J Mol Evol (2010) 71:401–414
DOI 10.1007/s00239-010-9394-3

A Tandem Duplication of Manganese Superoxide Dismutase in *Nosema bombycis* and Its Evolutionary Origins

Heng Xiang · Guoqing Pan · Charles R. Vossbrinck · Ruizhi Zhang · Jinshan Xu · Tian Li · Zeyang Zhou · Cheng Lu · Zhonghuai Xiang

Received: 1 March 2010 / Accepted: 17 September 2010 / Published online: 23 October 2010

Abstract Microsporidia are a group of obligate intracellular eukaryotic parasites with small genomes. They infect animals from a wide variety of phyla, including humans. Two manganese superoxide dismutase (MnSOD) genes, designated *NbMnSOD1* and *NbMnSOD2*, were found to be organized in a tandem array within the *Nosema bombycis* genome. The genes, both 678 bp in length, were found to be more similar to each other than they are to homologous genes of other Microsporidia, suggesting that the tandem duplication occurred subsequent to the development of this lineage. Reverse transcript PCR shows that mRNA for both genes is present in the spores. Analysis of the primary structure, hydrophobic cluster analysis, target signal analysis, and phylogenetic analysis all indicate that *NbMnSOD1* is dimeric and targeted to the cytosol. *NbMnSOD2* seems to have changed more rapidly and is under less evolutionary constraint than *NbMnSOD1* suggesting that *NbMnSOD2* may function under different conditions or in different tissues of its host rather than simply resulting in an increase in expression. A phylogenetic analysis of MnSOD sequences from eukaryotes, Archaea, and bacteria shows the microsporidial MnSODs to be grouped with the bacteria suggesting a possible horizontal gene transfer.

Keywords Manganese superoxide dismutase · Microsporidia · Gene duplication · Horizontal gene transfer · Structure · Subcellular localization

Electronic supplementary material The online version of this article (doi:10.1007/s00239-010-9394-3) contains supplementary material, which is available to authorized users.

H. Xiang · G. Pan · T. Li · Z. Zhou (✉) · C. Lu · Z. Xiang
Institute of Sericulture and Systems Biology, Southwest University, 2 Tiansheng Rd., Beibei District, Chongqing 400715, China
e-mail: zyzhou@swu.edu.cn

H. Xiang · C. Lu
College of Animal Science and Technology, Southwest University, Beibei, Chongqing 400715, China

J. Xu · Z. Zhou
College of Life Sciences, Chongqing Normal University, Chongqing 400047, China

R. Zhang
Biotechnology Research Center, Southwest University, Beibei, Chongqing 400715, China

C. R. Vossbrinck
The Connecticut Agricultural Experiment Station, 123 Huntington Street, PO Box 1106, New Haven, CT 06504, USA

Introduction

Superoxide dismutases (SODs, EC 1.15.1.1) are a group of ubiquitous metalloenzymes that catalyze the dismutation of toxic superoxide anions to molecular oxygen and hydrogen peroxide (McCord and Fridovich 1969; Fridovich 1995). In concert with catalase, SOD eliminates reactive oxygen species from cells. SODs have been reported to possess a potentially important role in host invasion and development of pathogenic fungi (Belinky et al. 2002; Giles et al. 2005) and Microsporidia (El-Taweel et al. 2007), but the exact mechanism has not been clearly demonstrated. SODs are classified into three main groups according to the type of metal cofactor: manganese SODs (MnSODs) are found in prokaryotes and in the mitochondrial matrices of eukaryotes; iron SODs (FeSODs) are found in prokaryotes, protists and plants; and copper/zinc SODs (Cu/ZnSODs) are found in bacteria and the cytosol and extracellular

compartments of eukaryotes (Bordo et al. 1994; Brouwer et al. 2003). FeSODs and MnSODs resemble each other in their sequence and structure (Parker et al. 1987), suggesting a common ancestry (Hassan 1989), but they are unrelated to Cu/ZnSODs (Tainer et al. 1982; Smith and Doolittle 1992). Cambialistic SODs are those that function efficiently with both manganese and iron (Martin et al. 1986) and are regarded as a transitional form between FeSODs and MnSODs.

Similarity among MnSODs suggests that these enzymes have a common ancestry (Fitch and Ayala 1994), but their distribution among species is variable. Studies have shown that most organisms contain only a single MnSOD gene (Bagnoli et al. 1998; Kliebenstein et al. 1998; Babitha et al. 2002; Fang et al. 2002; Regelsberger et al. 2002), and only a small number of organisms, such as *Zea mays* (Zhu and Scandalios 1993), *Nicotiana tabacum* (Van Camp et al. 1997), *Callinectes sapidus* (Brouwer et al. 1997), *Caenorhabditis elegans* (Hunter et al. 1997), and *Prunus persica* (Bagnoli et al. 2002), have been found to have multiple MnSODs. To date, however, there has been no report of a tandem duplication of the MnSOD gene.

Each MnSOD monomer adopts an α/β fold, and combines to form an oligomeric structure in solution (Stallings et al. 1984; Jackson and Cooper 1998). MnSODs in eukaryotes are generally tetrameric (Natvig et al. 1996; Palma et al. 1998; Alscher et al. 2002) while in bacteria, MnSODs are typically dimeric (Natvig et al. 1996; Alscher et al. 2002), although tetramers (Wagner et al. 1993) have been reported. In prokaryotes, MnSODs are usually associated with the cytosol (Lynch and Kuramitsu 2000), though they have also been found in the thylakoid and plasma membranes of cyanobacteria (Regelsberger et al. 2002). In eukaryotes, MnSODs are usually synthesized in the cytosol and imported post-translationally into the mitochondrial matrix (Weisiger and Fridovich 1973; Zhu and Scandalios 1993; Natvig et al. 1996; Alscher et al. 2002). Some MnSODs, such as those from *Callinectes sapidus* (Brouwer et al. 1997), *Ganoderma microsporum* (Pan et al. 1997), *Penicillium chrysogenum* (Diez et al. 1998), and *Candida albicans* (Lamarre et al. 2001) are targeted to the cytosol.

Microsporidia are a group of obligate intracellular eukaryotic parasites that infect a wide variety of species, including humans (Desportes et al. 1985; Canning 1993; Snowden 2004). Members of this phylum possess many unusual characteristics, including a prokaryotic type rRNA gene arrangement with the 5.8S rRNA homologue attached to the large subunit rRNA (Vossbrinck and Woese 1986) and a highly divergent small subunit rRNA (ssrRNA) gene. Comparative analysis of the ssrRNA led to the view that Microsporidia may represent an early branch of the eukaryotes (Vossbrinck et al. 1987). However, some recent studies have suggested that Microsporidia emerge within the fungi (Thomarat et al. 2004; Gill and Fast 2006; James et al. 2006), and that long-branch attraction (LBA) (Felsenstein 1978) explains the original placement of Microsporidia at the base of the eukaryotes (Germot and Philippe 1999). Another possible "primitive" feature of the Microsporidia is the lack of mitochondria. It has been shown, however, that Microsporidia retain a mitosome, which is a double-membrane-bound structure that is thought to be a remnant of the mitochondria (Williams et al. 2002). The very small genome of Microsporidia (the smallest known genome among eukaryotes) could also be considered a primitive trait but, based on an analysis of the completed genome of another Microsporidia, *Encephalitozoon cuniculi*, it has been suggested that the small size of the microsporidial genome is a recently acquired feature (Keeling et al. 2005).

Genomic analysis of *E. cuniculi* has shown that this species uses a unique MnSOD to deal with oxidative stress (Katinka et al. 2001). Our analysis shows that the MnSOD from the microsporidian *Nosema bombycis*, a parasite of the silk moth *Bombyx mori*, is closely related to that of *E. cuniculi*. In addition, the MnSOD shows a tandem duplication event in *N. bombycis*. Here we predict the structure and the subcellular localization of the two MnSODs from *N. bombycis* and perform a comparative phylogenetic analysis on a wide array of MnSODs.

Materials and Methods

Obtaining SOD Sequence Data

Nearly two hundred thousand random shotgun reads have resulted in a 7.8-fold genomic database as part of the *N. bombycis* genomic sequencing project in our laboratory. *N. bombycis* isolate CQ1, purified from infected silkworms in Chongqing, China, is contained in the China Veterinary Culture Collection Center (CVCC number 102059) (Xu et al. 2006). All initial reads were assembled using the RePS program (Wang et al. 2002). The *N. bombycis* manganese superoxide dismutase (*NbMnSOD*) genes were located by the GLIMMER 3.0 gene sequence prediction program (Delcher et al. 2007). Two MnSOD genes were identified as a tandem repeat and designated here as *NbMnSOD1* and *NbMnSOD2*.

Using the two *NbMnSOD*s in a BLAST (Altschul et al. 1990) search, 53 homologous amino acid sequences were retrieved from the NCBI database. In addition, 18 fungal sequences were obtained from NCBI based on previously published information (Frealle et al. 2006). Accession numbers of all amino acid sequences are indicated in the Supplementary Table S1. The nucleotide sequences determined in this study for the *N. bombycis* SODs have been

deposited in Genbank under accession numbers FJ377710 for *NbMnSOD1* and FJ377711 for *NbMnSOD2*.

