# The Molecular Evolution of Catalatic Hydroperoxidases: Evidence for Multiple Lateral Transfer of Genes Between Prokaryota and from Bacteria into Eukaryota

*Martin G. Klotz* and Peter C. Loewen†*

*Department of Biology and Center for Genetics and Molecular Medicine, University of Louisville; and †Department of Microbiology, University of Manitoba, Winnipeg, Canada

The past decade has produced an increasing number of reports on horizontal gene transfer between prokaryotic organisms. Only recently, with the flood of available whole genome sequence data and a renewed intensity of the debate about the universal tree of life, a very few reports on lateral gene transfer (LGT) from prokaryotes into the Eukaryota have been published. We have investigated and report here on the molecular evolution of the gene families that encode catalatic hydroperoxidases. We have found that this process included not only frequent horizontal gene transfer among prokaryotes but also several lateral gene transfer events between bacteria and fungi and between bacteria and the protistan ancestor of the alga/plant lineage.

## Introduction

Virtually all organisms that survive in oxic environments, whether they are capable of aerobic or anaerobic lifestyles or both, contain enzymes that convert reactive oxygen intermediates (ROI) to innocuous compounds. If not dismutated, ROI such as superoxide, hydrogen peroxide, and hydroxyl radical will interact with macromolecules, their derivatives and cellular structures thereby leading to conformational changes and loss of integrity (Storz and Imlay 1999). Fridovich and associates (McCord, Keele, and Fridovich 1971) were first to propose a correlation between the lack of antioxidative enzymes and obligate anaerobiosis, and this hypothesis has been solidified by numerous discoveries. Organisms are exposed to ROI in two different scenarios: (1) external ROI, in particular $H_2O_2$, generated by abiotic processes or produced by other organisms and (2) internal production of ROI by metabolic redox reactions. An *Escherichia coli* cell, for instance, can generate approximately 14 μmol $H_2O_2$ per second during exponential growth in glucose minimal medium (Seaver and Imlay 2001*a*, 2001*b*), which, if it accumulated, would be more than sufficient to kill the organism. Thus, single-celled organisms such as the prokaryotes and protista that did not live in well-balanced controlled microenvironments evolved with versatile mechanisms of ROI detoxification. Multicellular eukaryotes successfully avoid ROI stress by terminal differentiation of tissues and by compartmentalizing reactions that involve free molecular oxygen and iron (which is crucial for the spontaneous interconversion of ROI). Furthermore, eukaryotes have evolved dedicated structures and mechanism that recruit the oxidative burst for defense against infectious causal agents of disease (i.e., in plants [Sutherland 1991]). In vertebrate animals, these functions have been folded into the immune response, and in most organisms, ROI production is an essential outcome of genetic programs for self-destruction (apoptosis, or programmed cell death). Failure to control the deliberate oxidative burst is often the cause of accelerated aging and disease, including cancer (Ames 1995). In certain stress scenarios such as during pathogenesis or environmental onslaughts, demand often exceeds available activity of antioxidative enzymes. For example, diabetic OVE26 mice overexpressing monofunctional catalase in the heart were protected from the diabetic cardiomyopathy usually observed in diabetic animals (Epstein, Overbeek, and Means 1989; Kang, Chen, and Epstein 1996). Catalase also prevented the reduced contractility observed in cardiomyocytes isolated from diabetic mice (P. N. Epstein, personal communication). Vector-borne expression of a cyanobacterial catalase-peroxidase (CPX) in guinea pig cell cultures conferred to these transfectants significantly higher resistance to hydrogen peroxide and paraquat than the parental cells (Ishikawa et al. 1998). This shows that the acquisition of catalase activity, whether by paralogy or by lateral acquisition of genes, will usually lead to a better performance of aerobic organisms.

During a long coevolution of pathogens and their hosts, the pathogens amassed an amazing complement of antioxidative functions to maintain their potential for infection. Whereas superoxide (degraded by superoxide dismutase to hydrogen peroxide) and hydroxyl radical are highly reactive and have thus a limited radius of action, hydrogen peroxide can easily penetrate cellular structures (cell walls, plasma membranes, etc.) and form adducts with sugars and amino acids that are transported by numerous dedicated permeases (Schubert and Wilmer 1991). Hence, hydrogen peroxide can inflict damage anywhere in a cell or tissue through oxidation of proteins (sulfhydryl groups) and single strand breakage in nucleic acids (Richter and Loewen 1982; Imlay, Chin, and Linn 1988). It is not surprising to find a very high level of functional redundancy among enzymes involved in the detoxification of hydrogen peroxide.

Enzymes that reduce hydrogen peroxide or use it as a reductant are generally termed hydroperoxidases (HP). Catalatic hydroperoxidases (CHPs) primarily dismutate hydrogen peroxide to water and dioxygen by two-electron transfer redox reactions, and there is considerable diversity

M.G.K. dedicates his contribution to Dr. rer. nat. habil. Gerhard Franz Klotz, professor emeritus at the University of Jena, Germany, on the occasion of his 75th birthday.

Key words: catalase, catalase-peroxidase, nonheme catalase, lateral gene transfer.

E-mail: martin.klotz@louisville.edu.

*Mol. Biol. Evol.* 20(7):1098–1112. 2003
DOI: 10.1093/molbev/msg129
*Molecular Biology and Evolution*, Vol. 20, No. 7

among the enzymes that exhibit catalatic activity. Generally, these proteins can be placed into four main groups: (1) the ''classic'' heme-containing monofunctional catalases for which hydrogen peroxide is both electron donor and acceptor, (2) the heme-containing bifunctional CPXs in which the catalatic activity is much higher than the peroxidatic activity, (3) the nonheme-containing catalases, and (4) a miscellaneous group containing proteins with minor catalatic but no peroxidatic activities (Jones and Wilson 1978). CHPs have a variety of subunit sizes, a number of quaternary structures, at least two different heme prosthetic groups, and the reductant for bifunctional CPXs can vary. This article set out to investigate the evolutionary history of the first three categories, that is, enzymes with a considerable catalatic activity.

### Monofunctional Catalases

Monofunctional catalases (hydrogen peroxide oxidoreductase E.C. 1.11.1.6) degrade hydrogen peroxide to oxygen and water ($2H_2O_2 \rightarrow 2H_2O + O_2$). They are present in both empires of life and are the most extensively studied, beginning with the first report of a biochemical characterization of the enzyme in 1900 (Loew 1900). The enzyme's subsequent history has been well documented (Loewen 1997; Zamocky and Koller 1999). Investigations on catalases from a variety of sources, including tobacco, yeast, blood, and bovine liver (BLC), resulted in catalase being one of the first enzymes to be crystallized in 1937 (Sumner and Dounce 1937), and the enzyme has continued to be an object of interest and of intense study ever since. Monofunctional catalases use hydrogen peroxide alternately as oxidizing and reducing species. The oxidation involves a two-electron oxidation of heme iron, which in turn is reduced in a two-electron transfer from a second molecule of hydrogen peroxide to regenerate the resting enzyme. Catalases can also undergo two sequential one-electron reductions involving organic reductants, similar to peroxidases; however, the peroxidatic reaction is weak and usually restricted to small organic substrates because of limited accessibility to the active site. Biochemically, it is possible to divide the monofunctional catalases into two subgroups based on subunit size. One group contains small-subunit enzymes of 55 to 69 kd, and a second group contains large-subunit enzymes of 75 to 84 kd. All small-subunit enzymes so far characterized have heme b (or a degradation product of heme b) associated, and some have NADPH bound (Kirkman and Gaetani 1984; Almarsson et al. 1993; Hillar et al. 1994). With the exception of some mutant variants, all large-subunit enzymes so far characterized have heme d associated and none have been found with NADPH bound. The monofunctional catalases characterized in greatest detail, either by crystal structure analysis or electrospray mass spectrometry, have all been found to be homotetramers, although dimeric, hexameric, and even heterotrimeric enzymes have been reported (Nicholls, Fita, and Loewen 2001). There are currently over 250 known sequences of monofunctional catalases. Whereas the evolutionary relationships of subsets of enzymes have been described previously (Mayfield and Duvall 1996; Klotz, Klassen, and Loewen 1997; Scandalios, Guan, and Polidoros 1997; Frugoli et al. 1998; Kim, Sha, and Mayfield 2000; Loewen, Klotz, and Hassett 2000), we have set out to investigate the phylogeny of the entire monofunctional catalase gene family based on this prior literature and a comprehensive analysis of all presently available catalase gene and protein sequences, a number that has more than tripled since our last review (Klotz, Klassen, and Loewen 1997).

### Heme-Containing Catalase-Peroxidases

Catalase-peroxidases exhibit a predominant catalatic activity but biochemically differ from the monofunctional catalases in exhibiting a significant peroxidatic activity (Sun et al. 1994; Thomas, Morris, and Hager 1970). Of the more than 60 currently available CPX sequences, most have been identified in prokaryotes (including Negibacteria, Posibacteria, and Archaebacteria), but there is recent evidence for their presence in fungi. CPXs have sequence similarity to class I peroxidases in plants (e.g., plastid ascorbate peroxidase) and fungi (e.g., cytochrome C peroxidase); however, there is no resemblance in sequence to the monofunctional catalases. CPXs are active as either dimers or tetramers, and the large size of the subunits composed of two domains with similar sequences has lead to the hypothesis that the CPXs may have arisen through a gene duplication and fusion (Welinder 1991, 1992). The N-terminal domain has retained activity along with greater sequence similarity to other CPXs. The C-terminal domain has evolved with greater sequence deviation, does not bind heme, and is presumably inactive. The evolutionary relationships of CPXs have been described briefly in the past, including a critical discussion of hypothetical lateral transfer between Archaebacteria and Eubacteria (Klotz, Klassen, and Loewen 1997; Faguy and Doolittle 2000; Loewen, Klotz, and Hassett 2000; Zamocky et al. 2001; Cavalier-Smith 2002*a*).

### Nonheme Catalases

A very different catalatic mechanism has evolved in the nonheme catalases in which the reaction center is a manganese complex. The nonheme catalases make up a small group of enzymes, so far only of prokaryotic origin, with only three examples purified and characterized, but a larger number sequenced. A manganese-rich reaction center was identified as the active site in each of the three characterized enzymes (Kono and Fridovich 1983; Allgood and Perry 1986; Whittaker et al. 1999), and this has been confirmed to be a bridged binuclear manganese cluster by crystal structure analysis (Waldo, Fronko, and Penner-Hahn 1991). The mechanism of catalytic action in these enzymes is currently under investigation.

## Material and Methods

DNA and protein sequences of catalatic hydroperoxidases were obtained from our sequence library (Klotz, Klassen, and Loewen 1997; Loewen, Klotz, and Hassett 2000) and by searching available databases (GenBank/EMBL/DBBJ) and servers with sequences

**Table 1**
**Taxonomic Classification, Host Organisms, and Sources of All Monofunctional Catalases, Catalase-Peroxidases, and Nonheme Catalases Presently Published or Available in Genome Sequence Data Bases (as Outlined in the Text)**

