Molecular Evolution of Peptide Methionine Sulfoxide Reductases (MsrA and MsrB): On the Early Development of a Mechanism That Protects Against Oxidative Damage

Luis Delaye,[1] Arturo Becerra,[1] Leslie Orgel,[2] Antonio Lazcano[1]

[1] Facultad de Ciencias, UNAM, Apartado Postal 70-407, Cd. Universitaria, 04510 Mexico D.F., México
[2] Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

Abstract. Methionine sulfoxide reductases, enzymes that reverse the oxidation of methionine residues, have been described in a wide range of species. The reduction of the diastereoisomers of oxidized methionine is catalyzed by two different monomeric methionine sulfoxide reductases (MsrA and MsrB) and is best understood as an evolutionary response to high levels of oxygen either in the Earth's atmosphere or possibly in more localized environments. Phylogenetic analyses of these proteins suggest that their distribution is the outcome of a complex history including many paralogy and lateral gene transfer events.

Key words: Methionine sulfoxide reductases — Lateral gene transfer — Methionine oxidation

Introduction

It is generally accepted that the accumulation of free oxygen during the Precambrian era, both in the atmosphere and in local environments, must have exerted a severe selection pressure on early microbial species that had to cope with the resulting novel stress conditions. It has been argued that oxygen-induced damage must have led to a number of protection and

Correspondence to: Antonio Lazcano; *email:* alar@correo.unam.mx

detoxification mechanisms, some of which may have led to oxygen utilization as in aerobic respiration (Margulis 1993; Falkowski 2006). Although the significance of a number of molecular mechanisms that protect against oxygen-induced damage such as those involving superoxide dismutases and catalases has been acknowledged, little or no attention has been given to the evolutionary significance of the repair of oxidized methionine.

The oxidation of methionine residues in proteins leads to a diastereomeric mixture of methionine *S*-sulfoxide and methionine *R*-sulfoxide (Sharov et al. 1999; Sharov and Schoneich 2000). The reduction of the diastereoisomers of oxidized methionine can be explained by sulfenic acid chemistry (Boschi-Muller et al. 2005) catalyzed by two different monomeric methionine sulfoxide reductases (Kumar et al. 2002). The activity of methionine sulfoxide reductase A (MsrA) is restricted to methionine *S*-sulfoxide residues in proteins, while methionine sulfoxide reductase B (MsrB) reduces methionine *R*-sulfoxide residues.

The two methionine sulfoxide reductases are broadly distributed nonhomologous enzymes (Table 1) that have been known for a long time (cf. Grimaud et al. 2001). MsrA is a well-studied protein that is an essential regulator of longevity in a wide range of microbial and animal species (Ruan et al. 2002). Its role as a physiological virulence determinant in *Mycoplasma genitalium* (Dhandayuthapani et al. 2001) and the plant pathogen *Erwinia chrysanthemi* (El Hassouni et al. 1999) has been established,

Table 1. Phylogenetic distribution of MsrA and MsrB sequences in fully sequenced cellular genomes (December 2004)

				msrA	msrB	msrAB	msrBA	Complete	Lifestyle
Eukarya	Mammals	Homo sapiens	hsa	1	2			No	
		Mus musculus	mmu	1	2			No	
	Insect	Drosophila melanogaster	dme	3	1			Yes	
	Nematode	Caenorhabditis elegans	cel	1	1			Yes	
	Plant	Arabidopsis thaliana	ath	5	3			Yes	
	Protoctist	Dictyostelium discoideum	ddi	1	1			No	
	Fungi	Saccharomyces cerevisiae	sce	1	1			Yes	
		Schizosaccharomyces pombe	spo	1	1			Yes	
		Encephalitozoon cuniculi	ecu					Yes	
Bacteria	Proteobacteria γ Enterobacteriales	Escherichia coli K-12 MG 1655	eco	1	1			Yes	Facultative
		Escherichia coli K-12 W3110	ecj	1	1			Yes	
		Escherichia coli O157:H7 EDL993	ece	1	1			Yes	Facultative
		Escherichia coli O157:H7 Sakai	ecs	1	1			Yes	Facultative
		Escherichia coli CFT073	ecc	1	1			Yes	Facultative
		Salmonella typhi CT 18	sty	1	1			Yes	
		Salmonella typhi Ty2	stt	1	1			Yes	Facultative
		Salmonella typhimurium LT2	stm	1	1			Yes	Facultative
		Yersinia pestis CO92	ype	1	1			Yes	
		Yersinia pestis KIM	ypk	1	1			Yes	
		Shigella flexneri 301 (serotype 2a)	sfl	1	1			Yes	
		Shigella flexneri 2457T (serotype 2a)	sfx	1	1			Yes	
		Buchnera sp. Aps (symbiont of Acyrthosiphon pisum)	buc	1				Yes	Endosymbiont
		Buchnera aphidicola Sg (symbiont of Schizaphis graminum)	bas					Yes	Endosymbiont
		Buchnera aphidicola (symbiont of Baizongia pistaciae)	bab					Yes	Endosymbiont
		Wigglesworthia brevipalpis (symbiont of tsetse fly)	wbr	1				Yes	Endosymbiont
		Blochmannia floridanus (symbiont of carpenter ant)	bfl					Yes	Endosymbiont
		Photorhabdus luminescens laumondii TTO1	plu	1	1			Yes	Endosymbiont
	Proteobacteria γ Pasteurellales	Haemophilus influenzae Rd	hin			1		Yes	Microaerophilic
		Pasteurella multocida PM70	pmu	1	1			Yes	Facultative
		Haemophilus ducreyi 35000HP	hdu	1	1			Yes	
	Proteobacteria γ Xanthomonadales	Xylella fastidiosa 9a5c	Xfa	1	1			Yes	Aerobic
		Xylella fastidiosa Temecula 1	xft	1	1			Yes	
		Xanthomonas campestris pv. campestris ATCC 33913	xcc	1	1			Yes	Aerobic
		Xanthomonas axonopodis pv. citri 306	xac	1	1			Yes	Aerobic
	Proteobacteria γ Vibrionales	Vibrio cholerae El Tor N16961 (serotype O1)	vch	1	2	1		Yes	Facultative
		Vibrio vulnificus CMCP6	vvu	2	2	1		Yes	
		v.vulnificusy_y016.pep	vvy	2	2	1		Yes	
		Vibrio parahaemolyticus RIMD 2210633	vpa	2	1	1		Yes	
	Proteobacteria γ Pseudomonadales	Pseudomonas aeruginosa PA01	pae	1	1			Yes	Aerobic
		Pseudomonas putida KT2440	ppu	1	1			Yes	
		Pseudomonas syringae pv. tomato DC3000	pst	1	1			Yes	
	Proteobacteria γ Alteromondaceae	Shewanella oneidensis MR-1	son	1			1	Yes	Aerobic

(Continued)

Table 1. Continued

		msrA	msrB	msrAB	msrBA	Complete	Lifestyle
	Streptococcus pneumoniae TIGR4	spn			2		Yes	Facultative
	Streptococcus pneumoniae R6	spr			2		Yes	Facultative
	Streptococcus agalactiae 2603 (serotype V)	sag	1	1			Yes	
	Streptococcus agalactiae NEM316	san	1	1			Yes	
	Streptococcus mutans UA159 (serotype C)	smu	1				Yes	
	Lactobacillus plantarum WCFS1	lpl	3	1			Yes	
	Enterococcus faecalis V583	efa	1	1			Yes	
Firmicutes Clostridia	*Clostridium acetobutylicum* ATCC824	cac	1	1			Yes	Anaerobic
	Clostridium perfringens 13	cpe	1	1			Yes	
	Clostridium tetani E88	ctc	1				Yes	
	Thermoanaerobacter tengcongensis MB4T	tte					Yes	Anaerobic
Firmicutes Mollicutes	*Mycolasma genitalium* G-37	mge	1	1			Yes	Facultative
	Mycoplasma pneumoniae M129	mpn	1	1			Yes	Facultative
	Mycoplasma pulmonis	mpu	1	1			Yes	Facultative
	Mycoplasma penetrans HF-2	mpe	1				Yes	Microaerophilic
	Ureaplasma urealyticum (serovar 3)	uur	1				Yes	
	Mycoplasma gallisepticum R	mga		1			Yes	Anaerobic
	Phytoplasma sp. Onion yellows	poy					Yes	
Actinobacteria	*Mycobacterium tuberculosis* H37Rv (lab strain)	mtu	1	1			Yes	Aerobic
	Mycobacterium tuberculosis CDC1551	mtc	1	1			Yes	
	Mycobacterium leprae TN	mle	1	1			Yes	Aerobic
	Corynebacterium glutamicum ATCC 13032	cgl	1	1			Yes	
	Corynebacterium efficiens YS-314	cef	1	1			Yes	
	Streptomyces coelicolor A3(2)	sco	1	1			Yes	
	Streptomyces avermitilis	sma	1	1			Yes	
	Bifidobacterium longum NCC2705	blo			1		Yes	
	Tropheryma whipplei Twist	twh					Yes	Endosymbiont
	Tropheryma whipplei TW08/27	tws					Yes	Endosymbiont
	Corynebacterium diphtheriae NCTC 13129	cdi	1	1			Yes	
	Mycobacterium bovis AF2122/97	mbo	1	1			Yes	
Fusobacteria	*Fuscobacterium nucleatum* ATCC 25586	fnu			2		Yes	
Chlamydia	*Chlamydia trachomatis* (serovar D)	ctr					Yes	Obligatory parasite
	Chlamydia muridarum (*Chlamydia trachomatis* MoPn)	cmu					Yes	Obligatory parasite
	Chlamydophila pneumoniae CWL029	cpn					Yes	Obligatory parasite
	Chlamydophila pneumoniae AR39	cpa					Yes	Obligatory parasite
	Chlamydophila pneumoniae J138	cpj					Yes	Obligatory parasite
	Chlamydophila pneumoniae tw183	cpn					Yes	Obligatory parasite
	Chlamydophila caviae GPIC	cca					Yes	Obligatory parasite
Spirochete	*Borrelia burgdorferi* B31	bbu					Yes	Obligatory parasite
	Treponema pallidum Nichols	tpa				1	Yes	Microaerophillic
	Leptospira interrogans 56601 (serovar lai)	lil	2	1			Yes	
Bacteroid	*Bacteroides thetaiotaomicron* VPI-5482	bth		1	1		Yes	

(Continued)

