A large scale, multiple genome comparison of acidophilic Archaea (pH ≤ 5.0) extends our understanding of oxidative stress responses in polyextreme environments

Gonzalo Neira¹, Eva Vergara¹, Diego Cortez¹ and David S. Holmes* ¹,²
¹Center for Bioinformatics and Genome Biology, Fundación Ciencia & Vida, Santiago, Chile
²Universidad San Sebastián, Santiago, Chile
*Correspondence: dsholmes2000@yahoo.com
**Abstract**

Acidophilic Archaea thrive in anaerobic and aerobic low pH environments (<pH 5) rich in dissolved heavy metals that exacerbate stress caused by the production of reactive oxygen species (ROS) such as hydrogen peroxide (H₂O₂), hydroxyl radical (‘OH) and superoxide (O₂⁻). ROS react with lipids, proteins and nucleic acids causing oxidative stress and damage that can lead to cell death. Herein, genes and mechanisms potentially involved in ROS mitigation are predicted in over 200 genomes of acidophilic Archaea with sequenced genomes. These organisms can be subjected to simultaneous multiple stresses such as high temperature, high salinity, low pH and high heavy metal loads. Some of the topics addressed include: (1) the phylogenomic distribution of these genes and what can this tell us about the evolution of these mechanisms in acidophilic Archaea; (2) key differences in genes and mechanisms used by acidophilic versus non-acidophilic Archaea and between acidophilic Archaea and acidophilic Bacteria and (3) how comparative genomic analysis predicts novel genes or pathways involved in oxidative stress responses in Archaea and possible Horizontal Gene Transfer (HGT) events.

**Keywords:** Comparative Genomics, Catalase, Peroxiredoxin, Superoxide dismutase (SOD), Superoxide reductase (SOR), Rubrerythrin, Antioxidant enzymes, Extremophiles, Oxidative stress, Reactive oxygen species (ROS).
1. Introduction

Reactive oxygen species (ROS) are formed as by-products of aerobic metabolism including, for example, Fenton chemistry [1,2] and can involve molecules such as the superoxide anion (O$_2^-$), hydrogen peroxide (H$_2$O$_2$), and the hydroxyl radical (OH•). The accumulation of these molecules can lead to oxidative stress and produce damage to many cellular macromolecules such as DNA, RNA and proteins that, in turn can impair the function of important cellular structures such as membranes [3–6]. To mitigate these possible damages, different mechanisms have been developed by prokaryotes. Examples are superoxide scavengers such as superoxide dismutase (SOD) and superoxide reductase (SOR) that can diminish the quantity of superoxide anion [7] and a number of peroxide scavengers like peroxiredoxins (Prx) and catalases as examples that reduce the levels of hydrogen peroxide [8].

Whereas extensive research has been carried out on oxidative stress mechanisms in the bacterial domain [9,10], including the report of an inventory of genes across the whole domain [11], less is known about stress responses in Archaea. We aim to identify the potential mechanisms of oxidative stress response in extremely acidophilic Archaea and to analyze the phylogenetic distribution of predicted response mechanisms across the domain.

Since Archaea have been proposed to be linked with early Earth environments [12,13], their study could provide useful insight into the evolution of oxidative stress responses. Early evolution most likely proceeded in anaerobic conditions [14] but Archaea could be exposed to “whiffs” of oxygen that could lead to the need of the first oxidative stress response mechanisms [15,16], then additional mechanisms could have evolved as they transitioned from the anaerobic world to an aerobic atmosphere product of the great oxidation event (GOE) 2.4 billion year ago [17].

The archaeal domain comprises a wide range of organisms, many of which thrive in environments that are considered polyextreme, such as extremely high temperatures, low and high pHs and high salinity conditions [18–20]. Within these groups, stress response mechanisms have been described for hyperthermophiles [21] that constitute the majority of all identified Archaeal organisms. Acidophiles, are commonly defined as organisms that grow optimally at pH lower than 5 [22,23]. Another useful classification includes the categories of moderate acidophiles and extreme acidophiles, with the first group including organisms that grow optimally between pH 3 and pH 5 [24], and the second emphasizes organisms that grow optimally at pH values lower than 3 [25] as the environmental constraints that they face are particular to this pH range [26].

One particularly challenging constraint that acidophiles face is the unusually high concentration of dissolved metals and metalloids in their environments including Pb, Zn, Cu, Ni, Cr, Cd, Mn, Mo, and Fe [27–30], where the concentration of bioavailable Fe(II) in econiches at pH 3 can be 18 orders of magnitude greater than that encountered in circumneutral environments [31].