| Taxon | *Organism* | Group I catalase | Group II catalase | Group III catalase | CPX | non-heme (Mn)-cat | TPX |
|---|---|---|---|---|---|---|---|
| **BACTERIA** | | | | | | | |
| **Negibacteria, "Eobacteria"** | | | | | | | |
| Eobacteria, Chlorobacteria | *Chloroflexus aurantiacus J-10-fl* | none | none | none | none | none | >CloaurTPX[TIGR_115(13)] |
| Eobacteria, Hadobacteria | *Deinococcus radioduranis* strain "R1" | >DeiradA [AE002037; D63898] | >DeiradE [AE001863] | none | none | none | >DeiradTPX[gi_15807234] |
| | *Thermus aquaticus* | none | none | none | none | | >TheaquAhpC[AF276071] |
| | *Thermus spec. "YS 8-13"* | none | none | none | none | >ThespeM [AB008786] | |
| | *Thermus thermophilus* HB27 | none | none | none | none | >ThetheM [Ugo_0.12709] | >ThetheTPX [Ugo_0.24624] |
| **Negibacteria, "Glycobacteria"** | | | | | | | |
| Cyanobacteria, Hormogonaea | *Nostoc punctiforme* | none | none | >NospunA[JGI_554] | none | >NospunM1[JGI_621(47)] | >NospunTPX[JGI_536(5)] |
| | *Nostoc punctiforme* | | | | | >NospunM2[JGI_512(7)] | |
| | *Nostoc* sp. PCC7120 | none | none | none | none | >Nos7120M1[TIGR 358(9 | >NosPCC7120[NC_003272] |
| | *Nostoc* sp. PCC7120 | | | | | >Nos7120M2[TIGR 274(9) | |
| Cyanobacteria, Chroobacteria | *Synechocystis* sp. PCC 6803 | none | none | none | >Svn6803G[P6803; D83990] | none | >Svn6803TPX[NC 000911] |
| | *Synechococcus* sp. PCC 7942 | none | none | none | >Svn7942G[P7942; D61378] | none | |
| | *Synechococcus* sp. PCC 6301 | none | none | none | >Svn6301G[P6301; AF197161] | none | |
| | *Synechococcus* sp. WH 8102 | none | none | none | none | none | >SvnWH8102TPX |
| | *Prochlorococcus marinus* MED4 | none | none | none | none | none | >PromarED4TPX |
| | *Prochlorococcus marinus* MIT93 | none | none | none | none | none | >PromarMIT93TPX |
| Spirochaetae, Spirochaetes | *Borrelia burgdorferi* | none | none | none | none | none | none |
| | *Leptospira interrogans* serovar lai str. 56601 | none | none | >LepintA[AE011360] | none | none | |
| | *Treponema pallidum* Nichols | none | none | none | none | none | >TrepalAhpC[AE001227] |
| Planctobacteria, Chlamydiae | *Chlamydia pneumoniae* CWL029 | none | none | none | none | none | >ChlpneAhpC[gi 4377087] |
| Planctobacteria, Planctomycea | *Pirellula sp. strain 1* | none | none | >PirspeA[REGX_orf1340] | >PirspeCPX[REGX_orf2707] | >PirspeM[REGX_orf3656] | >PirspeTPX[REGX_orf4586] |
| Sphingobacteria, Flavobacteria | *Bacteriodes fragilis* | none | none | >BacfraB [U18676] | none | none | >BacfraAhpC[AF129406] |
| | *Fusobacterium nucleatum (AE009951)* | none | none | none | none | none | >FusnucAhpC[AE009951] |
| | *Cytophaga hutchinsonii* | none | >CythutE[JGI_125(23)] | none | none | none | >CythutTPX[TIGR_163(86)] |
| Sphingobacteria, Chlorobea | *Chlorobium tepidum* TLS | none | none | none | none | none | >ClotepTPX[TIGR_orf254] |
| Proteobacteria, Epsilobacteria | *Campylobacter jejuni* | none | none | >CamjejA[X85130] | none | none | >CamjejAhpC[AL139074] |
| | *Helicopacter pylori* | none | none | >HelpylA[AE001510, U67458] | none | none | >HelpylAhpC[gi_4156103] |
| | *Aquifex aeolicus* VF5 | none | none | none | none | none | >AquaeoAhpC1[AE000692] |
| Proteobacteria, Deltabacteria | *Desulfovibrio vulgaris* | none | none | >DesvulA[AB020341] | none | none | |
| | *Desulfovibrio vulgaris* | | | >DesvulB[TIGR_1369] | | | |
| | *Geobacter sulfurreducens* | none | none | none | >GeosulCPX[TIGR_1283] | none | |
| Proteobacteria, Alphabacteria | *Agrobacterium tumefaciens* C58 | none | >AgrtumC[AE007917.AE007872] | | >AgrtumCPX[AB033631.AE008223] | none | |
| | *Brucella abortis* | none | none | >BruaboA[U11439] | | | |
| | *Brucella melitensis* | none | none | >BrumelA[IG_RBME01716] | none | none | |
| | *Brucella suis* | none | none | >BrusuiA[TIGR_468] | | none | |
| | *Caulobacter crescentus* CB15 | none | none | | >CaucreG[AE005967; AF027168] | none | >CaucreAhpC[AE005956] |
| | *Mesorhizobium loti* | none | >MeslotE[AP002998] | | >MeslotCPX[AP003010] | none | |
| | *Methylobacterium extorquens* A1 | none | | >MetextA[L48340; TIGR 1544(9)] | >MetextCPX[TIGR 2173(9)] | none | >MetextAhpC[pedant orf8] |
| | *Rhodobacter capsulatus* | | | | >RhocapKatG[TIGR_p3(Orf60)] | none | |
| | *Rhodobacter capsulatus* "B10" | | | | >RhocapCPX[X71420] | | |
| | *Rhodospeudomonas palustris* | none | >RhopalE[JGI_55] | none | >RhopalCPX[JGI_] | none | |
| | *Rhodospirillum sphaeroides* | | >RhosphE[TIGR_782] | >RhosphA[TIGR_87] | | | |
| | *Rickettsia prowazekii* MadridE | none | none | none | none | none | >RicproAhpC[AJ235271] |
| | *(Sino)Rhizobium* spec. | | | >RhizspeA[U56239] | | | |
| | *(Sino)Rhizobium meliloti* | | >RhimelC[AF121348] | >RhimelA[U59271] | | | |
| | *(Sino)Rhizobium meliloti* 1021 | none | >RhimelC[AL603642] | >RhimelA[AL591688] | >RhimelCPX[AE007314] | none | >SinmelAhpC[AL603646] |
| | *Sphingomonas aromaticivorans* DSM12444 | none | >SpharoE[JGI 208] | none | >SpharoCPX[TIGR contig348(81)] | none | >SpharoAhpC[pedant 1292(5)] |
| | *Sphingomonas sp.* "A1" | | | | *>SphspecA1[AB011415] | | |
| | *Wolbachia spec.* | none | none | >WolsppA[AF069070] | none | none | >WolsspTPX[TIGR_21336] |
| Proteobacteria, Betabacteria | *Bordetella bronchiseptica* RB50 | none | none | >BorbroA[ 2544(10)] | none | none | >BorbroAhpC[TIGR 252(4)] |
| | *Bordetella pertussis* | | | >BorperA[U07800] | | | >BorperTPX[TIGR-2364] |
| | *Burkholderia fungarum (cepacia)* "C5424" | | | | >BurcepKatA[AF317697] | >BurcepM[TIGR 723(39)] | |
| | *Burkholderia fungarum (cepacia)* LB400 | none | >BurcepE[TIGR 681] | >BurcepA[TIGR 669] | >BurcepCPx1 [TIGR contiq691(10)] | | >BurcepTPX[TIGR contiq373(2 |
| | *Burkholderia fungarum (cepacia)* LB400 | | | | *>BurcepCPX2 [TIGR contiq757(36)] | | |
| | *Burkholderia mallei* | | | | >BurmalCPX[TIGR_472] | | |
| | *Burkholderia pseudomallei* | | | | >BurpseG[AY040244] | | |
| | *Neisseria gonnorrhoea* | | | >NeigonA[U35457; AF035966] | none | none | |
| | *Neisseria meningitiis* | none | none | >NeimenA[AE002379] | | | >NeimenTPX[gi_7379834] |
| | *Nitrosomonmas europaea* | none | none | >NiteurA[JGI_835(45)] | >NiteurCPX1[ JGI_0678] | none | >NiteurAhpC[JGI_652(1)] |
| | *Nitrosomonmas europaea* | | | | *>NiteurCPX2[JGI_0677] | | |
| | *Ralstonia metallidurans* CH34 *(R. eutropha)* | none | none | >RaleutA[JGI 219] | >RaleutCPX[JGI orf15@3696] | none | >RaleutAhpC[JGI 15] |
| | *Ralstonia solanacearum* GMI1000 | >RalsolA[gi 17432055] | none | none | >RalsolCPX1[gi 17427788/7] | none | >RalsolAhpC[AL646077] |
| Proteobacteria, Gammabacteria | *Actinobacillus actinomycetemcomitans* | none | none | >ActactA[AF162654] | none | none | >ActactTPX[pendant_orf1411] |
| | *Haemophilus influenzae* | none | none | >HaeinfA[L42023; U02682] | none | none | |
| | *Pseudomonas aeruginosa* PAO1 | >PseaerB[U89384] | >PseaerC[AE004642] | >PseaerA[AF047025] | | >PseaerM[G83372] | >PseaerAhpC[AE004452] |
| | *Pseudomonas fluorescens* | | | >PseflPA[U72068] | | | |
| | *Pseudomonas fluorescens* "2-79" | >PsefluB0[U83329] | | | | | |
| | *Pseudomonas fluorescens* "Pf01" | >PsefluB1[JGI 457(38)] | >PsefluC1[JGI 474(27)] | >PsefluA1[JGI 348(7)] | >PsefluCPX[JGI 500(25)] | >PsefluM[JGI 469(37)] | >PsefluAhpC[JGI 503(67)] |
| | *Pseudomonas fluorescens* "Pf01" | | | >PsefluA2[JGI 355(10)] | | | |
| | *Pseudomonas putida* "Corvallis" | | >PseputC1[U82622] | >PseputA1[U63511] | | | >PseputAhpC[AB010689] |
| | *Pseudomonas putida* KT2440 | | >PseputC2[TIGR] | >PseputA2[TIGR] | | | |
| | *Pseudomonas syringae* pv. syringae "61" | >PsesyrF[AF001355] | | | | | |
| | *Pseudomonas syringae* pv. tomato "DC3000" | >PsetomB[TIGR] | >PsetomC[TIGR] | | >PsetomCPX[TIGR 4247] | | |
| | *Shewanella putrifaciens* | | | >SheputA[TIGR_7833] | >SheputCPX[TIGR_] | | |
| | *Vibrio cholerae* | >VibchoA[AE004235] | | | >VibchoCPX[AE004233] | | >VibchoAhpC[NP_230380 ] |
| | *Vibrio fisheri* | | | >VibfisA [AF011784] | | | |
| | *Vibrio rumenoidis* | | | >VibrumA [AB030821] | | | |
| | *Xanthomonas oryzae pv. oryzae* "8707" | | >XanoryX [X97673] | | | | |
| | *Xanthomonas campestris* pv. phaseoli | >XanphaA[Mongolsuk, p.c.] | >XanphaE[AF170449] | | | | >XancamAhpC[U94336] |
| | *Xylella fastidiosa* 9a5c | none | none | none | >XylfasCPX[AE004035] | none | >XylfasAhpC[AE003983] |
| | *Xylella almond* | none | none | none | >XylalmCPX[TIGR_] | none | |
| | *Xylella oleander* | none | none | none | >XyloleCPX[TIGR_] | none | |
| | *Escherichia coli* | none | >EcoliE[M55161] | none | >EcoliG [M21516] | | >EcoliO157AhpC[AE005240] |
| | *Escherichia coli* | | | none | >EcoliP [X89017, AB011549] | | |
| | *Klebsiella pneumoniae* MGH 78578 | none | >KlepneE[TIGR_454(8)] | none | *>KlepneG[TIGR_515(21)] | >KlepneM[TIGR_899(44)] | |

from finished and unfinished genome sequencing projects (www.jgi.doe.gov; www.TIGR.org; pedant.gsf.de; sequence-www.stanford.edu; and www-genome.wi.mit.edu). The sequences are listed in table 1 as an acronym together with an accession number (GenBank or the relevant genome project) related to the source organisms, which have been organized alphabetically in phylogenetic groups using rDNA similarities (Woese 1987; Olsen, Woese, and Overbeek 1994) and the taxonomic system proposed by Cavalier-Smith (2002*a*, 2002*b*).

To summarize all available sequence information, full-length and core protein sequences were aligned with ClustalX (v.1.81 [Thompson et al. 1997]) by using the Gonnet 250 protein weight matrix and gap opening and gap extension penalties of 35/15 and 0.75/0.35, respectively, in the pairwise/multiple sequence alignments. In addition, alignments were produced using an online version of ClustalW (http://www.ddbj.nig.ac.jp/htmls/e-mail/clustalw-e.html) with Kimura correction and maximum fixed-gap opening and gap extension penalties. A total of 256 of more than 260 available heme catalases, 58 of available 61 CPXs and 29 of the available 32 nonheme catalase sequences were included into the subsequent separate alignments. Both ClustalX and ClustalW alignments for each sequence family were compared and used for manual refinement. The resulting alignments were then used for intron analysis as described previously (Johnson et al. 2002). In brief, intron residence in eukaryotic heme catalase genes was obtained (1) from annotated DNA sequence deposits, (2) by identification in nonannotated DNA deposits using the ''gt-ac rule,'' and (3) by identification in genes from unfinished genome databases (after contig alignment of sequence fragments using Sequencher v. 4.1 by applying the ''gt-ac rule''). Intron positions were mapped on the protein sequence alignment to generate a catalase intron positional matrix. Intron residence between different codon bases was treated as an independent intron position, and the position was assigned to the amino acid