RT-PCR

Total RNA was extracted from purified spores of *N. bombycis* using the Trizol reagent (Invitrogen, USA), according to the protocol provided by the manufacturer, and treated using RNase-free DNase I. Single-strand cDNA (sscDNA) was synthesized using M-MLV Reverse Transcriptase following the recommended protocols (Promega, USA). The RNA was subsequently digested using RNase H (MBI Fermentas).

The specific primers were designed by Oligo6 software (Molecular Biology Insights, Inc., Cascade, CO, USA). Primer sequences are as follows:

*NbMnSOD1*F: 5′-GGACTTATGAAAGAAATGGA-3′ and *NbMnSOD1*R: 5′-AATTACTAACGTATTCAGGC-3′, with a 268-bp predicted product;
*NbMnSOD2*F: 5′-ATTTGGCAGTTGATCTTAGG-3′ and *NbMnSOD2*R: 5′-AACTTCTTTCCTTCAACTAC-3′, with a 337-bp predicted product.

Gene amplification was carried out using both sscDNA and genomic DNA (as a control) templates.

Prediction of Subcellular Targeting

All superoxide dismutase proteins were analyzed for the presence of an amino-terminal extension of 20–40 residues targeting the protein for transport either within or out of the cell. Targeting sequence prediction programs: PSORT II (Nakai and Horton 1999), TargetP (Emanuelsson et al. 2000), and MitoProt II (Claros and Vincens 1996) were used.

Sequence Alignment

To define the structures and perform phylogenetic analysis, the amino acid sequences of 72 SODs were aligned using Muscle software (Edgar 2004). The alignment was then adjusted using the SOD structures determined from X-ray diffraction data for 3 of the MnSODs and 8 of the FeSODs (Wintjens et al. 2004). These structures are the following: MnSODs from *Homo sapiens* mitochondria (Protein Data Bank accession number 1n0j; X-ray resolution 2.20 Å; MnSOD; tetramer (Borgstahl et al. 1992)), *Aspergillus fumigatus* (1kkc; 2.00 Å; MnSOD; tetramer (Fluckiger et al. 2002)), *Escherichia coli* (1d5n and 1isa; 1.55 and 1.80 Å; MnSOD and FeSOD, respectively; dimer (Lah et al. 1995; Borgstahl et al. 2000)), *Porphyromonas gingivalis* (1qnn; 1.80 Å; FeSOD; dimer (Sugio et al. 2003)), *Sulfolobus acidocaldarius* (1b06; 2.20 Å; FeSOD; tetramer (Knapp et al. 1999)), *Sulfolobus solfataricus* (1sss; 2.30 Å; FeSOD; tetramer (Ursby et al. 1999)), *Aquifex pyrophilus* (1coj; 1.90 Å; FeSOD; tetramer (Lim et al. 1997)), *Propionium freudenreichii* subsp. *shermanii* (1bsm; 1.35 Å; FeSOD; tetramer (Schmidt 1999)), *Mycobacterium tuberculosis/vaccae* (1ids; 2.00 Å; FeSOD; tetramer (Cooper et al. 1995)), and *Pseudomonas ovalis/putida* (1dt0; 2.10 Å; FeSOD; dimer (Bond et al. 2000)).

Hydrophobic Cluster Analysis

Sequences of 12 SOD proteins (6 tetramers, 4 dimers, and the 2 *N. bombycis* SODs) were compared using Hydrophobic Cluster Analysis (HCA) and an HCA plot was drawn with the drawhca program (Woodcock et al. 1992). HCA was developed to show similarities among protein structures when amino acid sequence similarities are low by comparing the placement of hydrophobic regions in the protein. For proteins which have a greater than 50% sequence homology to one with a known X-ray structure, direct structural inferences can be made. At the same time, it has been shown that three-dimensional structural similarities based on X-ray structural analysis can be seen among proteins which have as little as 18% amino acid similarity. HCA can show similarities among proteins for which there is no X-ray data. HCA compares hydrophobic regions between two proteins while marking regions with prolines and glycines as representing possible loops and positions cystines to allow disulphide bonds. In addition to representing similarities visually, HCA can be used to obtain a numerical homology score between two proteins. This is done by comparing the number of hydrophobic residues that are in correspondence between two sequences as follows:

$$\text{HCA homology score } (\%) = \frac{2\text{CR} \times 100}{\text{RC1} + \text{RC2}}$$

where RC1 and RC2 are the number of hydrophobic residues in each protein and CR is the number of hydrophobic residues which are in correspondence between the two proteins (Gaboriaud et al. 1987). HCA shows an 80% homology in the hydrophobic regions between human hemoglobin α-chain and lupine leghemoglobin, although the two proteins have less than 15% sequence identity.

Analysis of Selection

We were interested in knowing whether the two *NbMnSOD*s have evolved under different selective constraints. Ka/Ks analysis is based on the fact that synonymous (Ks) or silent mutations in the coding regions (exons) of genes are much more frequent than non-synonymous (Ka) mutations which result in an amino acid change in the

protein. The analysis involves comparing the same genes from two organisms. Ka values smaller than Ks values (Ka/Ks < 1) imply that a gene is under purifying selection and may be functional and under constraint not to change at the amino acid level. Ka values closer to Ks values (Ka/Ks $= 1$) implies that the gene in question is either not a functional gene or that it is not constrained and is undergoing more rapid change at the structural level. Ka values greater than Ks values (Ka/Ks > 1), indicate that positive selection may be acting on the gene and that a new functionality may have evolved. Nucleotide sequences were aligned to correspond to the amino acid alignment described above and shown in Fig. 3. Ka/Ks ratios were calculated by the 2p method of (Kimura 1980) using the K-Estimator software (Comeron 1999). Because the *A. locustae* MnSOD sequence was incomplete, nucleotide sequences from the C-domain of the *N. bombycis*, *E. cuniculi*, and *A. locustae* MnSODs were compared. We also analyzed the selection of each individual codon for three complete MnSOD sequences (both *NbMnSOD*s and *EcMnSOD*) using both the modified branch-site model A and site models M1a, M2a, M7, and M8 of the CODEML program from the PAML package (Yang 1997). Posterior probabilities were calculated using the Bates empirical Bates (BEB) method (Yang 1997). In order to investigate whether the rate of evolutionary change was greater for one *N. bombycis MnSOD* gene than for the other, relative-rate tests were carried out using the RRTree program (Robinson-Rechavi and Huchon 2000), with the MnSOD sequence of *E. coli* (GenBank accession number P00448) as an outgroup.

Phylogenetic Analysis

The phylogenetic analysis included all 72 SOD amino acid sequences from the alignment described above and given in Supplemental Table S2. The most suitable substitution model, WAG with an alpha of 0.80, was determined by ProtTest software (Abascal et al. 2005). Phylogenetic analysis was done using all sites of SOD sequences using the maximum-likelihood Phyml software (Guindon and Gascuel 2003). Bootstrap values were obtained using 500 replicates. In order to check for LBA artifacts, the method of long-branch abstractions (Siddall and Whiting 1999) was used.

Results

Tandem *Nosema bombycis* MnSODs and Their Genomic Context

A search of the *N. bombycis* genomic dataset revealed a tandem array of two MnSOD genes. Each gene has a length of 678 bp, coding for 225 amino acid residues, with an intergenic region of 788 bp separating the two genes. Figure 1 shows a 9,227 bp region at one end of the superscaffold, which contains *NbMnSOD1* and *NbMnSOD2* and four genes of unknown function. There is a possible transposition-mediated insertion region, directly downstream of the *NbMnSOD*s relative to the arrangement of the MnSOD gene found on chromosome XI of *E. cuniculi* (Fig. 1).

