Single Eubacterial Origin of Eukaryotic Sulfide:Quinone Oxidoreductase, a Mitochondrial Enzyme Conserved from the Early Evolution of Eukaryotes During Anoxic and Sulfidic Times

Ursula Theissen, Meike Hoffmeister,* Manfred Grieshaber,† and William Martin**

*Institute of Botany III and †Institut für Zoophysiologie, University of Düsseldorf, Düsseldorf, Germany

Mitochondria occur as aerobic, facultatively anaerobic, and, in the case of hydrogenosomes, strictly anaerobic forms. This physiological diversity of mitochondrial oxygen requirement is paralleled by that of free-living α-proteobacteria, the group of eubacteria from which mitochondria arose, many of which are facultative anaerobes. Although ATP synthesis in mitochondria usually involves the oxidation of reduced carbon compounds, many α-proteobacteria and some mitochondria are known to use sulfide (H_2S) as an electron donor for the respiratory chain and its associated ATP synthesis. In many eubacteria, the oxidation of sulfide involves the enzyme sulfide:quinone oxidoreductase (SQR). Nuclear-encoded homologs of SQR are found in several eukaryotic genomes. Here we show that eukaryotic SQR genes characterized to date can be traced to a single acquisition from a eubacterial donor in the common ancestor of animals and fungi. Yet, SQR is not a well-conserved protein, and our analyses suggest that the SQR gene has furthermore undergone some lateral transfer among prokaryotes during evolution, leaving the precise eubacterial lineage from which eukaryotes obtained their SQR difficult to discern with phylogenetic methods. Newer geochemical data and microfossil evidence indicate that major phases of early eukaryotic diversification occurred during a period of the Earth's history from 1 to 2 billion years before present in which the subsurface ocean waters contained almost no oxygen but contained high concentrations of sulfide, suggesting that the ability to deal with sulfide was essential for prokaryotes and eukaryotes during that time. Notwithstanding poor resolution in deep SQR phylogeny and lack of a specifically α-protebacterial branch for the eukaryotic enzyme on the basis of current lineage sampling, a single eubacterial origin of eukaryotic SQR and the evident need of ancient eukaryotes to deal with sulfide, a process today germane to mitochondrial quinone reduction, are compatible with the view that eukaryotic SQR was an acquisition from the mitochondrial endosymbiont.

Introduction

ATP synthesis in most mitochondria involves the generation of a proton gradient with the help of the electron transport chain through respiratory complexes I to IV. Yet, numerous exceptions to that rule occur among anaerobic mitochondria, which can bypass complex IV or use alternative terminal acceptors other than O_2, such as nitrate or fumarate (Tielens et al. 2002). Further exceptions are hydrogenosomes, anaerobic mitochondria that generate ATP through substrate level phosphorylation without the help of a proton gradient (Martin and Müller 1998). In all hydrogenosomes and in most mitochondria, ATP synthesis involves the use of reduced carbon compounds as the electron donor in one or a series of redox reactions. Yet some mitochondria can synthesize ATP chemolithotrophically, using proton gradients that are generated with the help of electrons taken not from carbon compounds, but rather from sulfide (H_2S, HS^-, and S^{2-}), which is abundant in many anaerobic environments, such as marine sediments (Völkel, Hauschild, and Grieshaber 1995). In mitochondria of marine invertebrates from such environments, electrons from sulfide can be transferred to quinones as their entry point into the mitochondrial respiratory chain for ATP synthesis (Doeller et al. 1999; Doeller, Grieshaber, and Kraus 2001).

The terminal acceptor for electrons stemming from sulfide in mitochondria depends upon the sulfide concentration itself. This is because sulfide is a strong inhibitor of oxygen respiration (National Research Council 1979; Grieshaber and Völkel 1998). At sulfide concentrations below approximately 20 μM, electrons from sulfide can be donated to O_2 as the terminal acceptor. Sulfide concentrations in the range of 10 to 50 μM inhibit the electron transfer from cytochrome *c* to complex IV (Bagarinao and Vetter 1990; Grieshaber and Völkel 1998). At higher sulfide concentrations, mitochondrial complex IV is blocked and the electrons are donated to alternative, yet unknown, acceptors, possibly involving a similar alternative oxidase as is found in plants (Völkel and Grieshaber 1996*a*; Parrino, Kraus, and Doeller 2000). Since sulfide concentrations can reach approximately 20 mM in environments inhabited by marine invertebrates (Fenchel and Riedel 1970; Völkel, Hauschild, and Grieshaber 1995), and since sulfide is a potent toxin, both the energy-producing and detoxifying functions of sulfide oxidation are essential to mitochondrial function in these organisms.

In eubacteria, sulfide oxidation is commonly catalyzed by the flavoprotein, sulfide:quinone oxidoreductase (SQR) (also called sulfide quinone reductase). The biochemistry of eubacterial SQR has been characterized in some detail (Reinartz et al. 1998; Griesbeck, Hauska, and Schütz 2000). Eubacterial SQR catalyzes the reaction H_2S + Ubiquinone $\leftrightarrow$ $[S^{\pm 0}]$ + UbiquinoneH_2 (Griesbeck et al. 2002). The enzyme has been purified and cloned from *Rhodobacter capsulatus* (Schütz et al. 1997) and *Oscillatoria limnetica* (Arieli et al. 1994) and biochemically characterized in *Chlorobium limicola* (Shahak et al. 1992), *Rhodobacter capsulatus* (Shahak et al. 1994), *Paracoccus denitrificans* (Schütz et al. 1998), *Allochromatium vinosum* (Reinartz et al. 1998), and *Aquifex aeolicus* (Nübel et al. 2000) (reviewed in Griesbeck, Hauska, and Schütz 2000). Bacterial SQR is a single polypeptide with an

Key words: sulfide-quinone reductase, endosymbiosis, hydrogenosomes, electron transport chain, anaerobiosis.

E-mail: w.martin@uni-duesseldorf.de.

Mol. Biol. Evol. 20(9):1564–1574. 2003
DOI: 10.1093/molbev/msg174
Molecular Biology and Evolution, Vol. 20, No. 9,

apparent molecular mass of 48 to 55 kDa, is possibly active as a dimer, is membrane associated, belongs to the glutathione reductase family of flavoproteins, and is inhibited by quinone analogs at micromolar or nanomolar concentrations (Arieli et al. 1994; Schütz et al. 1997; Griesbeck et al. 2002). *Rhodobacter* SQR was shown to reside in the periplasm (Schütz et al. 1997). In *Chlorobium* and *Rhodobacter*, electrons from sulfide enter into the electron transport chain of anaerobic photosynthesis through SQR (Griesbeck, Hauska, and Schütz 2000; Griesbeck et al. 2002). In the nonphotosynthetic α-proteobacterium *Paracoccus denitrificans* (Schütz et al. 1998) and in *Aquifex aeolicus* (Nübel et al. 2000), SQR introduces electrons from sulfide into the respiratory chain (reviewed in Griesbeck, Hauska, and Schütz 2000).

Mitochondrial SQR activity has been indirectly measured in many organisms through the mitochodrion-dependent formation of thiosulfate from sulfide, for example, in the annelids *Heteromastus filiformis* and *Arenicola marina* as well as in the molluscs *Solemya reidi* and *Geukensia demissa* (Oeschger and Visman 1994; reviewed in Grieshaber and Völkel 1998), but has not been purified to homogeneity from any multicellular eukaryote. A functional mitochondrial SQR was, however, recently cloned and characterized from the ascomycete *Schizosaccharomyces pombe* (Vande Weghe and Ow 1999). *S. pombe* SQR showed marked sequence similarity to the SQR purified and extensively characterized at the biochemical level from the α-proteobacterium *Rhodobacter capsulatus* (Schütz et al. 1997) and was furthermore shown to be imported into and functional in *S. pombe* mitochondria (Vande Weghe and Ow 1999). Mitochondrial SQR activity has been most extensively studied in marine invertebrates that inhabit sulfide-rich intertidal sediments, most notably the annelid lugworm *Arenicola marina* and the ribbed mussel *Geukensia demissa* (Völkel and Grieshaber 1996*a*; Parrino, Kraus, and Doeller 2000). Both in those organisms (Doeller et al. 1999; Doeller, Grieshaber, and Kraus 2001; Völkel and Grieshaber 1996*b*) and in recent biochemical studies of SQR from chicken mitochondria (Yong and Searcy 2001) it was shown that mitochondrial sulfide consumption was coupled to ATP synthesis. Since chickens do not inhabit sulfide-rich environments, the role of SQR in their mitochondria is probably not ATP production, but may involve detoxification. The primary oxidation product of sulfide produced by mitochondrial SQR is still not known with certainty. The two-electron reaction would yield either elemental sulfur ($S^{\pm 0}$) or sulfanes (HSS_nH) as the primary oxidation product. The four-electron reaction would yield thiosulfate ($S_2O_3^{2-}$), which is the most commonly detected oxidation product (O'Brien and Vetter 1990; Völkel and Grieshaber 1992; Johns et al. 1997). Recent results by Yong and Searcy (2001) suggest that sulfanes might be produced during mitochondrial sulfide oxidation, but sulfanes have still not been directly detected.