**Table 1**
**Continued**

| Taxon | Organism | Group I catalase | Group II catalase | Group III catalase | CPX | non-heme (Mn)-cat | TPX |
|---|---|---|---|---|---|---|---|
| Proteobacteria, Gammabacteria | *Legionella pneumophila* | none | none | none | >LegpneCPx1 [AF078110] | | |
| | *Legionella pneumophila* | | | | >LegpneCPx2 [AB017595] | | |
| | *Pasteurella multocida* PM70 | none | none | >PasmulE[gi 12720234] | | | >PasmulTPX[AE006118] |
| | *Proteus mirabilis* | none | none | >PromirA[AJ400965] | | | |
| | *Salmonella paratyphi* | none | >SalparE[TIGR_558(9)] | none | | >SalparM [TIGR_829(12)] | |
| | *Salmonella typhi* CT18 | none | >SaltviEITIGR 1576(2091)] | none | >SaltoiG [TIGR 3312] | >SaltviM [TIGR 3696] | >SaltviAhpC [TIGR orf578] |
| | *Salmonella typhimurium "SL1344"* | none | >SaltypE[AJ289167] | none | >SaltypG[X53001] | >SaltypM [Y19180] | |
| | *Salmonella enterica* serovar *Typhimurium* LT2 | none | none | none | | | >SaltypAhpC[AE008724] |
| | *Yersinia pseudotuberculosis* | none | none | >YerpseA[_1745(6)] | | | |
| | *Yersinia pestis* CO92 | none | none | >YerpesA[_4050(1002)] | >YerpesKatY[AF135170, AL590842] | | >YerpesAhpC[NC_003143] |
| **Unibacteria** | | | | | | | |
| Posibacteria, Endobacteria | *(Geo)Bacillus stearothermophilus* | none | none | none | >BacsteCPX[Cpe, M29876] | >BacsteaM1 [_479(7)] | |
| | *(Geo)Bacillus stearothermophilus* | | | | | >BacsteaM2 [_679(22)] | |
| | *Bacillus anthracis* | >BacantA[TIGR_1719] | >BacantE[TIGR_1484] | none | none | | |
| | *Bacillus anthracis* | >BacantX[TIGR_1830] | | | | | |
| | *Bacillus cereus* | >BaccerX[IG_RZC00577] | >BaccerE[IG_RZC01424] | >BaccerA[IG_RZC02087] | none | none | >BaccerTPX[IG_RZC02683] |
| | *Bacillus firmus* | | >BacfirA[L02551; AH000868] | | | | |
| | *Bacillus halodurans* | >BachalX[AP001511] | >BachalB[AP001513] | none | >BachalCPX[AP001510] | >BachalM[AP001514] | >BachalCTPX[AP001516] |
| | *Bacillus subtilis* | >BacsubX[NC_000964] | >BacsubE[X85182] | >BacsubA[M80796] | none | >Bacsub-yjqC[Z99110] | >BacsubCTPX[NC_000964] |
| | *Bacillus subtilis* | | | | | >BacsubGSP80[AB001484 | >BacsubAhpC[D78193, Z99124 |
| | *Bacillus subtilis* | | | | | >BacsubCOT J[Q45538] | |
| | *Clostridium acetobutylicum* | | | | | >CloaceM1[(2198)] | |
| | *Clostridium acetobutylicum* | | | | | >CloaceM2[(932)] | |
| | *Clostridium pasteurianum* | | | | | >ClopasM[AF064550] | |
| | *Clostridium tetani* | none | none | none | none | >ClotetM[UGo_1.0.31229] | >ClotetTPX[UGo_1.0.31229] |
| | *Listeria seeligeri* | >LisseeA[M75944] | | | | | |
| | *Listeria monocytogenes* strain EGD | >LismonA[AL591984,AL591824] | | | | | >LismonTPX[AL591979] |
| | *Listeria monocytogenes* strain EGD | | | | | | >LismonAhpC[AL591979] |
| | *Listeria innocua* Clip11262 | >LisinnA[AL596174, AL592022] | | | | | >LisinnTPX[AL596169] |
| | *Listeria innocua* Clip11262 | | | | | | >LisinnAhpC[AL596169] |
| | *Desulfitobacterium hafniense* "DCB-2" | none | >DeshafE1[JGI 2571] | none | >DeshafCPX[DCB2:orf1] | >DeshafM[JGI 2527(3)] | >DeshafTPX[DCB2:orf16] |
| | *Desulfitobacterium hafniense* "DCB-2" | | >DeshafE2[JGI 2732] | | | | |
| | *Enterococcus faecalis* JH2-7 | | | >EntfaeA[TIGR(300)] | | | EntfaeAhpC[AF016233] |
| | *Enterococcus faecium* | | | | | >EntfacM[TIGR_967(22)] | |
| | *Lactobacillus sake* | | | >LacsakA[M84015] | | | |
| | *Lactobacillus plantarum* | | | | | >LacplaM[D87070] | |
| | *Lactococcus lactis* subsp. *Lactis* IL1403 | none | none | none | none | none | >LaclacAhpC[AE006270] |
| | *Staphylococcus aureus* | | | >StaaurA[AJ000471] | | | |
| | *Staphylococcus aureus* EMRSA 16-252 | none | none | >StaaurA2[contiq244:orf511] | none | | |
| | *Staphylococcus aureus* subsp. *aureus* Mu50 | none | none | >StaaurA3[AP003362] | none | | >StaaurAhpC[AP003130] |
| | *Staphylococcus epidermidis* | | | >StaepiA[TIGR_gse_6] | | | |
| | *Staphylococcus warneri* | | | >StawarA[AB045340] | | | |
| | *Staphylococcus xylosis* | | | >StaxylA[AJ295151] | | | |
| | *Streptococcus mutans* NCIB 11723 | none | none | none | none | | >StrmutAhpC[AB010712] |
| | *Streptococcus pyogenes* M1 GAS | | | | | | >StrpyoAhpC[AE006628] |
| Posibacteria, Actinobacteria | *Micrococcus luteus (M. leisodykticus)* | none | | >MicleiA[P29422] | | | |
| | *Thermobifida fusca* | none | none | >ThefusA[JGI_54(197)] | none | >ThefusM [JGI_58(141)] | >ThefusTPX[pedant_orf90] |
| | *Streptomyces coelicolor* A3(2) | none | >StrcoeB1[AL132973] | >StrcoeA1[7H9; AL450223] | >StrcoeCPX[AL121746, AF126956 | none | >StrcoeAhpC[AF186371] |
| | *Streptomyces coelicolor* A3(2) | none | | >StrcoeA2[F62; AL121855] | | | >StrcoeAhpD[AF186371] |
| | *Streptomyces coelicolor* A3(2) | none | | >StrcoeA3[2G5; AL035478] | | | |
| | *Streptomyces coelicolor* ATCC 10147 | none | >StrcoeB0[AF000419] | >StrcoeA0[X96981] | | | |
| | *Streptomyces reticuli* | none | | | >StrretCPX[ Y14317] | | |
| | *Streptomyces violaceus* ATCC 10712 | none | | >StrvioA[bca; X74791] | | | |
| | *Corynebacterium diphteriae* NCTC 13129 | none | none | >CordipA[TIGR orf293] | none | none | >CordipTPX[TIGR orf1451] |
| | *Mycobacterium* sp. PYR1 | none | | | >MycPYR1CPx[AF207899] | | |
| | *Mycobacterium avium* | none | >MycaviE[L41246] | >MycaviA1[TIGR_110] | >MycaviCPX[TIGR_385] | | >MycaviAhpC[A43858] |
| | *Mycobacterium avium* | none | | >MycaviA2[TIGR_110] | | | |
| | *Mycobacterium bovis* | none | | | >MycbovG[X38277] | | >MycbovAhpC[PIR:S71013] |
| | *Mycobacterium fortuitum* | none | | | >MycforG1[Y07865] | | |
| | *Mycobacterium fortuitum* | none | | | >MycforG2[Y07866] | | |
| | *Mycobacterium intercellulare* | none | | | >MycintG [M86741] | | |
| | *Mycobacterium leprae* TN | none | none | none | none | none | >MyclepAhpC[gi 13093656] |
| | *Mycobacterium smegmatis* | >MycsmeB[TIGR_2889] | none | >MycsmeA[TIGR_2951] | >MycsmeG[U46844] | >MycsmeM[TIGR_3269] | >MycsmeAhpC[TIGR_3267] |
| | *Mycobacterium tuberculosis* H37Rv | none | none | | >MyctubG[X68081,Z97193,AL1234 | none | >MyctubTPX[NC_000962] |
| | *Mycobacterium tuberculosis* H37Rv | none | none | | none | none | >MyctubAhpC[NP 216944] |
| | *Mycobacterium tuberculosis* CDC1551 | none | none | | >MyctubCPX[gi_13881615] | none | >MyctubAhpC[gi_13882221] |
| Archaebacteria, Crenarchaeota | *Aeropyrum pernix* strain K1 | none | none | none | none | none | >AerpyrAhpC[APE2278] |
| | *Pyrobaculum calidifontis* VA1 | none | none | none | none | >PyrcalM[AB076020] | |
| | *Sulfolobus solfataricus* P2 | none | none | none | none | none | >SulsolTPX[AE006819] |
| | *Sulfolobus tokodaii* strain 7 | none | none | none | none | none | >SultokAhpC [gi 15623565] |
| Archaebacteria, Euryarchaeota | *Pyrococcus abyssi* GE5/strain Orsay | none | none | none | none | none | >PyrabvTPX[PAB1673] |
| | *Pyrococcus horikoshii* OT3 | none | none | none | none | none | >PyrhorTPX[gi 3256395] |
| | *Thermoplasma volcanium* GSS1 | none | none | none | none | none | >ThevolAhpC[AP000991] |
| | *Thermoplasma acidophilum* DSM1728 | none | none | none | none | none | >TheaciAhpC[AL445063] |
| | *Archaeoglobus fulgidus* | none | none | none | >ArcfulCPX[AE000951] | none | >ArcfulAhpC[AE001087] |
| | *Halobacterium salinarum* | none | none | none | >HalsalCPX[AF069761] | none | |
| | *Halobacterium* sp. NRC-1 pNRC200 | none | none | none | >HalNRC1CPX1[AE005159] | none | >HalNRC1TPX1[AE005084] |
| | *Haloarcula marismortui* | none | none | none | >HalmarCPX[Y16851] | none | |
| | *Methanosarcina mazii Gö1* | none | >MetmazC[AE013432] | >MetmazA[AE013499] | none | none | >MetmazA[UGo_21.34469] |
| | *Methanosarcina barkeri* | none | none | >MetbarA[JGI_1930] | none | none | >MetbarTPX[JGI_1541(2)] |
| | *Methanobrevibacter arboriphilus* | none | none | >MetarbA[AJ300838] | | | |
| | *Methanobacterium thermoautotrophicum* delta | none | none | none | none | none | >MetheTPX[AE000804] |
| | *Methanococcus jannaschii* | none | none | none | none | none | >MetjanTPx[H64391] |
| | *Methanococcus maripaludris* LL | none | none | none | none | none | >MetmarAhpC [contig90(orf6)] |

residue that contained the codon base upstream of the intron insertion site. To date, there are only five sequences of eukaryotic (fungal) CPXs available, and intron analysis has been postponed until more sequences are available. An intron analysis for nonheme-containing catalase genes was not applicable because these enzymes reside exclusively in prokaryotic cells.

The refined three alignments were also used for inference of phylogeny. Phylogenetic relationships were investigated by using a distance algorithm and character-based tree-searching methods with maximum-parsimony (MP) or maximum-likelihood (ML) object functions. Distance neighbor-joining trees were constructed by using the BioNJ function in PAUP* (v. 4.10b [Swofford 1999]). The obtained trees were used as guide trees and for comparison with phylograms obtained with MP and ML methods. MP (50% majority consensus; random taxon-addition order replicates with Tree-Bisection-Reconnection branch-swapping, Mulpars and steepest decent functions in effect) trees were built independently from all three alignments by using the PAUP* program. The quality of the branching patterns was assessed by bootstrap resampling of the data sets using 100 replications. Because inclusion or exclusion of a few characters can highly affect bootstrap proportions of MP trees derived from limited data sets, we also conducted ML inferences for the CPX and nonheme catalase gene families. The sequence alignments of both gene families were each subjected to a Bayesian (ML) inference of phylogeny by using the program MrBayes v. 2.01 (written by Huelsenbeck and Ronquist; http://morphbank.ebc.uu.se/mrbayes/). The nonheme catalase and CPX protein sequence alignments were subjected to Metropolis-Coupled Monte Carlo Markov Chain sampling of 30,000 and 50,000 generations, respectively. Four equally heated Markov chains were utilized to build a sufficient number of reliable trees after the likelihoods of the trees have converged on a stable value and to allow successful swapping between chains. The searches were

**Table 1**
**Continued**

| Taxon | *Organism* | Group II catalase | Group II catalase | Group III catalase | CPX | TPX | TPX |
|---|---|---|---|---|---|---|---|
| EUKARYOTA, FUNGI (fungi lack Group I catalase and non-heme catalase) | | | | | | | |
| Basidiomycota, Hymenomyceles | *Cryptococcus (Filobasidiella) neoformans* | >CryneoA[SU_cneo011005] | >CryneoB[SU_cneo011005] | >CryneoP[SU_cneo011005] | | | |
| | *Cryptococcus (Filobasidiella) neoformans* | | | >CryneoX[Cat1, J.A. Alspaugh] | | | |
| | *Pleurotus ostreatus* | | >PleostA [U75451] | | | | |
| | *Pleurotus djamar* | | | *>PledjaA[U75450, partial] | | | |
| | *Pleurotus sajor-caju* | | | >PlesajA[AF286097] | | | |
| Ascomycota, Pezizomycotina | *Ajellomyces (Histoplasma) capsulatum* | >AjecapA[AF189368] | >AiecapB[AF139985, AF026268] | >AjecapP[AF189369] | | | |
| | *Aspergillus fumigatus* | >AspfumA[AFU87630] | >AspfumB[AFU87850] | | >AspfumG [TIGR_contig_316] | | |
| | *Aspergillus niger* | | >AspnigR [Z23138] | | | | |
| | *Blumeria graminis* | none | >BlugraA[(Cat1), AF327335] | | >BlugraCPX[AF329396] | | |
| | *Botrytolinia (Botrytis) fuckeliana* SAS56 | | >BotfukB[(cat2); AF2438531 | >BotfukA[(cat1); Z54346] | | | |
| | *Cladosporium fulvum* | | >ClafulB[AF222056] | >ClafulA[AF222055] | | | |
| | *Claviceps purpurea* | | >ClapurA[AJ001386] | | | | |
| | *Emericella (Aspergillus) nidulans* | >EmenidA[U37803] | >EmenidB[U80672] | >EmenidC[AF316033] | >EmenidCPX[CpeA; AJ305225] | | |
| | *Glomerella cingulata (C. gloeosporioides)* | | | >GlocinP[AJ311709] | | | |
| | *Neurospora crassa* | >NeucraA[(cat-1); AY027545] | >NeucraC[ (cat-3); AY027544] | >NeucraP[(cat-4);WHI] | >NeucraCPX[(cat-2), AF459787] | >NeuCraTPX[ped_349] | |
| | *Paracoccidioides brasiliensis* | | | >ParbraA[AF428076, AF395816] | | | |
| | *Penicillium vitale* | | >PenvitA[Williams, p.c.] | | | | |
| | *Podospora anserina* | >PodansA[AJ011298] | >PodansB[AJ011309] | | | | |
| Ascomycota, Saccharomycotina | *Candida albicans* | none | none | >CanalbP[AB006327] | none | | |
| | *Candida boidinii* | | | >CanboP[AB064338] | | | |
| | *Candida tropicalis* | | | >CantroA[M18832] | | | |
| | *Hansenula polymorpha (Pichia angusta)* | | | >HanpolA [X56501;AL435126] | | | |
| | *Kluyweromyces lactis* 230 | | | >Kl CTA1[AL425628, AL427393] | | | |
| | *Saccharomyces cerevisiae* | none | none | >SaccerA[X13028] | none | >SaccerTPX[ped_yml028w | >SaccerTPX[ped_ydr453c] |
| | *Saccharomyces cerevisiae* | | | >SaccerT[X04625] | | | |
| | *Yarrowia lipolytica* CLIB 89 | | | *>Yl CTA1(AL414117, AL413458] | | | |
| | *Zygosaccharomyces rouxii* | | | *>Zr_CTA1[<AL393974, AL394670>] | | | |
| Ascomycota, Schizosaccharomycet | *Schizosaccharomyces pombe* "PR745" | none | none | >SchpomA[D89126] | none | | |
| | *Schizosaccharomyces pombe* "PR745" | | | >SchpomX[D55675] | | | |
| EUKARYOTA, PROTISTA (protista lack Group I catalase and non-heme catalase) | | | | | | | |
| Alveolata, Apicomplexa | *Toxoplasma gondii* | | | >ToxgonA[AF161267; AF136344] | | | |
| Mycetozoa, Dictyostelida | *Dictyostelium discoideum* | | >DicdisB[AF183177] | >DicdisA[AF090443] | | | |
| EUKARYOTA, ANIMALIA (animals lack Group I catalase and non-heme catalase) | | | | | | | |
| Metazoa, Nematoda, Chromadorea | *Ascaris suum* | | | >AscsuuA[Y10611] | | >AscsuuTPX[AB022045] | |
| | *Caenorhabditis elegans* | none | none | >CaeeleA[AL034488] | none | >CaeeleTPX1[CE02609] | >CaeeleTPX2[CE00657] |
| | *Caenorhabditis elegans* | | | >CaeeleB[AL034488] | | | |
| | *Caenorhabditis elegans* | | | >CaeeleC[AL034488] | | | |
| | *Onchocerca volvulus* | | | | | >OncvolTPX[AF029247] | |
| Metazoa, Arthropoda, Tracheata | *Anpheles gambiae* strain "PEST" | | | >AnogamA[EAA10411] | | | |
| | *Drosophila melanogaster* | none | none | >DromelA[U00145] | none | >DromelTPX2[NP_525002 | >DromelTPX3[AF311747] |
| Metazoa, Chordata, Craniata | *Danio rerio* | | | >DanreA[AF170069; AJ007505] | | | |
| | *Rana rugosa* | | | >RanrugA[AB031872] | | | |
| | *Mus musculus* | none | none | >MusmusA[L25069] | none | >MusmusTPX1[U20611] | >MusmusTPX2[BC002034] |
| | *Mus musculus* | | | | | >MusmusTPX3[AF211933] | |
| | *Rattus norvegicus* | none | none | >RatnorA[M11670] | none | >RatnorTPX1[NM_057114 | >RatnorTPX2[U06099] |
| | *Rattus norvegicus* | | | | | >RatnorTPX3[AF106944] | |
| | *Cavia porcellus* | | | >CavporA[AJ005111] | | | |
| | *Sus scrofa* | | | >SusscrA[D89812] | | | |
| | *Canis familiaris* | | | >CanfamA[AB0129181, AB038231] | | | |
| | *Bos taurus* | | | >BostauA [P00432] | | >BostauTPX2[AF305562] | >BostauTPX3[D82025] |
| | *Homo sapiens* | none | none | >HomsapA[AL035079] | none | >HomsapTPX2[BC000452 | >HomsapTPX3[NM_006793] |