Table 1. Continued

			msrA	msrB	msrAB	msrBA	Complete	Lifestyle
Planctomyces	*Porphyromonas gingivalis W83*	pgi	2		1		Yes	
	Rhodopirellula baltica (Pirellla sp.)	rba	2	2			Yes	
Cyanobacteria	*Synechocystis sp. PCC6803*	syn	2	1			Yes	Aerobic
	Thermosynechococcus elongatus BP-1	tel	1	1			Yes	
	Anabaena sp. PCC7120 (Nostoc sp. PCC7120)	ana	2	1			Yes	
	Gloeobacter violaceus PCC7421	gvi	2	2			Yes	
	Prochlorococcus marinus SS120 (CCMP1375)	pma	2	1			Yes	
	Prochlorococcus marinus MED4 (CCMP1378)	pmm	2	1			Yes	
	Prochlorococcus marinus MIT9313	pmt	2	2			Yes	
	Synechococcus sp. WH8102	syw	2	2			Yes	
Green sulfur bacteria	*Chlorobium tepidum TLS*	cte	2	1			Yes	Anaerobic
Radioresistant bacteria	*Deinococcus radiodurans R1*	dra	1	1		1	Yes	Aerobic
Hyperthermophilic bacteria	*Aquifex aeolicus VF5*	aae					Yes	Microaerophilic
	Thermotoga maritima MSB8	tma					Yes	Anaerobic
Archaea Euryarchaeota	*Methanococcus jannaschii DSM2661*	mja	1	1			Yes	Anaerobic
	Methanosarcina acetivorans C2A	mac	1	1			Yes	Anerobic
	Methanosarcina mazei Goe1	mma	1	1			Yes	Anerobic
	Methanobacterium thermoautotrophicum deltaH	mth	1	1			Yes	Anerobic
	Methanopyrus kandleri AV19	mka					Yes	Anerobic
	Archaeoglobus fulgidus DSM4304	afu					Yes	Anerobic
	Halobacterium sp. NRC-1	hal	1	1			Yes	Aerobic
	Thermoplasma acidophilum	tac					Yes	Facultative
	Thermoplasma volcanium GSS1	tvo					Yes	Facultative
	Pyrococcus horikoshii OT3	pho					Yes	Anerobic
	Pyrococcus abyssi	pab					Yes	Anerobic
	Pyrococcus furiosus DSM3638	pfu					Yes	Anerobic
Crenarchaeota	*Aeropyrum pernix K1*	ape					Yes	Aerobic
	Sulfolobus solfataricus P2	sso	1				Yes	Aerobic
	Sulfolobus tokodaii strain 7	sto					Yes	Aerobic
	Pyrobaculum aerophilum IM2	pai					Yes	Facultative
Nanoarchaeota	*Nanoarchaeum equitans*	neq					Yes	Obligate symbiont

Fig. 1. a Minimum evolution phylogeny of naturally concatenated MsrA and MsrB proteins, calculated on the basis of 100 bootstrap replications (259 amino acid sites and 26 sequences). Only branches with more than 70% bootstrap are shown in this tree. Root is placed on midpoint. Naturally fused MsrA-MsrB sequences are indicated by a black-white rectangle. Species codes: Proteobacteria γ Pasteurellales—*Haemophilus influenzae* Rd (hin); Proteobacteria γ Vibrionales—*Vibrio cholerae* El Tor N16961 (serotype O1) (vch), *Vibrio vulnificus* CMCP6 (vvu), *Vibrio vulnificus* yj016 (vvy), *Vibrio parahaemolyticus* RIMD 2210633 (vpa); Proteobacteria β—*Neisseria meningitidis* MC58 (serogroup B) (nme), *Neisseria meningitidis* Z2491 (serogroup A) (nma), *Nitrosomonas europaea* ATCC 19718 (neu); Proteobacteria ∊—*Helicobacter pylori* 26695 (hpy), *Helicobacter pylori* J99 (hpj); Firmicutes Bacillales—*Bacillus anthracis* Ames (ban), *Bacillus cereus* ATCC 14579 (bce), *Oceanobacillus iheyensis* HTE831 (oih); Firmicutes Lactobacillales—*Streptococcus pyogenes* SF370 (serotype M1) (spy), *Streptococcus pyogenes* MGAS8232 (serotype M18) (spm), *Streptococcus pyogenes* MGAS315 (serotype M3) (spg), *Streptococcus pyogenes* SSI-1 (serotype M3) (sps), *Streptococcus pneumoniae* TIGR4 (spn), *Streptococcus pneumoniae* R6 (spr); Actinobacteria—*Bifidobacterium longum* NCC2705 (blo); Fusobacteria—*Fusobacterium nucleatum* ATCC 25586 (fnu); Bacteroid—*Bacteroides thetaiotaomicron* VPI-5482 (bth), *Porphyromonas gingivalis* W83 (pgi). b Minimum evolution phylogeny of in silico concatenated MsrA and MsrB proteins from those species whose genomes are endowed with only one single copy of each gene. 100 bootstrap replications (225 amino acid sites and 67 sequences). Only branches with more than 50% bootstrap are shown in this tree. Black rhombus, MsrB sequence with zinc motif; white rhombus, MsrB sequence without zinc motif. Root is placed on midpoint. Species codes: Eukarya—*Caenorhabditis elegans* (cel), *Dictyostelium discoideum* (ddi), *Saccharomyces cerevisiae* (sce), *Schizosaccharomyces pombe* (spo); Proteobacteria γ Enterobacteriales—*Escherichia coli* K-12 MG1655 (eco), *Escherichia coli* K-12 W3110 (ecj), *Escherichia coli* O157:H7 EDL933 (ece), *Escherichia coli* O157:H7 Sakai (ecs), *Escherichia coli* CFT073 (ecc), *Salmonella typhi* CT18 (sty), *Salmonella typhi* Ty2 (stt), *Salmonella typhimurium* LT2 (stm), *Yersinia pestis* CO92 (ype), *Yersinia pestis* KIM (ypk), *Shigella flexneri* 301 (serotype 2a) (sfl), *Shigella flexneri*

2457T (serotype 2a) (sfx), *Photorhabdus luminescens* laumondii TTO1 (plu); Proteobacteria γ Pasteurellales—*Pasteurella multocida* PM70 (pmu); Proteobacteria γ Xanthomonadales—*Xylella fastidiosa* 9a5c (xfa), *Xylella fastidiosa* Temecula1 (xft), *Xanthomonas campestris* pv. campestris ATCC 33913 (xcc), *Xanthomonas axonopodis* pv. citri 306 (xac); Proteobacteria γ Pseudomonadales—*Pseudomonas aeruginosa* PA01 (pae), *Pseudomonas putida* KT2440 (ppu), *Pseudomonas syringae* pv. tomato DC3000 (pst); Proteobacteria δ proteobacteria—*Geobacter sulfurreducens* PCA (gsu); Proteobacteria β proteobacteria—*Ralstonia solanacearum* GMI1000 (rso), *Bordetella bronchiseptica* (bbr), *Bordetella parapertussis* 12822 (bpa), *Bordetella pertussis* Tohama I (bpe), *Chromobacterium violaceum* ATCC 12472 (cvi); Proteobacteria ∊—*Campylobacter jejuni* NCTC11168 (cje), *Helicobacter hepaticus* ATCC 51449 (hhe); Proteobacteria α—*Agrobacterium tumefaciens* C58 (UWash/Dupont) (atu), *Agrobacterium tumefaciens* C58 (Cereon) (atc), *Brucella melitensis* 16M (bme), *Brucella suis* 1330 (bms); Firmicutes Bacillales—*Bacillus subtilis* 168 (bsu), *Listeria monocytogenes* EGD-e (lmo), *Listeria innocua* CLIP 11262 (lin); Firmicutes Lactobacillales—*Streptococcus pyogenes* SF370 (serotype M1) (spy), *Streptococcus pyogenes* MGAS8232 (serotype M18) (spm), *Streptococcus pyogenes* MGAS315 (serotype M3) (spg), *Streptococcus pyogenes* SSI-1 (serotype M3) (sps), *Streptococcus agalactiae* 2603 (serotype V) (sag), *Streptococcus agalactiae* NEM316 (san), *Enterococcus faecalis* V583 (efa); Firmicutes Clostridia—*Clostridium acetobutylicum* ATCC824 (cac), *Clostridium perfringens* 13 (cpe); Firmicutes Mollicutes—*Mycoplasma genitalium* G-37 (mge), *Mycoplasma pneumoniae* M129 (mpn), *Mycoplasma pulmonis* (mpu), *Mycoplasma gallisepticum* R (mga); Actinobacteria—*Mycobacterium tuberculosis* H37Rv (lab strain) (mtu), *Mycobacterium tuberculosis* CDC1551 (mtc), *Corynebacterium glutamicum* ATCC 13032 (cgl), *Corynebacterium efficiens* YS-314 (cef), *Streptomyces coelicolor* A3(2) (sco), *Streptomyces avermitilis* (sma), *Corynebacterium diphtheriae* NCTC 13129 (cdi), *Mycobacterium bovis* AF2122/97 (mbo); Cyanobacteria—*Thermosynechococcus elongatus* BP-1 (tel); radioresistant bacteria—*Deinococcus radiodurans* R1 (dra); Euryarchaeota—*Methanosarcina acetivorans* C2A (mac), *Methanosarcina mazei* Goe1 (mma), *Methanobacterium thermoautotrophicum* deltaH (mth), *Halobacterium* sp. NRC-1 (hal).

and the enzyme is known to play a role in protecting eye lens cells against cataracts (Kantorow et al. 2004). Sequence comparisons have indicated that MsrB, which are absent in some species that are endowed with MsrA (Table 1), can be divided into two families distinguished by the presence or absence of zinc (Kryukov et al. 2002; Kumar et al. 2002).