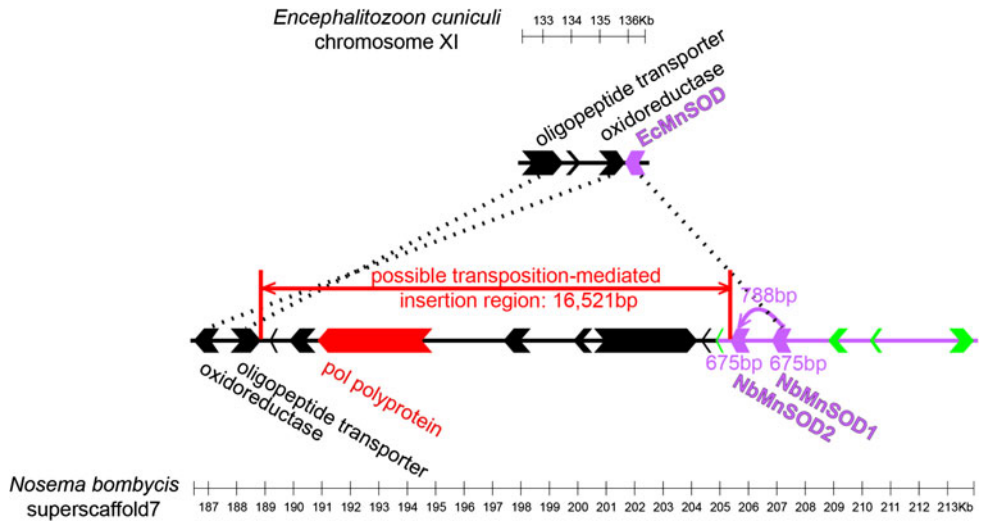

Fig. 1 Genomic context of the tandem duplication of *NbMnSOD* with a weak synteny between the *N. bombycis* superscaffold7 and that of *E. cuniculi* chromosome XI. The orientation and position of *NbMnSOD1*, *NbMnSOD2*, and *EcMnSOD* are shown in *gray* (*purple*). The pol polyprotein and possible transposition-mediated insertion region are illustrated in *gray* (*red*). The four coding sequences in *light gray* (*green*) are functionally unknown genes and have no homology to any other genes, based on a BLASTX search. *Arrows* represent gene positions and transcriptional orientation in the genome. The *scale bar* is in kilobases (Color figure online)

A complete structure containing the N-domains and C-domains for both proteins is predicted based on our analysis using the SMART (http://smart.embl-heidelberg.de/) and PDB (http://www.rcsb.org/pdb/) databases. We have also found two independent candidate promoter regions upstream from each gene through SIB-EPD (http://www.epd.isb-sib.ch/), PROSCAN (http://www-bimas.cit.nih.gov/molbio/proscan/), and PLACE (http://www.dna.affrc.go.jp/PLACE/signalscan.html) analyses. The amino acid identity between the two *NbMnSOD*s is 55.66%. BLASTP searches of the genomes of *E. cuniculi* and *A. locustae* revealed that both organisms have only one copy of the MnSOD gene and their identities were 45.54 and 45.69%, respectively, to *NbMnSOD1* and 36.20 and 37.61% to *NbMnSOD2*.

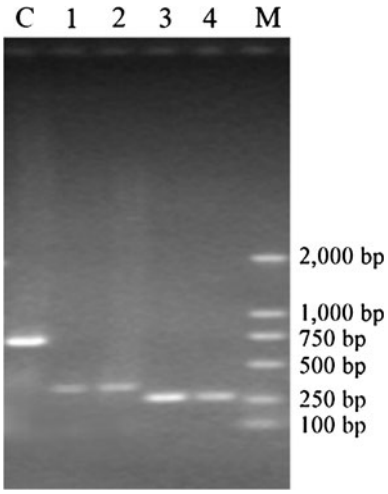

Fig. 2 Transcriptional activity of *NbMnSOD1* and *NbMnSOD2* in vivo. *Lane C* amplification of *N. bombycis* β-tubulin using sscDNA as the template as an internal control (739 bp). *Lanes 1*, *2* amplification of *NbMnSOD2* using genomic and sscDNA as template, respectively; *Lanes 3*, *4* amplification of *NbMnSOD1* using genomic and sscDNA as template, respectively; *Lane M* DL2000 marker

Transcription of MnSODs and Selection Analysis

Primers specific for *NbMnSOD1* and *NbMnSOD2*, respectively, were used to amplify *N. bombycis* cDNA and genomic DNA isolated from *N. bombycis* spores. Figure 2 shows a single band of the expected size (268 bp for *NbMnSOD1*; 337 bp for *NbMnSOD2*) for both the cDNA and genomic DNA for each of the *NbMnSOD*s.

Table 1a shows that the Ka/Ks ratios calculated for the C-domain of each of the *NbMnSOD*s in a comparison with the C-domain of two other Microsporidia (*E. cuniculi* and *A. locustae*) are less than 1. This is indicative of purifying selection to eliminate deleterious mutations to keep the protein from changing (Hurst 2002). This agrees with the evidence from the cDNA amplification experiments (Fig. 2) indicating that the two *NbMnSOD*s are transcribed and function in vivo. The bigger Ks values and smaller Ka values for *NbMnSOD1* than for *NbMnSOD2* (Table 1a) may indicate that *NbMnSOD2* is less constrained and evolving more rapidly than *NbMnSOD1*.

The relative-rate tests of microsporidial MnSODs are shown in Table 1b. Although relative rates could not be computed for the Ks values, the relative-rate tests for the Ka and amino acid values of *NbMnSOD2* compared to the other microsporidial MnSODs are significant ($P < 0.05$), suggesting that *NbMnSOD2* evolves faster than *NbMnSOD1*. Positive selection tests and posterior probabilities were calculated for the two *NbMnSOD*s and *EcMnSOD* using the PAML package (Table 2). Branch-site model A reveals several positively selected codons within the foreground branch of *NbMnSOD2* (Table 2).

Table 1 Non-synonymous (Ka) and synonymous (Ks) substitution rates (a) and the *P* value of relative-rate tests (b) for two *NbMnSOD*s

	EcMnSOD			*AlMnSOD*		
	Ka	Ks	Ka/Ks	Ka	Ks	Ka/Ks
(a)						
NbMnSOD1	0.50339	0.83575	0.60232	0.48256	1.40801	0.34272
NbMnSOD2	0.65128	0.82947	0.78518	0.68938	1.01282	0.68065
	EcMnSOD			*AlMnSOD*		
	Ka	Ks	Amino acid	Ka	Ks	Amino acid
(b)						
NbMnSOD1	0.54923	NC	1.00000	0.34364	NC	0.57728
NbMnSOD2	0.01149*	NC	0.03467*	0.02669*	NC	0.11967

The values are the comparisons of each *NbMnSOD* with *EcMnSOD* and *AlMnSOD*, respectively, using their nucleotide sequences from the C-domain, because the sequence of *AlMnSOD* is incomplete

NC not computable

* $P < 0.05$

Table 2 Positive selection tests for three microsporidian MnSODs (*NbMnSOD1*, *NbMnSOD2*, and *EcMnSOD*), calculated using the CODEML program from the PAML package

Model	*l*	Parameters	Positively selected codons
Site models			
M0: one ratio	−2277.61	$\omega = 0.14$	Not allowed
M1a: neutral	−2255.41	$p_0 = 0.57$, $p_1 = 0.43$	Not allowed
M2a: selection	−2255.41	$p_0 = 0.57$, $p_1 = 0.29$, $p_2 = 0.14$, $\omega_2 = 1.00$	None
M7: beta	−2246.45	$p = 0.57$, $q = 1.76$	Not allowed
M8: beta and ω	−2246.45	$p_0 = 1.00$, $p = 0.57$, $q = 1.76$, ($p_1 = 0.00$), $\omega_s = 3.95$	None
Branch-site models			
Null model A ($\omega_2 = 1$ fixed)	−2246.60	$p_0 = 0.41$, $p_1 = 0.25$, ($p_2 + p_3 = 0.34$)	Not allowed
Model A	−2244.77	$p_0 = 0.51$, $p_1 = 0.33$, ($p_2 + p_3 = 0.16$), $\omega_2 = 40.15$	12F 17H 31F 41W 69F 112V 114L 116S 126D 133H **136Q** **140A** 141L **144I** 160N 178I 182L 189H 197N 200L **203E**

Posterior probabilities were calculated using the Bayes empirical Bayes (BEB) method. All codons included in the table have posterior probabilities >0.50, and codons in bold have posterior probabilities >0.95

Alignment of Sequences

The metal specificity and oligomeric state of 72 SOD enzymes was predicted based the known X-ray structures of 11 SODs included in the alignment. Fifty-five MnSODs (28 tetramers, 25 dimers, and 2 archaebacterial SODs of unknown oligomeric state), 1 dimeric cambialistic SOD, and 16 FeSODs (2 tetramers and 14 dimers) were analyzed (Supplementary Table S1).

Figure 3 shows an amino acid alignment of 12 of the above sequences, including three MnSODs for which the structure has been determined based on X-ray diffraction studies (*H. sapiens* (tetramer; mitochondria), *A. fumigatus* (tetramer), and *E. coli* (dimer)). Figure 3 also includes three microsporidial sequences (the two MnSODs from *N. bombycis* and the *E. cuniculi* MnSOD) and pairs of MnSODs from *Phytophthora nicotianae* (*PnMnSOD1* (dimer) and *PnMnSOD2* (tetramer)) and *Nematostella vectensis* (*NvMnSOD1* (dimer) and *NvMnSOD2* (tetramer)). Two protistan FeSODs from *Tetrahymena thermophila* (*TtFeSOD*) and *Tetrahymena pyriformis* (*TpFeSOD*) (Barra et al. 1990; Takao et al. 1991; Wintjens et al. 2004), which show a sister group relationship to the MnSODs (Fig. 5), are also included in Fig. 3.

The 12 SODs shown in Fig. 3 all have the conserved metal binding motif DxWEHAYYx (box) as well as 4 additional constant amino acid positions (Gly_{79}, Gly_{80}, Gln_{160} (with the exception of *NbMnSOD2,* which has Asn_{160}), and Asp_{161}) common to MnSODs and the *Tetrahymena* FeSODs (Parker and Blake 1988). Two additional conserved amino acids (Met_{24}, Phe_{87}) are specific to the manganese-containing SODs (Parker and Blake 1988; Jackson and Cooper 1998; Wintjens et al. 2004).