The ability of mitochondria to perform sulfide oxidation for ATP synthesis raises the question as to the evolutionary origin of eukaryotic SQR genes, particularly from the standpoint of endosymbiotic theory (recently reviewed in Martin et al. 2001). In general, there are four simple possibilities for the origin of mitochondrial sulfide oxidation in eukaryotes. (1) The host that acquired the mitochondrion could have possessed an SQR enzyme that was retargeted to the mitochondrion to become functional there; in this case eukaryotic SQR should be related to archaebacterial SQR because the DNA replication (Tye 2000), translation (Lecompte et al. 2002), transcription (Reeve 2003), and chromatin-packaging systems (Reeve 2003) of eukaryotes are specifically related to their archaebacterial homologs. (2) The mitochondrial symbiont could have possessed the SQR enzyme, whereby the gene must have been transferred to the host's chromosomes, since SQR is not encoded in any mitochondrial DNA; in this case eukaryotic SQR should reveal a single eubacterial origin. (3) Neither host nor symbiont may have possessed SQR, and the SQR gene could have been acquired through horizontal gene transfer in organisms that inhabit sulfidic environments; in this case eukaryotic SQR should reveal multiple origins from diverse prokaryotic donors. (4) Eukaryotic SQR is an invention specific to the eukaryotic lineage; in this case eukaryotic SQR should be unrelated to prokaryotic SQR. The phylogenetic distribution of SQR-related enzymes among various genomes has been previously studied (Vande Weghe and Ow 1999; Griesbeck, Hauska, and Schütz 2000), but the phylogeny of SQR itself has not.

SQR was almost certainly an essential and possibly ubiquitous enzyme during the phase of eukaryotic evolution from 2 to 1 billion years ago, since newer geochemical evidence indicates that the Earth's ocean waters were anoxic and very sulfidic during that time (Canfield 1998; Shen, Buick, and Canfield 2001; Anbar and Knoll 2002), findings that underscore the evolutionary importance of anaerobic biochemistry in both ancient and modern eukaryotes (Tielens et al. 2002; Embley et al. 2003). Here, we report the occurrence of SQR and SQR-related enzymes among genomes of eubacteria, archaebacteria, and eukaryotes and examine the phylogeny of SQR, with particular attention to the evolutionary origin of the eukaryotic nuclear genes for mitochondrial sulfide:quinone oxidoreductase.

Materials and Methods

Database searching was performed using the sequences for the functionally characterized SQR from *S. pombe* (Vande Weghe and Ow 1999) and from *Rb. capsulatus* (Schütz et al. 1997) as queries against GenBank and against unfinished microbial genomes listed by the Institute for Genome Research (http://tigrblast.tigr.org/ufmg/). Sequence handling, data formatting, and alignment were performed with programs of the GCG Package version 10.3 (Genetics Computer Group, Madison, Wis.) and with ClustalW (Thompson, Higgins, and Gibson 1994). Alignments were reinspected and manually adjusted. For phylogenetic analyses, programs of the PHYLIP package (Felsenstein 1998), the MOLPHY package (Adachi and Hasegawa 1996), and Puzzle (Strimmer and von Haeseler 1996) were used. Uncorrected proportions of differences between sequences (p-distances) were

Source genome	Cysteine 159 *		FAD-binding domain III Glycine *		Cysteine 353 *		Group	Accession No.
Rhodobacter capsulatus SQR	AAQGAS**C**FGPAYEF	159	GFVIVDQ-HQQNPTFKN---VFAV**G**VCVAIPP	299	SWNAV**C**LADFGDK	353	I	CAA66112
Rhodobacter sphaeroides	AAQGAS**C**FGPAYEF	201	GFTLVDP-HQRNPTFPN---IFAV**G**VCVAIPP	340	TWNAV**C**LADFGDR	394	I	ZP_00005326
Magnetococcus sp.	AVQGAS**C**FGPAYEF	160	GFVKVDK-HQRNTVWPN---IYSA**G**VCVAIPP	301	TWNAI**C**LADMGDT	355	I	ZP_00044757
Aphanothece halophytica SQR	AVPGAS**C**MGPAYEF	160	GFLPVLD-TYQHPDYPS---IYSA**G**VITQLAA	291	SLEAI**C**MADFGDT	346	I	AAF72963
Nostoc sp. PCC 7120	ALPKTS**C**LGPAYEF	158	GFIPVLP-TYRHPEYAS---IYAV**G**VVVEIKP	289	TLDAI**C**FADFGNS	344	I	NP_488552
Oscillatoria limnetica SQR	AVPGAS**C**FGPAYEF	159	GFVPVTN-TYQHPKYES---VYSA**G**VIVEINP	290	TLEAI**C**IADFGDT	345	I	AAF72962
Thermosynechococcus elongatus	AAPGAS**C**FGPAYEF	156	GFLPLLP-TLQHPNYPE---IYGV**G**ISVHLPP	286	TLDAL**C**FADFGNT	341	I	NP_681079
Thiobacillus denitrificans	AVQGAS**C**YGPAYEF	160	GFITIDP-YQRNPKYPN---VYSV**G**VCVAIPP	300	TWNAI**C**LADFGDT	354	I	AAM52227
Aquifex aeolicus	AIPGVS**C**FGPAYEF	156	NMVIVNR-CFQNPTYKN---IFGV**G**VVTAIPP	293	RLSAI**C**IADFGED	347	I	NP_214500
Anopheles gambiae	PNSPVK**C**PGAPQKV	172	GFVDVNKDTLQHQKFSN---VFAI**G**DCSSSPN	307	DGYAS**C**PLVTGYN	351	II	EAA08424
Drosophila melanogaster	PNCPIK**C**AGAPQKI	203	GFVDVDQLTLQHNKYSN---VFAI**G**DSASSPN	338	DGYSS**C**PIVTGYS	382	II	NP_647877
Caenorhabditis elegans	PNTPIK**C**AGAPQKA	169	GFMDVDGGSLQSKKYPN---VFGV**G**DCMNTPN	356	DGYAS**C**PLVVSTN	400	II	NP_502729
Homo sapiens	PNTPVK**C**AGAPQKI	201	GWVDVDKETLQHRRYPN---VFGI**G**DCTNLPT	335	DGYTS**C**PLVTGYN	379	II	NP_067022
Mus musculus	PNTPVK**C**AGAPQKI	201	GWVDVDKETLQHKKYPN---VFGI**G**DCTNLPT	335	DGYTS**C**PLVTGYN	379	II	NP_067482
Schizosaccharomyces pombe SQR	PSGVLK**C**AGAPQKI	204	GFVAVDQSTTQSTKFPN---VFAI**G**DCSGLPT	339	NGYTS**C**PLLTGYG	383	II	NP_596067
Dictyostelium discoideum	PTTGVK**C**GGAPQKI	77	GFVNVDKGTLQHVKYNN---VFSL**G**DTSNLPT	209	DGYTS**C**PIKTSYS	253	II	see text
Ralstonia solanacearum	PAMPIK**C**AGAPQKA	174	GWCEVDPATLQHVRHAD---VFSL**G**DACSTPN	310	DGYGG**C**PLTVERG	354	II	NP_519663
Pseudomonas aeruginosa	PAMPIK**C**AGAPQKA	168	GWCEVDPATLQHVRHGE---IFAL**G**DVCGTAN	304	DGYGG**C**PLTVERG	348	II	NP_251035
Pseudomonas syringae	PPMPIK**C**AGAPQKA	166	GWVDVDPATLKHRQFAN---VHGL**G**DATNTSN	302	DGYGS**C**PLTVEKG	346	II	ZP_00124760
Pseudomonas fluorescens	PPIPIK**C**AGAPQKA	169	GWCEVNPHTLQHPRYPE---VFAL**G**DICGTTH	305	DGYGS**C**PLTVEKG	349	II	ZP_00084090
Synechocystis PCC6803	PATPIK**C**AGAPQKI	163	GWVDVDKFTLQHNRYDN---VFSL**G**DASSLPT	296	GGYTC**C**PLVTGYG	340	II	NP_440916
Chloroflexus aurantiacus	PSTPIK**C**GGAPQKI	178	GWVEVDKHTLQHVRYPN---VFSL**G**DCSNLPT	311	DGYTS**C**PLVTGYG	355	II	ZP_00018906
Staphylococcus aureus	PNTPIK**C**GGAPMKI	164	GWVDVNPTTLQHKSYSN---VFAL**G**DASNVPT	297	DGYTS**C**PIVTGYN	341	II	NP_370612
Pasteurella multocida	PATEMK**C**AGAPLKY	211	GWVEVEKHTLRHRRYAN---VFAV**G**DVAGVPK	348	NGYTS**C**PLITQLG	392	II	NP_246416
Rhodospirillum rubum	PITEMK**C**AGAPLKY	213	GWMEVDKASLRHRRYPN---IFGV**G**DVAGVPK	350	NGYTS**C**PLITKLG	394	II	ZP_00016380
Nitrosomonas europaea *	PPMPIK**C**AGAPQKA	334	GYCEVNPKTLQHTRFAN---IFSL**G**DACSSPN	470	DGYGA**C**PLTVENG	514	II	ZP_00004121
Burkholderia fungorum *	PPMPIK**C**AGAPQKA	303	GWVDVDQTTLRHKRFHN---VYGL**G**DVTNTPN	439	DGYGS**C**PLTVERG	483	II	ZP_00032335
Ralstonia metallidurans *	PPMPIK**C**AGAPQKA	307	GWVDVDPSSLRHKKYAN---VFAL**G**DCTNTTN	443	DGYGS**C**PLTVERG	487	II	ZP_00023899
Archaeoglobus fulgidus	MGIPHK**C**PVAPIEV	160	GWVPTDRYTLKAEGLEN---VYVV**G**DATNLPV	288	FGKAM**C**FIETGFS	332	III	NP_069393
Magnetospirillum magnetotacticum	VNAPHK**C**PVAPLEV	160	GWVPTNTKTLHREGSTN---VFVV**G**DTTNIPI	287	DGKVF**C**FVETGLG	331	III	ZP_00055086
Chlorobium tepidum	AELPFK**C**PVAPIEF	159	GYVPTHHNTLKALKHDG---VYVI**G**DATNVPT	287	DGHST**C**FIVYSKG	331	III	NP_661917
Desulfovibrio desulfuricans	CESPIK**C**PVAPLEF	159	GFLETDKGTLKSKRWEN---MYIL**G**DGTNVPT	287	DGHST**C**FILTGYE	331	III	ZP_00129507
Sulfolobus solfataricus	AKLPHR**C**PVAPLEV	164	RWVPTDKFTLRMKDHSN---VYVM**G**DATDLPV	292	GGDVL**C**YIATGTD	336	III	NP_343961
Pyrobaculum aerophilum	TSTPYK**C**PVAPYEF	160	GWVPVDRHTLQIQAQGATGAEYAI**G**DATNLPV	287	DGRVI**C**FILTGFE	331	III	NP_560139
Ferroplasma acidarmanus	ASTPFQ**C**PVAPGEF	157	GYVDVDKFTLEYSDYDN---VFAV**G**DAANFPM	285	DGFMG**C**SSIYAEQ	328	III	ZP_00000829
Sulfolobus tokodaii	YFGIIK**C**PMAPFEF	145	GFIPVDKEKLNYKDYSD---VFVV**G**DATNITL	266	DGKAM**C**AGYAGYN	310	III	NP_376504
Thermoplasma volcanium	SSQWYK**C**PPVPWEM	155	GWASTNLKDFRNPKYDD---IFAI**G**DVIAPTI	286	NKNAV**C**IIATGSS	331	III	NP_110992