Previous studies of oxidative stress responses of acidophilic Archaea demonstrated the presence of several mechanisms that have also been reported for other Archaea, such as superoxide dismutase, peroxiredoxin and ruberythrin [32–34]. Interestingly, a specific alkyl hydroperoxide reductase has been described in some thermoacidophilic archaea, such as the strict aerobic, hyper-acidophilic (<pH1) Picrophilus torridus (phylum Euryarchaeota) that has not been detected in non-acidophilic Archaea [35].
However, despite this progress, a large scale, more comprehensive understanding of stress-associated mechanisms in acidophilic Archaea has not been undertaken. Our report begins to fill in this lacuna in our knowledge, providing a more extensive inventory of oxidative stress responses and helping to identify mechanisms provoked by multiple simultaneous stresses such as high temperature, extreme acidity, and high metal loads. Since Archaea have been proposed to root at the base of the Tree of Life, their study could provide useful insight into the evolution of oxidative stress responses on exposure to regular oxygen intrusion as the world transitioned from an anaerobic to aerobic atmosphere.

2. Materials and Methods

2.1 Genome collection

A dataset comprising 234 acidophilic (optimal growth <pH 5) archaeal genomes was obtained from the database AciDB [36]. All genomes were downloaded from the National Center for Biotechnology Information (NCBI) and the Joint Genome Institute (JGI). Only high-quality genomes (draft or complete) were considered for the analyses, defined as a >80% completeness and <5% contamination calculated by CheckM v1.0.12 with the standard lineage workflow [37]. Existing genome annotations in Refseq [38] were used when available and, for genomes with no annotation, Prokka v1.13.3 was used to identify genomic features [39]. Growth condition metadata for each genome was obtained from AciDB using the optimum values for pH and temperature. Oxygen requirement information was collected from bibliographic research associated with each species. If the data for a specific species was not available, it was extrapolated from the genus description. This leads to a classification of all organisms into aerobic, anaerobic, or facultative. A summary of genomic information is provided in Table S1.

2.2 Phylogeny

Phylogenetic reconstruction was performed using Phylophlan 3.0 using the provided dataset of 400 conserved proteins [40]. The parameters were set as following: diversity as high (as recommended for phylum wide phylogenies); marker search performed using diamond (v2.0.2) [41]; sequence alignment by MAFFT (v7.310) with the L-INSI iterative refinement option [42] and alignment trimming by trimal (v1.2) [43]. Tree construction was performed using IQTREE v1.6.1. The best evolutionary model was identified by ModelFinder according to the Bayesian (BIC) and Akaike information criterion (AIC) [44]. To assess the support of each branch, the non-parametric ultrafast bootstrap method (with 1000 replicates) was used [45]. Phylogenetic trees were visualized and annotated using iTOL [46]. A species clustering analysis was performed using Fast-ANI with a value of 95% identity to classify genomes as the same or different species [47].

2.3 Functional annotation and orthologous groups identification

A set of 52 proteins associated with oxidative stress responses was identified from the literature. Functional annotations based on Pfam [48] and Interprofamilies (domain and superfamily) [49] for each protein were obtained (Table S2). All genomes in the study were functionally annotated using the InterProScan pipeline with default parameters [50]. Orthogroups defined as proteins that have evolved from a common ancestor and include both orthologs (homologous protein from two species) and paralogs (a set of proteins that have a common origin in the same genome) were identified using OrthoFinder v2.3.3 [51]. Protein sequences were compared using Diamond in an all-versus-all search and orthogroups were inferred [41]. DendroBLAST [52] was used to generate unrooted gene trees for each of the
identified orthogroups. MAFFT was used for multiple sequence alignments and FastTree [53] was used for the tree inference. All other steps were performed with default parameters. An in-house Python script was used to identify oxidative stress response proteins in the acidophilic genomes based on the functional annotation. Pfam annotations provided a broad scope view of candidates which were subsequently hand curated by InterPro family and domain annotation to obtain additional specific information. Subcellular locations were assigned to predicted proteins and classified as cytoplasmatic, inner membrane, exported, outer membrane or periplasmic using PSORTb v3.0 [54] and accompanied by signal peptide identification performed by SignalP v5.0 [55].

2.4 Evolutionary trajectory analysis

Phylogenetic trees were constructed using the 50 best hits from a blastp search against the non-redundant database from all identified proteins in their respective orthogroups. Proteins from the blastp search were downloaded using the Batch Entrez web tool in NCBI [56]. Sequence alignments were performed using MAFFT and were visualized with Aliview v1.2.6 [57] to identify the conservation of key amino acids when needed. The phylogenetic tree was then constructed using IQTREE with an ultrafast bootstrap of 1000 replicates. Genome neighborhoods were analyzed using Gene Graphics with a region size of 5000bp [58].