The oligomeric structure for three of the MnSODs shown in the Fig. 3 alignment has been determined by X-ray diffraction studies (*A. fumigatus* and *H. sapiens,* tetrameric; *E. coli*, dimeric:). *NbMnSOD1*, *PnMnSOD1*, *NvMnSOD1*, and *EcMnSOD* have several amino acid residues found in dimeric SODs, while *PnMnSOD2*, *NvMnSOD2*, *TtFeSOD*, and *TpFeSOD* have several amino acid residues found in the tetrameric SODs. In MnSOD dimers, the Asn_{76} residue and the aromatic Phe_{136} residue are linked to form an interchain aromatic–polar interaction across the dimer interface situated not far from the entrance of the main substrate channel. This interaction plays an important role in the structure and function of MnSODs (Wintjens et al. 2004). The *NbMnSOD2* protein shows an interesting co-variation at these key positions: Asn_{76} is replaced by the aromatic amino acid Tyr_{76}, while the aromatic Phe_{136} is replaced by Gln_{136}, allowing for a Tyr_{76}–Gln_{136} interaction similar to the Phe_{76}–Gln_{136} aromatic–polar interaction seen in tetramers.

Hydrophobic Cluster Analysis

Figure 4 is an HCA alignment of the 12 SODs shown in Fig. 3. The tetrameric SODs are on the left and the dimeric SODs are on the right. The presence of helix 2 between residue numbers 50 and 60 (highlighted) in the microsporidial forms implies that they are dimeric (Lah et al. 1995); in general, this helix is absent from tetrameric MnSODs (Borgstahl et al. 1992; Cooper et al. 1995).

Our calculations of HCA homology scores are presented in Table 3. Table 3a illustrates a greater similarity among MnSODs with similar quaternary structure. Tetramers show significantly higher HCA values when compared with

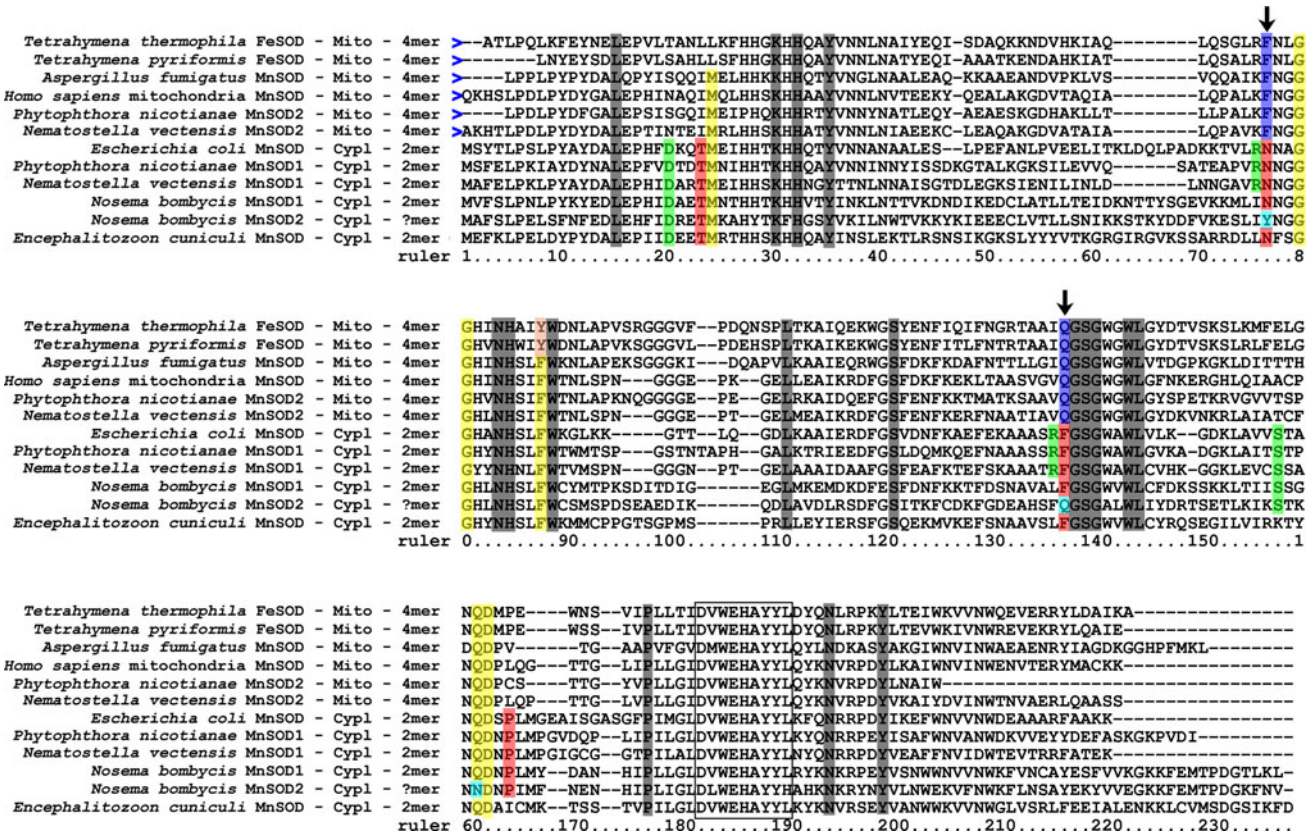

Fig. 3 Structural alignment of 12 SOD proteins. Amino acid sequences from *NbMnSOD1*, *NbMnSOD2*, *EcMnSOD*, *PnMnSOD1*, *PnMnSOD2*, *NvMnSOD1*, *NvMnSOD2*, *TtFeSOD*, and *TpFeSOD* were aligned with 3 MnSODs for which the X-ray structure has been determined (*Homo sapiens* mitochondria, *Aspergillus fumigatus*, and *Escherichia coli*). Residues conserved in 100% of the SOD sequences are in *dark gray*. Residues common to MnSODs (Met_{24}, Gly_{79}, Gly_{80}, Phe_{87}, Gln_{160}, and Asp_{161}) are in *yellow*, while Tyr_{87}, found in the two FeSODs, is in *pink*; residues common to dimers (Asp_{20}, Arg_{75}, Arg_{135}, and Ser_{156}) are in *green*. Residues Thr_{23}, Asn_{76}, Phe_{136}, and Pro_{163} in *red* are systematically encountered in dimers and never in tetramers, whereas Phe_{76} and Gln_{136} (in *dark blue*) are conserved among tetramers. *NbMnSOD2* residues Tyr_{76} and Gln_{136} (*light blue*) show a covariant deviation compared to other dimeric SODs, while Asn_{160} (also in *light blue*) shows a deviation from all other SODs presented. The *forward arrows* at position 0 represent the omitted amino-terminal extensions found in 6 of the SODs shown here. These extensions were identified by our analysis as mitochondrial target sequences. An interchain aromatic–polar interaction formed by the amino acid residues at sites 76 and 136 is indicated by the *black vertical arrows*. The conserved metal binding site is *boxed*. Residue numbers are indicated *below* the sequence (Color figure online)

other tetramers (P value of T test = 0.0084, <0.01) and dimers show higher values when compared with other dimers (P value of T test = 0.0327, <0.05).

The HCA homology scores for *NbMnSOD1* compared to the other SODs shown in Table 3b are significantly higher (P value of T test = 0.0033, <0.01) than those for *NbMnSOD2* compared to the other SODs of Table 3b. This agrees with the phylogenetic and Ka/Ks analyses, indicating that *NbMnSOD2* has changed more rapidly than *NbMnSOD1* since the time of duplication and may have developed a slightly different function.

Subcellular Localization

Subcellular localization of 72 SODs was predicted using PSORT II (Nakai and Horton 1999), TargetP (Nakai and Horton 1999) and MitoProt II (Claros and Vincens 1996). Target sequences were identified for 24 of the MnSODs (12 from fungi, 7 from metazoa, 2 from Plantae, and 3 from cyanobacteria) and 6 of the protistan FeSODs (Supplementary Table S1). The signals for *PnMnSOD2* and *TpFeSOD* could not be determined in our analysis due to incomplete sequence information but previous studies have shown that these proteins are located in the mitochondria (Barra et al. 1990; Blackman et al. 2005).