FIG. 1.—Conserved regions from the alignment of SQR sequences (accession numbers given) corresponding to Cys 159, the FAD binding domain III, and Cys 353 using the numbering of Griesbeck et al. (2002). The residue number of the cysteines and the conserved glycine in FAD binding domain III (boldface type) in each sequence is given. Group designations (I, II, and III) refer to families of sequence similarity. Species names designated as SQR indicate that the sequence is known to encode SQR acitvity; those marked with an asterisk possess the N-terminal translational fusion indicated in figure 2. Gaps are indicated as dashes. *Nostoc sp*. PCC 7120 is synonymous with *Anabaena* PCC 7120.

calculated with ClustalW (Thompson, Higgins, and Gibson 1994). Logdet distances were calculated with the LDDist program available at the Web site http://artedi.ebc.uu.se/molev/software/LDDist.html. NeighborNet (Bryant and Moulton 2002) networks for representing the data were constructed with the software available at the Web site http://www.mcb.mcgill.ca/~bryant/NeighborNet/. For ProtML and Puzzle, the JTT-F matrix was used. Quartet-puzzling (QP) was employed using a gamma distribution and eight categories of rate heterogeneity.

Results

Conservation and Fusions Among SQR-Related Proteins

Using Blast, SQR homologs were identified in genomes of eubacteria, archaebacteria, and eukaryotes and retrieved from the databases. Because of poor sequence conservation, the alignment of these sequences is challenging. For example, the functionally characterized SQR proteins from *S. pombe* and *Rb. capsulatus* share only 24% amino acid identity in the pairwise Needleman-Wunsch alignment. As a consequence, the automatic alignment programs ClustalW and Pileup align neither the FAD binding domains nor strictly conserved cysteine residues identified by Bronstein et al. (2000) and Griesbeck et al. (2002) in their analyses of SQR and related sequences from prokaryotes. Using those conserved domains as a guide, and using the terminology of Griesbeck et al. (2002), we manually identified the conserved regions corresponding to SQR fingerprint 2, containing Cys 159, and SQR fingerprint 5, containing Cys 353, as well as FAD binding domain III in all sequences shown in figure 1. These anchor points of sequence conservation were used for further manual refinement of the alignment. Residues Cys 159 and Cys 353, originally identified by Bronstein et al. (2000), refer to their positions in the SQR sequence from *Rb. capsulatus* (Griesbeck et al. 2002). The spacing of these conserved motifs is generally uniform throughout the alignment. Two further sequence motifs identified by Griesbeck et al. (2002) in the analysis of seven prokaryotic SQR homologs surrounding the residues Cys 127 and His 196 are not conserved across all 37 SQR homologs shown in figure 1 and hence are not shown. Patterns of sequence similarity visible in the alignment around Cys 159, Cys 353, and the FAD binding domain suggested the presence of three distinct groups of sequence diversity which we term groups I, II, and III. Although not thoroughly exhaustive with respect to all unfinished genomes, this database search provides a representative overview of sequence diversity for the SQR family, including all functionally characterized members. Across groups I, II, and III, amino acid identity in pairwise Needleman-Wunsch comparisons of unaligned sequences was approximately 25% on average. Of the sequences indicated in

Source genome	|		*		Accession No.	
Nitrosomonas europaea	IIC**C**NRPDGEG	40	PVLAF**C**NSGNR	101	ZP_00004121	*
Burkholderia fungorum	VIC**C**NRPDGEG	34	PVLAY**C**RSGMR	94	ZP_00032335	*
Ralstonia metallidurans	VIC**C**NRPDGEA	34	PVFAY**C**RTGTR	94	ZP_00023899	*
Brucella melitensis	IIC**C**NRPDGEG	34	PVLAY**C**RSGAR	95	NP_541565	
Desulfitobacterium hafniense	VIC**C**NRPDGEA	34	PVFAY**C**RTGTR	94	ZP_00101383	
Mesorhizobium loti	VIC**C**NRPDDEQ	34	PVFAY**C**RSGAR	94	NP_108432	
Neisseria meningitidis	VIC**C**NRPDREE	47	PVLAY**C**RTGTR	108	NP_284686	
Agrobacterium tumefaciens	IVC**C**HRPDGEE	34	PMLGY**C**RSGAR	94	NP_532805	
Sinorhizobium meliloti	IVC**C**HRPDFEA	34	PMLGY**C**RSGAR	94	NP_386312	
Caulobacter crescentus	VINNRPDGED	-	PVLAF**C**RSGTR	94	NP_419740	
Magnetospirillum magnetotacticum	LINNRPDGEE	-	PVVAH**C**RSGTR	94	ZP_00051966	
Nostoc punctiforme	VLNLRSPDEP	-	PILIH**C**AAGAR	95	ZP_00107777	
Novosphingobium aromaticivorans	IVNNRPEGES	-	PVLAY**C**RSGTR	137	ZP_00093528	
Rhodobacter capsulatus	LINNRPD-EE	-	PVLAW**C**RSGTR	93	T10511	
Rhodobacter sphaeroides	IIDNRPDGEI	-	PVLAY**C**ASGNR	94	ZP_00006823	
Agrobacterium tumefaciens	IINARPDGEE	-	PVFAH**C**KGGTR	94	NP_357144	**
Mesorhizobium loti	MINARPDGEE	-	PVFAH**C**KTGTR	94	NP_103568	**
Sinorhizobium meliloti	LINNRPHKED	-	PVLAH**C**QTGGR	94	NP_435818	**
Xylella fastidiosa	IINLRPDGEE	-	SVLVH**C**KSGTR	145	ZP_00041131	**

FIG. 2.—Conserved regions in an ORF N-terminally fused present to several SQR homologs (indicated with an asterisk next to the accession number; see also fig. 1) and present as an independent ORF in several genomes (no asterisk). Sequences marked with two asterisks possess at their C-terminus not SQR, but rather an unrelated ORF that is annotated in several genomes on the basis of weak sequence similarity as a β-lactamase (see text). The strictly conserved cysteine is marked with ''*,'' and the partially conserved cysteine is marked with ''|.'' Their positions are indicated as in figure 1.