There is increasing recognition of the medical and biochemical significance of MsrA and MsrB. However, as of today, no attempt has been made to ana-
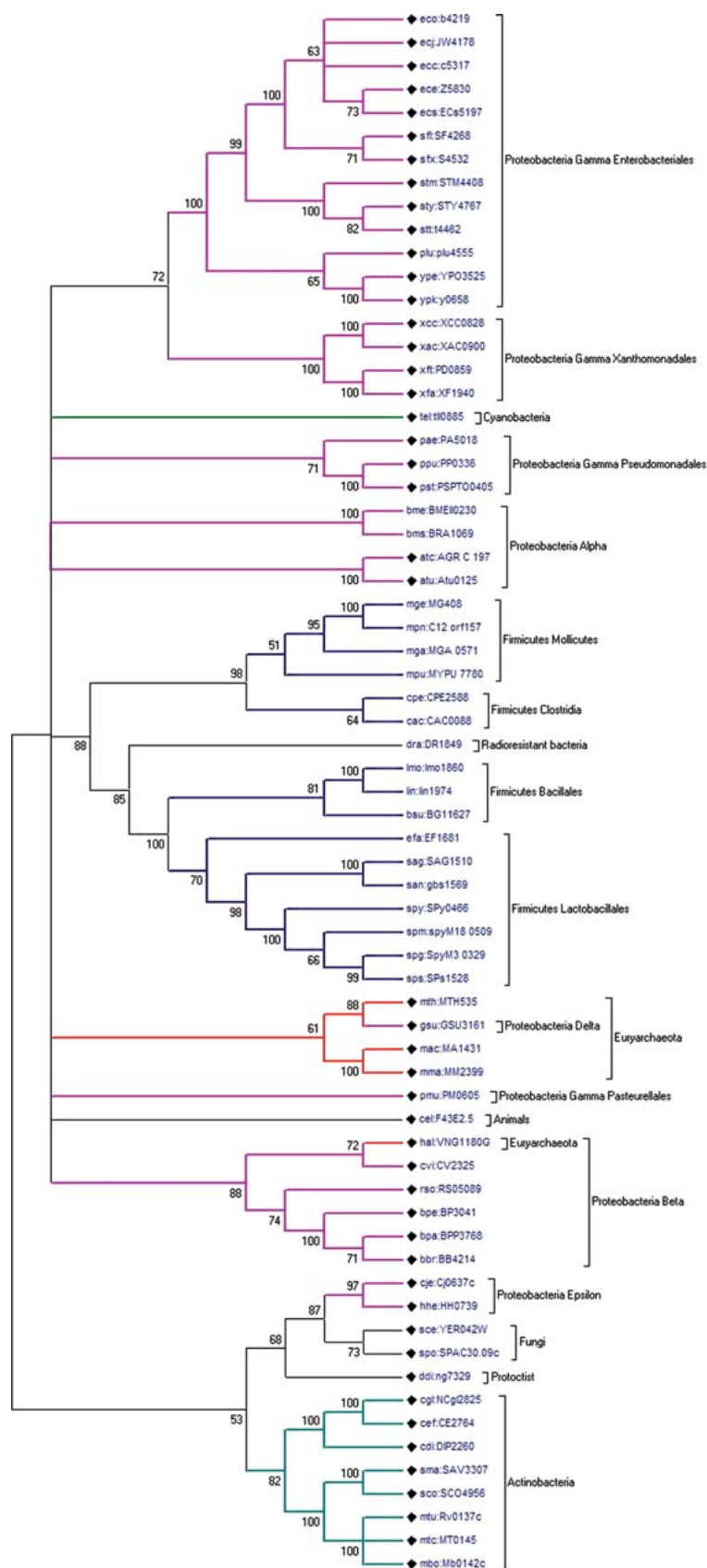
Fig. 1 (b). Continued

lyze these enzymes from the viewpoint of microbial evolution during the Precambrian or to discuss their significance as a response to the accumulation of oxygen in the terrestrial environment. The current availability of a large number of completely sequenced genomes now allows us to discuss the rela-

Fig. 2. **a** Minimum evolution MsrA phylogeny, 100 bootstrap replications (134 amino acid sites and 106 sequences). Only branches with more than 70% bootstrap are shown in this tree. Clades with lesser values have been collapsed. Root is placed on midpoint. **b** Minimum evolution MsrB phylogeny, 100 bootstrap replications (103 amino acid sites and 125 sequences). Only branches with more than 70% bootstrap value are shown. Root is placed on midpoint. Zinc-bearing MsrB sequences are indicated with a black rhombus. Naturally MsrA-MsrB fused sequences are indicated by a black-white rectangle, while naturally fused MsrB-MsrA sequences are indicated by a white-black rectangle. Species codes: Proteobacteria γ Enterobacteriales—*Escherichia coli* K-12 MG1655 (eco), *Escherichia coli* K-12 W3110 (ecj), *Escherichia coli* O157:H7 EDL933 (ece), *Escherichia coli* O157:H7 Sakai (ecs), *Escherichia coli* CFT073 (ecc), *Salmonella typhi* CT18 (sty), *Salmonella typhi* Ty2 (stt), *Salmonella typhimurium* LT2 (stm), *Yersinia pestis* CO92 (ype), *Yersinia pestis* KIM (ypk), *Shigella flexneri* 301 (serotype 2a) (sfl), *Shigella flexneri* 2457T (serotype 2a) (sfx), *Photorhabdus luminescens* laumondii TTO1 (plu); Proteobacteria γ Pasteurellales—*Pasteurella multocida* PM70(pmu), *Haemophilus influenzae* Rd (hin), *Haemophilus ducreyi* 35000HP (hdu); Proteobacteria γ Xanthomonadales—*Xylella fastidiosa* 9a5c (xfa), *Xylella fastidiosa* Temecula1 (xft), *Xanthomonas campestris* pv. campestris ATCC 33913 (xcc), *Xanthomonas axonopodis* pv. citri 306 (xac); Proteobacteria γ Vibrionales—*Vibrio cholerae* El Tor N16961 (serotype O1) (vch), *Vibrio vulnificus* CMCP6 (vvu), *Vibrio vulnificus* yj016 (vvy), *Vibrio parahaemolyticus* RIMD 2210633 (vpa); Proteobacteria γ Pseudomonadales—*Pseudomonas aeruginosa* PA01 (pae), *Pseudomonas putida* KT2440 (ppu), *Pseudomonas syringae* pv. tomato DC3000 (pst); Proteobacteria γ Alteromonadaceae—*Shewanella oneidensis* MR-1 (son); Proteobacteria γ Legionellales—*Coxiella burnetii* SA 493 (cbu); Proteobacteria δ proteobacteria—*Geobacter sulfurreducens* PCA (gsu); Proteobacteria β proteobacteria—*Ralstonia solanacearum* GMI1000 (rso), *Nitrosomonas europaea* ATCC 19718 (neu), *Bordetella bronchiseptica* (bbr), *Bordetella parapertussis* 12822 (bpa), *Bordetella pertussis* Tohama I (bpe), *Chromobacterium violaceum* ATCC 12472 (cvi), *Nitrosomonas europaea* ATCC 19718 (neu); Proteobacteria ε—*Helicobacter pylori* 26695 (hpy), *Helicobacter pylori* J99 (hpj), *Campylobacter jejuni* NCTC11168 (cje), *Helicobacter hepaticus* ATCC 51449 (hhe); Proteobacteria α—*Mesorhizobium loti* MAFF303099 (mlo), *Sinorhizobium meliloti* 1021 (sme), *Agrobacterium tumefaciens* C58 (UWash/Dupont) (atu), *Agrobacterium tumefaciens* C58 (Cereon) (atc), *Brucella melitensis* 16M (bme), *Brucella suis* 1330 (bms), *Bradyrhizobium japonicum* USDA110 (bja), *Caulobacter crescentus* (ccr), *Rhodopseudomonas palustris* CGA009 (rpa); Firmicutes Bacillales—*Bacillus subtilis* 168 (bsu), *Bacillus halodurans* C-125 (bha), *Bacillus anthracis* Ames (ban), *Bacillus cereus* ATCC 14579 (bce), *Oceanobacillus iheyensis* HTE831 (oih), *Staphylococcus aureus* N315 (meticillin-resistant) (sau), *Staphylococcus aureus* Mu50 (vancomycin-resistant) (sav), *Staphylococcus aureus* MW2 (sam), *Staphylococcus epidermidis* ATCC 12228 (sep), *Listeria monocytogenes* EGD-e (lmo), *Listeria innocua* CLIP 11262 (lin); Firmicutes Lactobacillales—*Lactococcus lactis* IL1403 (lla), *Streptococcus pyogenes* SF370 (serotype M1) (spy), *Streptococcus pyogenes* MGAS8232 (serotype M18) (spm), *Streptococcus pyogenes* MGAS315 (serotype M3) (spg), *Streptococcus pyogenes* SSI-1 (serotype M3) (sps), *Streptococcus pneumoniae* TIGR4 (spn), *Streptococcus pneumoniae* R6 (spr), *Streptococcus agalactiae* 2603 (serotype V) (sag), *Streptococcus agalactiae* NEM316 (san), *Streptococcus mutans* UA159 (serotype C) (smu), *Lactobacillus plantarum* WCFS1 (lpl), *Enterococcus faecalis* V583 (efa); Firmicutes Clostridia—*Clostridium acetobutylicum* ATCC824 (cac), *Clostridium perfringens* 13 (cpe), *Clostridium tetani* E88 (ctc); Firmicutes Mollicutes—*Mycoplasma genitalium* G-37 (mge), *Mycoplasma pneumoniae* M129 (mpn), *Mycoplasma pulmonis* (mpu), *Mycoplasma penetrans* HF-2 (mpe), *Ureaplasma urealyticum* (serovar 3) (uur), *Mycoplasma gallisepticum* R (mga); Actinobacteria—*Mycobacterium tuberculosis* H37Rv (lab strain) (mtu), *Mycobacterium tuberculosis* CDC1551 (mtc), *Mycobacterium leprae* TN (mle), *Corynebacterium glutamicum* ATCC 13032 (cgl), *Corynebacterium efficiens* YS-314 (cef), *Streptomyces coelicolor* A3(2) (sco), *Streptomyces avermitilis* (sma), *Bifidobacterium longum* NCC2705 (blo), *Corynebacterium diphtheriae* NCTC 13129 (cdi), *Mycobacterium bovis* AF2122/97 (mbo); Fusobacteria—*Fusobacterium nucleatum* ATCC 25586 (fnu); Spirochete—*Treponema pallidum* Nichols (tpa), *Leptospira interrogans* 56601 (serovar lai) (lil); Bacteroid—*Bacteroides thetaiotaomicron* VPI-5482 (bth), *Porphyromonas gingivalis* W83 (pgi); Planctomyces—*Rhodopirellula baltica* (Pirellula sp.) (rba); Cyanobacteria—*Synechocystis* sp. PCC6803 (syn), *Thermosynechococcus elongatus* BP-1 (tel), *Anabaena* sp. PCC7120 (*Nostoc* sp. PCC7120) (ana), *Gloeobacter violaceus* PCC7421 (gvi), *Prochlorococcus marinus* SS120 (CCMP1375) (pma), *Prochlorococcus marinus* MED4 (CCMP1378) (pmm), *Prochlorococcus marinus* MIT9313 (pmt), *Synechococcus* sp. WH8102 (syw); green sulfur bacteria—*Chlorobium tepidum* TLS (cte); radioresistant bacteria—*Deinococcus radiodurans* R1 (dra); Eukarya—*Homo sapiens* (hsa), *Mus musculus* (mmu), *Drosophila melanogaster* (dme), *Caenorhabditis elegans* (cel), *Arabidopsis thaliana* (ath), *Dictyostelium discoideum* (ddi), *Saccharomyces cerevisiae* (sce), *Schizosaccharomyces pombe* (spo); Euryarchaeota—*Methanosarcina acetivorans* C2A (mac), *Methanosarcina mazei* Goe1 (mma), *Methanobacterium thermoautotrophicum* deltaH (mth), *Halobacterium* sp. NRC-1 (hal); Crenarchaeota—*Sulfolobus solfataricus* P2 (sso).

tionship between organism lifestyles and the presence or absence of MsrA and MsrB, as well as the phylogenetic patterns exhibited by these enzymes. Although the early evolutionary history of these enzymes has been partially obscured by a number of lateral gene transfer events, the results presented here may be interpreted as suggesting that the sequences encoding MsrA and MsrB originated in the bacterial domain and spread from it into the Archaea and the Eukarya.