3. Results

3.1 Phylogeny and Species clustering

High quality draft and complete genomes from acidophilic Archaea were obtained from AciDB [36] with a total number of 180 different genomes that fulfill the over 80% completeness and less than 5% of contamination defined by CheckM [37]. Of these genomes, 47 are complete and 133 are permanent drafts (Table S1). A phylogenetic tree was inferred using a set of 400 conserved markers as explained in material and methods. Three organisms from the ARMAN clade (Ca. Mancarchaeum acidophilum Mia14, Thermoplasmales archaeon A_DKE and Ca. Micrarchaeum sp. AZ1) were excluded from the tree inference as these organisms lack the majority of the conserved markers. Four different phyla are represented in the phylogenetic tree Crenarchaeota, Euryarchaeota, Marsarchaeota and Thaumarchaeota (Figure 1). At the genus level, Crenarchaeota and Euryarchaeota are the most diverse phyla in the dataset with ten and six different genera respectively. The species clustering analysis based on ANI values reveals that Cuniculiplasma divulgatum is the Euryarchaeota with the most genomes associated (19 in total), as a series of genomes annotated as “Thermoplasmales archaeon” were identified as part of this species. Sulfolobus acidocaldarius from Crenarchaeota has the most genomes of any species with 54 genomes associated with it. These results are summarized in the Table S3. The lack of markers in the ARMAN clade organisms is not surprising as this clade is characterized by their small genome sizes derived from their symbiotic lifestyle [59]. The species clustering highlights the importance for the development of new standards in the naming of prokaryotic genomes [60,61] that are not from isolated sources, as the adoption of higher rank taxonomic names could lead to confusion in subsequent analysis.
Figure 1. Phylogenetic tree reconstruction of acidophilic Archaea with genomes available in AciDB using Phylophlan 400 markers. Branches are collapsed at genus level, except for *Sulfobulbus* which is collapsed at species level. Colors represent different phyla associations with yellow for Crenarchaeota, red for Thaumarchaeota, green for Marsarchaeota and blue for Euryarchaeota. Bootstrap values over 60% are represented by black dots. Abbreviations: Uncultured (Un.); Candidatus (Ca.). Scale bar represent 0.1 amino acid substitution per site.

3.2 Superoxide defense proteins

Superoxide anion (O$_2^•$) is a highly reactive species that can cause severe damage inside the cell. The first system used in the process of removing O$_2^•$ is superoxide dismutase (SOD) that produces hydrogen peroxide and oxygen. Superoxide dismutase is classified in four different categories depending on the cofactor used (Mn, Fe, Cu/Zn and Ni). The second system is associated with the superoxide reductase (SOR) which produces hydrogen peroxide and protons in the reduction of superoxide. Details on how both mechanisms work are reviewed by Sheng and colleagues [62].

3.2.1 Superoxide dismutase (SOD)

We found a wide distribution of Fe-SOD throughout several acidophilic archaeal genera from all phyla in study, despite being previously identified in only a few organisms such as...
Saccharolobus solfataricus [63], Ferroplasma acidiphilum [64] and Metallosphaera sedula [65] (Figure S1). The abundance of this category of SOD (Fe-SOD type) could be of evolutionary importance given that the iron containing isoform has been suggested to be an ancient type of SOD [7]. The higher availability of iron in low pH environments, together with the notion that Archaea are ancestral organisms [12], highlights these acidophiles as prime candidates to be the original source of one of the most relevant oxidative stress mechanisms used across all domains of life for evolutionary studies and biotechnological applications. Phylogenetic analysis of the predicted sequences was performed to study the evolution of SOD in these organisms. The sequences were divided into four different clades, two clades that grouped most of the sequences associated with the phylum Crenarchaeota and Euryarchaeota and two smaller clades formed by sequences of organisms from the phylum Marsarchaeota and Thaumarchaeota, the latter including outliers of the Crenarchaeota phylum from the Acidilobus and Caldiphaera genera. The distribution in each of the separated clades appears to be associated with phylogenetic relatedness, as each predicted protein branch is formed mainly by organisms from the same genus with a clear distinction between the sequences from Euryarchaeota and those from the phyla Crenarchaeota, Thaumarchaeota and Marsarchaeota which are phylogenetically associated as part of the TACK superphylum [66] (Figure 2).

An exception to the phylogenetic association can be observed, specifically, Ca. Micrarchaeum sp. AZ1 and Ca. Mancarchaeum acidiphilum Mia14 sequences, which are both inside the Crenarchaeota clade with the former close to Thermoproteus and the latter to Acidianus (Figure 2, red asterisks). This difference could be explained by the fact that Micrarchaeota organisms have reduced genome size and limited metabolic capabilities, relying on other members of the community to supply these deficiencies by a symbiotic approach [67]. The observation that all SODs from these organisms are closely related with organisms that thrive in the same environments suggest that this oxidative stress mechanism probably was gained by HGT, a process which has been observed to be essential in the evolution of other members of the DPANN clade [68].
Figure 2. Superoxide dismutase phylogenetic tree from predicted protein sequences for gene sodA. Two Candidatus Micrarchaeota leaves are marked with red asterisks which are in the bigger clade of Crenarchaeota. Colors represent different phyla associations with yellow for Crenarchaeota, red for Thaumarchaeota, green for Marsarchaeota and blue for Euryarchaeota. Scale bar represents 1 amino acid substitution per site.