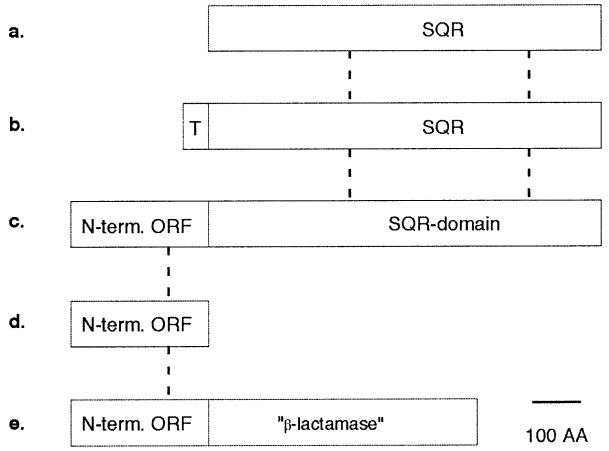


FIG. 3.—Schematic overview of fusions involving SQR. The positions of conserved cysteine residues is indicated by dotted lines. (*a*) SQR as it occurs in most prokaryotes. (*b*) SQR as it occurs in *S. pombe*. T indicates the mitochondrial transit peptide identified (Vande Weghe and Ow 1999). (*c*) SQR as it occurs in sequences marked with an asterisk in figures 1 and 2. (*d*) Unfused ORF present in many prokaryotic genomes as indicated in figure 2. (*e*) Fusion of the ORF with an SQR-unrelated sequence as it occurs in sequences marked with two asterisks in figure 2.

figure 1, only the homologs from *Rb. capsulatus* (Schütz et al. 1997), from *S. pombe* (Vande Weghe and Ow 1999), and from the cyanobacteria *Aphanonthece* and *Oscillatoria* (Bronstein et al. 2000) have been shown to represent active SQR enzymes.

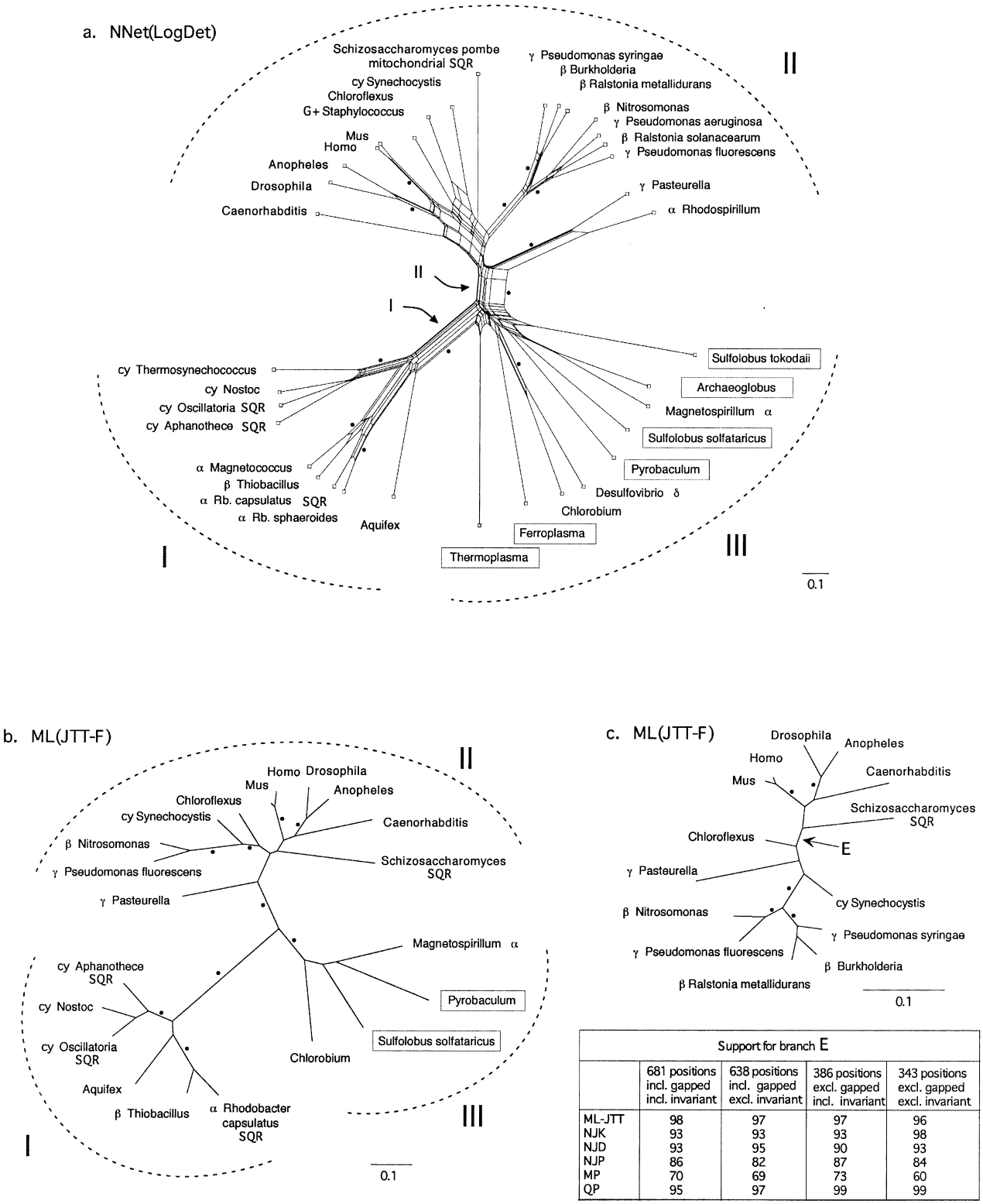
The partial sequence of an SQR homolog from *Dictyostelium discoideum* that contained the three strictly conserved motifs was assembled from unannotated sequence data (http://db.dictybase.org/) but lacked 125 N-terminal residues relative to the *S. pombe* sequence (fig. 1). One additional SQR homolog each was identified in the *Caenorhabditis elegans* (GenBank accession number NP_500688) and *S. pombe* (GenBank accession number T43278) genomes, but these sequences appeared to be C-terminally truncated and lacked the Cys 353 region. The *Caenorhabditis* homolog shown (GenBank accession number NP_502729) possesses a unique approximately 50–amino acid insertion at about position 230 that might represent a translated intron, but we were unable to identify possible intron donor and acceptor sites in the nucleotide sequence that would justify its removal. Searches among the genome sequence data for the anaerobic protists *Giardia intestinalis* and *Entamoeba histolytica* revealed no identifiable SQR homologs. No SQR homologs were detected among photosynthetic eukaryotes, although SQR detects glutathione reductase as a distant relative in several eukaryotic genomes in BLAST searches (data not shown).

The SQR homologs from *Nitrosomonas europaea*, *Burkholderia fungorum*, and *Ralstonia metallidurans* possess a well-conserved approximately 180–amino acid open reading frame (ORF) N-terminal to the SQR domain that is present as an independent ORF of unknown function in several prokaryotic genomes (fig. 2) but does not always co-occur with SQR. This N-terminal ORF contains one strictly conserved cysteine and an additional cysteine conserved in some sequences (fig. 2). Homologs of the N-terminal ORF from *Agrobacterium*, *Sinorhizobium*, *Mesorhizobium*, and *Xylella* are fused with yet another approximate 240–amino acid ORF of unknown function (fig. 3) that shares no similarity with SQR and that is annotated in several entries as a member of the metallo-β-lactamase family by virtue of its similarity to the PFAM (Bateman et al. 2002) lactamase B family (PF00753; http://pfam.wustl.edu/), which includes several thiolesterases.

SQR Phylogeny

Phylogenetic analyses started with the complete alignment containing 36 OTUs (operational taxonomical units; sequences) and 715 positions, only seven of which were constant. Recalling that these sequences are highly divergent (<25% identity in many comparisons), and recalling that neither Pileup using the PAM250 matrix nor ClustalW using the Blosum matrices recovered the conserved cysteines (fig. 1) for all sequences, we started analyses with a NeighborNet network of protein logdet distances to obtain graphical representation of the data (fig. 4*a*). Neighbor-joining (NJ) (Saitou and Nei 1987) trees using the uncorrected proportion of amino acid differences (p-distance or Hamming distance) as a distance measure (NJP) gave a very similar picture. Although the p-distance is a rough measure, it has been shown to perform well in computer simulations when sequence divergence is high because it has a lower variance than correction procedures (Nei 1996; Nei and Kumar 2000). At the 95% bootstrap proportion (BP) level, the NJ tree of p-distances recovered all branches that protein logdet did plus only one additional branch—that joining *Archaeoglobus* and *Magnetospirillum*, which was found at BP = 93% using protein logdet.