| Taxon | *Organism* | Group I catalase | Group I catalase | Group I catalase | Group I catalase | TPX |
|---|---|---|---|---|---|---|
| EUKARYOTA, PLANTAE (plants lack group II or group III catalases, CPX and non-heme catalase) | | | | | | |
| Viridiplantae, Chlorophyta | *Chlamydomonas reinhardtii* | >ChirheA[AF016902] | | | | >ChireiTPX[AF312025] |
| Viridiplantae, Streptophyta | *Arabidopsis thaliana* | >ArathaA[AF021937] | >ArathaB[X64271, AL022023] | >ArathaC[AF021937] | none | >ArathaCTPX1[X94218] |
| | *Avicennia marina* | >AvimarA[AF328861] | | | | |
| | *Brassica juncea* | >BrajunA[AF104451] | >BrajunB[AF104452] | >BrajunC[AF104453] | >BrajunD[AF104454] | |
| | *Brassica napus* | >BranapA[U68219] | | | | >BranapCTPX[AF311863] |
| | *Capsicum annum* | >CapannA[AF227952] | | | | |
| | *Cucurbita pepo* | >CucpepA[D55645] | >CucpepB[D55646] | >CucpepC[D55647] | | |
| | *Digitalis lantharum* | *>DiglanA[partial, AJ297612] | | | | |
| | *Gossypium hirsutum* | >GoshirA[X52135] | >GoshirB[X56675] | | | |
| | *Glycine max* | >GlymaxA[AF035252] | >GlymaxB[AF035253] | >GlymaxC[AF035254] | >GlymaxD[AF035255] | |
| | *Helianthus annus* | >HelannA[L28740] | >HelannB[AF243517] | >HelannC[AF243518] | >HelannD[AF243519] | |
| | *Hevea brasiliensis* | >HevbraA[AF151368] | | | | |
| | *Hordeum vulgare* | >HorvulA[U20777,AF021938] | >HorvulB[U20778,AF021939] | | | >HorvulCTPX[Z34917] |
| | *Ipomoea batatas* | >IpobatA[X05549] | | | | |
| | *Lycopersicon esculentum* | >LycescA[M93719] | >LycescB[AF112368] | | | |
| | *Manihot esculenta* | >ManescA[AF170272] | | | | |
| | *Mesembryanthemum crystallinum* | >MescryL[AF069319] | >MescryR[AF069320] | | | |
| | *Nicotiana glutinosa* | >NicgluA[AF006067] | | | | |
| | *Nicotiana plumbaginifolia* | >NicpluA[Z36975] | >NicpluB[Z36976] | >NicpluC[Z36977] | | |
| | *Nicotiana sylvestris* | >NicsylA[U07626] | | | | |
| | *Nicotiana tabacum* | >NictabSR1A[U07627] | | | | |
| | *Oryza sativa* | >OrysatA[D29966,X61626] | >OrysatB[D64013,D26484] | >OrysatC[AB020502,D86611] | | |
| | *Phaseolus vulgaris* | >PhavulA[partial AF149283] | | | | >PhavulCTPX[AJ288895] |
| | *Pisum sativum* | >PissatA[X60169] | | | | |
| | *Prunus persica* | >PruperA[AJ132444] | | | | |
| | *Raphanus sativus* | >RaosatA[AF031318,AF24849] | >RapsatB[AF139538] | >RapsatC[AF139539] | | |
| | *Ricinus communis* | >RiccomA[D21161] | >RiccomB[D21162] | | | |
| | *Secale cereale* | >SeccerA[Z54143] | | | | >SeccerTPX[AF076920] |
| | *Soldanella alpina* | >SolalpA[Z99633] | | | | |
| | *Solanum melongena* | >SolmelA[X71653] | | | | |
| | *Solanum tuberrosum* | >SoltubA[U27082] | | | | |
| | *Suaeda maritima subsp. Salsa* | >SuamarA[AF390210] | | | | |
| | *Triticum aestivum* | >TriaesA[D86327,Q43206] | >TriaesB[X94352] | | | |
| | *Vigna radiata* | >VigradA[D13557] | | | | |
| | *Zea mays* | >ZeamayA[X12538,M33104] | >ZeamayB[X54819,Z54358] | >ZeamayC[X12539,L05934] | | |
| | *Zantedeschia aethiopica* | >ZaneatA[AF207906] | >ZaneatB[partial; AF332973] | | | |

NOTE.—Sources include Genbank accession number or respective sequence data base contig number. Thioredoxin peroxidase (= alkyl hydroperoxide reductasease, peroxiredoxin) genes were also listed when accessible. Entries were organized at the class level following the taxonomic system proposed by Cavalier-Smith (2002*a*, 2002*b*).

conducted assuming equal rates across sites and using the JTT empirical amino acid substitution model (Jones, Taylor, and Thornton 1992). In a postrun analysis, MrBayes summarized the results concerning tree topology and branch lengths. By ignoring the trees generated before the search converged on stable likelihood values (removed as ''burn-in''), a 50% majority rule consensus phylogram was constructed that displayed the mean branch lengths and posterior probability values of the observed clades. These probability values were comparable to the bootstrap proportions calculated for the branches in the MP consensus trees. Additionally, programs contained in the PHYLIP package (Felsenstein 1994) were used as before (Klotz, Klassen, and Loewen 1997; Loewen, Klotz, and Hassett 2000) for additional analysis of the much larger data set of the monofunctional heme-catalase sequences.

Sequence alignments, contig alignments for intron analysis (Sequencher 4.1, Gene-Codes, Madison, Wis.), phylogenetic inferences (PAUP*, MrBayes, PHYLIP 3.5), tree constructions (TreeViewPPC, v. 1.6.5, Roderick Page, personal communication), and graphics were accomplished by using Apple Macintosh G3 computers. The SEQBOOT, PROTPARS, PROTDIST, CONSENSE, NEIGHBOR, FITCH, and TREE programs of the PHYLIP package were also run on a SGI Impact R10000

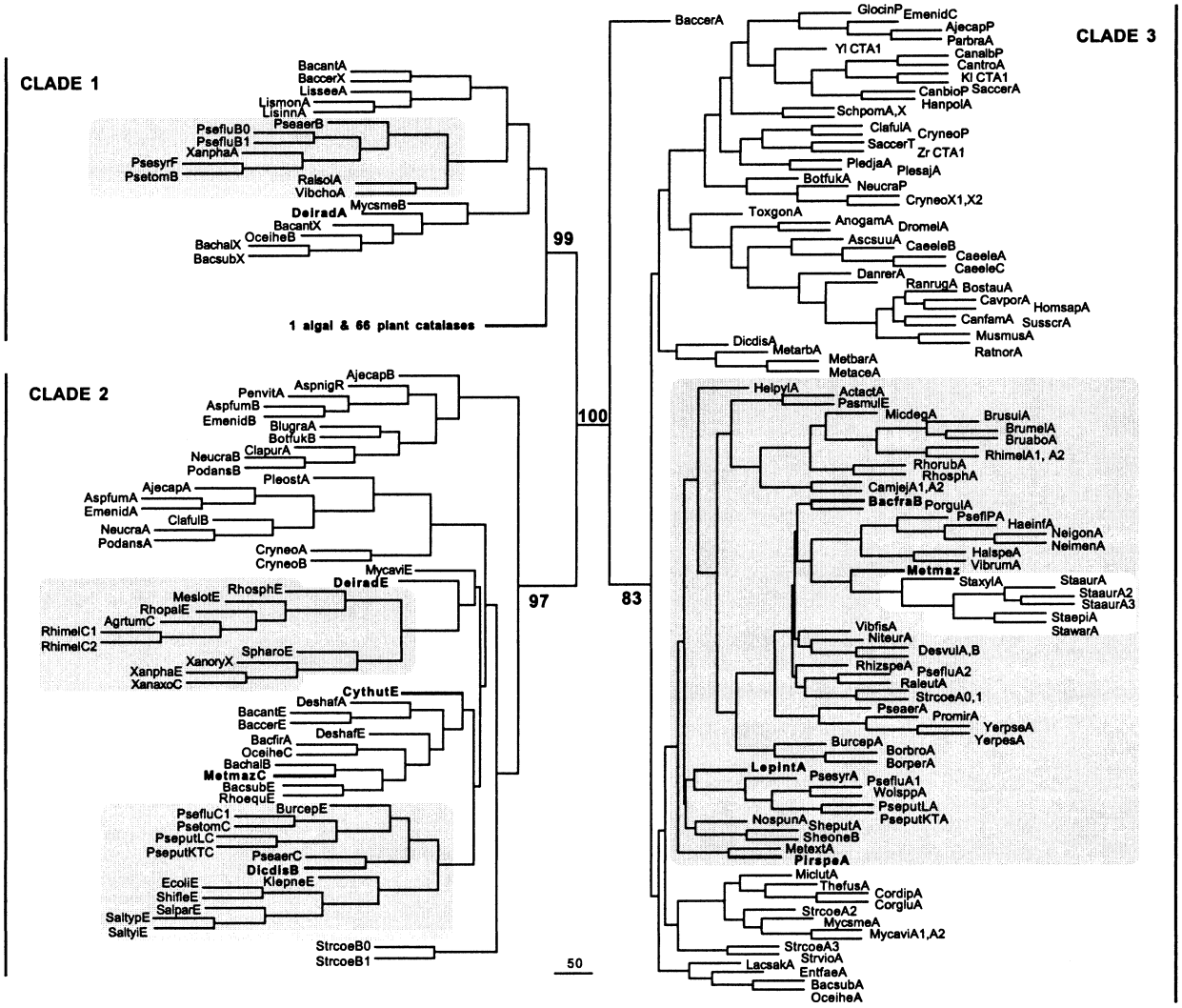


FIG. 1.—Phylogenetic analysis of the heme-containing monofunctional catalase gene family. Distance neighbor-joining tree constructed from 256 protein sequences aligned using ClustalX or ClustalW. Bootstrap support of the tree is very high. Most values were omitted for better readability of the figures; however, readers are referred to several recent publications where bootstrap values were included in the figure (Klotz, Klassen, and Loewen 1997; Loewen, Klotz, and Hassett 2000; Johnson et al. 2002). Bold lines and font was used to indicate suspected horizontal gene transfer. The abbreviations used for the catalases are explained in table 1.

workstation. Alignments generated from full-length and core sequences resulted in similar or identical phylogenetic trees.

## Results and Discussion

### The Evolution of Monofunctional Heme Catalase

There are accessible records of over 260 sequences of monofunctional catalases (table 1), most of them in form of cDNA. The evolutionary relationships of subsets of these enzymes have been discussed in the past (Mayfield and Duvall 1996; Klotz, Klassen, and Loewen 1997; Scandalios, Guan, and Polidoros 1997; Frugoli et al. 1998; Kim, Sha, and Mayfield 2000; Loewen, Klotz, and Hassett 2000), and differences in conclusion were likely due to limitations in sample size. For instance, frequent lateral shuttling of catalase genes between eukaryotic hosts and bacteria (Mayfield and Duvall 1996; Kim, Sha, and Mayfield 2000) had been proposed using selected subsets of enzymes not large enough to reveal the trifurcation of the phylogenetic tree. Similarly, the evolution of plant catalases by intron loss from an intron-rich ancestral eukaryotic catalase gene (Frugoli et al. 1998) had been proposed using only plant catalases, which did not allow a determination of the polarity of the process. In contrast, the following discussion of the evolution of the monofunctional catalase gene family is based on a comprehensive analysis of all available catalase gene and protein sequences, a number that has more than tripled since our last review (Klotz, Klassen, and Loewen 1997).