3.2.2 Superoxide reductase (SOR)

The superoxide reductase (SOR) was predicted only in the genus Aciduliprofundum, which is the only genus that lacks any SOD sequence. Phylogenetic analysis of the sequences identified in Aciduliprofundum and archaeal proteins from SORGoDB [69] shows a clade formed with other strict anaerobes from Euryarchaeota (non-acidophiles), suggesting that vertical transmission could play a key role in the evolution of this protein (Figure S2). Previous studies have shown that SOR is mainly found in anaerobic organisms [70] which is the case of Aciduliprofundum, increasing the interest for other anaerobic organisms in the study (for example Vulcanisaea and Acidilobus) which use Fe-SOD instead as main medium to get rid of $O_2^\cdot$. Rubredoxin (Rb) has been identified as an electron donor for the reduction of superoxide via SOR in Pyrococcus furiosus (also from the Euryarchaeota phylum) [71] and was also identified only in Aciduliprofundum, supporting the idea that this organism is the only one in dataset able to reduce superoxide. One possible explanation for this contradiction with previous studies that claim that the main mechanism of coping with superoxide in anaerobic organisms is the use of SOR (avoiding the production of $O_2$ resulting from SOD) are the special...
conditions in which acidophiles thrive. SOR in the reduction of O2• produces extra protons (which instead are not produced by SOD), leading to additional stress for acidophilic organisms that need to maintain a neutral intracellular pH and balance the extremely high concentration of protons in the environment. This produces a disadvantage to the specific use of SOR in low pH environments that could favor the use of SOD, which has been previously observed as one of the main oxidative stress response mechanisms in acidophiles [72], resulting in the loss of the SOR gene through adaptation to the harsh environmental conditions.

3.3 Peroxide Scavengers

3.3.1 Catalase

Catalases convert hydrogen peroxide into water and oxygen and work as a common oxidative stress mechanism in organisms from all domains of life [11]. In acidophilic Archaea, only genomes from *Thermoplasmatales* and *Aciduliprofundum* sp. MAR08-339 contain predicted catalase-peroxidase (with the same functional annotation as KatG). Phylogenetic analyses with proteins from NCBI show that the sequences from *Thermoplasmatales* form a clade close to extremely acidophilic bacteria from the genus *Methylococcus* which are also moderate thermophilic organisms. In the case of the KatG identified for *Aciduliprofundum* sp. MAR08-33, this forms a separated clade with other Euryarchaeota organisms (not acidophiles) and some Deltaproteobacteria (Figure S3). The low presence in Archaea has been reported previously, as organisms that have SOD usually do not have catalases [21] and as previously discussed we identified SOD in most acidophilic genomes with the exception of *Aciduliprofundum*. The result of mixed origin clades in the phylogenetic reconstruction where Archaea and Bacteria share the same clade suggests that HGT between domains could play a major role in the evolution of catalases. This goes in line with studies that show that the phylogeny of catalases is complex and often does not follow a pattern of vertical inheritance [73,74].

3.3.2 Osmotically inducible protein C (OsmC)

Another protein with peroxide scavenging activity is the osmotically inducible protein C (OsmC) which has been shown to be involved in the elimination of organic peroxides in Bacteria [75]. OsmC was found to be ubiquitous in acidophilic Archaea with most organisms in Crenarchaeota, Euryarchaeota and Marsarchaeota predicted to have one or more copies, even reaching 3 different copies in *Metallosphaera* organisms. Phylogenetic tree analysis of the identified sequences reveals a clear division between the sequences that evolved from Euryarchaeota and Crenarchaeota. Nine different orthologous groups were formed by the predicted OsmC sequences. The orthogroups mostly match the clade as observed in Figure S4, where each orthogroup contains proteins from the same phylum except for OG2461 and OG3045 which also contain sequences from Marsarchaeota. Protein sequences from Marsarchaeota were observed to be distributed across clades with sequences from both Crenarchaeota and Euryarchaeota, forming separated clades suggesting HGT between these phyla. KAD1 in *Thermococcus kodakaraensis*, is an OsmC protein that can reduce inorganic peroxides as well as organic ones [76] and it is interesting to note that the different orthogroups identified across acidophilic Archaea could be related with functional distinction to the classic OsmC. Further experimental investigation is needed to unravel the functional activities of these different identified predicted proteins.

3.3.3 DNA-Binding Protein from Starved Cells
Another way to prevent the creation of ROS is sequestering iron, thus decreasing the Fenton reactions. A protein that fulfills this function is the DNA-Binding Protein from Starved Cells (Dps), a member from the ferritin-like superfamily which can also reduce hydrogen peroxide in the process of oxidation of free ferrous iron [77]. Dps in Archaea can be divided into two different functional annotations, the first one with the InterPro family IPR002177 group from the model protein for Dps in Bacteria and a second InterPro family IPR014490 domain that represents Dps-like proteins from Archaea first identified in *Saccharolobus solfataricus* [78]. Using InterPro family IPR002177 domain as a probe (Bacteria associated), we predicted Dps in Marsarchaeota and Sulfolobales MK5 as part of a single orthogroup. Unlike the previous case, the Dps-like family (Archaea associated) was identified in three different phyla, Crenarchaeota (*Metallosphaera, Saccharolobus* and *Sulfobolus*), Euryarchaeota (*Aciduliprofundum*) and Thaumarchaeota (*Ca. Nitrosotalea*). To study the possible evolutionary origin for the Dps protein, a phylogenetic tree was constructed using the best hits from a blastp search. The Dps phylogeny shows their closest sequences include other acidophiles coming from Acidobacteria and several Proteobacteria suggesting a likely cross domain HGT as the main evolutionary force for this protein in Archaea that could be associated with a specific eco-niche adaptation [79] (Figure 3).