Materials and Methods

MsrA (b4219) and MsrB (b1778) sequences from *Escherichia coli* were downloaded from the Kyoto Encyclopedia of Genes and Genomes, KEGG database (www.genome.jp/kegg/) (Kanehisa and Goto 2000). Homologues to these two sequences were searched in a database of 159 completely sequenced genomes from the three domains of life available as of December 2004, using BLAST search (Altschul et al. 1997). Results are shown in Table 1 (http://bacteria.fciencias.unam.mx/Msr/Msr.html). The lifestyles recorded in Table 1 follow the description given in the original publication of each genome sequence. Sequences were aligned using ClustalX (Thompson et al. 1997), and the result was visually edited using Bioedit software (Hall 1999). In order to perform phylogenetic analysis, minimum evolution trees using MEGA3 software version 3.1 (Kumar et al. 2004) were calculated for MsrA and MsrB separately (100 bootstrap replications each). Uniform rate variation was assumed along all sites. A Poisson correction was employed to calculate distances between the sequences compared in this analysis. All trees were rooted using the midpoint value using the MEGA3.1 software. The complete deletion option available in the MEGA3.1 program was employed to avoid considering positions exhibiting gaps. In the resulting trees, only clades with > 70% bootstrap support were depicted, and branches harboring species from the same taxonomic group were pulled together.

Figure 1 shows cladograms of naturally (Fig. 1a) and artificially (Fig. 1b) concatenated MsrA and MsrB sequences. The tree in Figure 1b was constructed by concatenating in silico the MsrA and MsrB sequences from those species whose genomes are endowed with only a single copy of each gene. The hypothesis underlying this approach is that concatenated sequences with significant
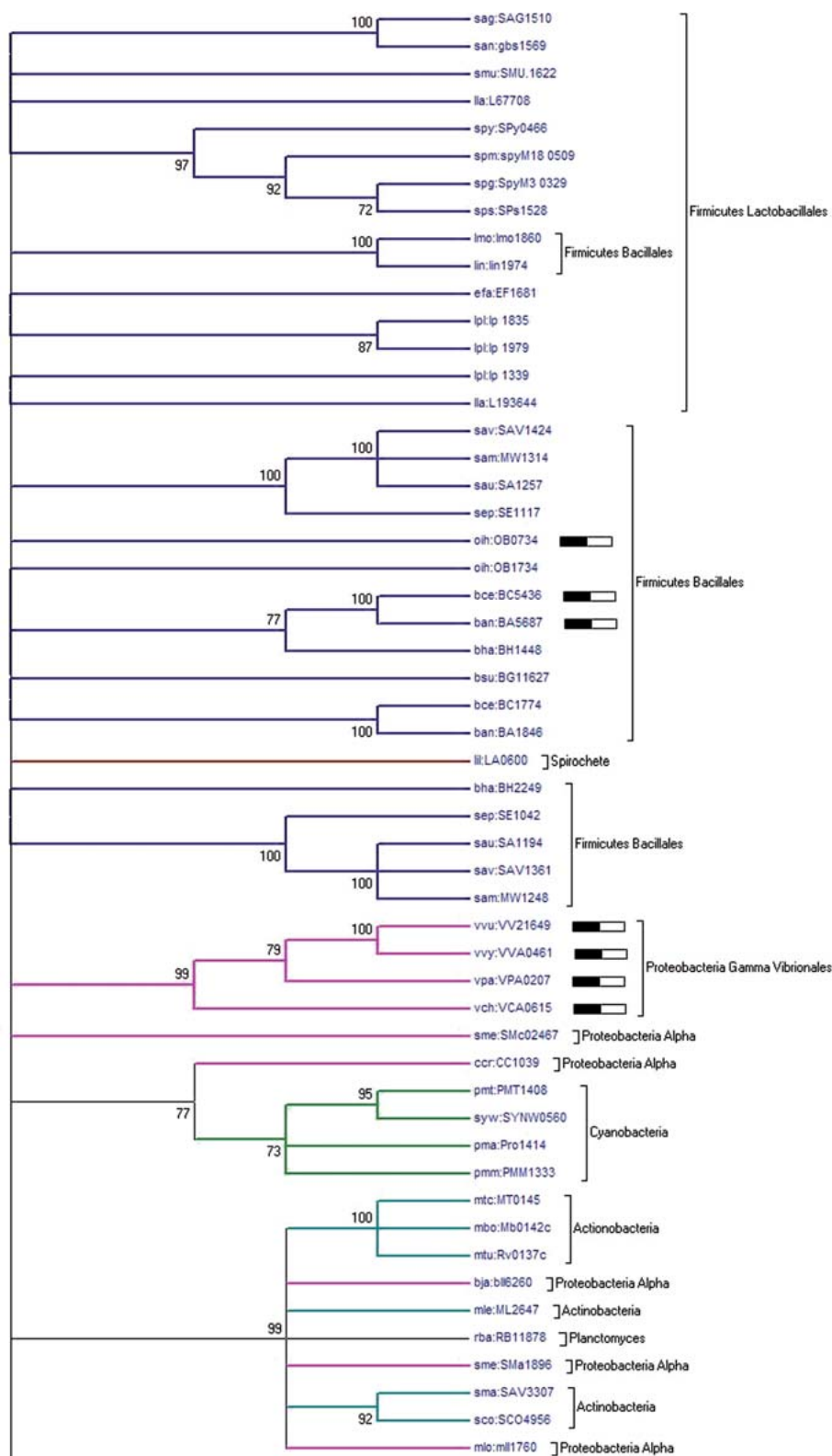
Fig. 2 (a). Continued

bootstrap values may have been inherited jointly (Fig. 1b). Concatenation of sequences has been recognized as enhancing phylogenetic signals (Brown et al. 2001; Rokas et al. 2003). Although leaving out of the sample under consideration all the genomes endowed with paralogous sequences reduces the number of species from 159 to 72, a significant gain in phylogenetic signals is achieved, as indicated by the bootstrap values. Zn-bearing MsrB sequences were identified on the basis of the conserved $Cys^{81}XXCys^{84}...Cys^{101}XXCy^{s104}$ motif described by Kumar et al (2002) for the *Drosophila melanogaster* enzyme. Species endowed with Zn-bearing MsrB sequences are indicated by a black rhombus in the trees presented here.

Quantitative estimates of the levels of agreement between the MsrA and the MsrB phylogenies were performed by comparing the
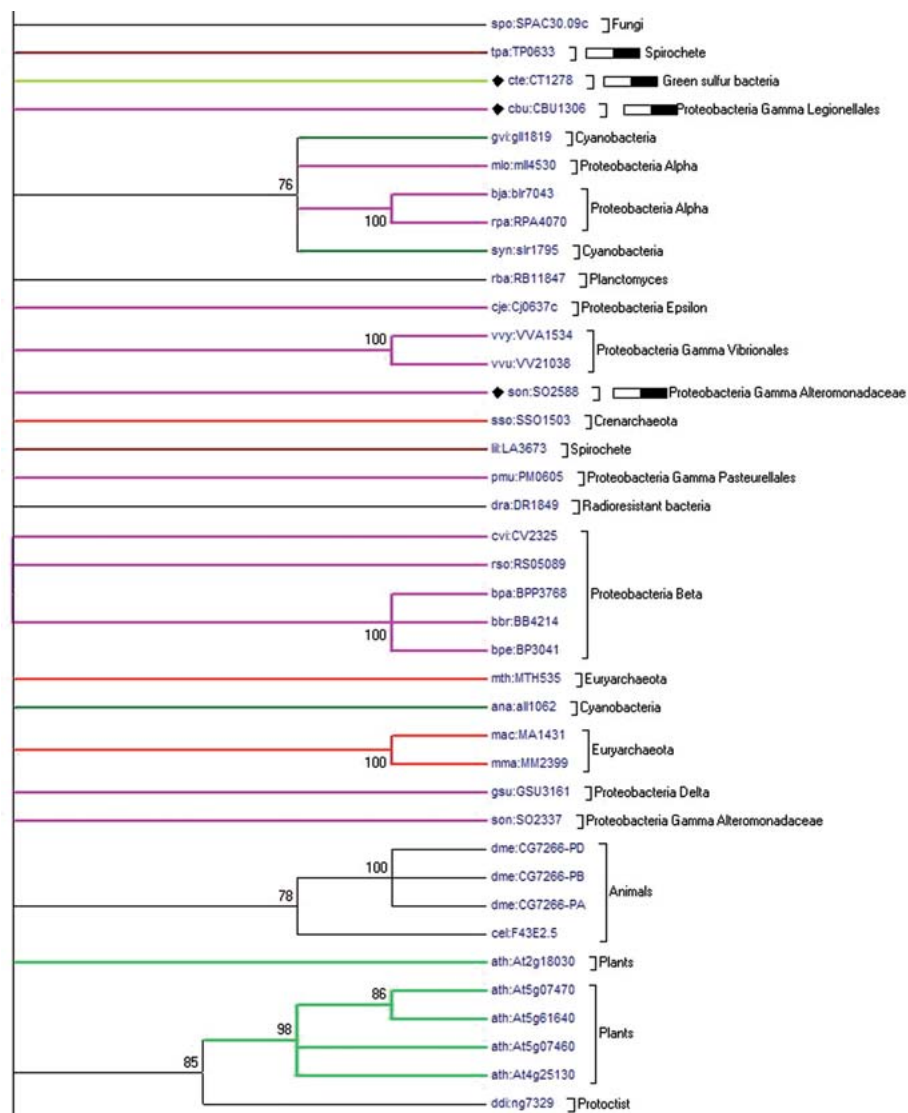
spo:SPAC30.09c] Fungi
tpa:TP0633] ☐■☐ Spirochete
♦ cte:CT1278] ☐■☐ Green sulfur bacteria
♦ cbu:CBU1306] ☐■☐ Proteobacteria Gamma Legionellales
gvi:gll1819] Cyanobacteria
mlo:mll4530] Proteobacteria Alpha
bja:blr7043 ⌐ Proteobacteria Alpha
rpa:RPA4070 ⌐ 100
syn:slr1795] Cyanobacteria
76
rba:RB11847] Planctomyces
cje:Cj0637c] Proteobacteria Epsilon
vvy:VVA1534 ⌐ 100 Proteobacteria Gamma Vibrionales
vvu:VV21038 ⌐
♦ son:SO2588] ☐■☐ Proteobacteria Gamma Alteromonadaceae
sso:SSO1503] Crenarchaeota
lil:LA3673] Spirochete
pmu:PM0605] Proteobacteria Gamma Pasteurellales
dra:DR1849] Radioresistant bacteria
cvi:CV2325
rso:RS05089
bpa:BPP3768 Proteobacteria Beta
bbr:BB4214
100 bpe:BP3041
mth:MTH535] Euryarchaeota
ana:all1062] Cyanobacteria
mac:MA1431 Euryarchaeota
mma:MM2399 100
gsu:GSU3161] Proteobacteria Delta
son:SO2337] Proteobacteria Gamma Alteromonadaceae
dme:CG7266-PD
dme:CG7266-PB 100
dme:CG7266-PA Animals
78 cel:F43E2.5
ath:At2g18030] Plants
ath:At5g07470 86
ath:At5g61640
98 ath:At5g07460 Plants
ath:At4g25130
85 ddi:ng7329] Protoctist