The microsporidial MnSODs *NbMnSOD1*, *NbMnSOD2*, and *EcMnSOD* show no terminal extensions. The fungi typically possess two MnSODs, one targeted for the mitochondria and one without a target sequence. Similarly, both the oomycete *Phytophthora nicotianae* and the cnidarian *Nematostella vectensis* have two MnSODs, one targeted for the mitochondria and one lacking a target sequence. Interestingly, two of the cyanobacterial MnSODs with target sequences showed mitochondrial-type target

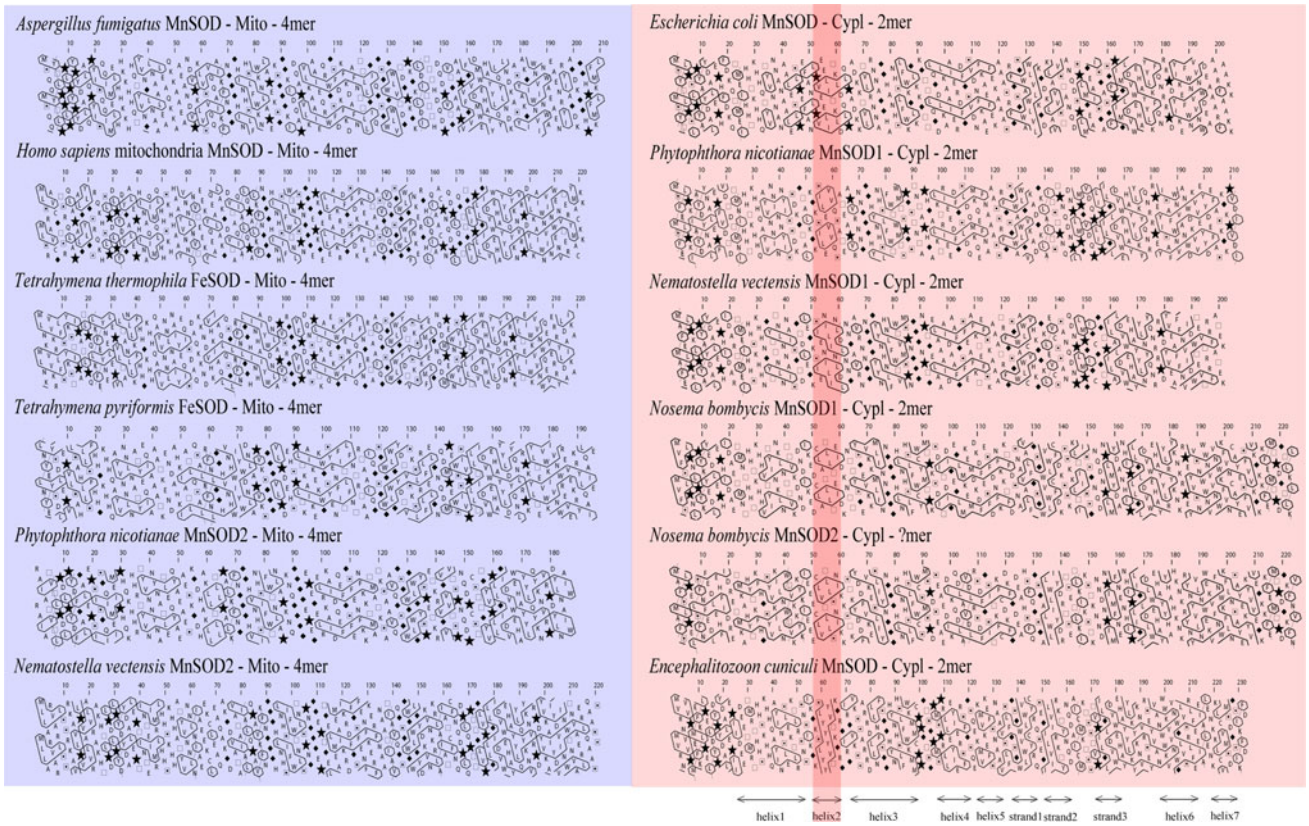

Fig. 4 This is an HCA alignment of the 12 SODs shown in Fig. 3, including *Homo sapiens* mitochondria, *Aspergillus fumigatus*, and *Escherichia coli* which have known X-ray structures. Tetrameric forms are on left, while dimeric forms are on right. Amino acids are symbolized by their single letter code with the following exceptions: threonine (*open square*), serine (*square with dot*), glycine (*black diamond*), and proline (*black star*). The HCA plot is created by placing the sequence on a cylinder of 3.6 amino acids per turn. After 5 turns, residue i and i + 18 have the same position on the cylinder. The cylinder is then cut lengthwise and unrolled (see Fig. 4). Since this would make some adjacent residues widely separated, the cylinder is duplicated to allow visualization of adjacent residues (Gaboriaud et al. 1987). Groups of *encircled* amino acids represent clusters of hydrophobic residues (*W*, *Y*, *M*, *F*, *I*, *L*, *V* are considered as hydrophobic amino acids). Seven helices and three strands are shown. Dimer-specific helix 2 is shaded. Amino acid residue numbers are indicated *above* each sequence

sequences and one showed a chloroplast-type target sequence.

Phylogenetic Analysis

Figure 5 shows the most likely tree inferred from the 72 SOD amino acid sequences. The analysis indicates that *NbMnSOD1* and *NbMnSOD2* sequences are more closely related to each other than to any other sequence (bootstrap values of 100%). The monophyly of Microsporidia is strongly supported by the bootstrap values (85% of the replicates).

Interestingly, the Microsporidia do not cluster with fungi but resolve as a sister taxon to the alpha- and delta/epsilon-proteobacteria albeit with low bootstrap values (43%). Our analysis indicates that LBA artifact is not a factor and therefore would not explain the divergent nature of the microsporidial MnSODs.

The MnSODs from the Archaea form a separate group and thus the MnSOD phylogeny separates these taxa (with some exceptions) into the three domains of life (Woese et al. 1990).

Figure 5 shows a fundamental dichotomy between the FeSODs and the MnSODs. In addition, we can see a clear distinction between the bacterial MnSODs, which are dimeric, and those from eukaryotes, which are usually tetrameric. *Phytophthora nicotianae* and *N. vectensis* each have two MnSODs. *PnMnSOD1* and *NvMnSOD1* branch within the bacteria while *PnMnSOD2* and *NvMnSOD2* cluster together with Plantae and metazoa. Thus, these two organisms have both the "prokaryotic" dimeric and "eukaryotic" tetrameric forms of MnSOD.

Figure 5 also shows two distinct clades of fungal MnSODs, one targeted to the cytosol and the other to the mitochondria. It appears that for these SODs, phylogenetic position correlates reasonably well with protein structure, subcellular localization, and metal cofactor.

Table 3 The HCA homology scores and BLASTP identity scores (in parentheses) of the 12 SODs shown in Fig. 4

HCA scores (identity scores)	*Homo sapiens* mitochondria MnSOD—Mito—4mer	*Aspergillus fumigatus* MnSOD—Mito—4mer	*Escherichia coli* MnSOD—Cypl—2mer
(a) A comparison of the 3 MnSODs with known X-ray structure to the remaining 9 SODs			
Nematostella vectensis MnSOD2—Mito—4mer	91% (76%)	76% (50%)	73% (44%)
Tetrahymena thermophila FeSOD—Mito—4mer	83% (46%)	78% (44%)	73% (38%)
Phytophthora nicotianae MnSOD2—Mito—4mer	82% (64%)	78% (50%)	78% (43%)
Tetrahymena pyriformis FeSOD—Mito—4mer	79% (45%)	77% (44%)	74% (38%)
Nematostella vectensis MnSOD1—Cypl—2mer	79% (48%)	75% (40%)	86% (53%)
Phytophthora nicotianae MnSOD1—Cypl—2mer	80% (47%)	76% (38%)	82% (52%)
Nosema bombycis MnSOD1—Cypl—2mer	73% (46%)	72% (37%)	79% (45%)
Nosema bombycis MnSOD2—Cypl—?mer	68% (36%)	68% (31%)	74% (33%)
Encephalitozoon cuniculi MnSOD—Cypl—2mer	70% (41%)	70% (36%)	74% (43%)

HCA scores (identity scores)	*Nosema bombycis* MnSOD1—Cypl—2mer	*Nosema bombycis* MnSOD2—Cypl—?mer
(b) A comparison of *NbMnSOD1* and *NbMnSOD2* with the remaining 10 SODs		
Nematostella vectensis MnSOD2—Mito—4mer	74% (45%)	64% (36%)
Tetrahymena pyriformis FeSOD—Mito—4mer	71% (38%)	66% (29%)
Homo sapiens mitochondria MnSOD—Mito—4mer	73% (46%)	68% (36%)
Phytophthora nicotianae MnSOD2—Mito—4mer	70% (41%)	64% (31%)
Aspergillus fumigatus MnSOD—Mito—4mer	72% (37%)	68% (31%)
Tetrahymena thermophila FeSOD—Mito—4mer	68% (37%)	65% (31%)
Encephalitozoon cuniculi MnSOD—Cypl—2mer	83% (46%)	75% (36%)
Nematostella vectensis MnSOD1—Cypl—2mer	80% (47%)	71% (34%)
Escherichia coli MnSOD—Cypl—2mer	79% (45%)	74% (33%)
Phytophthora nicotianae MnSOD1—Cypl—2mer	78% (45%)	71% (35%)

Discussion

A tandem duplication of the manganese superoxide dismutase gene in *N. bombycis* has been identified and seems to have occurred at some time since *N. bombycis* diverged from the common ancestor with *E. cuniculi* (Fig. 1). Both *NbMnSOD1* and *NbMnSOD2* transcripts are present in the spores, indicating that both isozymes are expressed. Ka/Ks calculations and relative-rate tests obtained from a comparison of the nucleotide sequences of the two *NbMnSOD*s with the nucleotide sequences of the *A. locustae* and *E. cuniculi* MnSODs indicate that *NbMnSOD2* is under less constraint from evolutionary change and evolves faster than *NbMnSOD1* (Table 1). Hydrophobic cluster analysis (Fig. 4) shows that both *NbMnSOD1* and *NbMnSOD2* have the alpha helix (between positions 50 and 60) typically observed for dimeric MnSODs. Pairwise comparison of HCA homology scores (Table 3), however, shows larger differences between *NbMnSOD2* and the other SODs than between *NbMnSOD1* and the other SODs, providing further evidence that *NbMnSOD2* is under less evolutionary constraint than *NbMnSOD1*.