**Figure 3.** Phylogenetic tree from predicted protein sequences of DNA-binding protein from starved cells (Dps). In yellow background are shown the protein sequences from acidophilic Archaea. In red letters are the sequences associated with acidophiles. Collapsed branches show the number of genomes in each clade in parenthesis after the names. Bar represents 1 amino acid substitution per site.

In contrast, Dps-like protein is widely found in Archaea from Crenarchaeota and Thaumarchaeota. The functionality of the Dps-like protein as an actual Dps protein (instead of other ferritin-like superfamily protein) could be inferred from the high similarity of the sequences with the functionally active protein previously identified in *Saccharolobus solfataricus* [78]. In addition, we could not identify any homologous proteins to either Dps or
Dps-like to be widely present in Euryarchaeota, suggesting that this mechanism could be replaced in the oxidative stress response by another protein of similar function.

### 3.3.4 Ruberythrin (Rbr)

Ruberythrin (Rbr, rbr), another member of the ferritin-like superfamily that consists of varied non-heme di-iron proteins, has an important role in oxidative stress response by reducing hydrogen peroxide [80]. Rbr was identified in almost all acidophilic Archaea both in aerobic and anaerobic organisms. The identified proteins with ruberythrin domains are annotated in several different orthogroups, forming distinct clades in the phylogenetic analysis. Five different domain architectures were identified across the different proteins. The first group (associated with OG0767) is characterized by a single and isolated Rbr domain with conserved di iron metal motifs, this group was identified across all major phyla in study. The second group (associated with OG0732) is characterized by a domain architecture of a full Rbr domain with di-iron metal motifs together with a VIT1 domain [81]. This group is also conserved across all major phyla in study. A third domain (associated with OG6126, OG7923 and OG3502) similar to the first group identified contains a small truncated Rbr domain at the C-terminal end and is exclusive to Euryarchaeota. The fourth group (associated with OG1026 and OG3704), presents a truncated Rbr domain with missing di-iron metal motifs and with the addition of Linocin M18 domain. The final domain architecture group (associated with OG1052) presents a small Rbr domain without any di-iron metal motifs. The presence of the di-iron metal motifs is important in the context of oxidative stress response, as these are essential in the action of Rbr as peroxidase [82]. From these analyses, we propose that group 1 and 2 could be directly related with hydrogen peroxide reduction. In particular, the addition of the domain VIT1 in group 2 has been previously reported to aid in iron detoxification via transportation of the metals, a process that prevents oxidative stress via Fenton reaction products in plants, yeast, and fungi [83,84]. The presence of a Linocin M18 domain in group 3 could help to the formation of nano compartments that encapsulate peroxidases [85]. Previously, this variant was only identified in the extreme thermophile *Pyrococcus furiosus* [80], where the specific action has been hypothesized to be related with the hyperthermophilic trait of these organisms, in our analysis this variant appears not to be exclusive to hyperthermophiles; *Vulcanisaeta* and *Caldivirga* also have this domain architecture and grow at a more moderate temperature (45-50 ºC). (Figure 4).

**Figure 4.** Ruberythrin domain architecture in acidophilic Archaea. The phylogenetic tree is collapsed by clades representing specific protein domain architectures (groups 1-5). Specific orthogroups with phylum association are indicated by each clade with the number of sequences.
in parentheses. Rbr domains are represented in green, VIT1 domain in purple and Linocin_M18 in blue. Di iron metal motifs identified are represented as red markings on Rbr domains. Scale represents protein length in residues.

3.3.5 Peroxiredoxins

Peroxiredoxins (Prx) have been identified as the major mechanism of hydrogen peroxide elimination in Archaea as reviewed in Pedone et al. 2020. These enzymes have been classified in 6 different subfamilies (AhpC-Prx1, Bcp-PrxQ, Tpx, Prx5, Prx6 and AhpE) using structural information near the active site [86]. A previous study in *Saccharolobus solfataricus* characterized four different peroxiredoxins from these organisms, three of them (described as Bcp1-3-4) belonging to the Bcp-PrxQ subfamily and one (Bcp2) to the Prx6 [87]. All these sequences have their crystal structure solved, providing experimentally proved and classified examples of peroxiredoxins in Archaea [88]. Peroxiredoxins are widely distributed across acidophilic Archaea with a varying number of copies depending on the organisms (number of copies mostly conserved at genus level). The phylogenetic distribution of the different enzymes identified across the acidophilic organisms is presented in Figure 5a showing four larger clades that include the proteins characterized in *Saccharolobus* (Prx6/Bcp2, Bcp1, Bcp3 and Bcp4) and a fifth new clade not previously reported in acidophilic Archaea with sequences corresponding to the Prx1 subfamily. The clades formed by Bcp1 (PrxQ family) and Bcp 2 (Prx6 family) sequences are widely distributed across organisms from all the phyla in study, which shows a high degree of conservation of both families of Prx across all acidophilic Archaea. In contrast, Bcp3 is exclusive to organisms of the Crenarchaeota phylum. The Bcp4 clade is divided into three different sub-clades with the first divergent clade showing a similar distribution to the Bcp3 clade with *Vulcanisaeta, Thermoproteus* and *Caldivirga*, a second clade containing sequences from Euryarchaeota (that includes the ARMAN organism *Thermoplasmatales archaeon B_DKE*) and finally a clade that include the rest of Crenarcheota and Marsarchaeota. The unique divergence of this clade in the phylogenetic tree alongside the identification of the specific motifs confirms the idea that this may be a Prx1 type previously not reported in acidophilic Archaea, and further experimental validation is needed to confirm this finding.
Figure 5. Peroxiredoxin (Prx) family proteins identified across acidophilic Archaea. (a) Phylogenetic tree of all identified Prx sequences. Branches are collapsed to phyla or genus level when it is possible, with the number of sequences that formed the clade in parentheses. In purple are the sequences from PrxQ type (bcp1-3-4 like), in blue Prx6 type (bcp2) and darker blue Prx1. Bar represents 1 amino acid substitution per site. (b) Multiple sequence alignment of the Prx1 sequences (N-terminal extract showed). Conserved motifs associated with the
characterization of Prx1 are marked in blue and shown in LOGOS graphic in the upper right corner.