The phylogenetic trees constructed by distance neighbor-joining (Phylip 3.5) or MP (PAUP*) inferences yielded nearly identical trees, showing that the monofunctional catalase gene family falls into three member clades (fig. 1). While clade 1 and 3 contain small-subunit catalase sequences, clade 2 contains exclusively large-subunit catalase sequences. The observed branching pattern

implies an evolution of the three sequence lineages after a minimum of two gene duplications (Klotz, Klassen, and Loewen 1997; Loewen, Klotz, and Hassett 2000). Catalases are found in both empires of life, the Prokaryota and Eukaryota (Mayr 1998; Cavalier-Smith 2002*b*), but their abundance and diversity in some taxa is limited. In the Eukaryota, catalases are present in all major taxa, the Protista, Animalia, Fungi, and Planta, but overall, catalases do not group according to ssu-rRNA–based species-phylogenetic relationships (Olsen, Woese, and Overbeek 1994). For example, all three catalase clades contain bacterial enzymes, and analysis of the individual clades shows that each clade has a base of bacterial sequences that branches off into a bacterial branch and an eukaryotic branch (fig. 1).

In bacteria with multiple catalase genes, the enzymes usually (with the exception of some small subunit catalases) group in different clades (Klotz, Klassen, and Loewen 1997; Loewen, Klotz, and Hassett 2000). This is fundamentally different from the case of catalases in animals and plants. Higher animals seem to have only one clade 3–type catalase gene. In lower animals with multiple catalase genes (e.g., *Caenorhabditis elegans*), the genes subgroup tightly in the clade and have likely arisen by an initial duplication of one clade 3–type catalase gene, after which each of the loci have evolved independently. The worm *Onchocerca* has no functional catalase gene at all, and it may use the catalase of its obligate α-proteobacterial endosymbiont (Henkle-Dührsen et al. 1998) to degrade hydrogen peroxide.

All algal/plant catalases reside in clade 1 together with a subset of small-subunit catalases from Posibacteria and Negibacteria. Multiple catalase genes in the plant genome have likely arisen by an initial duplication of a clade 1–type catalase gene. Plant genomes harbor small catalase gene subfamilies of up to four member genes, and their phylogenetic relationships (Frugoli et al. 1996; Guan and Scandalios 1996) and gene structures (Frugoli et al. 1996) have been analyzed recently. The expression products of these genes can assemble to various different tetrameric isozymes, and the regulation of this allows the synthesis of organ-specific catalases (Frugoli et al. 1998). Clade 1–type catalases have not yet been found in animals or fungi. Fungi generally have multiple catalase genes per genome. Some fungi express catalases that group in a single clade (either clade 2 or clade 3), and others show distribution of their catalases into both clade 2 and clade 3. Clade 2 catalases are also found in bacteria, whereas clade 3 catalases are found in all phylogenetic groups except for plants. Eukaryotic clade 3–type catalases are usually peroxisomal.

To refine the relative groupings of eukaryotic catalases, the number and locations of introns in each gene was analyzed as previously described (Johnson et al. 2002), and the results were used to construct a phylogenetic tree inferring maximum parsimony (fig. 2). Whereas a grouping of plant (clade 1) and animal catalases (clade 3) in the intron tree implies the existence of single-rooted intron lineages, multiple independent intron lineages are evident for protozoan (clade 3) and fungal (clades 2 and 3) catalase genes. Prokaryotic catalase genes naturally lack introns.

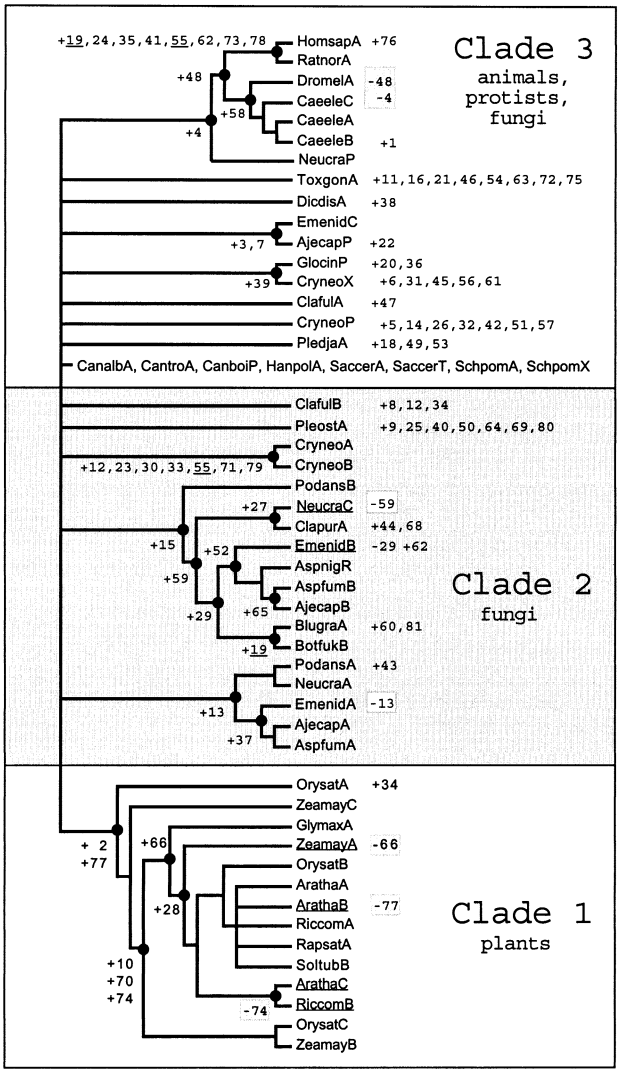


FIG. 2.—Maximum-parsimony tree constructed from analyses of intron residence in eukaryotic catalase genes. The positions of intron residence in individual catalase genes were mapped onto the alignment of the catalase protein sequences. Intron positions were numbered consecutively beginning at the N-terminus of the alignment. The resulting groups reflect common and unique acquisition (+) and loss (–) events of introns in an individual intron lineage (numbers given before the node). Numbers given behind protein definitions refer to acquisition (+) and loss (–) events of introns in the individual encoding gene. Intron lineages were grouped according to the clade structure in the tree constructed from phylogenetic analysis of aligned catalase sequences (see fig. 1). Abbreviations for the catalases are defined in table 1.

An evolutionary discussion of a gene family should attempt to identify the ancestral gene and put the molecular evolution of the enzyme in perspective. All currently discussed hypotheses (Martin and Müller 1998; Mayr 1998; Embley and Martin 1998; Müller and Martin 1999; Rotte et al. 2000; Henze and Martin 2001; Bansal and Meyer 2002; Cavalier-Smith 2002*a*, 2002*b*; Hartman and Fedorov 2002) are compatible with a bacterial origin of the monofunctional catalase gene family (as per fig. 1). For this communication, we mapped catalase residence on a composite tree (fig. 3) schematically constructed by integration of several recently published universal trees of life (Gupta 1998; Cavalier-Smith 2002*a*, 2002*b*; Wolf et al. 2002).

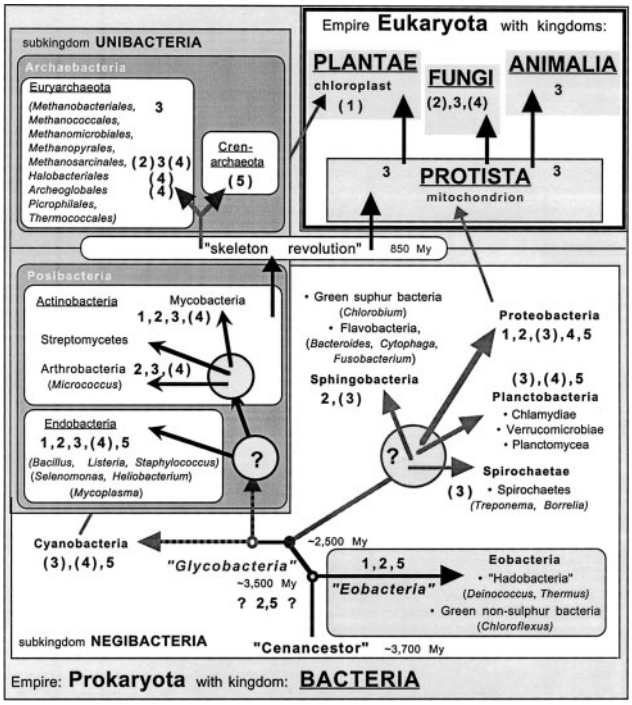


FIG. 3.—Distribution of catalatic hydroperoxidases on a conceptual universal tree of life and using taxonomic groups proposed by Cavalier-Smith (2002*a*, 2002*b*). Monofunctional catalases of the clade 1–type, clade 2–type, or clade 3–type were listed using the respective numbers. Catalase-peroxidase (CPX) distribution is indicated with number 4, whereas the distribution of the nonheme (Mn) catalase is indicated by using number 5. Numbers in parenthesis indicate that the genes were highly likely laterally transferred into the respective taxon.

## Prokaryotic Monofunctional Catalases

In general, all recent monofunctional heme-containing catalases are significantly similar in sequence and are true homologs as the result of divergent evolution. First, we will discuss the evolution of catalases in the bacteria. Some gram-negative and gram-positive bacteria (here called Negibacteria and Posibacteria, respectively) contain functional catalase genes from all three clades (table 1), which could be interpreted as evidence that the catalase gene duplication events occurred early in bacteria after their evolution from the cenancestor, the Woesean progenote (Olsen, Woese, and Overbeek 1994), approximately 3,500 MYA and before the delineation of the ''Glycobacteria'' (Cavalier-Smith 2002*a*) approximately 2,500 MYA (fig. 3). A closer look, however, suggests that clade 3–type catalases were ''invented'' much later (see below). The archetypal monofunctional catalase gene is likely a clade 2–type, large-subunit catalase gene. Large-subunit catalases are tetramers with a remarkable stability that remain functional after exposure to extreme temperatures and pH environments (Switala, O'Neil, and Loewen 1999). Not surprisingly, clade 2–type catalase genes have been identified in extant species of the Eobacteria (Cavalier-Smith 2002*a*) such as *Deiniococcus radiodurans*, and they have divergently evolved in the evolutionarily younger negibacterial divisions of the ''Glycobacteria'' and in the Posibacteria (fig. 3). This conclusion is also based on the suggestion that the genome of *D. radiodurans* has been subjected to unusually few horizontal gene transfer events when compared with other prokaryotic genomes (Makarova et al. 2001; White et al. 1999). According to most universal tree phylogenies, the Cyanobacteria and unibacterial Posibacteria are paraphyletic sister groups. Clade 2–type catalase genes are absent from the Cyanobacteria (table 1 and fig. 1), thus the Cyanobacteria seem to have lost the clade 2–type catalase gene in the process of their emergence from the common ancestor of Cyanobacteria and Posibacteria, whereas Posibacteria retained the gene (fig. 3). A plausible explanation for this gene loss is that the usually aquatic Cyanobacteria lived in more stable habitats with respect to environmental stress than the mostly soilborne Posibacteria. In addition, the complement of hydrogen peroxide–degrading enzymes had increased in the ''Glycobacteria'' with the emergence of the unrelated nonheme (Mn) catalases (see fig. 3 and below). As obvious from table 1 and figure 1, clade 2–type, large-subunit catalase genes were retained by the abundant environmental genera in both subdivisions of the Posibacteria (e.g., *Bacillus*, *Streptomyces*, and *Mycobacterium*), whereas mostly host-associated genera (e.g., *Staphylococcus* and *Listeria*) have lost the gene by genome economization likely due to the reduction of environmental stresses. Clade 2–type catalase genes have also holophyletically evolved in the other branch of the negibacterial ''Glycobacteria,'' as representatives of the Sphingobacteria (*Cytophaga*) and Proteobacteria (e.g., pseudomonads and Enterobacteria) harbor the gene (table 1 and figs. 1 and 3). Interpretation of the branching pattern in clade 2 (fig. 1) suggests that the posibacterial lineage evolved independently from the ''glycobacterial'' lineage, which appears to be substructured into α-bacterial and β/γ-proteobacterial sublineages (fig. 1 and table 1). Clade 2–type catalases are absent from the usually catalase-free Spirochaetea (see below for an exception), the Planktobacteria, and the δ-Proteobacteria and ε-Proteobacteria. This is not surprising, as most of these bacteria are host-associated and have suffered significant gene losses in their evolution (Cavalier-Smith 2002*a*). The recent identification of a clade 2–type catalase gene in the genome of the Euryarchaebacterium *Methanosarcina mazei* was very surprising, and it constitutes the first finding of a prokaryotic large-subunit catalase gene outside the Eubacteria (Deppenmeier et al. 2002). Its sequence is most closely related to those of large-subunit enzymes from bacilli (fig. 1), and the high abundance and wide distribution of bacilli that overlap with those of the methanogens makes an acquisition of this clade 2–type catalase gene by *M. mazei* from bacilli very likely.

Fairly soon after the emergence of clade 2–type catalases, a gene duplication event accompanied by losses of sequence at both the 5′ and the 3′ ends resulted in the ancestral gene encoding small-subunit catalases. Such a ''sequence loss upon duplication'' scenario is supported by recent experimental data in that an N-terminally and C-terminally digested clade 2–type catalase was found to retain catalase activity (Loewen, in preparation). The first

small-subunit catalase gene was likely a clade 1–type gene, as it is present in the Eobacteria in which a clade 3–type gene has not yet been found (fig. 3). Furthermore, the orientation of heme in clade 2–type and clade 1–type catalases is identical but differs by 180° in clade 3–type catalases (Carpena et al. 2002). In addition, the evolutionary distance between clades 2 and 1 on the phylogenetic tree is always shorter than the distance between clade 3 and either one of clade 1 or clade 2 (fig. 1), independent of the method of inference. The tree in figure 1 also tells us that the clade 1–type catalase gene has further evolved paraphyletically and orthologously in the evolutionarily younger lineages of the Negibacteria and Posibacteria. Interestingly, like in the case of the clade 2–type catalases, the clade 1–type catalase genes of the Posibacteria are more similar in sequence to the eobacterial gene found in *Deinococcus* than the negibacterial clade 1–type genes (figs. 1 and 3).