Fig. 2 (a). Continued

number of symmetric differences that exist among the 16/18SrRNA, MsrA, MsrB, and concatenated MsrA-MsrB trees, with a randomly generated 16/18SrRNA tree that included all species under consideration, as well as with a tree built with randomly concatenated MsrA-MsrB sequences. The latter tree was constructed using a random number-based algorithm that generated chance concatenated MsrA-MsrB sequences from our database. The number of symmetric differences was calculated using the *treedist* algorithm from the PHYLIP 3.65 package (Felsenstein 1989). The statistical significance of the symmetric differences among these various trees was estimated by comparison with 100,000 randomly generated trees (Supplementary Table 3) built using a *perl* script (available upon request).

The three-dimensional structure classification of MsrA and MsrB used here follows the Structural Classification of Proteins, SCOP database (Murzin et al. 1995). Sequence alignments and all other supplementary information are available at http://bacteria.fciencias.unam.mx/Msr/Msr.html.

Results

The phylogenetic distributions of MsrA and MsrB are listed in Table 1, and extend and confirm previous reports by Kryukov et al. (2002) and Ezraty et al.

(2005). With the exception of well-characterized endosymbionts *(Buchnera, Tropheryma, Wigglesworthia*, and *Blochmannia)* and some endoparasites *(Phytoplasma, Rickettsia*, and *Chlamydia)*, almost all of the reported bacterial genomes, including the mycoplasma, are endowed with at least one copy of MsrA and one of MsrB. Two exceptions are *Mycoplasma penetrans* and *Ureaplasma urealyticum* (serovar 3), which have one MsrA copy each but lack MsrB. The only free-living bacterial species in the sample studied here that lack both sequences are the anaerobic pathogen *Wolinella succinogenes*, the hyperthermophilic microaerophile *Aquifex aeolicus*, and the anaerobe *Thermotoga maritima*.

Several bacterial species present more than one copy of MsrA and/or MsrB (Table 1). Phylogenetic analysis of this set of possible paralogous genes (Supplementary Fig. 4) does not indicate any obvious evolutionary trend. Although it could be argued that some of this redundancy is due to lateral gene transfer, the bootstrap values of these two trees
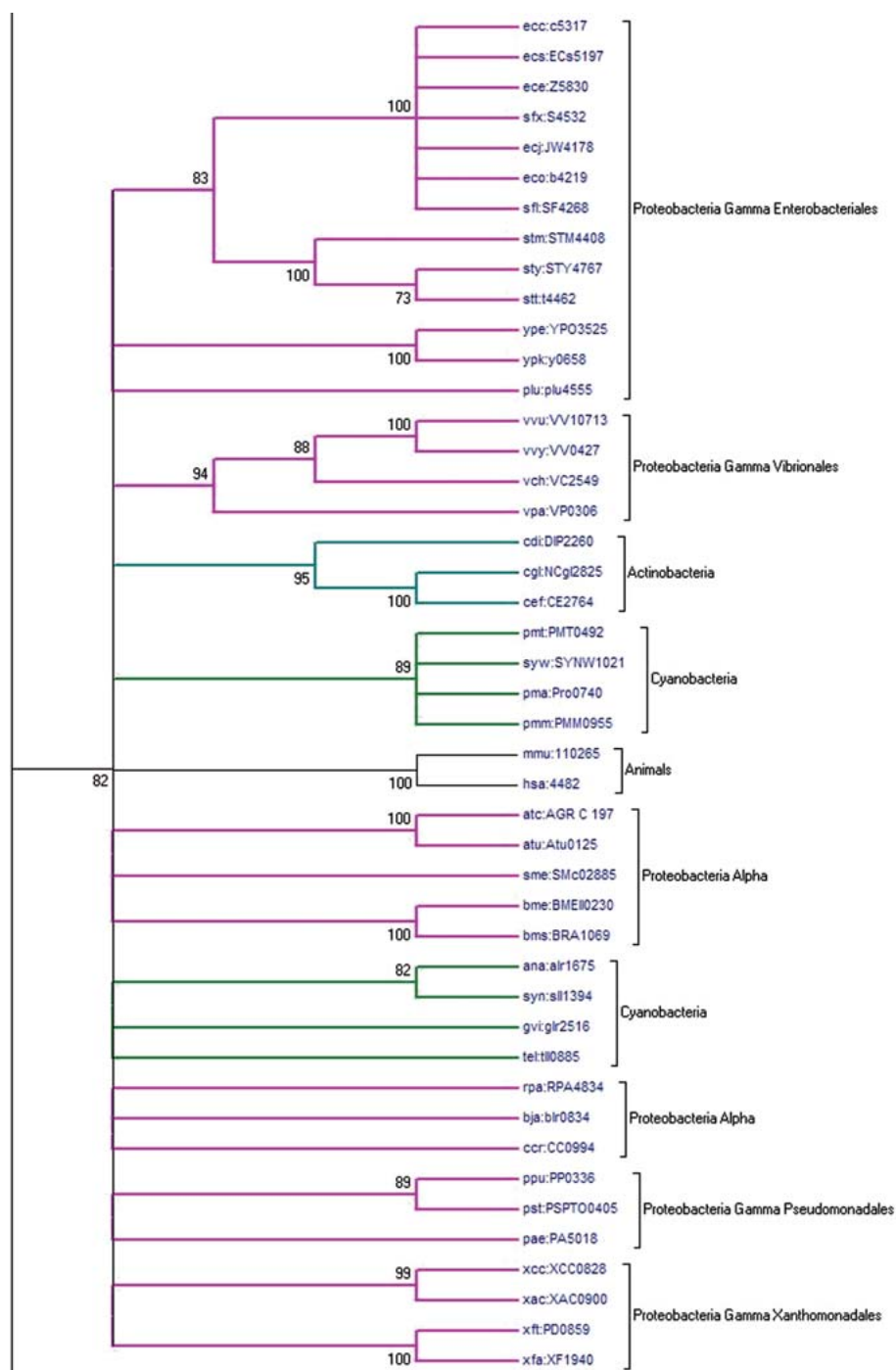
Fig. 2 (a). Continued

(Figs. 2a and b) do not allow any major evolutionary inference. Fused MsrAB or MsrBA sets appear to be the outcome of polyphyletic arrangements. However, in the tree shown in Fig. 2b the phylogenetic distribution of Zn-bearing MsrB sequences is not random, and it is interesting to note that no firmicutes or naturally fused MsrAB pairs are endowed with the Zn-binding motif. This suggests that MsrB that took part in the ancestral fusion that led to the MsrAB arrangements lacked Zn-binding abilities.

MsrA and MsrB homologs are absent in all cren-archaeotal species included in this study *(Pyrobacu-lum aerophilum, Aeropyrum pernix*, and *Sulfolobus tokodaii)*, with the exception of *S. solfataricus*, which has an ORF sequence homologous to the *E. coli* MsrA (Table 1). The parasitic microsporidium *En-cephalitozoon cuniculi* lacks both MsrA and MsrB, probably due to a secondary loss, but all the other eukaryotic genomes included in Table 1 exhibit at least one copy each of MsrA and MsrB. All eukary-otic MsrB sequences are endowed with the Zn-bind-ing motif, with the sole exception of a highly divergent copy of MsrB present in *Arabidopsis thali-ana* (Fig. 2b).
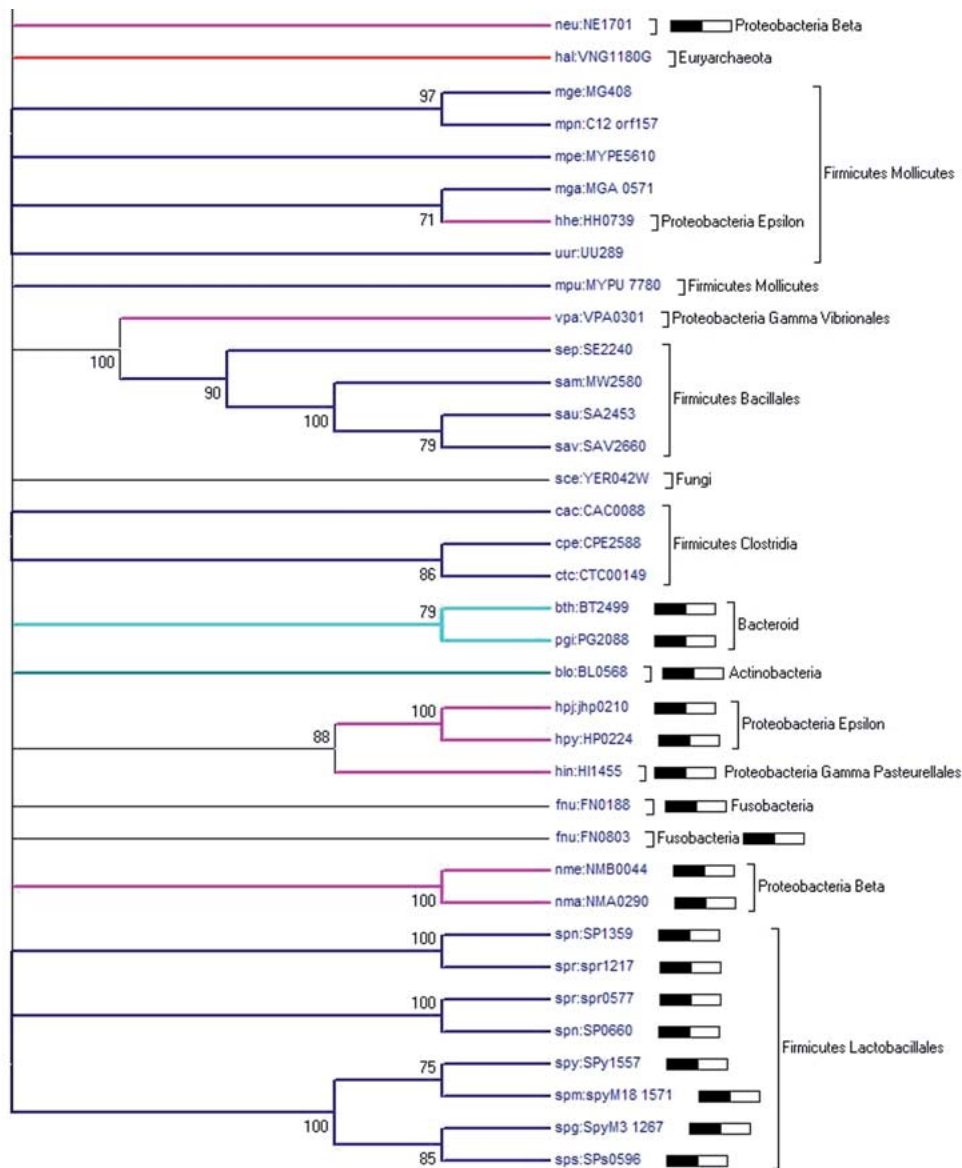
MsrA and MsrB have an ample biological distribution, but their phylogenies (Figs. 2a and b) resemble only in part the 16/18S rRNA-based cladograms with three well-defined monophyletic major domains (Woese et al. 1990). In some cases high bootstrap values are observed, but with few exceptions, discussed below, the clades have little if any biological significance. Some recently diverged bacterial groups, such as the γ-proteobacterial enterobacterial groups (the different *E. coli* strains, *Shigella*, *Salmonella*), exhibit similar topology in both 16/18S rRNA trees and MsrA (Fig. 2a) and MsrB (Fig. 2b) cladograms. MsrA and MsrB trees are characterized by very low bootstrap values, and a mixed distribution of actinobacteria, cyanobacteria, firmicutes, most proteobacteria, and archaeal and eukaryal groups is observed. However, quantitative estimates of the similarity between the MsrA and the Msr B