3.3.6 Thioredoxin

The mechanism of action of thioredoxins (Trx, trxA) and thioredoxin reductases (TR, trxB) is similar to that of peroxiredoxins as they are needed to complete the redox cycle and reduce the disulfide bonds in Prx to continue with the peroxide elimination process. Our results show that most organisms with Trx have two copies, with both copies being part of the same orthogroup (paralogs) suggesting a shared evolutionary story that provides clues of a gene duplication event that could have occurred at an ancestral level. An interesting finding is a hypothetical protein (defined by the orthogroup OG0923) that appears contiguous to trxA, and was identified in members of Crenarchaeota, Euryarchaeota and Marsarchaeota with high conservation (Figure S5). Sequences from this orthogroup were analyzed by subcellular localization prediction tools which predict these proteins are in either the cellular membrane or extracellular space (Table S4). This highly conserved hypothetical protein could be important for these organisms as proteins in the cellular membrane of acidophiles are exposed directly to the physicochemical constraints of a low pH environment. Experimental analyses of this kind of protein of unknown function could help us reveal new proteins or mechanisms of oxidative response that are unique to Archaea in extreme environments. Predicted TR were also identified widely across acidophilic Archaea as part of the same orthogroup, with most members in Acidianus, Saccharolobus, Sulfolobus and Marsarchaeota having an extra copy (Table S5). The presence of both Trx and TR in addition to different types of Prx suggests that the complete redox thiol cycle could play an important role in oxidative stress response in Archaea as has been previously demonstrated in Bacteria [91].

3.4 Methionine sulfoxide reductases

Another target sensitive to ROS oxidation is methionine, a sulfur-containing amino acid. The first stage of oxidation of this amino acid is the formation of methionine sulfoxide with two different stereospecific forms (Met-S-O and Met-R-O), process that could lead to the protein being aggregated and eliminated or even further oxidized to methionine sulfone (an irreversible state). MsrA can reduce Met-S-O and MsrB reduces the Met-R-O stereospecific form. The first protein, MsrA, was identified in several genera of the acidophilic Archaea. Previous studies have reported this to be mainly absent in hyperthermophiles [92]. In contrast, our results include hyperthermophiles such as Metallosphaera, Acidianus and Sulfolobus, all of which have optimal growth over 70°C, pointing out a previous underestimation of the abundance of MsrA in Archaea. The second protein, MsrB, was only identified in Thermoplasmatales archaeon and Cuniculiplasma in Euryarchaeota, Marsarchaeota and Thaumarchaeota, completely absent from Crenarchaeota. The protein fRMSR also can reduce the same isoform as MsrB, but in free amino acids (not as part of a protein) and is characterized by a GAF-like domain, identified in Picrophilus and Thermoplasm, both organisms that also have MsrB. All these proteins share similarity with the fRMSR from Thermoplasma acidophilum, where structural and biochemical analysis has been performed to confirm its functional activity [93]. Given the high degree of similarity of the sequences, we suggest that all other copies in the same orthogroup could also be functionally active. Finally, both stereospecific forms can also be reduced by the molybdopterin-dependent sulfite oxidase family MrsP. We identified this protein in acidophiles from all phyla. Some organisms from Crenarchaeota (Saccharolobus, Metallosphaera, Vulcanisaeta and Thermocladium) have more than one copy, which could act as counterbalance to the fact they lack the other common mechanism for reducing Met-R-O. These results contrast with previous observations that found MrsP to be mostly absent in
Archaea [92]. MsrQ is commonly used as an electron donor associated with MsrP in Bacteria [94], however, it was not identified in any of the organisms under study, suggesting that an alternative pathway to obtain electrons could be in use by Archaea. This may have special relevance when considering the differences in cell wall components compared with Bacteria. An overview of the described system is shown in Figure 5.