The groups I, II, and III identifiable in the alignment were recovered in the NJP topology, although the BP for group III was only 73% using NJP (48% using logdet). Local rearrangement using ProtML starting from the NJ tree of ML distances and using the JTT-F matrix strongly separated groups I, II, and III at BP greater than 95% each



Support for branch E				
	681 positions incl. gapped incl. invariant	638 positions incl. gapped excl. invariant	386 positions excl. gapped incl. invariant	343 positions excl. gapped excl. invariant
ML-JTT	98	97	97	96
NJK	93	93	93	98
NJD	93	95	90	93
NJP	86	82	87	84
MP	70	69	73	60
QP	95	97	99	99

FIG. 4.—Phylogenetic analyses of SQR sequences. (*a*) NeighborNet (NNet) network of protein logdet distances among SQR sequences. The NNet network depicts splits in the data as series of parallel lines, conflicting or non–tree-like signals for a given taxon or group of taxa can thus be represented. For example, in addition to the split linking *Rhodospirillum* and *Pasteurella*, there is also one linking *Rhodospirillum* and *Sulfolobus tokadaii* (a conflicting or non–tree-like signal). The splits joining members of group I and group II, respectively, are indicated. *Rhodobacter* is abbreviated as "*Rb.*" (*b*) ProtML tree using the JTT-F matrix for sequences that pass the significance test for amino acid compositional equilibrium at $P = 0.95$. (*c*) ProtML tree using the JTT-F matrix for sequences from group II that pass the significance test for amino acid compositional equilibrium at $P = 0.95$. Bootstrap support for the monophyly of eukaryotic SQR in with various methods and for various subsets of the data is summarized in the box below the tree in (*c*) (see text). Branches supported at a BP $\geq$95% with the method indicated are labeled with a dot; for the NNet graph, dots indicate

and furthermore grouped the eukaryotic sequences together (not shown) but with a low BP (59%). However, the amino acid composition of 14 sequences in the 36 OTU data deviated from the expectation at $P = 0.95$ as estimated with Puzzle: *Thermosynechococcus*, *Magnetococcus*, *Rb. sphaeroides*, *P. syringae*, *P. aeruginosa*, *R. metallidurans*, *R. solanacearum*, *Staphylococcus*, *Rhodospirillum*, *Archaeoglobus*, *Desulfovibrio*, *Ferroplasma*, *S. tokodaii*, and *Thermoplasma*. Removing these from the alignment (22 OTU data) and rechecking revealed that *Burkholderia* deviated, upon removal of which all remaining 21 sequences passed the significance test for amino acid composition. ProtML analysis of that data after excluding gapped sites (21 OTUs, 343 positions) produced the topology in figure 4*b*, which had a very low BP (< 50%) for the monophyly of the eukaryotic sequences, support for which increased to 86% when gapped sites were included (not shown).

However, the 21 OTU data (fig. 4*b*) was still highly divergent, with between group (I, II, and III) amino acid identity of sequences from the multiple alignment in the range of only 20%. To examine the possible monophyly of eukaryotic SQR more closely, we investigated the phylogeny of the 18 sequences belonging to group II in figure 4*a*. This data set (18 OTUs, 681 sites including gaps) was inspected for amino acid compositional bias, whereby *P. aeruginosa*, *R. solanacearum*, and *Staphylococcus* failed the frequency distribution test at $P = 0.95$. Removal of these sequences from the data and rechecking revealed that *Rhodospirillum* failed, leaving 14 OTUs, all of which passed the amino acid frequency distribution test. The resulting alignment (14 OTUs, 681 sites) contained better sequence conservation with 43 invariant sites and all aligned sequences being at least 30% identical in all comparisons, although this is still generally poor sequence conservation. Nonetheless, analysis with ML, NJ, and QP, also after exclusion of gapped positions (14 OTUs, 386 sites), provided good support for the monophyly of eukaryotic SQR (branch E in figure 4*c*) with all methods except MP (fig. 4*c*).

Discussion

Sulfide:quinone oxidoreductase belongs to the larger family of disulfide oxidoreductase (DiSR) flavoproteins that includes glutathion reductases, the lipoamid dehydrogenase (E3) subunit of pyruvate dehydrogenase, thioredoxin reductase, and—importantly—flavocytochrome *c* (Fcc or FCC) (Schütz et al. 1997; Griesbeck et al. 2002). FCC is used by numerous prokaryotes for sulfide oxidation as an alternative to SQR (Schütz et al. 1997; Griesbeck et al. 2002). However, sequence comparisons revealed that all SQR sequences examined here are more similar to each other than they are to FCC or other members of the disulfide oxidoreductase flavoprotein family (fig. 5). For example, in Blast comparisons to GenBank, FCC members detect each other at E-values roughly less than 10^{-30} and usually share more than 25% sequence identity in pairwise comparisons, but FCC members detect SQR members at E-values roughly greater than 10^{-10}, with which they usually share less than 20% sequence identity in pairwise comparisons, and vice versa. Other members of the DiSR family were detected at much lower similarity levels. This indicates that SQR and FCC are distinct but specifically related subfamilies within the larger family of DiSR flavoproteins. Notably, many of the prokaryotic genomes surveyed here encode both SQR and FCC homologs: *Chlorobium* (one SQR and two FCCs), *Magnetospirillum*, *R. solanacaerum*, *R. metallidurans*, *Magnetococcus*, *Aquifex*, *Rb. sphaeroides*, *Paracoccus*, *Solfolobus*, and *Thermoplasma*.

Sequence Conservation and Motifs in SQR

Sequences such as SQR that share only about 20% identity in many pairwise comparisons pose a challenge to phylogenetic analysis. In such cases, phylogeny is strongly aided by information from three-dimensional structures (Schütz et al. 2000; Baymann et al. 2003). Crystal structures are available for several members of the DiSR flavoprotein family, including of FCC from *Chromatium vinosum* (Chen et al. 1994), but there are currently no three-dimensional structures available for SQR. In studies of ancient protein phylogeny or poorly conserved sequences, as with the present SQR data, the discrepancy between sequence conservation and structure conservation can become severe, as underscored by Rost (1997), who found that the majority of pairwise comparisons among protein sharing a common structure reveal only about 8% to 9% sequence identity. At low sequence identity, many assumptions of phylogenetic methods are almost certainly violated. For example, recent computer simulations showed that the fraction of sites determined to be neutral by the criterion of protein folding thermodynamics fluctuates in a manner that depends upon the randomly chosen neutral mutations accumulated by the sequence as it mutates through sequence space (Bastolla et al. 2002). Such findings suggest that at levels of sequence similarity where structural constraints can cause different fractions of sites in a protein to become neutral, current phylogenetic inference methods, which are founded in neutral theory, will reflect patterns of shared sequence similarity but will unlikely recover the true tree. Despite these problems, if one wishes to study the evolution of SQR, one has to work with the degree sequence conservation that SQR has to offer.

Based on spectroscopic and mutational analyses, Griesbeck et al. (2002) proposed a mechanism for the SQR that involved the participation of three cysteine residues,

←

splits with a BP ≥95% in NJ trees of logdet distances. Eubacterial taxon designations are indicated for major recognized groups: α, β, γ, δ, α-proteobacteria, etc.; cy indicates cyanobacteria, and G+ indicates gram-positives. Sequence groups I, II, and III are indicated at the periphery of the tree in (*a*) and (*b*). Species names designated as SQR indicate that the sequence is known to encode SQR activity. Archaebacterial sequences are boxed. Eukaryotic sequences are indicated in bold. The scale bar indicates 0.1 substitution per site with the respective method.

Cys 127, Cys 159, and Cys 353; however, they also noted that a catalytic mechanism involving only two cysteine residues instead of three would be compatible with the available data. We found that Cys 159 and Cys 353 are strictly conserved in SQR homologs investigated here, but we found no evidence for a conservation of Cys 127 outside of group I. Cysteine residues were lacking in several SQR sequences within 50 amino acids N-terminal of Cys 127 to Cys 159 (alignment available upon request) and also in the *S. pombe* SQR, the only sequence outside of group I that has been shown to be directly involved in sulfide oxidation.

Noting that Cys 127 is missing in *S. pombe* SQR, Griesbeck et al. (2002) pointed out that the K_m values of roughly 2 mM each for sulfide and quinone measured for *S. pombe* SQR are 1000-fold higher than for the eubacterial enzymes. This low substrate affinity could, in principle, cast doubt on the functional identity of *S. pombe* SQR (HMT2, the product of the *hmt2* gene) as a functional SQR enzyme. However, Vande Weghe and Ow (1999) showed (1) that isolated mitochondria from *hmt2*$^+$ *S. pombe* cells could reduce exogenous quinones with sulfide, whereas *hmt2*$^-$ could not, (2) that *hmt2*$^+$ *S. pombe* cells could oxidize endogenously produced sulfide, whereas *hmt2*$^-$ could not, (3) that HMT2 produced in *E. coli* is a flavoprotein, (4) that his-tagged HMT2 purified from *E. coli* reduces quinones in a sulfide-dependent manner in vitro, albeit with poor kinetic constants, and (5) that HMT2 resides in mitochondria. Thus, despite the high K_m values measured for *S. pombe* HMT2 produced in *E. coli* (which lacks SQR), the brunt of evidence indicates that *S. pombe* HMT2 is an active mitochondrial SQR, but there remains the possibility that it requires an additional subunit or factor not required by group I SQR for full activity. This possibility and the lack of Cys 127 in several SQR sequences drew our attention to the ORF of unknown function translationally fused to the N-terminus of group I SQR from *Nitrosomonas europaea*, *Burkholderia fungorum*, and *Ralstonia metallidurans*. These possess a strictly conserved cysteine residue at position 40 of the *Nitrosomonas* sequence (fig. 2). Database searching revealed that this ORF is present in numerous eubacterial genomes and that Cys 40 is strictly conserved in all homologs, perhaps suggesting that it might be able to assume the function of Cys 127 in SQR from group II. However we could not identify this ORF in all genomes whose group II SQR lacks Cys 127, for example, *Pasteurella multocida*, leaving the question of whether two or three cysteine residues are involved in the SQR catalytic mechanism (Griesbeck et al. 2002) open from this standpoint.