cladograms based on orthologue sequences show that despite the low preservation of phylogenetic signals, these two trees are more similar to each other, and to a 16/18S rRNA tree, than to randomly generated cladograms constructed as described above (Table 2).

Based on the similarities between the MsrA and the MsrB orthologue trees (Table 2), an attempt to recover phylogenetic signals that may be susceptible to evolutionary analysis was perfomed by an in silico artificial joining of MsrA and MsrB sequences, as described under Materials and Methods. No naturally concatenated archaeal or eucaryal MsrA and MsrB sequences are present in our database. Minimum evolution trees for naturally (Fig. 1a) and artificially (Fig. 1b) concatenated MsrA and MsrB sequences have been calculated. In the minimum evolution tree constructed with naturally concatenated sequences shown in Fig. 1a, several groups
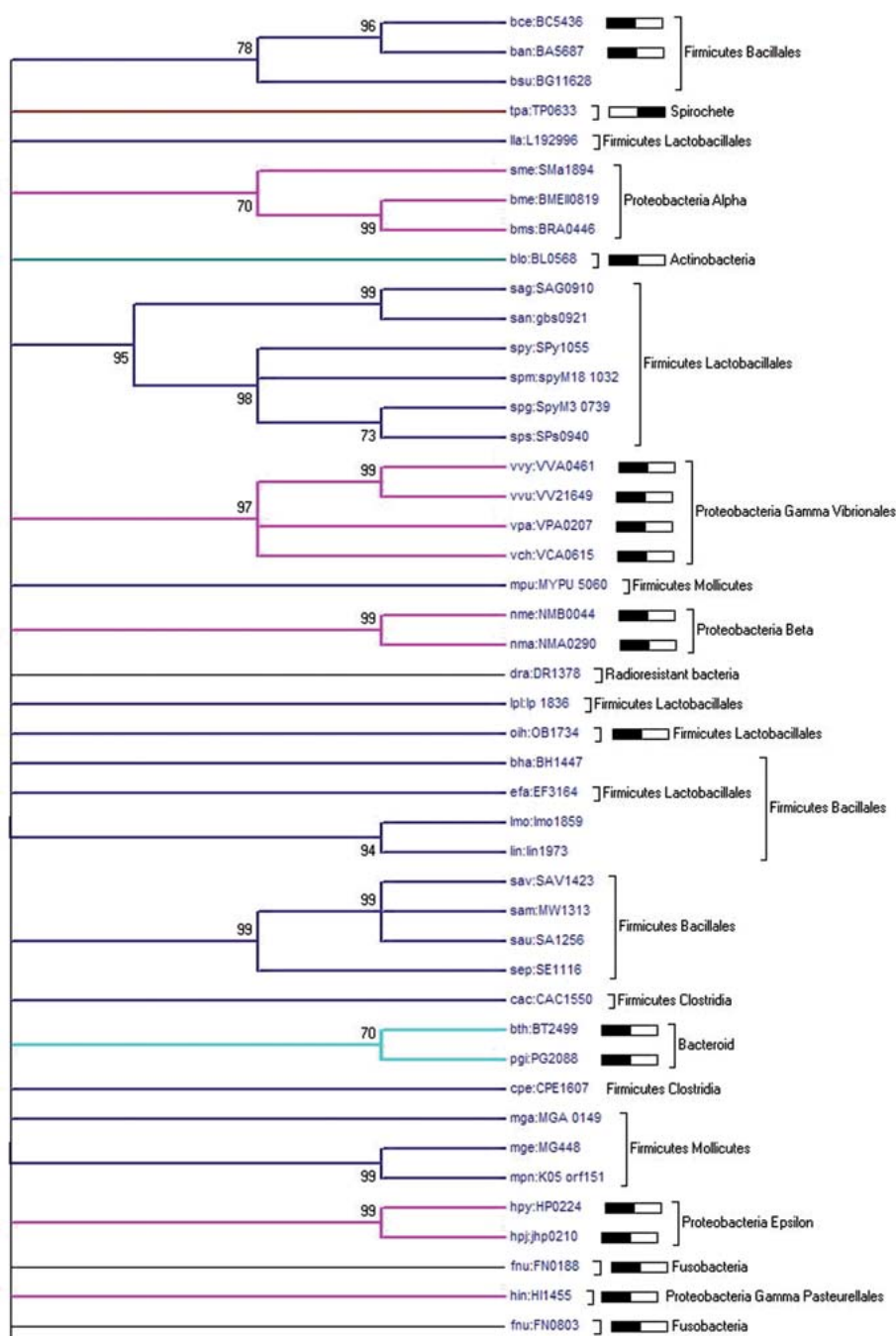
Fig. 2 (b). Continued

form clades whose branching order is comparable to those observed in rRNA phylogenies. Although the firmicutes are clearly separated in two well-defined monophyletic clades formed by the bacillales and lactobacillales groups, their internal branching orders are quite similar to that observed in trees based on 16/18S rRNA and other canonical markers. The same is true for the the γ-vibrionales clade, which has significant bootstrap values, although other proteobacterial branches are dispersed throughout the tree. The presence of the sequences of *Nitrosomonas europaea* in a separate branch can be explained as a long-branch attraction phenomenon, but because of the low bootstrap values, little can be said about the distribution of other species in this sample.

A significant reduction of the differences between the concatenated MsrA-MsrB tree and the canonical 16/18S rRNA tree is observed when the two are compared with the corresponding random trees (Table 2). The tree shown in Fig. 1b was constructed by concatenating in silico the sequences from those species whose genomes are endowed with only a single copy of the gene encoding MsrA and a single copy of the gene encoding MsrB. Concatenation of sequences has been recognized as a way of enhancing phylogenetic signals (Brown et al. 2001; Rokas et al. 2003). Although
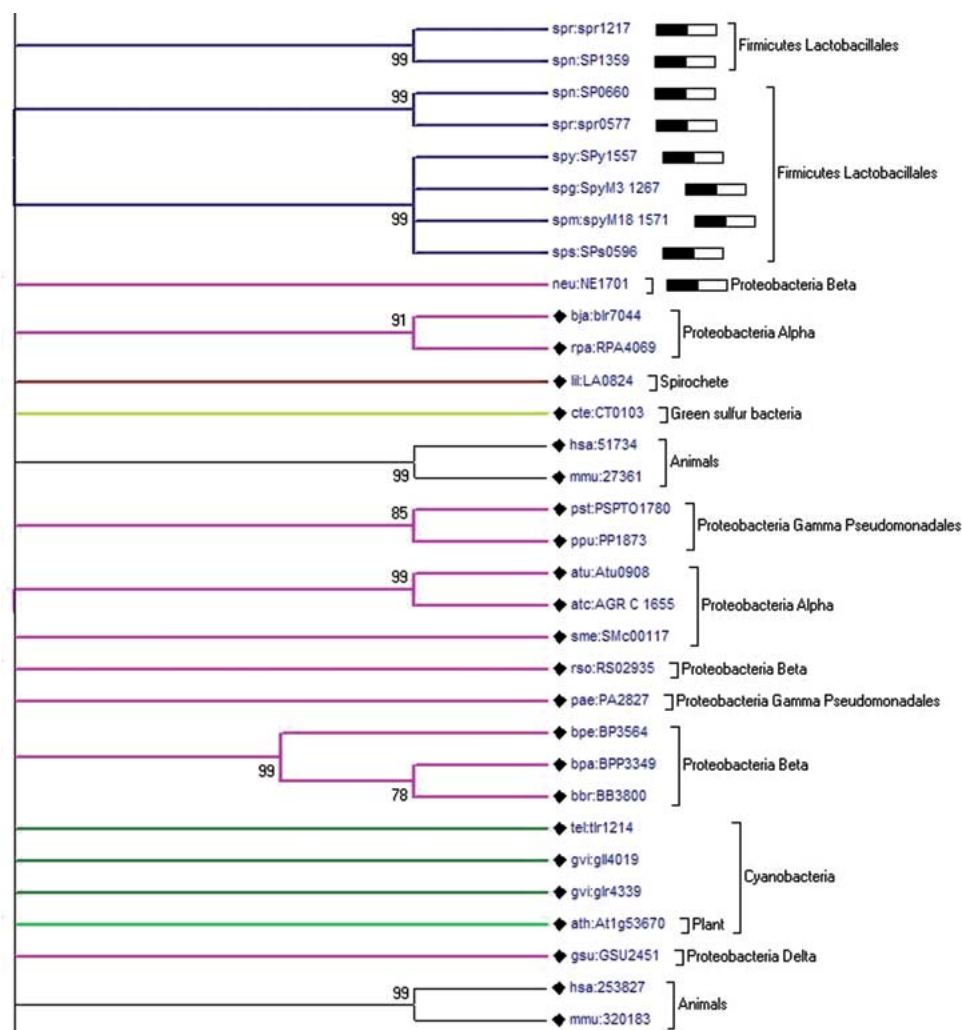
Fig. 2 (b). Continued

leaving out of the sample under consideration all the genomes endowed with paralogous sequences reduces the number of species from 159 to 72, a significant gain in phylogenetic signals is achieved, as indicated by the higher bootstrap values. The rationale behind the in silico concatenation is based on the possibility that naturally concatenated sequences may have been inherited jointly (Fig. 1a). In order to evaluate the significance of the in silico concatenation used here, a phylogeny based on randomly concatenated sequences was constructed. The resulting tree (not shown) exhibits very low bootstrap values.