**Figure 5.** Methionine sulfoxide reductases in acidophilic Archaea. Methionine is oxidized by ROS to two different stereospecific forms, where Met-S-O is reduced by MsrA, Met-R-O is reduced by MsrB and fRMsr at the periplasm and both forms could be reduced by MsrP obtaining electrons from an unknown source protein. The presence of the proteins in each phylum are represented by circles with the color purple for Micrarchaeota, pink for Thaumarchaeota, green for Marsarchaeota, blue for Euryarchaeota and yellow for Crenarchaeota.

### 3.5 General Discussion

Our study provides valuable clues for the prediction of the mechanisms used by acidophilic Archaea to cope with oxidative stress and represents one of the most extensive comparative genomics analyses on this topic to date. An overview of all the analyzed mechanisms identified is presented in Figure 6. A comparison of the predicted proteins with those from non-acidophilic Archaea, indicates that most of the mechanisms are conserved widely across the different acidophilic and no-acidophilic phyla in this study. However, there are some interesting deviations from this observation. Superoxide dismutase (SOD) constitutes an interesting case, where all the structurally described Fe-SOD are from acidophilic Archaea, while other non-acidophilic Archaea have either Mn-SOD or the cambialistic Mn/Fe-SOD. This difference could be associated with acidophiles maintaining the ancestral version of Fe-SOD [7], given the high availability of iron in acidic environments even in current times, while non-acidophilic Archaea prefer SOD that could use other elements as catalytic centers [95]. Peroxiredoxins are also predicted both in acidophilic and non-acidophilic Archaea, but one specific family identified in this study (Prx1) has not been previously described in the literature in acidophilic Archaea, which highlights the importance of the use of comparative genomics strategies to help unravel the distribution of different proteins across a big dataset of organisms. Lastly, a hypothetical protein described by OG0923 was identified to be highly conserved in the genomic context of Trx and to be restricted to acidophiles. OG0923 is predicted to be transported to the membrane. We hypothesize that it carries out a role in acid stress as membrane proteins are directly confronted by the extremely low pH of the environment. This
serves as an example of the use of bioinformatics in predicting potential new proteins associated with the oxidative stress response that could be particular to Archaea.

A notable difference that we detect between Archaea and Bacteria is that both aerobic and anaerobic representatives of acidophilic Archaea use SOD to deal with superoxide whereas anaerobic Bacteria use SOR [96]. The Archaeal preference for SOD could be related to the issue of the increased availability of Fe in acidophilic environments.

It has been observed that HGT events play an important role in adaptation and speciation in Archaea, and that many of these events imply the acquisition of genes from Bacteria (mostly acidophilic Bacteria) to Archaea [97,98]. The high similarity between the sequences from Archaea and acidophilic Bacteria also points out that the HGT event could be associated with specific niche adaptations that confer ecological advantage to these organisms. The importance of HGT events in acidophiles has been previously highlighted in several works [99–101]. Cross domain events of HGT were also predicted and could play a role in the evolution of catalases and Dps. In both cases of HGT, we identify a restricted presence in specific lineages of Archaea in contrast with what is found in Bacteria, where both proteins are widely distributed across the domain [102,103], which makes us suggest that the most probable case for direction of this HGT event is from Bacteria to Archaea.

The association of Archaea and early anoxic Earth environments and their deep rooting in the tree of life is commonly discussed in literature in the context of Archaea as candidates to harbor ancestral mechanisms and the possibility that they may illuminate the origin of different traits [104,105]. The wide conservation across organisms of superoxide dismutase, peroxiredoxin, and ruberythrin in both anaerobic and aerobic organisms, could suggest an ancestral origin of these proteins [82,106,107], and that they were evolved in similar environments to the early LUCA in response to local concentrations of oxygen before the GOE [108]. It has been previously reported that multiple mechanisms for oxidative stress response are found in hyperthermophiles [21] and we identified a similar pattern in acidophiles with no major differences across lifestyles (such as aerobic versus anaerobic or thermophiles versus mesophiles). Instead, the distribution of oxidative stress response mechanisms present in each organism seems to be largely determined by phylogeny. Interestingly, the same type of distribution was observed in an extensive study of the stress enzymes in Bacteria [11], showing that this could be a trend globally associated with oxidative stress response.
Figure 6. Summary of predicted oxidative stress response proteins in acidophilic Archaea. The heatmap depicts the presence of each protein as a percentage of the genomes it was identified in (see insert scale-bar). Genomes were grouped at the genus level. Colors correspond to phylum level with yellow for Crenarchaeota, red for Thaumarchaeota, green for Marsarchaeota and blue for Euryarchaeota.

4. Conclusions

Through comparative genomics of more than 200 genomes from acidophilic Archaea, we identified a wide range of mechanisms to cope with oxidative stress mostly conserved across all 5 different phyla showing no clear difference between aerobic and anaerobic lifestyles. Nonetheless, with this study we could unravel previously unidentified protein families, missed in individual studies, as the Prx1 family present in several acidophiles. On the other hand, the conservation of two proteins with iron coordinating centers, SOD and Rbr, across almost all organisms in study, suggests an early evolutionary origin that could have arisen in an anaerobic early Earth environments and be maintained in acidic environments. Furthermore, HGT events from Bacteria to Archaea, exemplified in the case of catalases and Dps highlight the importance of these events in prokaryotic adaptation to oxidative stress.