SQR Sequence Diversity and Lateral Gene Transfer Among Prokaryotes

SQR homologs encompass three groups of sequence diversity that are nonuniformly distributed across eubacteria, archaebacteria, and eukaryotes (figs. 1 and 4*a* and *b*). Group I contains the functionally characterized eubacterial SQR enzymes from cyanobacteria and *Rhodobacter capsulatus* and furthermore contains only eubacterial homologs. Group III contains functionally uncharacterized

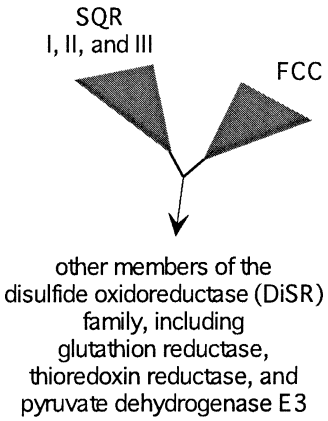


FIG. 5.—Schematic phylogeny depicting sequence similarity shared between SQR and FCC relative to other members of the DiSR family as estimated by Blast results and sequence identity in pairwise comparisons.

SQR homologs detected in sequenced archaebacterial genomes in addition to eubacterial homologs from the sulfate reducer *Desulfovibrio*, from the anaerobic, photosynthetic, green sulfur bacterium *Chlorobium*, and from the α-proteobacterium *Magnetospirillum*. SQR sequences in group II (fig. 4*a* and *b*) comprise eubacterial and eukaryotic homologs.

Lateral gene transfer (LGT) exists among prokaryotes and has become a major issue in gene and genome evolution (Gogarten, Doolittle, and Lawrence 2002). The present analysis suggests that also SQR genes may have transferred among prokaryotes during evolution, judging from the interleaving of eubacterial and archae distribution of α-proteobacterial homologs (fig. 4*a*). However, when sequences possessing significant amino acid bias are removed from the data (fig. 4*b*), the degree of interleaving also decreases, suggesting that a phylogenetic argument for unrestricted LGT of SQR genes among prokaryotes cannot be made for these data. However, the presence of four robustly clustering cyanobacterial SQR genes in group I and the presence of a single cyanobacterial SQR in group II (*Synechocystis*) suggests that the *Oscillatoria*-type SQR (group I) might represent the endogenous cyanobacterial gene, whereas *Synechocystis* may have picked up its SQR from a proteobacterial donor (fig. 4*c*). The interleaving of β-proteobacterial and γ-proteobacterial homologs in figure 4*c* indicates further probable workings of LGT for SQR among prokaryotes.

A Single Origin of Eukaryotic SQR, a Eubacterial Relict from the Anoxic and Sulfidic Past

The present analyses provide evidence for a single eubacterial origin of eukaryotic SQR, indicating that eukaryotes sampled here acquired the gene for mitochondrial SQR once in evolution from a eubacterial donor (fig. 4*c*). The nature of that eubacterial donor is highly relevant to the issue of mitochondrial evolution and eukaryote origins. There are two simple possiblities: the donor of the SQR gene either was the ancestor of mitochondria or was not.

On the one hand, arguing meekly against the view that the SQR donor was the mitochondrial endosymbiont is a single finding, namely that eukaryotic SQR does not specifically branch with α-proteobacterial SQR. Indeed, it has recently been argued that any eukaryotic nuclear gene that is to be inferred to be of mitochondrial origin must be shown to branch specifically with α-proteobacterial homologs (Kurland and Andersson 2000; Canback, Andersson, and Kurland 2002). However, that view is probably too simplistic for several reasons. First, SQR is not a highly conserved protein, such that the early evolution of this eukaryotic gene as viewed from the perspective of phylogenetics may have simply been obscured by mutation. Sequence conservation in SQR permits one to trace the origins of the eukaryotic gene to eubacteria, but tracing it to any particular eubacterial lineage on the basis of 30% sequence identity is probably asking too much of phylogenetic inference methods. Such loss of phylogenetic signal among poorly conserved genes has been well documented in genome-wide phylogenies involving genes acquired from chloroplasts (Martin et al. 2002). Second, the overall pattern of sequence similarity in figure 4*a* suggests that contemporary α-proteobacteria may themselves have acquired their SQR genes from different sources, as evidenced by the presence of α-proteobacterial SQR homologs in group I (*Rhodobacter*), group II (*Rhodospirillum*, very close to a γ-proteobacterial homolog), and group III (*Magnetospirillum*, branching among archaebacterial homologs). It is reasonable to assume that α-proteobacteria were undergoing LGT, also for SQR genes, at the time of mitochondrial origins and subsequently. Incorporating LGT into evolutionary thinking thus makes it difficult to pinpoint exactly which genes the ancestor of mitochondria possessed and/or contributed to eukaryotes on the basis of today's sequence comparisons. In other words, allowing for the existence of LGT during prokaryotic evolution (Gogarten, Doolittle, and Lawrence 2002), no single contemporary α-proteobacterium can be expected to contain exactly the same set of orthologous genes as the ancestral mitochondrial endosymbiont did (Rotte et al. 2001). Third, the current sampling of eubacterial lineages is currently quite sparse; in time, eubacterial homologs that are more closely related to eukaryotic SQR might be found. Thus, the lack of an α-proteobacterial branch for this poorly conserved and laterally transferred gene (SQR) does not constitute clear evidence against its mitochondrial origin.

On the other hand, several findings argue in favor of the view that the eukaryotic SQR gene was acquired from the ancestor of mitochondria. First, eukaryotic SQR functions in the mitochondrial membrane the same way that α-proteobacterial SQR functions in the eubacterial membrane, donating electrons from sulfide to quinones. Hence, a eubacterium with a diversified (facultatively) anaerobic electron transport chain would be the most likely SQR gene donor, for example one that could use fumarate as an electron acceptor. Eubacteria such as the α-proteobacterium *Rhodospirillum rubrum*, which possesses SQR, commonly use rhodoquinone (Okayama et al. 1968) alternatively to ubiquinone in their anaerobic electron transport chain, just like anaerobic mitochondria do today (Tielens et al. 2002). The inheritance by mitochondria of preexisting and functioning aerobic and anaerobic components in the same electron transport chain from a facultatively anaerobic ancestor of mitochondria that was perhaps similar in overall physiology to facultatively anaerobic α-proteobacteria, such as *Rhodospirillum*, *Paracoccus denitrificans* (John and Whatley 1975), or *Rhodobacter*, which possess SQR (Schütz et al. 1998; Griesbeck, Hauska, and Schütz 2000), seems much more likely than the piece-by-piece addition during eukaryotic evolution of anaerobic components (SQR, rhodoquinone, etc.) to an (hypothetical) ancestrally aerobic mitochondrial electron transport chain as envisaged by those who argue for an origin of mitochondria from strictly aerobic *Rickettsia*-like parasites (Kurland and Andersson 2000).

Second, the nuclear gene for SQR was apparently acquired once in eukaryotic evolution, not several times as would be predicted under models that envisage lateral gene acquisition from food bacteria as the major source of eubacterial, but apparently non–α-proteobacterial, genes in eukaryotes (Doolittle 1998). Such single acquisition with a phylogenetically unresolved eubacterial origin as seen for SQR is also observed for several other proteins involved in anaerobic ATP synthesis in eukaryotes, for example, [Fe]-hydrogenase (Horner, Foster, and Embley 2000; Horner et al. 2002), pyruvate:ferredoxin oxidoreductase (Horner, Hirt, and Embley 1999; Rotte et al. 2001; Embley et al. 2003), many glycolytic enzymes (Hannaert et al. 2000), and NADH oxidase (Nixon et al. 2002), not to mention many other eukaryotic proteins that are not involved directly in ATP synthesis, such as proteasome homologs HslV and HslU (Couvreur et al. 2002). Notably, [Fe]-hydrogenase and pyruvate:ferredoxin oxidoreductase possess several FeS clusters (Chabriere et al. 1999; Peters 1999), as do several proteins of the mitochondrial respiratory chain (Burger et al. 1996; Friedrich and Schiede 2000). Recent findings indicate that many proteins required for the assembly of FeS clusters are localized in mitochondria (Lill and Kispal 2000) and related organelles such as hydrogenosomes (Tachezy, Sanchez, and Müller 2001) and mitosomes (Katinka et al. 2001; Williams et al. 2002). The emerging monophyly of FeS cluster assembly in eukaryotes suggests that it was acquired en bloc from the ancestor of mitochondria (Huynen et al. 2001), as we suggest here for an SQR-containing mitochondrial respiratory chain.