Clade 3–type catalase genes are absent from modern representatives of the phylogenetically old taxonomic groups and abundant only in Posibacteria, Proteobacteria, and Eukaryota (fig. 3 and table 1). If clade 3–type catalase genes were as old as the other two members of the gene family, then the gene must have been lost massively in all these divergently evolving taxa after the emergence of the Posibacteria and Proteobacteria. Although the scantiness of molecular traces (pseudogenes) in bacterial genomes after gene loss can be explained theoretically (Lawrence, Hendrix, and Casjens 2001), a selective massive simultaneous gene loss in some but not in other taxa is very unlikely. We reason instead, that the clade 3–type, small-subunit catalase gene has likely developed last by a second gene duplication event in either the Posibacteria or Proteobacteria and was subsequently horizontally acquired by the other taxon and a few species in other bacterial taxa (see below). We propose that the clade 3–type catalase gene has evolved in the Posibacteria. This is concluded from the branching pattern in figure 1 because posibacterial sequences group near the root of the clade, forming two subclades, one being exclusively Actinobacteria and the other being a mix of actinobacterial and endobacterial sequences. The latter subclade contains predominantly proteobacterial catalase genes, which subgroup into α-proteobacterial and β/γ-proteobacterial lineages. Hence we propose that (a) posibacterial clade 3–type catalase gene(s) were (was) laterally acquired by Proteobacteria. This laterally acquired clade 3–type catalase gene was retained in addition to other types of catalase genes in environmental genera such as *Pseudomonas*, *Ralstonia*, and *Rhizobium*, whereas animal host–associated genera such as *Neisseria*, *Brucella*, and *Bordetella* have kept this gene at the expense of all other catalase gene types. Interestingly, the animal hosts also harbor clade 3–type catalases, suggesting that the similar physiological selection pressures resulted in the selection of the same enzyme in both parasite and host. The proteobacterial subclusters in clade 3 (fig. 1) are ''spiked'' with taxonomic ''exceptions'': a sphingobacterial gene (*Bacteroides*), a planktobacterial gene (*Pirellula*), a spirochaeteal gene (*Leptospira*), a cyanobacterial gene (*Nostoc*), and an archaebacterial gene (*Methanosarcina mazei*). The *Bacteroides* and *Methanosarcina* and *Nostoc* and *Pirullela* catalases group closely together on the phylogenetic tree (fig. 1), although the organisms are rather distantly related (fig. 3). A clade 3–type catalase gene was isolated just recently from the spirochaete *Leptospira interrogans* (GenBank accession number AE011360). In addition to its close grouping with clade 3–type catalases from pseudomonads and *Wolbachia* on the phylogenetic tree (fig. 1), the *Leptospira* catalase gene is also in a tandem arrangement with ankyrin as found for several pseudomonad catalases (Ma et al. 1999). Thus, these clade 3–type catalase genes likely arrived in their respective host cells by lateral transfer, as they share habitats with many of the prospective gene-donating eubacterial cousins. Extant representatives of the Spirochaetae, Planctobacteria, Sphingobacteria, and Cyanobacteria otherwise lack clade 3–type catalase genes, and the other detected clade 3–type catalase genes in the Euryarcheota are too distant from the *Methanosarcina mazei* gene on the tree to propose a direct common ancestor for all three archaebacterial genes (fig. 1). Hence, clade 3–type catalase genes have likely been horizontally transferred in several events between Eubacteria and at least once between Eubacteria and the euryarchaebacterium *Methanosarcina mazei*. The latter organism seems to have acquired both of its catalase genes horizontally from different donors: the clade 2–type catalase gene from bacilli and the clade 1–type catalase gene from a proteobacterial donor (fig. 1). Horizontal gene transfer has been addressed recently as a widespread and evolutionary important event among bacteria (e.g., Doolittle 1999; Ochman, Lawrence, and Groisman 2000; Nesbø et al. 2001; Gogarten, Doolittle, and Lawrence 2002).

The other three documented occurrences of a clade 3–type catalase in Euryarcheota (Methanobacteriales and Methanosarcinales [figs. 1 and 3]) represent likely vertical holophyletic molecular evolution from a mycobacterial/streptomycetal clade 3–type catalase gene in the transitional state of exoskeleton revision (Cavalier-Smith 2002*a*, see fig. 3 and below). A careful reexamination of the sequencing data by the *Methanosarcina acetivoran*s genome project staff (chad@genome.wi.mit.edu) confirmed a frame-shift mutation immediately downstream of the catalytic region. The sequence corrected by us for this mutation and included in the analysis (fig. 1) represents therefore the hypothetical expression product of a pseudogene. We hypothesize that similar loss of function mutations occurred also in other methanogens whose molecular traces have been erased (Lawrence, Hendrix, and Casjens 2001). All other crenarchaeal and euryarchaeal descendents of the ''neoexoskeletal ancestor'' likely have lost catalase genes among many other genes early in the process of adaptation to extreme environments (acidic, high temperature, etc.) and ongoing genome economization (Cavalier-Smith 2002*a* and references therein).

## Eukaryotic Monofunctional Catalases

There are only a few reports on LGT from bacteria into the Eukaryota. Some genes were suspected of having been transferred as a result of acquisition of endo-

symbionts by the eukaryotic cell, and a few others have been reported between bacteria and protozoa (Andersson and Roger 2002; Nixon et al. 2002 and references therein). Eukaryotic cells and Archaebacteria have likely acquired catalase genes from eubacteria both vertically and laterally after the three types of catalase genes evolved as outlined above, and we are going to discuss in the following how the Eukaryota have acquired them. It will be helpful in this discussion to also consider the physiological roles of individual catalases in their host organisms and to ask whether they are housekeeping enzymes or induced or secreted enzymes that confer advantageous properties to their hosting cell. Furthermore, our analysis of intron residence (number and location) in eukaryotic catalase genes, summarized in figure 2, provides further information about the directedness of catalase gene evolution and—if lateral transfer is suspected—how many times such transfer might have had taken place. Although the phylogenetic tree derived from alignments of catalase protein sequences reveals at least one lateral transfer event of a clade 3–type catalase gene from eubacteria to methanogens, the neoexoskeletal ancestors of the Archaebacteria and Eukaryota seem to have acquired their housekeeping catalase gene from descendents of the Actinobacteria. This is in line with the numerous recently reviewed similarities between and lines of descent of many structures and functions in the Actinobacteria and Archaebacteria (Cavalier-Smith 2002*a*) and lower Eukaryota such as the Protozoa (Cavalier-Smith 2002*b*).

It is almost certain that acquisition of the type 3 catalase gene by eukaryotic cells from bacteria succeeded both catalase gene duplication events. If this were not the case, residence of all three types of catalases or the molecular remnants of their encoding genes should be expected in most eukaryotic taxa. Instead, catalase-positive Protista, Animalia, and Fungi contain a clade 3–type catalase as a housekeeping enzyme, whereas modern algae and plants lack type 3 catalases. It is well documented (Cavalier-Smith 2002*b* and references therein) that plants evolved from ancestral protista other than the ancestors of the fungi and animals. Therefore, absence of clade 3–type catalases in plants can be well explained by a single gene loss event early in the alga/plant evolutionary lineage. Modern plants and algae contain exclusively small clade 1–type catalase gene families (Frugoli et al. 1996). These likely did not originate from the endosymbiontic α-bacterium ancestral to the mitochondria or the oxygenic photosynthetic endosymbiotic cyanobacterium ancestral to chloroplasts because α-Proteobacteria and Cyanobacteria did not holophyletically acquire clade 1–type catalase genes (table 1 and fig. 3). Instead and because clade 1–type catalase genes are not abundant among the Actinobacteria (table 1), the protistan cell(s) ancestral to the alga/plant lineage likely acquired their clade 1–type catalase gene directly via LTG from an endobacterial/ eobacterial or proteobacterial donor whose recent descendants carry clade 1–type genes (figs. 1 and 3). The single lineage of intron residence in plant catalase genes (fig. 2) supports the singularity of the gene loss (clade 3–type gene) and acquisition (clade1–type gene) events in the alga/plant ancestral protistan cell.

Protist, animal, and fungal lineages of clade 3–type catalase genes delineate from separate bacterial branches that contain actinobacterial sequences (fig. 1). Hence, it is plausible to propose that protozoa, animals, and fungi acquired their catalase genes as a result of holophyletic molecular evolution from their Actinobacteria-derived ancestors (Cavalier-Smith 2002*b*). However, it is also clear from figure 1 that clade 3–type catalase genes in the protists and animals and in the fungi have likely originated from different protozoan ancestors. Although there is currently no catalase sequence from a modern protozoon available that groups with the fungal clade 3–type catalases, the sequence of catalase A from *Dictyostelium* subgroups with archaebacterial catalases, and the sequence of catalase A from *Toxoplasma* subgroups with the animal catalases. The clade 3–type catalases from *Dictyostelium*, *Toxoplasma*, and the animals represent different lineages of gene development in that they constitute different lineages of intron acquisition (fig. 2). Clade 3–type catalases from the fungi represent multiple lineages of eukaryotic gene evolution in that they have acquired from one to as many as seven introns, or they lack introns all together (fig. 2).

In contrast to most of the housekeeping clade 3–type catalases, many of the clade 2–type catalases in fungi are secreted and/or inducible by environmental cues (Johnson et al. 2002; Kawasaki et al. 1997). These properties suggest that clade 2–type catalases have been acquired as supplements to constitutive housekeeping enzymes and play a role in the virulence of their recipients (Kawasaki et al. 1997). A recent characterization of the catalase complement of the fungus *Histoplasma capsulatum* led Johnson et al. (2002) to propose that the known large-subunit, clade 2–type catalase genes in fungi were acquired from bacteria by at least two independent LGT events. The protein sequences of the two clade 2 catalase lineages were most similar to those of large-subunit catalases from Posibacteria such as *Bacillus*, *Mycobacterium*, and *Streptomyces* (fig. 1), which are abundant in aerobic soil habitats that overlap with those of the fungi. Although the description of *Histoplasma* catalase evolution is one of the first reports of direct LGT from bacteria into fungi, evidence for lateral transfer of genes from bacteria to fungi has been documented in laboratory experiments (Sprague 1991; Hayman and Bolen 1993). In this context, it is helpful to mention that there is also evidence for recent LGT of CPX genes from bacteria into fungi (see below).

The data of figure 2 support our hypothesis that the clade 1–type, clade 2–type, and clade 3–type catalases evolved in the Eukaryota by intron acquisition. The number of introns in catalase genes correlates with general intron-richness of their host genome; thus the absence of introns in the Hemiascomycotina is not surprising. The small number of intron loss versus gain events (only three of the plant, two of the fungal, and two of the animal catalase introns were lost) supports the proposed polarity of the process: invasion of introns into intronless genes. This is in contrast to the model of intron loss (Frugoli 1998) based on the analysis of selected plant catalase sequences.

Taken together, we conclude that monofunctional catalase genes are of bacterial origin and that the clade 3–type genes have holophyletically evolved in the Archaebacteria, Protista, Animalia, and Fungi. Catalase genes have also been laterally acquired by Archaebacteria (clade 3), plants (clade 1), and fungi (clade 2) more than once (direct LGT and/or phagocytosis), and they have further evolved in eukaryotes orthologously and paralogously by intron acquisition into formerly intron-free genes. In addition to analyses of evolutionary distance and intron residence, emerging structural/crystallographic information on catalases from all three clades (Carpena et al. 2002) supports a model of divergent evolution rather than convergent evolution of the monofunctional catalase gene family.

## The Evolution of Bifunctional Catalase-Peroxidase

Bifunctional CPXs are encoded by a second family of catalase genes that are sequence-unrelated to the monofunctional catalases discussed above. In a recent article, Faguy and Doolittle (2000) aligned 19 CPX protein sequences (nine proteobacterial, five posibacterial, three archaebacterial, and two cyanobacterial sequences) and concluded from the derived trees that CPX genes had been laterally transferred from Archaebacteria into pathogenic Proteobacteria. Two years later and discounting a handful of short sequences that are likely the result of incomplete gene duplication events, we count a total of 58 usable sequences: 30 proteobacterial, 14 posibacterial, five archaebacterial, three cyanobacterial, and one planctobacterial sequences. Also five eukaryotic CPX gene sequences from pezizomycotinal fungi are reported. The tree resulting from alignment of these 58 sequences (fig. 4) is much less robust than previous trees that were based on a much smaller set of selected sequences (Klotz, Klassen, and Lowen 1997; Faguy and Doolittle 2000; Loewen, Klotz, and Hassett 2000). The majority-consensus tree obtained from Bayesian inference of CPX phylogeny shown in figure 4 allows for several different interpretations of CPX gene evolution; however, it is evident that several LGT processes contributed to CPX gene distribution. An interpretation of the phylogenetic tree (fig. 4) in conjunction with figure 3 suggests that the evolution of the CPX gene family has likely occurred much later than that of the heme-containing monofunctional catalases discussed above. The groupings of the CPX sequences from *Desulfitobacterium* (Posibacteria) and *Geobacter* (Negibacteria); *Shewanella*, *Legionella*, and *Vibrio* (Proteobacteria) and Cyanobacteria and *Bacillus* (Posibacteria); diverse Proteobacteria and *Pirellula* (Planctobacteria) and Archaebacteria; and Proteobacteria and Archaebacteria underline the frequency of lateral exchange of the CPX genes among prokaryotes. The claim by Faguy and Doolittle (2000) of directed LGT of CPX genes from Archaebacteria to pathogenic Eubacteria is incongruent with phylogenetic trees obtained with the larger data set (fig. 4), as the proposed ''recipient group'' contains nonpathogenic bacteria in the genera *Nitrosomonas*, *Shewanella*, and *Pirellula*. The CPX proteins found in fungi are functional (Johnson et al. 2002; Kawasaki and Aguirre 2001) and their sequences group closely on the phylogenetic tree as a subcluster of a larger cluster of proteobacterial sequences. Thus, it seems feasible to propose a single recent lateral acquisition event of a proteobacterial CPX gene by an ancestor of the Pezizzomycotina.