An amino acid alignment of 6 tetrameric and 6 dimeric SODs (Fig. 3) shows several significant changes in the primary structure of *NbMnSOD2* compared to *NbMnSOD1* and the other MnSODs. *NbMnSOD2* has an asparagine at position 160, unlike the other MnSODs shown in Fig. 3, which all have a glutamine at this position. Particularly notable amino acids changes occur at residues 76 and 136, which form an interchain aromatic–polar interaction in both dimers (Asn_{76} and Phe_{136}) and tetramers (Phe_{76} and Gln_{136}). Unlike the other MnSODs in Fig. 3, *NbMnSOD2* has a tyrosine at position 76, and a glutamine at position 136. The codons for residue 76 in *EcMnSOD*, *NbMnSOD1*, and *NbMnSOD2* are AAC, AAT, and TAC, respectively. We presume that AAC is the ancestral codon (present in most dimeric MnSODs), and that a synonymous mutation to AAT occurred in *NbMnSOD1*, while a nonsynonymous mutation to TAC occurred in *NbMnSOD2*. The change from Gln_{136} to Phe_{136} in *NbMnSOD2* would allow the important interchain aromatic-polar interaction between residues 76 and 136 to be conserved, and as Table 2 shows, Gln_{136} in *NbMnSOD2* is under strong positive selection, with a posterior probability >0.95.

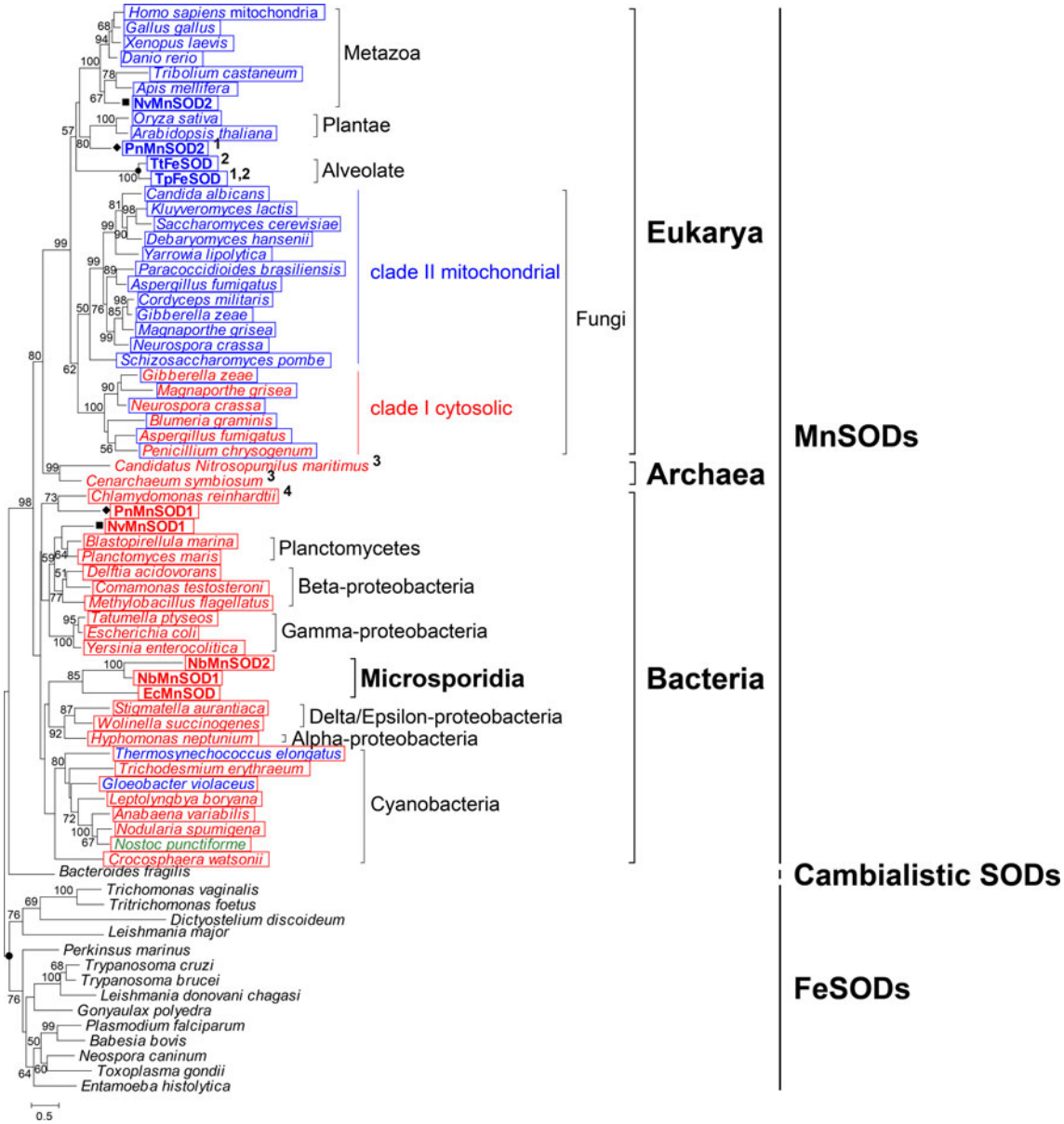

Fig. 5 A phylogenetic tree of 72 SOD proteins. Accession numbers of the SOD sequences retrieved from NCBI are shown in the Supplementary Table S1. *Numbers* near the branching nodes indicate bootstrap values given as percentages. Nodes with values of <50% are not shown. *Names in bold* are MnSODs from the Microsporidia *N. bombycis* and *E. cuniculi*. *Black circles* indicate the protistan FeSODs, *diamonds* indicate the two MnSODs of *P. nicotianae*, and *squares* indicate the two MnSODs of *N. vectensis*. Predicted cytosolic, mitochondrial, and chloroplast SODs are shown in *red*, *blue*, and *green letters*, respectively. Predicted dimeric or tetrameric structure of MnSODs is indicated by *red* or *blue boxes*, respectively. [1] *PnMnSOD2* and *TpFeSOD* are located in the mitochondria (Barra et al. 1990; Blackman et al. 2005). [2] *TtFeSOD* and *TpFeSOD* are two iron-containing SOD enzymes, but they are similar to MnSODs in amino acid sequence. [3] Quaternary not determined for the two Archaea MnSODs. [4] This organism is a green algae which belong to the Kingdom Plantae (Color figure online)

Based on these analyses, a more detailed investigation of *NbMnSOD2* to determine its physical and enzymatic properties is warranted. This would allow us to understand whether the *NbMnSOD* gene duplication represents an added dose effect to enhance antioxidant efficiency in *N. bombycis,* or if the duplication of this gene has resulted in the acquisition of a new gene function (Ohno 1970). In plants multiple forms of SODs are present which vary based on their location within cells, tissues and organelles. Changes in plant SOD isozymes occur during development and in response to environmental factors, suggesting different roles for these SOD isozymes (Scandalios 1993). In animals different SODs are expressed at different levels depending on tissue type (Marklund 1984), indicating a

need for SODs with different functions. The two *NbMn-SOD* isozymes may be expressed at different times in spore and meront development, in different insect host tissues, or at different temperatures.

Phylogenetic analysis reveals that MnSODs group according to subcellular localization, quaternary structure, and taxonomic relationship. There are clear groupings separating those proteins that are targeted for the cytosol (Fig. 5, red lettering) from those that are targeted for the mitochondria (Fig. 5, blue lettering). The ascomycote fungi presented in Fig. 5 have tetrameric MnSODs targeted for the cytosol and the mitochondria. Both types of MnSOD (cytosolic and mitochondrial) have been identified in several of the fungi shown, suggesting that gene duplication occurred before the radiation of these species. We also see a sharp dichotomy between the dimeric (primarily bacterial) MnSODs (Fig. 5, red boxes) and the tetrameric (eukaryotic) MnSODs (Fig. 5, blue boxes). In addition, it is clear from Fig. 5 that many of the eukaryotic MnSODs are targeted for the mitochondria (blue lettering), while most of the bacterial and a number of the fungal MnSODs presented here are located in the cytosol (red lettering). Several exceptions are seen in the cyanobacteria, two of which (*Thermosynechococcus elongatus* and *Gloeobacter violaceus*) have mitochondrial target sequences and one of which (*Nostoc punctiforme*) has a chloroplast target sequence. The oomycete *P. nicotianae* and the cnidarian *N. vectensis* each have both a tetrameric mitochondrial and a dimeric "prokaryotic" MnSOD (PnMnSOD1, 2 and NvMnSOD1, 2, respectively). The final noteworthy exception is the Microsporidia, which group weakly with the bacteria and have dimeric MnSODs.