Third, newer evidence suggests that during the period of Earth's history from 2 billion years ago to 1 billion years ago (2 to 1 Ga) SQR must have been very important, if not essential, for most, if not all, eukayrotes, at least the ones that inhabited the oceans. This is because the sulfur isotope record indicates that biological sulfate reduction, which produces sulfide, was highly active and globally widespread during that time (Canfield 1998; Shen, Buick, and Canfield 2001; Anbar and Knoll 2002). The consequence is that Earth's oceans subsurface water would have been both anoxic (without oxygen) and sulfidic (laden with sulfide) during that time. Anbar and Knoll (2002) discussed this anoxic, sulfidic marine environment in the context of low resulting copper and molybdenum concentrations, which they argued to have

possibly impaired eukaryotic diversity, because these are important trace elements for eukaryotes. However, from the standpoint of our present findings, the more immediate problem posed by such environments for early eukaryotes would have been (1) ATP production without oxygen and (2) dealing with high concentrations of sulfide. Put another way, only osmotrophic eukaryotes such as fungi would have been limited by trace element availability—phagocytosing eukaryotes would have been able to obtain their trace elements from ingested prey, but all subsurface eukaryotes during the period from 1 to 2 Ga would have been confronted with high sulfide concentrations. SQR is the mechanism that contemporary eukaryotes use to deal with high sulfide concentrations today, both in terms of detoxification and in terms of utilizing sulfide for mitochondrial ATP synthesis (Grieshaber and Völkel 1998). It is therefore reasonable to assume that ancient eukaryotes dealt with sulfide the same way as contemporary eukaryotes do, namely with mitochondrial SQR. Hence, the sulfidic and anoxic phase of Earth's history revealed by the sulfur isotope record does not lead to the prediction of limited eukaryotic diversity during the period from 1 to 2 Ga as suggested by Anbar and Knoll (2002), rather it leads to the prediction that eukaryotes diversified during anaerobic times and therefore that they should have preserved abundant traces of that anaerobic past—which they have, particularly in their mitochondria and hydrogenosomes (Martin and Müller 1998; Tielens et al. 2002; Embley et al. 2003) and also in the form of mitochondrial SQR.

Unicellular eukaryotes are at least 1.5 Ga old (Javaux, Knoll, and Walter 2001) and multicellular red algae are at least 1.2 Ga old (Butterfield 2000), meaning that differentiation of eukaryotic lineages below the plant lineage occurred in an anoxic and sulfidic world. Thus, eukaryotes that today inhabit anoxic and sulfidic marine environments did not necessarily have to become especially adapted to such conditions, nor did they need to acquire SQR genes by lateral transfer to do so. Rather, it seems that they ''grew up'' in an anoxic and sulfidic world and that mitochondrial SQR is simply a relic retained from that phase of eukaryotic history, whereby it still fulfills those same essential functions in modern eukaryotes from sulfidic habitats. Newer data indicate the fungal-animal divergence to be among the deepest branches in the eukaryotic tree (Stechmann and Cavalier-Smith 2002), such that the animal and fungal lineages sampled here cover much of the depth but not the breadth of eukaryotic diversity. In eukaryotes from aerobic and/or nonsulfidic habitats, such as *S. pombe*, the SQR gene and activity have nonetheless been retained (Vande Weghe and Ow 1999), perhaps for detoxification functions, and the SQR gene has apparently been lost in many lineages, among them *Arabidopsis* and *Saccharomyces*. No SQR homologs have yet been sequenced from those eukaryotes in which mitochondrial SQR has been most extensively characterized at the biochemical level: marine invertebrates (Grieshaber and Völkel 1998; Doeller, Grieshaber, and Kraus 2001) and chicken (Yong and Searcy 2001). However, work on the marine invertebrates is ongoing. Clearly, our prediction is that SQR from these eukaryotes will share the same origin as *S. pombe* SQR.

Acknowledgments

We thank the Deutsche Forschungsgemeinschaft for financial support, Carmen Rotte for critical reading, and Günter Hauska for many helpful comments on the manuscript.

Literature Cited

Adachi, J., and M. Hasegawa. 1996. Computer science monographs, No. 28. MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. Institute of Statistical Mathematics, Tokyo.

Anbar, A. D., and A. H. Knoll. 2002. Proterozoic ocean chemistry and evolution: a bioinorganic bridge. Science **297**:1137–1142.

Arieli, B., Y. Shahak, D. Taglicht, G. Hauska, and E. Padan. 1994. Purification and characerization of sulfide-quinone reductase, a novel enzyme driving anoxygenic photosynthesis in *Oscillatoria limnetica*. J. Biol. Chem. **269**:5705–5711.

Bagarinao, T., and R. D. Vetter. 1990. Oxidative detoxification of sulfide by mitochondria of the California killifish *Fundulus parvipinnis* and the speckled sanddab *Citharichthys stigmaeus*. J. Comp. Physiol. **160B**:519–527.

Bastolla, U., M. Porto, H. E. Roman, and M. Vendruscolo. 2002. Lack of self-averaging in neutral evolution of proteins. Phys. Rev. Lett. **89**: art. no. 208101.

Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S. R. Eddy, S. Griffiths-Jones, K. L. Howe, M. Marshall, and E. L. L. Sonnhammer. 2002. The *Pfam* protein families database. Nucleic Acids Res. **30**:276–280.

Baymann, F., E. Lebrun, M. Brugna, B. Schoepp, M.-T. Guidici-Oritconi, and W. Nitschke. 2003. The redox protein construction kit: pre-last universal common ancestor evolution of energy conserving enzymes. Philos. Trans. R. Soc. Lond. B Biol. Sci. **358**:267–274.

Bronstein, M., M. Schütz, G. Hauska, E. Padan, and Y. Shahak. 2000. Cyanobacterial sulfide-quinone reductase: cloning and heterologous expression. J. Bacteriol. **182**:3336–3344.

Bryant, D., and V. Moulton. 2002. NeighborNet: an agglomerative method for the construction of planar phylogenetic networks. Proceedings of the 2nd Workshop in Algorithms for Bioinformatics. (http://www.mcb.mcgill.ca/~bryant/NeighborNet/).

Burger, G., B. F. Lang, M. Reith, and M. W. Gray. 1996. Genes encoding the same three subunits of respiratory complex II are present in the mitochondrial DNA of two phylogenetically distant eukaryotes. Proc. Natl. Acad. Sci. USA **93**:2328–2332.

Butterfield, N. J. 2000. *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology **263**:386–404.

Canback, B., S. G. Andersson, and C. G. Kurland. 2002. The global phylogeny of glycolytic enzymes. Proc. Natl. Acad. Sci. USA **99**:6097–6102.

Canfield, D. E. 1998. A new model for Proterozoic ocean chemistry. Nature **396**:450–453.

Chabriere, E., M. H. Charon, A. Volbeda, L. Pieulle, E. C. Hatchikian, and J. C. Fontecilla-Camps. 1999. Crystal structures of the key anaerobic enzyme pyruvate:ferredoxin oxidoreductase, free and in complex with pyruvate. Nature Struct. Biol. **6**:182–190.

Chen, Z. W., M. Koh, G. Van Driessche, J. J. Van Beeumen, R. G. Bartsch, T. E. Meyer, M. A. Cusanovich, and F. S. Mathews. 1994. The structure of flavocytochrome c sulfide

dehydrogenase from a purple phototrophic bacterium. Science **266**:430–432.

Couvreur, B., R. Wattiez, A. Bollen, P. Falmagne, D. Le Ray, and J. C. Dujardin. 2002. Eubacterial hslV and hslU subunits homologs in primordial eukaryotes. Mol. Biol. Evol. **19**:2110–2117.

Doeller, J. E., B. K. Gaschen, V. Parrino, and D. W. Kraus. 1999. Chemolithoheterotrophy in a metazoan tissue: sulfide supports cellular work in ciliated mussel gills. J. Exp. Biol. **202**:1953–1961.

Doeller, J. E., M. K. Grieshaber, and D. W. Kraus. 2001. Chemolithoheterotrophy in a metazoan tissue: thiosulfate production matches ATP demand in ciliated mussel gills. J. Exp. Biol. **204**:3755–3764.

Doolittle, W. F. 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trend. Genet. **14**:307–311.

Embley, T. M., M. van der Giezen, D. S. Horner, P. L. Dyal, and P. Foster. 2003. Hydrogenosomes and mitochondria: phenotypic variants of the same fundamental organelle. Philos. Trans. R. Soc. Lond. B Biol. Sci. **358**:191–203.

Felsenstein, J. 1998. PHYLIP (phylogeny inference package). Distributed by the author, Department of Genetics, University of Washington, Seattle.

Fenchel, T. M., and R. J. Riedl. 1970. The sulfide system: a new biotic community underneath the oxidized layer of marine sand bottoms. Mar. Biol. **7**:255–268.

Friedrich T, and D. Schiede. 2000. The respiratory complex I of bacteria, archaea, and eukarya and its module in common with membrane-bound multisubunit hydrogenases. FEBS Lett. **479**:1–5.

Gogarten, P. J., W. F. Doolittle, and J. G. Lawrence. 2002. Prokaryotic evolution in light of lateral gene transfer. Mol. Biol. Evol. **19**:2226–2238.

Griesbeck, C., G. Hauska, and M. Schütz. 2000. Biological sulfide oxidation: sulfide-quinone reductase (SQR), the primary reaction. Pp. 179–203 *in* S. G. Pandalai, ed. Recent research developments in microbiology, Vol 4. Research Signpost, Trivadrum, India.

Griesbeck, C., M. Schütz, T. Schödl, S. Bathe, L. Nausch, N. Mederer, M. Vielreicher, and G. Hauska. 2002. Mechanism of sulfide-quinone reductase investigated using site-directed mutagenesis and sulfur analysis. Biochemistry **41**:11552–11565.

Grieshaber, M. K., and S. Völkel. 1998. Animal adaptations for tolerance and exploitation of poisonous sulfide. Ann. Rev. Physiol. **60**:30–53.

Hannaert, V., H. Brinkmann, U. Nowitzki, J. A. Lee, M. A. Albert, C. W. Sensen, T. Gaasterland, M. Muller, P. Michels, and W. Martin. 2000. Enolase from *Trypanosoma brucei*, from the amitochondriate protist *Mastigamoeba balamuthi*, and from the chloroplast and cytosol of *Euglena gracilis*: pieces in the evolutionary puzzle of the eukaryotic glycolytic pathway. Mol. Biol. Evol. **17**:989–1000.