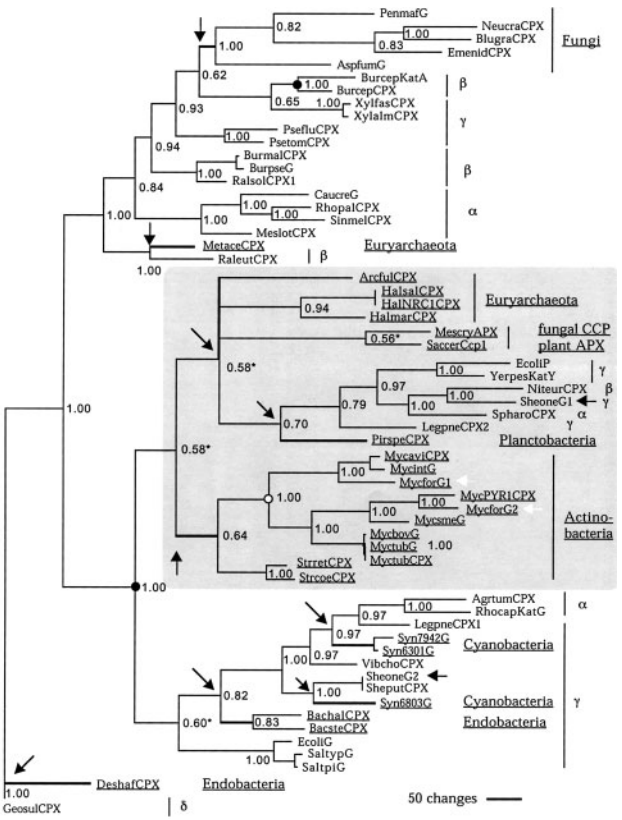


FIG. 4.—Phylogenetic analysis of 58 protein sequences in the heme-containing bifunctional catalase-peroxidase (CPX) gene family (aligned using ClustalX): 50% majority consensus tree of all credible topologies sampled by MrBayes over 50,000 generations using a maximum-likelihood approach with empirical substitution frequencies (JTT) and assuming equal rates across sites. Posterior probability values for the clades are shown at the branch points. Branch points indicating lateral gene transfer events were marked with an arrow. Shading was used to divide the tree to reflect an early gene duplication event (filled circle) after which the two genes have evolved independently in either the same or in different host cells. Open circles were used to indicate more recent gene duplication events. Relatively low posterior probability values were marked with an asterisk.

CPX is presently most abundant in proteobacterial genera, which represent a holophyletic direction of eubacterial evolution (fig. 3). Therefore, we propose that an ancestral CPX gene has likely arisen in an ancestral proteobacterium (note that the α-proteobacterial and β/γ-proteobacterial sequences group separately) and diversified before it was acquired in independent LGT events by selected genera in the divisions of the Planctobacteria, Cyanobacteria, Endobacteria, Actinobacteria, Archaebacteria, and Pezizzomycotina. At the beginning of this paper we mentioned the hypothesis that CPXs may have arisen through a gene duplication and fusion after which the C-terminal domain likely lost its functionality (Welinder 1991, 1992; Zamocky et al. 2001). The existence of shortened and likely nonfunctional CPX gene duplicates in

several Proteobacteria only (e.g., *Burkholderia fungorum* [*cepacia*] and *N. europaea*) and the tree (showing two functional CPX gene copies for *Legionella* and *Shewanella*) reveal that paralogy in the Proteobacteria likely preceded many of the LGT events. The N-terminal domain of CPX has considerable sequence similarity with plant ascorbate peroxidases (APX) of which seven types have been identified (Jespersen et al. 1997). Several recent articles have discussed the relationships among the members in the class I plant APX family, which consists of fungal cytochrome C peroxidases (CCPs), plant cytosol, and chloroplast APX and the bacterial CPXs (Welinder 1992; Jespersen et al. 1997, Zamocky et al. 2001). An inclusion of plant cytosolic and chloroplast APX sequences into our phylogenetic analysis of CPX (both full-length and N-terminus only) generated trees in which all APX delineated from the archaebacterial subcluster (in the shaded box in fig. 4) and not in the cluster with cyanobacterial sequences. A high sequence similarity between APX and cyanobacterial CPX would have supported an endosymbiont-mediated transfer of CPX genes into algae and plants. In contrast, the obtained result indicates that the modern APX genes and the genes encoding the CPX in the shaded part of the tree in figure 4 are likely sharing the same ancestral CPX gene that diverged from the other gene copy found in modern Cyanobacteria after an early duplication (filled circle in fig. 4). We propose that this proteobacterial CPX gene arrived by LGT in an eukaryotic ancestor cell, and CPX evolved in the descendents of this ancestor into the class I enzymes in the plant peroxidase superfamily (Welinder 1992). Whereas the fungal and algae/plant lineages retained the gene, it was obviously lost in modern protozoa and the animals. Subsequent intron acquisition into the acquired CPX gene ancestral to APX and CCP might have led to a separation of the functional N-terminus from the less or nonfunctional C-terminus during evolution of the fungi, algae, and plants, thereby leading to the shorter modern genes that encode APX and CCP. In addition, shortening of the gene had likely consequences regarding protein folding, which together with crucial base substitutions is likely responsible for the loss of catalase activity by APX and CCP enzymes (Zamocky et al. 2001). Plant peroxidases in the class III family, such as horseradish peroxidase, seem to have retained catalase activity (Hernandez-Ruiz et al. 2001; Hiner et al. 2001), but more research is needed before class III plant peroxidases might be considered catalatic hydroperoxidases. Two domain-encoding CPX genes found in fungi have likely been laterally acquired fairly recently, as their genes should have evolved similar to the genes encoding APX and CCP as a result of intron acquisition (separation of the active and inactive domains representing the fused gene copies). Because coding genes are not interrupted in the prokaryotic Archaebacteria, CPX genes found in modern Archaebacteria retained the two domains.

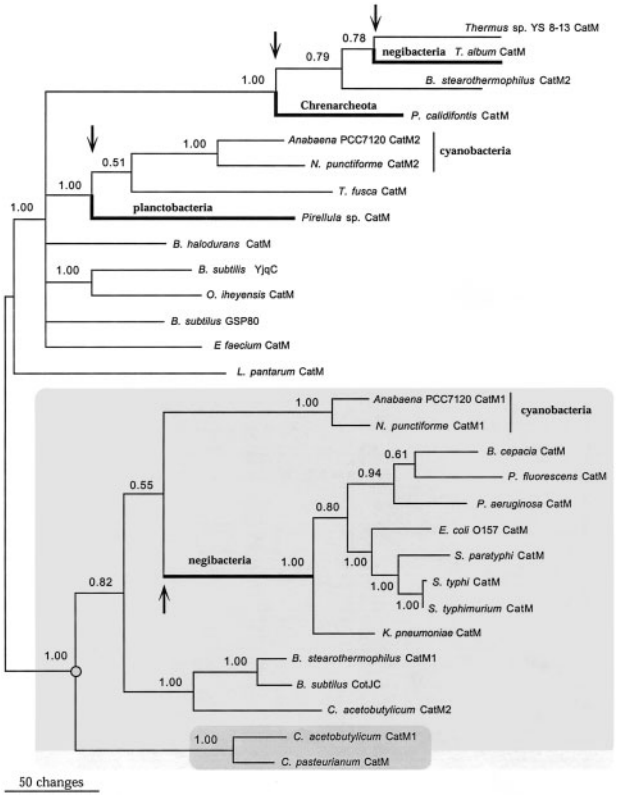


FIG. 5.—Phylogenetic analysis of 29 protein sequences in the nonheme monofunctional catalase gene family (aligned using ClustalX): 50% majority consensus tree of all credible topologies sampled by MrBayes over 30,000 generations using a maximum-likelihood approach with empirical substitution frequencies (JTT) and assuming equal rates across sites (Gamma-correction did not generate a different consensus tree). Posterior probability values for the clades are shown at the branch points. Branch points indicating lateral gene transfer events were marked with an arrow. Shading was used to divide the tree into two parts, which reflects an early gene duplication event after which the two genes have evolved independently in either the same or in different host cells. The circle in the shaded box indicates another gene duplication event.

## The Evolution of Nonheme Catalase

To summarize and discuss the evolution of the nonheme (Mn) catalase gene family, 29 of the 32 available sequences were aligned and used for the construction of phylogenetic trees. The depicted majority-rule consensus tree obtained by Bayesian inference of nonheme catalase phylogeny (fig. 5) indicated that nonheme catalase genes cluster in two major groups that have likely arisen by gene duplication. Nonheme catalases have apparently evolved later than monofunctional catalases but preceded the emergence if bifunctional CPX. Consideration of figures 3 and 5 led us to propose that the ancestral nonheme catalase gene evolved and was maintained in the common negibacterial ancestor of Cyanobacteria and Posibacteria but lost in the common ancestor of Spirochaetae, Sphingobacteria, Planctobacteria, and Proteobacteria. Present residence of Mn-catalase genes in the Planctobacterium *Pirellula* sp. and several β-Proteobacteria and γ-Proteobacteria is highly likely due to LGT from their consortial associates such as the nostocales or bacilli because the recipients contain only one or the other member gene of the family (fig. 5). Mn-catalase–positive Posibacteria and Cyanobacteria often contain representative genes from both members of the family. Whereas some of the Mn-

catalase–positive organisms are opportunistic pathogens, Mn-catalase genes have only been found in fairly widely distributed and abundant species (fig. 5 and table 1).

Mn-catalases have not been detected in Eukaryota (fig. 3) and were found mainly in Posibacteria (but not Actinobacteria) and in four of the six negibacterial divisions. The Actinobacteria are implicated as a crucial starting point for eukaryogenesis (Cavalier-Smith 2002*a*, 2002*b*), which may explain why Mn-catalase has largely remained restricted to the eubacteria. The recent finding of an Mn-catalase in a facultative aerobically respiratory, hyperthermophilic Chrenarcheabacterium, *Pyrobaculum calidifontis* VA1, represents a unique exception in the otherwise strictly eubacterial residence of Mn-catalases (Amo, Atomi, and Imanaka 2002). The catalatic activity of Mn-catalase is minor compared with heme-containing catalases, and it is likely that they only have significance in modern anaerobic bacteria such as the clostridia and some lactobacilli, whereas they were good candidates for gene loss in aerobic bacteria with diverse complements of catalases (e.g., bacilli and pseudomonads). Interestingly, specific catalase activity of the Mn-catalase in *P. calidifontis* was uncharacteristically high (23,500 units/mg protein) (Amo, Atomi, and Imanaka 2002).

## Concluding Remarks

The early atmosphere was reducing in nature and thus the concentration of molecular oxygen was low. Consequently, the concentration of ROI was also likely low. Interestingly, most modern taxa, including many anaerobes that lack catalatic hydroperoxidases, contain functional thioredoxin peroxide reductase (TPX, = peroxiredoxin, = alkyl hydroperoxide reductase) for defense against peroxides (table 1). Enzymes in this family have a 1,000-fold higher specificity to $H_2O_2$ than CHPs and are saturated at comparatively low (micromolar) peroxide concentrations (Seaver and Imlay 2001*a*, 2001*b*). This leads us to propose that TPX is likely the most ancient hydrogen peroxide–degrading enzyme that evolved before the rising atmospheric oxygen concentration favored the operation of the less specific (higher $K_m$ values) but more effective (higher turnover rates) catalatic hydroperoxidases.

Our discussion of the composite tree in figure 3 leads to the conclusion that monofunctional heme-containing catalase and Mn-catalase evolved likely in the ancestors of modern Eobacteria and Negibacteria, and they were followed by the emergence of bifunctional CPX in the Proteobacteria. Compared with the monofunctional heme-containing catalase and Mn-catalase families, distribution of CPX by LGT was significantly higher. The recent finding of an Mn-catalase with uncharacteristically high catalatic activity in an aerobic hyperthermophilic Chrenarcheabacterium is surprising; however, the phylogenetic analysis revealed that this gene arrived in this organism by LGT.

We propose that catalatic hydroperoxidases have evolved in the Eukaryota through vertical holophyletic evolution and several independent lateral transfer events of genes from eubacterial donors into eukaryotic recipients. The evolution of catalase genes in the Eukaryota is also characterized by genome-specific acquisition of introns into formerly intronless prokaryotic genes.

## Acknowledgments

Our general appreciation goes to all colleagues who openly share results and make sequencing data publicly available. Numerous preliminary sequence data were obtained from The Institute for Genomic Research (http://www.tigr.org), the Joint Genome Institute (http://www.jgi.org) and Integrated Genomes (http://www.integratedgenomes.com), whose URLs state the funding support for each individual project, including, but not limited to, the National Institute of Allergy and Infectious Diseases, the Department of Energy, and the National Science Foundation. In particular, we like to thank the REGX-consortium (www.regx.de) for providing us with the sequences from *Pirellula* sp. strain 1 before publication. Likewise, we are grateful to James A. Alspaugh (alspa001@mc.duke.edu) for providing us with the CryneoX (Cat1) sequence from *Cryptococcus neoformans* strain JEC21. We also thank Chad Nusbaum, Whitehead/MIT Center for Genome Research, for confirmation of the frame-shift mutation in the *katA* gene from *M. acetivorans*. Our thanks to David J. Schultz (University of Louisville) for a critical reading of the manuscript. Work reviewed in this paper was supported, in part, by an incentive grant to M.G.K. from the University of Louisville, College of Arts and Sciences and by an operating grant OGP9600 from the Natural Sciences and Engineering Research Council (NSERC) of Canada to P.C.L.