The minimum evolution tree based on in silico concatenated MsrA and MsrB sequences shown in Fig. 1b fails to resolve clearly the three domains of life (Bacteria, Archaea, and Eukarya) and exhibits many unresolved basal branches. Nonetheless, it represents a significant improvement over the phylogenies shown in Fig. 2 and has good bootstrap values for branches depicting recent speciation events, which are in good agreement with trees based on canonical phylogenetic markers including 16/18S rRNA cladograms. Although in this tree the prote-

obacteria do not group into a single clade, their different divisions are well defined and include branches with high bootstrap values ($>70\%$) that clearly differentiate (a) separate branches of α-, β-, and ε-proteobacteria, as well as three clades for γ-proteobacteria (enterobacteriales, xanthomonadales, pseudomonadales); (b) the four firmicutes subgroups (mollicutes, clostridia, bacillales, and lactobacillales), although the *Deinococcus radiodurans* branch separates early from the bacillales; and (c) the actinobacteria. All naturally concatenated sequences group within the Firmicutes clade, and their incorporation as a branch of the in silico concatenated sequence cladogram does not modify its overall topology (Supplementary Fig. 3). Quite interestingly, the euryarcheaota species included here cluster with *Geobacter sulfurreducens* (gsu), a δ-proteobacterium that shares ecological niches with some of the euryarchaeotal species, suggesting a possible case of lateral gene transfer in related environments. With the exception of the branch corresponding to the *C. elegans* MsrAB sequences, all other eukarya included in this sample are grouped together and branch with ε-
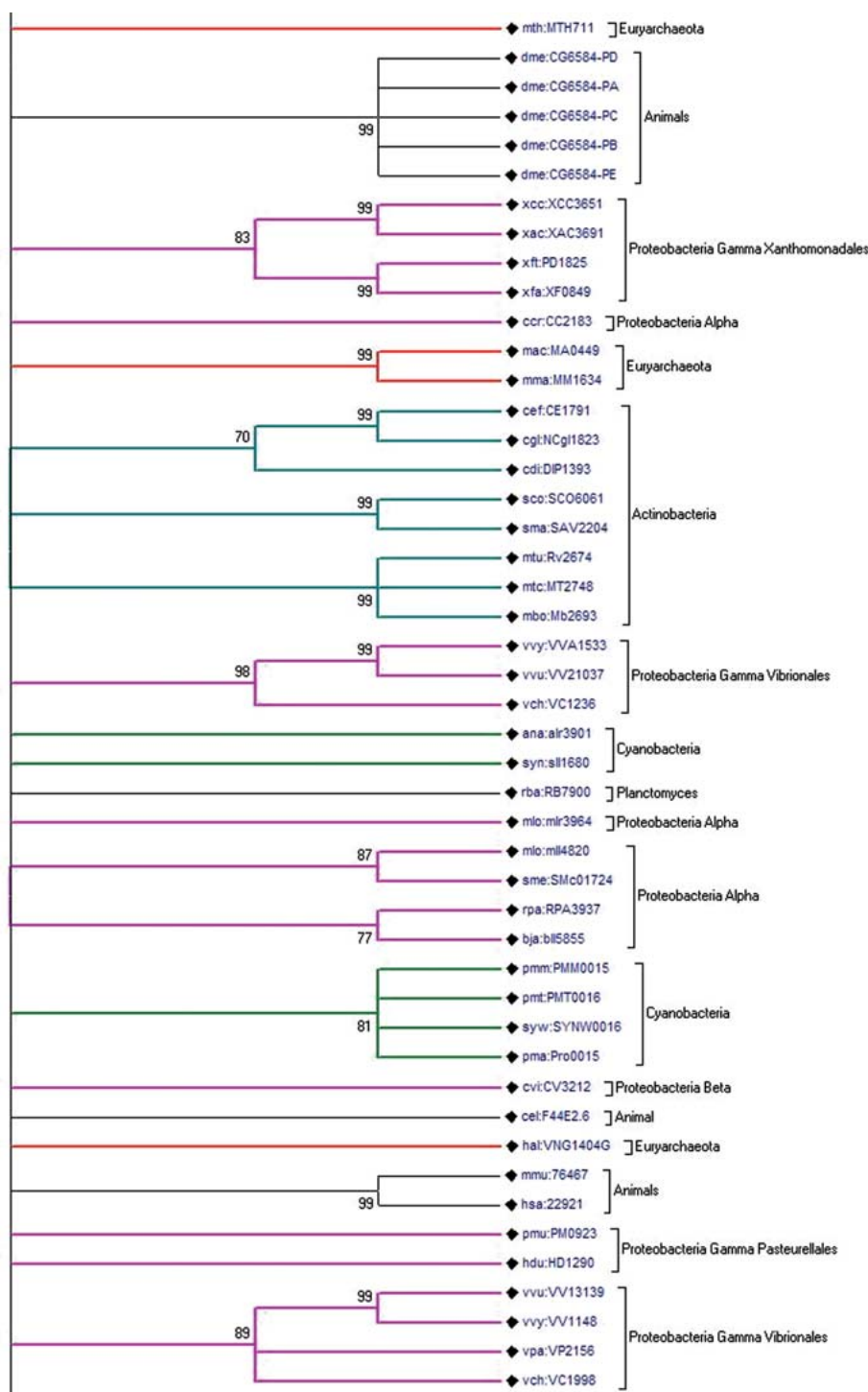
Fig. 2 (b). Continued

proteobacteria. Despite the large number of unresolved branches, the Zn-bearing motif does not exhibit a random biological distribution.

Discussion and Conclusions

Analysis of the three-dimensional structure of MsrA and MsrB has demonstrated the independent origin of these two enzymes (Lowther et al. 2002); their active sites exhibit approximate mirror symmetry.

However, comparisons of their crystal structures have shown that MsrB is formed by several coiled β-sheets, while MsrA is an α/β protein that has a β-sheet core surrounded by α-helices (Gladyshev 2002). Comparison of MsrA and B structures with available databases shows that these two enzymes do not have structural similarity to other known proteins (Gladyshev 2002; Lowther et al. 2002). Although both MsrA and thioredoxin are thiol oxidoreductases with similar reaction mechanisms and comparable
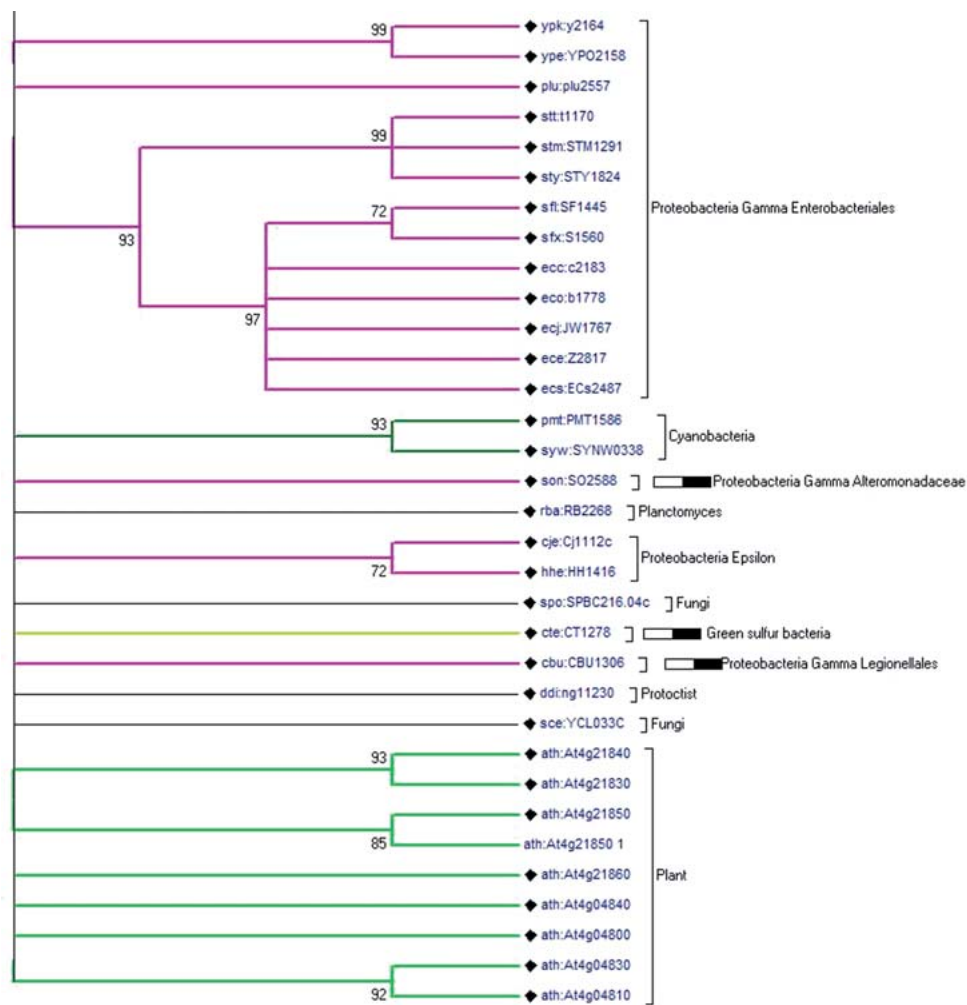
Fig. 2 (b). Continued

architectures (Gladyshev 2002), no sequence similarity or common structural folds are shared between them. MsrA has been included in the SCOP database classification within the ferredoxin-like fold group together with 48 other superfamilies, but this does not allow inferences about their evolutionary relationships, and depends on the likelihood of homology among superfamilies defined solely by one fold.