Horner, D. S., P. D. Foster, and T. M. Embley. 2000. Iron hydrogenase and the evolution of anaerobic eukaryotes. Mol. Biol. Evol. **17**:1695–1709.

Horner, D. S., B. Heil, T. Happe, and T. M. Embley. 2002. Iron hydrogenases—ancient enzymes in modern eukaryotes. Trends Biochem. Sci. **27**:148–153.

Horner, D. S., R. P. Hirt, and T. M. Embley. 1999. A single eubacterial origin of eukaryotic pyruvate: ferredoxin oxidoreductase genes: implications for the evolution of anaerobic eukaryotes. Mol. Biol. Evol. **16**:1280–1291.

Javaux, E. J., A. H. Knoll, and M. R. Walter. 2001. Morphological and ecological complexity in early eukaryotic ecosystems. Nature **412**:66–69.

John, P., and F. R. Whatley. 1975. *Paracoccus denitrificans* and the evolutionary origin of the mitochondrion. Nature **254**:495–498.

Johns, A. R., A. C. Taylor, R. J. A. Atkinson, and M. K. Grieshaber. 1997. Sulphide metabolism in thalassinidean crustacea. J. Mar. Biol. Ass. UK **77**:127–144.

Huynen M. A., B. Snel, P. Bork and T. J. Gibson. 2001. The phylogenetic distribution of frataxin indicates a role in iron-sulfur cluster protein assembly. Hum. Mol. Genet. **10**:2463–2468.

Katinka, M.D., S. Duprat, E. Cornillot et al. (17 co-authors). 2001. The genome of the intracellular parasite, *Encephalitho-zoon cuniculi*. Nature **414**:450–453.

Kurland, C. G., and S. G. Andersson. 2000. Origin and evolution of the mitochondrial proteome. Microbiol. Mol. Biol. Rev. **64**:786–820.

Lecompte, O., R. Ripp, J. C. Thierry, D. Moras, and O. Poch. 2002. Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic Acids Res. **30**:5382–5390.

Lill, R., and G. Kispal. 2000. Maturation of cellular Fe-S proteins: an essential function of mitochondria. Trend. Biochem. Sci. **25**:352–356.

Martin, W., M. Hoffmeister, C. Rotte, and K. Henze. 2001. An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondria and hydrogenosomes), and their heterotrophic lifestyle. Biol. Chem. **382**:1521–1539.

Martin, W., and M. Müller. 1998. The hydrogen hypothesis for the first eukaryote. Nature **392**:37–41.

Martin, W., T. Rujan, E. Richly, A. Hansen, S. Cornelsen, T. Lins, D. Leister, B. Stoebe, M. Hasegawa, and D. Penny. 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. USA **99**:12246–12251.

National Research Council. 1979. Hydrogen sulfide. University Park Press, Baltimore.

Nei, M. 1996. Phylogenetic analysis in molecular evolutionary genetics. Ann. Rev. Genet. **30**:371–403.

Nei, M., and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press, New York.

Nixon, J. E. J., A. Wang, J. Field, H. G. Morrison, A. G. McArthur, M. L. Sogin, B. J. Loftus, and J. Samuelson. 2002. Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*. Euk. Cell **1**:181–190.

Nübel, T., C. Klughammer, R. Huber, G. Hauska, and M. Schütz. 2000. Sulfide:quinone oxidoreductase in membranes of the hyperthermophilic bacterium *Aquifex aeolicus* (VF5) Arch. Microbiol. **173**:233–244.

O'Brien, J., and R. D. Vetter. 1990. Production of thiosulfate during sulphide oxidation by mitochondria of the symbiont-containing bivalve *Solemya reidi*. J. Exp. Biol. **149**:133–148.

Oeschger, R., and B. Vismann. 1994. Sulphide tolerance in *Heteromastus filiformis* (Polychaeta): mitochondrial adaptations. Ophelia **40**:147–158.

Okayama, S., N. Yamamoto, K. Nishikawa, and T. Horio. 1968. Roles of ubiquinone-10 and rhodoquinone in photosynthetic formation of adenosine triphosphate by chromatophores from Rhodospirillum rubrum. J. Biol. Chem. **243**:2995–2999.

Parrino, V., D. W. Kraus, and J. E. Doeller. 2000. ATP Production from the oxidation of sulfide in gill mitochondria of the ribbed mussel *Geukensia demissa*. J. Exp. Biol. **202**:2209–2218.

Peters, J. W. 1999 Structure and mechanism of iron-only hydrogenases. Curr. Opin. Struct. Biol. **9**:670–676.

Reeve, J. 2003. Archaeal chromatin and transcription. Mol. Microbiol. **48**:587–598.

Reinartz, M., T. Tschäpe, T. Brüser, H. G. Trüper, and C. Dahl. 1998. Sulfide oxidation in the phototrophic bacterium *Chromatium vinosum*. Arch. Microbiol. **170**:59–68.

Rost, B. 1997. Protein structures sustain evolutionary drift. Fold. Des. **2**:S19–S24.

Rotte, C., F. Stejskal, G. Zhu, J. S. Keithly, and W. Martin. 2001. Pyruvate:$NADP^+$ oxidoreductase from the mitochondrion of *Euglena gracilis* and from the apicomplexan *Cryptosporidium parvum*: a fusion of pyruvate:ferredoxin oxidoreductase and NADPH-cytochrome P450 reductase. Mol. Biol. Evol. **18**:710–720.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

Schütz, M., M. Brugna, E. Lebrun et al. (12 co-authors). 2000. Early evolution of cytochrome *bc* complexes. J. Mol. Biol. **300**:663–675.

Schütz, M., C. Klughammer, C. Griesbeck, A. Quentmeier, C. G. Friedrich, and G. Hauska. 1998. Sulfide-quinone reductase activity in membranes of the chemotrophic bacterium *Paracoccus denitrificans* GB17. Arch. Microbiol. **170**:353–360.

Schütz, M., Y. Shahak, E. Padan, and G. Hauska. 1997. Sulfide-quinone reductase from *Rhodobacter capsulatus*. J. Biol. Chem. **272**:9890–9894.

Shahak, Y., B. Arieli, E. Padan, and G. Hauska. 1992. Sulfide quinone reductase (SQR) activity in *Chlorobium*. FEBS Lett. **299**:127–130.

Shahak, Y., C. Klughammer, U. Schreiber, E. Padan, I. Herrmann, and G. Hauska. 1994. Sulfide-quinone and sulfide-cytochrome reduction in *Rhodobacter capsulatus*. Photosynthesis Res. **39**:175–181.

Shen, Y., R. Buick, and D. E. Canfield. 2001. Isotopic evidence for microbial sulphate reduction in the early Archaean era. Nature **410**:77–81.

Stechmann, A. and T. Cavalier-Smith. 2002. Rooting the eukaryote tree by using a derived gene fusion. Science **297**: 89–91.

Strimmer, K., and A. von Haeseler. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. Mol. Biol. Evol. **13**:964–969.

Tachezy, J., L. B. Sanchez, and M. Müller. 2001. Mitochondrial type iron-sulfur cluster assembly in the amitochondriate eukaryotes *Trichomonas vaginalis* and *Giardia intestinalis*, as indicated by the phylogeny of IscS. Mol. Biol. Evol. **18**:1919–1928.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

Tielens, A. G. M., C. Rotte, J. van Hellemond, and W. Martin. 2002. Mitochondria as we don't know them. Trends Biochem. Sci. **27**:564–572.

Tye, B. K. 2000. Insights into DNA replication from the third domain of life. Proc. Natl. Acad. Sci. USA **97**:2399–2401.

Vande Weghe, J. G., and D. W. Ow. 1999. A fission yeast gene for mitochondrial sulfide oxidation. J. Biol. Chem. **274**: 13250–13257.

Völkel, S., and M. K. Grieshaber. 1992. Mechanisms of sulphide tolerance in the peanut worm, *Sipunculus nudus* (Sipunculidae) and in the lugworm, *Arenicola marina* (Polychaeta). J. Comp. Physiol. B Biochem. Syst. Environ. Physiol. **162**: 469–477.

———. 1996*a*. Mitochondrial sulfide oxidation in *Arenicola marina*: evidence for alternative electron pathways. Eur. J. Biochem. **235**:231–237.

———. 1996*b*. Sulphide oxidation and oxidative phosphorylation in the mitochondria of the lugworm *Arenicola marina*. J. Exp. Biol. **200**:83–92.

Völkel, S., K. Hauschild, and M. K. Grieshaber. 1995. Sulfide stress and tolerance in the lugworm *Arenicola marina* during low tide. Mar. Ecol. Prog. Ser. **122**:205–215.

Williams, B.A., R. P. Hirt, J. M. Lucocq, and T. M. Embley. 2002 A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. Nature **418**:865–869.

Yong, R., and D. G. Searcy. 2001. Sulfide oxidation coupled to ATP synthesis in chicken liver mitochondria. Comp. Biochem. Physiol. B **129**:129–137.

Geoffrey McFadden, Associate Editor

Accepted May 8, 2003