Our phylogenetic analysis (Fig. 5) correlates well with what we know about the relationships among these organisms. We see a clear separation of the 14 protistan FeSODs which we use here as the outgroup for the remaining, primarily manganese SODs. Figure 5 shows a clear division of the 3 domains of life (Woese et al. 1990). Our analysis groups the metazoa together arranging the vertebrates in the expected order. The fungi for which the SOD gene has been sequenced all belong to the Phylum Ascomycota, and show relationships which vary only in minor respects to those previously published for a six-gene phylogeny (Thomarat et al. 2004; Gill and Fast 2006; James et al. 2006). The bacteria group together into previously established clades, and phylogenetic relationships among the cyanobacteria presented here are in agreement with those determined using ssrDNA (Tomitani et al. 2006) and whole genomes (Zhaxybayeva et al. 2006). We believe that the phylogeny presented here, based on MnSOD analysis, shows a remarkable amount of resolution among a wide range of taxa.

There are several possible explanations regarding the placement of the Microsporidia in Fig. 5. One possibility is that the Microsporidia obtained MnSOD through horizontal gene transfer (HGT) from a proteobacterium (perhaps from the common ancestor of the alpha and delta/epsilon group). As long as the Microsporidia are considered to be derived fungi (Fast et al. 2003; Keeling et al. 2005), HGT to the Microsporidia will remain an explanation for the numerous genes which are very different from those of the fungi. Another possibility is that the microsporidial MnSOD is a divergent eukaryotic gene. The bootstrap value of Fig. 5 linking this gene to the proteobacteria is not high (43%). In none of the analyses, with any of the alignments obtained, did the MnSODs from the Microsporidia group with those of the fungi, or show a sister relationship with the fungi presented.

Acknowledgments This study was supported by the project of Chongqing Science & Technology Commission (CSTC, 2006AA5019), National Basic Research Program of China under the Grant no. 2005CB121000, the Programme of Introducing Talents of Discipline to Universities (B07045) and State Development Fund at Risk of Callus silk (M012005-000Y-00070). The genomic database was obtained from the *N. bombycis* genomic sequencing project. The authors appreciate Profs. Bettina Debrunner-Vossbrinck and Ze Zhang for assistance in improving this manuscript.

References

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104–2105

Alscher RG, Erturk N, Heath LS (2002) Role of superoxide dismutases (SODs) in controlling oxidative stress in plants. J Exp Bot 53:1331–1341

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Babitha MP, Bhat SG, Prakash HS, Shetty HS (2002) Differential induction of superoxide dismutase in downy mildew-resistant and -susceptible genotypes of pearl millet. Plant Pathol 51: 480–486

Bagnoli F, Capuana M, Raccfii ML (1998) Developmental changes of catalase and superoxide dismutase isoenzymes in zygotic and somatic embryos of horse chestnut. J Plant Physiol 25:909–913

Bagnoli F, Giannino D, Caparrini S, Camussi A, Mariotti D, Racchi ML (2002) Molecular cloning, characterisation and expression of a manganese superoxide dismutase gene from peach (*Prunus persica* [L.] Batsch). Mol Genet Genomics 267:321–328

Barra D, Schinina ME, Bossa F, Puget K, Durosay P, Guissani A, Michelson AM (1990) A tetrameric iron superoxide dismutase from the eucaryote *Tetrahymena pyriformis*. J Biol Chem 265:17680–17687

Belinky PA, Goldberg D, Krinfeld B, Burger M, Rothschild N, Cogan U, Dosoretz CG (2002) Manganese-containing superoxide dismutase from the white-rot fungus *Phanerochaete chrysosporium*: its function, expression and gene structure. Enzyme Microb Tech 31:754–764

Blackman LM, Mitchell HJ, Hardham AR (2005) Characterisation of manganese superoxide dismutase from *Phytophthora nicotianae*. Mycol Res 109:1171–1183

Bond CJ, Huang J, Hajduk R, Flick KE, Heath PJ, Stoddard BL (2000) Cloning, sequence and crystallographic structure of recombinant iron superoxide dismutase from *Pseudomonas ovalis*. Acta Crystallogr 56:1359–1366

Bordo D, Djinovic K, Bolognesi M (1994) Conserved patterns in the Cu, Zn superoxide dismutase family. J Mol Biol 238:366–386

Borgstahl GE, Parge HE, Hickey MJ, Beyer WF Jr, Hallewell RA, Tainer JA (1992) The structure of human mitochondrial manganese superoxide dismutase reveals a novel tetrameric interface of two 4-helix bundles. Cell 71:107–118

Borgstahl GEO, Pokross M, Chehab R, Sekher A, Snell EH (2000) Cryo-trapping the six-coordinate, distorted-octahedral active site of manganese superoxide dismutase1. J Mol Biol 296:951–959

Brouwer M, Brouwer TH, Grater W, Enghild JJ, Thogersen IB (1997) The paradigm that all oxygen-respiring eukaryotes have cytosolic CuZn-superoxide dismutase and that Mn-superoxide dismutase is localized to the mitochondria does not apply to a large group of marine arthropods. Biochemistry 36:13381–13388

Brouwer M, Hoexum Brouwer T, Grater W, Brown-Peterson N (2003) Replacement of a cytosolic copper/zinc superoxide dismutase by a novel cytosolic manganese superoxide dismutase in crustaceans that use copper (haemocyanin) for oxygen transport. Biochem J 374:219–228

Canning EU (1993) Microsporidia. In: Kreier JP (ed) Parasitic protozoa. Academic Press, New York

Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem 241:779–786

Comeron JM (1999) K-Estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals. Bioinformatics 15:763–764

Cooper JB, McIntyre K, Badasso MO, Wood SP, Zhang Y, Garbe TR, Young D (1995) X-ray structure analysis of the iron-dependent superoxide dismutase from *Mycobacterium tuberculosis* at 2.0 Angstroms resolution reveals novel dimer-dimer interactions. J Mol Biol 246:531–544

Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673–679

Desportes I, Le Charpentier Y, Galian A, Bernard F, Cochand-Priollet B, Lavergne A, Ravisse P, Modigliani R (1985) Occurrence of a new microsporidan: *Enterocytozoon bieneusi* n.g., n. sp., in the enterocytes of a human patient with AIDS. J Protozool 32:250–254

Diez B, Schleissner C, Moreno MA, Rodriguez M, Collados A, Barredo JL (1998) The manganese superoxide dismutase from the penicillin producer *Penicillium chrysogenum*. Curr Genet 33:387–394

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797

El-Taweel HA, El-Zawawy LA, Said DE, Sharara GM (2007) Influence of the antioxidant drug (Antox) on experimental giardiasis and microsporidiosis. J Egypt Soc Parasitol 37:189–204

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300:1005–1016

Fang GC, Hanau RM, Vaillancourt LJ (2002) The SOD2 gene, encoding a manganese-type superoxide dismutase, is up-regulated during conidiogenesis in the plant-pathogenic fungus *Colletotrichum graminicola*. Fungal Genet Biol 36:155–165

Fast NM, Law JS, Williams BA, Keeling PJ (2003) Bacterial catalase in the microsporidian *Nosema locustae*: implications for microsporidian metabolism and genome evolution. Eukaryot Cell 2:1069–1075

Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. Syst Biol 27:401–410

Fitch WM, Ayala FJ (1994) The superoxide dismutase molecular clock revisited. Proc Natl Acad Sci USA 91:6802–6807

Fluckiger S, Mittl PRE, Scapozza L, Fijten H, Folkers G, Grutter MG, Blaser K, Crameri R (2002) Comparison of the crystal structures of the human manganese superoxide dismutase and the homologous *Aspergillus fumigatus* allergen at 2-A resolution. J Immunol 168:1267–1272

Frealle E, Noel C, Nolard N, Symoens F, Felipe MS, Dei-Cas E, Camus D, Viscogliosi E, Delhaes L (2006) Manganese superoxide dismutase based phylogeny of pathogenic fungi. Mol Phylogenet Evol 41:28–39

Fridovich I (1995) Superoxide radical and superoxide dismutases. Annu Rev Biochem 64:97–112

Gaboriaud C, Bissery V, Benchetrit T, Mornon JP (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett 224:149–155

Germot A, Philippe H (1999) Critical analysis of eukaryotic phylogeny: a case study based on the HSP70 family. J Eukaryot Microbiol 46:116–124

Giles SS, Batinic-Haberle I, Perfect JR, Cox GM (2005) Cryptococcus neoformans mitochondrial superoxide dismutase: an essential link between antioxidant function and high-temperature growth. Eukaryot Cell 4:46–54