It is not always easy to rationalize the presence or absence of the methionine sulfoxide reductases in different species. Their presence in anaerobic species (including *Methanobacterium thermoautotrophicum*, *Clostridium acetobutylicum*, *C. perfringens*, and *Chlorobium tepidum*) may reflect their role in repairing proteins during transient exposures to oxygen. The lack of genes encoding MsrA and MsrB in endosymbionts and endoparasites can perhaps be understood as the outcome of a secondary adaptation that makes them unnecessary under the anaerobic conditions of the intracellular cytoplasm of their hosts. This is consistent with the many other losses that the genomes of these organisms have undergone.

There are two possible explanations of the lack of detectable MsrA and MsrB genes in the bacteria

Aquifex aeolicus and *Thermotoga maritima*, and in the archaea *Thermoplasma acidophilum*, *T. volcanii*, *Aeropyrum pernix*, *Pyrobaculum aerophilum*, and *Sulfolobus tokodaii*. In these species the corresponding genes may have been lost secondarily, or their absence may indeed be an ancestral character. Given the low levels of free oxygen in which these species live (Jannasch and Mottl, 1985; Brock 1986; David E. Graham, personal communication), which may be comparable to those of the primitive environment, both explanations are feasible. The absence of MsrA and MsrB in these species could perhaps be explained by the early development of other protective mechanisms, including antioxidant ferritin-like di-iron-carboxylate proteins (Wiedenheft et al. 2005). The presence of the MsrA sequence in *S. solfataricus* is probably due to a horizontal gene transfer event.

Although MsrA and MsrB have an ample biological distribution, they are small proteins that exhibit low sequence conservation. The inconsistencies of MsrA and MsrB trees (Figs. 2a and b) with the 16/18S rRNA topology demonstrate the limited value of these enzymes as phylogenetic markers, due to a

Table 2. Number of symmetric differences among the following trees: (1) 16/18S rRNA, (2) MsrA, (3) MsrB, (4) *in silico* concatenated MsrAB, (5) randomly concatenated MsrAB sequences, and (6) randomly generated 16/18SrRNA tree

	1	2	3	4	5	6
1	0					
2	60	0				
3	58	74	0			
4	48	52	48	0		
5	81	73	85	71	0	
6	122	122	122	122	121	0

Note. These results are based solely on the comparison of trees constructed with species having one copy of MsrA and one of MsrB.

combination of their small sizes, their unequal rates of substitution, a number of paralogous duplications (as in the case of cyanobacteria), and lateral transfer events. However, several clades in the trees depicting naturally and artificially concatenated sequences (Figs. 1a and b) suggest that the phylogenies of MsrA and MsrB still record recent speciation events. This conclusion is supported by the nonrandom distribution of the Zn-bearing MsrB sequences in all trees depicted here. The similarity between the 16/18S rRNA tree and the concatenated MsrA-MsrB phylogeny suggests that, with the few exceptions described below, these two genes have undergone joint vertical inheritance, although this does not imply that both sequences first evolved in the same species.

It is likely that the dispersal of eukaryotic MsrA and MsrB sequences within proteobacterial branches reflects the ultimate bacterial origin of the sequences observed in nucleated cells. This is consistent with the idea that nucleated cells may have acquired the enzyme with the endosymbiotic bacteria that led to mitochondria. Since the mitochondria are closely related to the γ-proteobacteria, the eukaryal MsrA and MsrB genes should be more closely related to this group. However, the observed low bootstrap values make it difficult to support this conclusion.

The comparison of the complete genomes of *Haemophilus influenzae* and *Mycoplasma genitalium* led to an inventory of 256 genes that included MsrA, which was hypothesized to correspond to the genetic complement of the ancestor of the Gram-positive and Gram-negative bacteria and, perhaps, of even older entities (Mushegian and Koonin 1996). However, other searches have failed to confirm the presence of the genes encoding MsrA in sets of highly conserved orthologues (Daubin et al. 2002; Harris et al. 2003; Yang et al. 2005; Delaye et al. 2005).

It has been suggested that cellular stress response mechanisms, including methionine sulfoxide reductase activities, may have already been present in the last common ancestor of all extant living beings (Kültz 2005). Methionine sulfoxide reductase activity is probably very ancient (Hansel et al. 2005), but it is unlikely that MsrA and MsrB were broadly distributed before the evolution of oxygenic photosynthesis. Methionine oxidation is mediated by many highly reactive chemical species, including hydrogen peroxide, hydroxyl radicals, ozone, peroxynitrite, hypochlorite, and metal catalyzed oxidation systems (Moskovitz et al. 1997). It could argued that the fact that these two enzymes lack obvious homologues is evidence that they were present prior to the ancient separation among the three major domains, and that their uneven distribution is the outcome of a complex history of secondary losses. However, the ubiquity of methionine sulfoxide reductase activity is probably best understood as the evolutionary response to the high levels of atmospheric oxygen that accumulated during prokaryotic evolution. Although the broad dissemination of the genes encoding the MsrA and MsrB proteins was probably the result of the increasing levels of free oxygen in the atmosphere, their origin probably took place locally at an earlier time, for example, in cyanobacteria or their neighbors. Unfortunately, recognition of the prokaryotic group in which these proteins originated is clouded not only by the poor preservation of phylogenetic signal of these two sequences, but also by the complex history of duplications and lateral gene transfer events that MsrA and MsrB have undergone.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Boschi-Muller S, Olry A, Antoine M, Branlant G (2005) The enzymology and biochemistry of methionine sulfoxide reductases. Biochim Biophys Acta 1703:231–238

Brock TD (ed) (1986) Thermophiles: general, molecular, and applied microbiology. John Wiley & Sons, New York

Brown JR, Douady CJ, Italia MJ, Marshall WE, Stanhope MJ (2001) Universal trees based on large combined protein sequence data sets. Nature Genetics 28:281–285

Daubin V, Gouy M, Perriere G (2002) A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. Genome Res 12:1080–1090

Delaye L, Becerra A, Lazcano A (2005) The last common ancestor: What's in a name? Origins Life Evol Biosph 35:537–554

Dhandayuthapani S, Blaylock MW, Bebear CM, Rasmussen WG, Baseman JB (2001) Peptide methionine sulfoxide reductase (MsrA) is a virulence determinant in *Mycoplasma genitalium*. J Bacteriol 183:5645–5650

El Hassouni M, Chambost JP, Expert D, van Gijsegem F, Barras F (1999) The minimal gene set member msrA, encoding peptide methionine sulfoxide reductase, is a virulence determinant of the peptide of the plant pathogen (MsyA) *Erwinia chrysanthemi*. Proc Natl Acad Sci USA 96:887–892

Ezraty B, Aussel B, Barras F (2005) Methionine sulfoxide reductases in prokaryotes. Biochim Biophys Acta 1703:221–229

Falkowski PG (2006) Tracing oxygen's imprint on Earth's metabolic evolution. Science 311:1724–1725

Felsenstein J (1989) PHYLIP—phylogeny inference package (version 3.2). Cladistics 5:164–166

Gladyshev VN (2002) Thioredoxin and peptide methionine sulfoxide reductase: convergence of similar structure and function in distinct structural folds. Proteins 46:149–152

Grimaud R, Ezraty B, Mitchell JK, Lafitte D, Briand C, Derrick PJ, Barras F (2001) Repair of oxidized proteins. J Biol Chem 276:48915–48920

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hansel A, Heinemann SH, Hoshi T (2005) Heterogeneity and function of mammalian MSRs: enzymes for repair, protection and regulation. Biochim Biophys Acta 1703:239–247

Harris JK, Kelley ST, Spiegelman GB, Pace NR (2003) The genetic core of the universal ancestor. Genome Res 13:407–412

Jannasch HW, Mottl MJ (1985) Geomicrobiology of deep-sea hydrothermal vents. Science 229:717–725

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27–30

Kantorow M, Hawse JR, Cowell TL, Benhamed S, Pizarro GO, Redd VN, Hejtmancik JF (2004) Methionine sulfoxide reductase A is important for lens cell viability and resistance to oxidative stress. Proc Natl Acad Sci USA 101:9654–9659

Kryukov GD, Kumar RA, Koc A, Sun Z, Gladyshev VM (2002) Selenoprotein R is a zinc-containing stereo-specific methionine sulfoxide reductase. Proc Natl Acad Sci USA 99:4245–4250

Kültz D (2005) Molecular and evolutionary basis of the cellular stress response. Annu Rev Physiol 67:225–257

Kumar RA, Koc A, Cerny RL, Gladyshev VN (2002) Reaction mechanism, evolutionary analysis, and the role of zinc in *Drosophila* methionine-R-sulfoxide reductase. J Biol Chem 277:37527–37535

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform 5:150–163

Lowther WT, Weissbach H, Etienne F, Brot N, Matthews BW (2002) The mirrowed methionine sulfoxidase reductases of *Neisseria gonorrhoeae pilB*. Nat Struct Biol 9:348–352

Margulis L (1993) Symbiosis in cell evolution: microbial communities in the archean and proterozoic eons. Freeman, San Francisco

Moskovitz J, Berlett BS, Poston JM, Stadtman ER (1997) The yeast peptide-methionine sulfoxide reductase functions as an antioxidant in vivo. Proc Natl Acad Sci USA 94:9585–9589

Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536–540

Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci USA 93:10268–10273

Rokas A, Williams BL, King N, Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425:798–804

Ruan H, Tang XD, Chen ML, Joiner ML, Sun G, Brot N, Weissbach H, Heinemann SH, Iverson L, Wu CF, Toshinori H (2002) High-quality life extension by the enzyme peptide methionine sulfoxide reductase. Proc Natl Acad Sci USA 99:2748–2753

Sharov VS, Schoneich C (2000) Diastereoselective protein methionine oxidation by reactive oxygen species and diastereoselective repair by methionine sulfoxide reductase. Free Radic Biol Med 29:986–994

Sharov VS, Ferrington DA, Squier TC, Schoneich C (1999) Diastereoselective reduction of protein-bound methionine sulfoxide by methionine sulfoxide reductase. FEBS Lett 455:247–250

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Wiedenheft B, Mosolf J, Willits D, Yeager M, Dryden KA, Young M, Douglas T (2005) An archaeal antioxidant: characterization of a Dps-like protein from *Sulfolobus solfataricus*. Proc Natl Acad Sci USA 102:10551–10556

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 87:4576–4580

Yang S, Doolittle RF, Bourke PE (2005) Phylogeny determined by protein domain content. Proc Natl Acad Sci USA 102:373–378