. Rev. Biochem. 56 (1987) 945–993.
[42] K. Nakamura, N. Go, Function and molecular evolution of multicopper blue proteins, Cell Mol. Life Sci. 62 (2005) 2050–2066.
[43] F. Passardi, M. Zamocky, J. Favet, C. Jakopitsch, C. Penel, C. Obinger, C. Dunand. Phylogenetic distribution of catalase-peroxidase: are there patches of order in the chaos? Submitted for publication.
[44] M. van der Giezen, J. Tovar, C.G. Clark, Mitochondrion-derived organelles in protists and fungi, Int. Rev. Cytol. 244 (2005) 175–225.
[45] T.M. Embley, et al., Hydrogenosomes, mitochondria and early eukaryotic evolution, IUBMB Life 55 (2003) 387–395.
[46] M.W. Gray, Evolutionary biology: the hydrogenosome's murky past, Nature 434 (2005) 29–31.
[47] I. Hrdy, et al., Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I, Nature 432 (2004) 618–622.
[48] M.W. Gray, Evolution of organellar genomes, Curr. Opin. Genet. Dev. 9 (1999) 678–687.
[49] G.I. McFadden, Primary and secondary endosymbiosis and the origin of plastids, J. Phycol. 37 (2001) 951–959.
[50] T. Cavalier-Smith, Only six kingdoms of life, Proc. Biol. Sci. 271 (2004) 1251–1262.
[51] A. Stechmann, T. Cavalier-Smith, The root of the eukaryote tree pinpointed, Curr. Biol. 13 (2003) R665–R666.
[52] T.M. Embley, W. Martin, Eukaryotic evolution, changes and challenges, Nature 440 (2006) 623–630.
[53] F. Passardi, D. Longet, C. Penel, C. Dunand, The class III peroxidase multigenic family in rice and its evolution in land plants, Phytochemistry 65 (2004) 1879–1893.
[54] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[55] J. Felsenstein, PHYLIP (Phylogeny Inference Package), University of Washington, Seattle, 1993.
[56] N. Galtier, M. Gouy, C. Gautier, SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny, Comput. Appl. Biosci. 12 (1996) 543–548.
[57] S. Guindon, O. Gascuel, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood, Syst. Biol. 52 (2003) 696–704.