Gill EE, Fast NM (2006) Assessing the microsporidia-fungi relationship: combined phylogenetic analysis of eight genes. Gene 375:103–109

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Hassan HM (1989) Microbial superoxide dismutases. Adv Genet 26:65–97

Hunter T, Bannister WH, Hunter GJ (1997) Cloning, expression, and characterization of two manganese superoxide dismutases from *Caenorhabditis elegans*. J Biol Chem 272:28652–28659

Hurst LD (2002) The Ka/Ks ratio: diagnosing the form of sequence evolution. Trends Genet 18:486–487

Jackson SM, Cooper JB (1998) An analysis of structural similarity in the iron and manganese superoxide dismutases based on known structures and sequences. Biometals 11:159–173

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung GH, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schussler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lucking R, Budel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R (2006) Reconstructing the early evolution of fungi using a six-gene phylogeny. Nature 443:818–822

Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivares CP (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414:450–453

Keeling PJ, Fast NM, Law JS, Williams BA, Slamovits CH (2005) Comparative genomics of microsporidia. Folia Parasitol (Praha) 52:8–14

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Kliebenstein DJ, Monde RA, Last RL (1998) Superoxide dismutase in Arabidopsis: an eclectic enzyme family with disparate regulation and protein localization. Plant Physiol 118:637–650

Knapp S, Kardinahl S, Hellgren N, Tibbelin G, Sch fer G, Ladenstein R (1999) Refined crystal structure of a superoxide dismutase from the hyperthermophilic archaeon *Sulfolobus acidocaldarius* at 2.2 resolution1. J Mol Biol 285:689–702

Lah MS, Dixon MM, Pattridge KA, Stallings WC, Fee JA, Ludwig ML (1995) Structure-function in Escherichia coli iron superoxide dismutase: comparisons with the manganese enzyme from *Thermus thermophilus*. Biochemistry 34:1646–1660

Lamarre C, LeMay JD, Deslauriers N, Bourbonnais Y (2001) Candida albicans expresses an unusual cytoplasmic manganese-containing superoxide dismutase (SOD3 gene product) upon the entry and during the stationary phase. J Biol Chem 276:43784–43791

Lim JH, Yu YG, Han YS, Cho S, Ahn BY, Kim SH, Cho Y (1997) The crystal structure of an Fe-superoxide dismutase from the hyperthermophile *Aquifex pyrophilus* at 1.9 resolution: structural basis for thermostability1. J Mol Biol 270:259–274

Lynch M, Kuramitsu H (2000) Expression and role of superoxide dismutases (SOD) in pathogenic bacteria. Microbes Infect 2:1245–1255

Marklund SL (1984) Extracellular superoxide dismutase and other superoxide dismutase isoenzymes in tissues from nine mammalian species. Biochem J 222:649–655

Martin ME, Byers BR, Olson MO, Salin ML, Arceneaux JE, Tolbert C (1986) A *Streptococcus mutans* superoxide dismutase that is active with either manganese or iron as a cofactor. J Biol Chem 261:9361–9367

McCord JM, Fridovich I (1969) Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). J Biol Chem 244:6049–6055

Nakai K, Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem Sci 24:34–35

Natvig DO, Sylvester K, Dvorachek WH, Baldwin JL (1996) Superoxide dismutases and catalases. In: Brambl R, Marzluf GA (eds) Biochemistry and molecular biology. Springer-Verlag, Berlin

Ohno S (1970) Evolution by gene duplication. Springer-Verlag, New York

Palma JM, Lopez-Huertas E, Corpas FJ, Sandalio LM, Gomez M, del Rio LA (1998) Peroxisomal manganese superoxide dismutase: purification and properties of the isozyme from pea leaves. Physiol Plant 104:720–726

Pan SM, Ye JS, Hseu RS (1997) Purification and characterization of manganese superoxide dismutase from *Ganoderma microsporum*. Biochem Mol Biol Int 42:1035–1043

Parker MW, Blake CC (1988) Iron- and manganese-containing superoxide dismutases can be distinguished by analysis of their primary structures. FEBS Lett 229:377–382

Parker MW, Blake CC, Barra D, Bossa F, Schinina ME, Bannister WH, Bannister JV (1987) Structural identity between the iron- and manganese-containing superoxide dismutases. Protein Eng 1:393–397

Regelsberger G, Atzenhofer W, Ruker F, Peschek GA, Jakopitsch C, Paumann M, Furtmuller PG, Obinger C (2002) Biochemical characterization of a membrane-bound manganese-containing superoxide dismutase from the cyanobacterium Anabaena PCC 7120. J Biol Chem 277:43615–43622

Robinson-Rechavi M, Huchon D (2000) RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. Bioinformatics 16:296

Scandalios JG (1993) Oxygen stress and superoxide dismutases. Plant Physiol 101:7–12

Schmidt M (1999) Manipulating the coordination mumber of the ferric iron within the cambialistic superoxide dismutase of *Propionibacterium shermanii* by changing the pH-value. A crystallographic analysis. Eur J Biochem 262:117–127

Siddall ME, Whiting MF (1999) Long-branch abstractions. Cladistics 15:9–24

Smith MW, Doolittle RF (1992) A comparison of evolutionary rates of the two major kinds of superoxide dismutase. J Mol Evol 34:175–184

Snowden KF (2004) Zoonotic microsporidia from animals and arthropods with a discussion of human infections. In: Lindsay DS, Weiss LM (eds) World class parasites. Kluwer Academic Press, Boston

Stallings WC, Pattridge KA, Strong RK, Ludwig ML (1984) Manganese and iron superoxide dismutases are structural homologs. J Biol Chem 259:10695–10699

Sugio S, Hiraoka BY, Yamakura F (2003) Crystal structure of cambialistic superoxide dismutase from *Porphyromonas gingivalis*. Eur J Biochem 267:3487–3495

Tainer JA, Getzoff ED, Beem KM, Richardson JS, Richardson DC (1982) Determination and analysis of the 2 A-structure of copper, zinc superoxide dismutase. J Mol Biol 160:181–217

Takao M, Yasui A, Oikawa A (1991) Unique characteristics of superoxide dismutase of a strictly anaerobic archaebacterium *Methanobacterium thermoautotrophicum*. J Biol Chem 266: 14151–14154

Thomarat F, Vivares CP, Gouy M (2004) Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. J Mol Evol 59:780–791

Tomitani A, Knoll AH, Cavanaugh CM, Ohno T (2006) The evolutionary diversification of cyanobacteria: molecular–phylogenetic and paleontological perspectives. Proc Natl Acad Sci USA 103:5442–5447

Ursby T, Adinolfi BS, Al-Karadaghi S, De Vendittis E, Bocchini V (1999) Iron superoxide dismutase from the archaeon *Sulfolobus solfataricus*: analysis of structure and thermostability. J Mol Biol 286:189–205

Van Camp W, Inze D, Van Montagu M (1997) The regulation and function of tobacco superoxide dismutases. Free Radic Biol Med 23:515–520

Vossbrinck CR, Woese CR (1986) Eukaryotic ribosomes that lack a 5.8S RNA. Nature 320:287–288

Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR (1987) Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326:411–414

Wagner UG, Pattridge KA, Ludwig ML, Stallings WC, Werber MM, Oefner C, Frolow F, Sussman JL (1993) Comparison of the crystal structures of genetically engineered human manganese superoxide dismutase and manganese superoxide dismutase from *Thermus thermophilus*: differences in dimer-dimer interaction. Protein Sci 2:814–825

Wang J, Wong GKS, Ni P, Han Y, Huang X, Zhang J, Ye C, Zhang Y, Hu J, Zhang K (2002) RePS: a sequence assembler that masks exact repeats identified from the shotgun data. Genome Res 12:824–831

Weisiger RA, Fridovich I (1973) Mitochondrial superoxide simutase. Site of synthesis and intramitochondrial localization. J Biol Chem 248:4793–4796

Williams BA, Hirt RP, Lucocq JM, Embley TM (2002) A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. Nature 418:865–869

Wintjens R, Noel C, May AC, Gerbod D, Dufernez F, Capron M, Viscogliosi E, Rooman M (2004) Specificity and phenetic

relationships of iron- and manganese-containing superoxide dismutases on the basis of structure and sequence comparisons. J Biol Chem 279:9248–9254

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576–4579

Woodcock S, Mornon JP, Henrissat B (1992) Detection of secondary structure elements in proteins by hydrophobic cluster analysis. Protein Eng 5:629–635

Xu J, Pan G, Fang L, Li J, Tian X, Li T, Zhou Z, Xiang Z (2006) The varying microsporidian genome: existence of long-terminal repeat retrotransposon in domesticated silkworm parasite *Nosema bombycis*. Int J Parasitol 36:1049–1056

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Bioinformatics 13:555–556

Zhaxybayeva O, Gogarten JP, Charlebois RL, Doolittle WF, Papke RT (2006) Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. Genome Res 16:1099–1108

Zhu D, Scandalios JG (1993) Maize mitochondrial manganese superoxide dismutases are encoded by a differentially expressed multigene family. Proc Natl Acad Sci USA 90:9310–9314