## Literature Cited

Allgood, G. S., and J. J. Perry. 1986. Characterization of a manganese-containing catalase from the obligate thermophile *Thermophilum album*. J. Bacteriol. **168**:563–567.

Almarsson, O., A. Sinha, E. Gopinath, and T. C. Bruice. 1993. Mechanism of one-electron oxidation of NAD(P)H and function of NADPH bound to catalase. J. Am. Chem. Soc. **115**:7093–7102.

Ames, B. N. 1995. Understanding the causes of aging and cancer. Microbiologia **11**:305–308.

Amo, T., H. Atomi, and T. Imanaka. 2002. Unique presence of a manganese catalase in a hyperthermophilic Archaeon, *Pyrobaculum calidifontis* VA1. J. Bacteriol. **184**:3305–3312.

Andersson, J. O., and A. J. Roger. 2002. Evolutionary analyses of the small subunit of glutamate synthase: gene order conservation, gene fusions, and Prokaryote-to-Eukaryote lateral gene transfers. Eukaryotic Cell **1**:304–310.

Bansal, A. K., and T. E. Meyer. 2002. Evolutionary analysis by whole-genome comparisons. J. Bacteriol. **184**:2260–2272.

Carpena, X., M. Soriano, M. G. Klotz, L. Donald, I. Fita, and P. C. Loewen. 2002. Structure of the clade I catalase, CatF of *Pseudomonas syringae*, at 1.8 Å resolution. Proteins **50**:423–436.

Cavalier-Smith, T. 2002*a*. The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. Int. J. Syst. Evol. Microbiol. **52**:7–76.

———. 2002*b*. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int. J. Syst. Evol. Microbiol. **52**:297–354.

Deppenmeier, U., A. Johann, T. Hartsch et al. (22 co-authors). 2002. The genome of *Methanosarcina mazei*: Evidence for lateral gene transfer between Bacteria and Archaea. J. Mol. Microbiol. Biotechnol. **4**:453–461.

Doolittle, W. F. 1999. Phylogenetic classification and the universal tree. Science **284**:2124–2129.

Embley, T. M., and W. Martin. 1998. A hydrogen-producing mitochondrion. Nature **392**:37–41.

Epstein, P. N., P. A. Overbeek, and A. N. Means. 1989. Calmodulin-induced early-onset diabetes in transgenic mice. Cell **58**:1067–1073.

Faguy, D. M., and W. F. Doolittle. 2000. Horizontal transfer of catalase-peroxidase genes between Archaea and pathgogenic bacteria. Trends Genet. **16**:196–197.

Felsenstein, J. 1994. PHYLIP (phylogeny inference package). Version 3.5. User manual. Washington University, St. Louis, Mo.

Frugoli, J. A., M. A. McPeek, T. L. Thomas, and C. R. McClung. 1998. Intron loss and gain during evolution of the catalase gene family in angiosperms. Genetics **149**:355–365.

Frugoli, J. A., H. H. Zhong, M. L. Nuccio, P. McCourt, M. A. McPeek, T. L. Thomas, and C. R. McClung. 1996. Catalase is encoded by a multigene family in *Arabidopsis thaliana* (L.) Heynh. Plant Physiol. **112**:327–336.

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

Guan, L., and J. G. Scandalios. 1996. Molecular evolution of maize catalases and their relationship to other eukaryotic and prokaryotic catalases. J. Mol. Evol. **42**:570–579.

Gupta, R. S. 1998. Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukaryotes. Microbiol. Mol. Biol. Rev. **62**:1435–1491.

Hartman, H., and A. Fedorov. 2002. The origin of the eukaryotic cell: a genomic investigation. Proc. Natl. Acad. Sci. USA **99**: 1420–1425.

Hayman, G. T., and P. L. Bolen. 1993. Movement of shuttle plasmids from *Escherichia coli* into yeasts other than *Saccharomyces cerevisiae* using trans-kingdom conjugation. Plasmid **30**:251–257.

Henkle-Dührsen, K., V. H. O. Eckelt, G. Wildenburg, M. Blaxter, and R. D. Walter. 1998. Gene structure, activity and localization of a catalase from intracellular bacteria in *Onchocerca volvulus*. Mol. Biochem. Parasitol. **96**:69–81.

Henze, K., and W. Martin. 2001. How do mitochondrial genes get into the nucleus? Trends Genet. **17**.

Hernandez-Ruiz, J., M. B. Arnao, A. N. P. Hiner, F. Garcia-Canovas, and M. Acosta. 2001. Catalase-like activity of horseradish peroxidase: relationship to enzyme inactivation by $H_2O_2$. Biochem. J. **354**:107–114.

Hillar, A., P. Nicholls, J. L. Switala, and P. C. Loewen. 1994. NADPH binding and control of catalase compound II formation: comparison of bovine, yeast and *Escherichia coli* enzymes. Biochem. J. **300**:531–539.

Hiner, A. N. P., J. Hernandez-Ruiz, G. A. Williams, M. B. Arnao, F. Garcia-Canovas, and M. Acosta. 2001. Catalase-like oxygen production by horseradish peroxidase must predominantly be an enzyme-catalyzed reaction. Arch. Biochem. Biophys. **392**:295–302.

Imlay, J. A., S. M. Chin, and S. Linn. 1988. Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. Science **240**:640–642.

Ishikawa, T., Y. Ohta, T. Takeda, S. Shigeoka, and M. Nishikimi. 1998. Increased cellular resistance to oxidative stress by expression of cyanobacterium catalase-peroxidase in animal cells. FEBS Lett. **426**:221–224.

Jespersen, H. M., V. H. Kjaersgård, L. Østergaard, and K. G. Welinder. 1997. From sequence analysis of three novel ascorbate peroxidases from *Arabidopsis thaliana* to structure, function and evolution of seven types of ascorbate peroxidase. Biochem. J. **326**:305–310.

Johnson, C. H., M. G. Klotz, J. L. York, V. Kruft, and J. E. McEwen. 2002. Redundancy, phylogeny and differential expression of *Histoplasma capsulatum* catalases. Microbiology **148**:1129–1142.

Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8**:275–282.

Jones, P., and I. Wilson. 1978. Catalases and iron complexes with catalase-like properties, Vol. 7. Marcel Dekker, New York.

Kang, Y. J., Y. Chen, and P. N. Epstein. 1996. Suppression of doxorubicin cardiotoxicity by overexpression of catalase in the heart of transgenic mice. J. Biol Chem. **271**:12610–12616.

Kawasaki, L., and J. Aguirre. 2001. Multiple catalase genes are differentially regulated in *Aspergillus nidulans*. J. Bacteriol. **183**:1434–1440.

Kawasaki, L., D. Wysong, R. Diamond, and J. Aguirre. 1997. Two divergent catalase genes are differentially regulated during *Aspergillus nidulans* development and oxidative stress. J. Bacteriol. **179**:3284–3292.

Kim, J. A., Z. Sha, and J. E. Mayfield. 2000. Regulation of *Brucella abortus* catalase. Infect. Immun. **68**:3861–3866.

Kirkman, H. N., and G. F. Gaetani. 1984. Catalase: a tetrameric enzyme with four tightly bound molecules of NADPH. Proc. Natl. Acad. Sci. USA **81**:4343–4347.

Klotz, M., G. Klassen, and P. Loewen. 1997. Phylogenetic relationships among prokaryotic and eukaryotic catalases. Mol. Biol. Evol. **14**:951–958.

Kono, Y., and I. Fridovich. 1983. Isolation and characterization of the pseudocatalase of *Lactobacillus plantarum*: a new manganese containing enzyme. J. Biol. Chem. **258**:6015–6019.

Lawrence, J. G., R. W. Hendrix, and S. Casjens. 2001. Where are the pseudogenes in bacterial genomes? Trends Microbiol. **9**:535–540.

Loew, O. 1900. Physiological studies of Connecticut leaf tobacco. U.S. Dept of Agri Repts. **65**:5–57.

Loewen, P. C. 1997. Bacterial catalases. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Loewen, P. C., M. G. Klotz, and D. J. Hassett. 2000. Catalase—an ''old'' enzyme that continues to surprise us. ASM News **66**:76–82.

Ma, J. F., U. A. Ochsner, M. G. Klotz, V. K. Nanayakkara, M. L. Howell, Z. Johnson, J. E. Posey, M. L. Vasil, J. J. Monaco, and D. J. Hassett. 1999. Bacterioferritin A modulates catalase A (KatA) activity and resistance to hydrogen peroxide in *Pseudomonas aeruginosa*. J. Bacteriol. **181**:3730–3742.

Makarova, K. S., L. Aravind, Y. I. Wolf, R. L. Tatusov, K. W. Minton, E. V. Koonin, and M. J. Daly. 2001. Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics. Microbiol. Mol. Biol. Rev. **65**:44–79.

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

Mayfield, J. E., and M. R. Duvall. 1996. Anomalous phylogenies based on bacterial catalase gene sequences. J. Mol. Evol. **42**:469–471.

Mayr, E. 1998. Two empires or three? Proc. Natl. Acad. Sci. USA **95**:9720–9723.

McCord, J. M., J. Keele, B.B., and I. Fridovich. 1971. An enzyme-based theory of obligate anaerobiosis: the physiolog-

ical function of superoxide dismutase. Proc. Natl. Acad. Sci. USA **68**:1024–1027.

Müller, M., and W. Martin. 1999. The genome of *Rickettsia prowazekii* and some thoughts on the origin of mitochondria and hydrogenosomes. Bioessays **21**:377–381.

Nesbø, C. L., S. L'Haridon, K. O. Stetter, and W. F. Doolittle. 2001. Phylogenetic analysis of two ''archaeal'' genes in *Thermotoga maritima* reveal multiple transfers berween Archaea and bacteria. Mol. Biol. Evol. **18**:362–375.

Nicholls, P., I. Fita, and P. C. Loewen. 2001. Enzymology and structure of catalases. Adv. Inorg. Chem. **51**:51–106.

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*. Eukaryotic Cell **1**:181–190.

Ochman, H., J. G. Lawrence, and A. Groisman. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405**:299–304.

Olsen, G. J., C. R. Woese, and R. Overbeek. 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. **177**:1–6.

Richter, H. E., and P. C. Loewen. 1982. Rapid inactivation of bacteriophage T7 by ascorbic acid is repairable. Biochim. Biophys. Acta **697**:25–30.

Rotte, C., K. Henze, M. Müller, and W. Martin. 2000. Origins of hydrogenosomes and mitochondria. Curr. Opin. Microbiol. **3**:481–486.

Scandalios, J. G., L. Guan, and A. N. Polidoros. 1997. Catalase in plants: gene structure, properties, regulation, and expression. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schubert, J., and J. W. Wilmer. 1991. Does hydrogen peroxide exist ''free' in biological systems? Free Radic. Biol. Med. **11**:545–555.

Seaver, L. C., and J. A. Imlay. 2001*a*. Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. J. Bacteriol. **183**:7173–7181.

———. 2001*b*. Hydrogen peroxide fluxes and compartmentalization inside growing *Escherichia coli*. J. Bacteriol. **183**:7182–7189.

Sprague, J., G.F. 1991. Genetic exchange between kingdoms. Curr. Opin. Genet. Dev. **1**:530–533.

Storz, G., and J. A. Imlay. 1999. Oxidative stress. Curr. Opin. Microbiol. **2**:188–194.

Sumner, J. B., and A. L. Dounce. 1937. Crystalline catalase. J. Biol. Chem. **121**:417–424.

Sun, W., T. A. Kadima, M. A. Pickard, and H. B. Dunford. 1994. Catalase activity of chloroperoxidase and its interaction with peroxidase activity. Biochem. Cell. Biol. **72**:321–331.

Sutherland, M. W. 1991. The generation of oxygen radicals during host plant responses to infection. Mol. Physiol. Plant Pathol. **39**:79–93.

Switala, J., J. O. O'Neil, and P. C. Loewen. 1999. Catalase HPII from *Escherichia coli* exhibits enhanced resistance to denaturation. Biochemistry **38**:3895–901.

Swofford, D. L. 1999. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0.b8. Sinauer Associates, Sunderland, Mass.

Thomas, J. A., D. R. Morris, and L. P. Hager. 1970. Chloroperoxidase. VII. Classical peroxidatic, catalatic and halogenating forms of the enzyme. J. Biol. Chem. **245**:3129–3134.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**:4876–4882.

Waldo, G. S., R. M. Fronko, and J. E. Penner-Hahn. 1991. Inactivation and reactivation of manganese catalase: oxidation-state assignments using X-ray absorption spectroscopy. Biochemistry **30**:10486–10490.

Welinder, K. G. 1991. Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. Biochim. Biophys. Acta **1080**:215–220.

———. 1992. Superfamily of plant, fungal and bacterial peroxidases. Curr. Opin. Struct. Biol. **2**:388–393.

White, O., J. A. Eisen, J. F. Heidelberg et al. (33 co-authors). 1999. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. Science **286**:1571–1577.

Whittaker, M. M., V. V. Barynin, S. V. Antonyuk, and J. W. Whittaker. 1999. The oxidized (3,3) state of manganese catalase: comparison of enzymes from *Thermus thermophilus* and *Lactobacillus pantarum*. Biochemistry **38**:9126–9136.

Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. **51**: 221–271.

Wolf, Y. I., I. B. Rogozin, N. V. Grishin, and E. V. Koonin. 2002. Genome trees and the tree of life. Trends Genet. **18**: 472–478.

Zamocky, M., and F. Koller. 1999. Understanding the structure and function of catalases: clues from molecular evolution and in vitro mutagenesis. Prog. Biophys. Mol. Biol. **72**:19–66.

Zamocky, M., G. Regelsberger, C. Jakopitsch, and C. Obinger. 2001. The molecular peculiarities of catalase-peroxidases. FEBS Lett. **492**:177–182.

Anthony Dean, Associate Editor

Accepted March 11, 2003