Supplementary Materials

The following are available online Figure S1: Multiple sequence alignment of Crenarchaeota Fe-SOD, Figure S2: Unrooted phylogenetic tree constructed from the predicted amino acid sequences of SorA in *Aciduliprofundum* and Archaeal sequences from sorGOdb, Figure S3:
Unrooted phylogenetic tree constructed from the predicted amino acid sequences of KatG in *Aciduliprofundum*, *Cuniculiplasma* and *Thermoplasmatales* and their best hits in the NCBI, Figure S4: Unrooted phylogenetic tree constructed from the predicted amino acid sequences of OsmC in the different acidophilic Archaea, Figure S5: Genomic context of *trxA* and the conserved hypothetical protein OG0923 in several representative genomes, Table S1: Genomic features for acidophilic genomes in the study, Table S2: Functional annotations used for the identification of each protein associated with oxidative stress response, Table S3: Species clustering analysis for each of the genomes in the study done by Fast-Ani, Table S4: Subcellular location prediction results for the hypothetical protein OG0923 and MsrP, Table S5: Orthogroups for all the identified oxidative stress response associated protein with the corresponding protein id for each of the genomes in the study.

**Author Contributions**

GN designed the research. GN and EV performed the research. GN, EV, DC and DH analyzed the data. GN and DH wrote the paper. All authors participated in the construction of the final manuscript. All authors read and approved the final manuscript.

**Funding**

Fondecyt 1181717 to DSH and Centro Ciencia & Vida, FB210008, Financiamiento Basal para Centros Científicos y Tecnológicos de Excelencia de ANID

**Institutional Review Board Statement**

Not applicable.

**Informed Consent Statement**

Not applicable.

**Data Availability Statement**

The data presented in this study are available in article.

**Conflicts of Interest**

The authors declare no conflict of interest.
References

1. Imlay, J.A. Pathways of Oxidative Damage. *Annual Review of Microbiology* 2003, 57, 395–418.
2. Apel, K.; Hirt, H. Reactive Oxygen Species: Metabolism, Oxidative Stress, and Signal Transduction. *Annual Review of Plant Biology* 2004, 55, 373–399.
3. Hemmami, T.; Parihar, M.S. REACTIVE OXYGEN SPECIES AND OXIDATIVE DNA DAMAGE; 1998; Vol. 42;.
4. Cabiscol, E.; Tamarit, J.; Ros, J. Oxidative Stress in Bacteria and Protein Damage by Reactive Oxygen Species. *International Microbiology* 2000, 3, 3–8.
5. Lushchak, V.I. Oxidative Stress and Mechanisms of Protection against It in Bacteria. *Biochemistry (Moscow)* 2001, 66, 476–489.
6. Su, L.J.; Zhang, J.H.; Gomez, H.; Murugan, R.; Hong, X.; Xu, D.; Jiang, F.; Peng, Z.Y. Reactive Oxygen Species-Induced Lipid Peroxidation in Apoptosis, Autophagy, and Ferroptosis. *Oxidative Medicine and Cellular Longevity* 2019, 2019.
7. Miller, A.F. Superoxide Dismutases: Ancient Enzymes and New Insights. *FEBS Letters* 2012, 586, 585–595.
8. Zámocký, M.; Gasselhuber, B.; Furtmüller, P.G.; Obinger, C. Molecular Evolution of Hydrogen Peroxide Degrading Enzymes. *Archives of Biochemistry and Biophysics* 2012, 525, 131–144.
9. Farr2, S.B.; Kogoma2, T. Oxidative Stress Responses in Escherichia Coli and Salmonella Typhimurium; 1991; Vol. 55;.
10. Ezraty, B.; Gennaris, A.; Barras, F.; Collet, J.F. Oxidative Stress, Protein Damage and Repair in Bacteria. *Nature Reviews Microbiology* 2017, 15, 385–396.
11. Johnson, L.A.; Hug, L.A. Distribution of Reactive Oxygen Species Defense Mechanisms across Domain Bacteria. *Free Radical Biology and Medicine* 2019, 140, 93–102.
12. Gribaldo, S.; Brochier-Armanet, C. The Origin and Evolution of Archaea: A State of the Art. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2006, 361, 1007–1022.
13. Kozubal, M.A.; Romine, M.; Jennings, R.D.; Jay, Z.J.; Tringe, S.G.; Rusch, D.B.; Beam, J.P.; McCue, L.A.; Inskemp, W.P. Geoarchaeota: A New Candidate Phylum in the Archaea from High-Temperature Acidic Iron Mats in Yellowstone National Park. *ISME Journal* 2013, 7, 622–634, doi:10.1038/ismej.2012.132.
14. Lane, N. The Evolution of Oxidative Stress. In *Principles of Free Radical Biomedicine* ; Nova Biomedical Books, 2011; Vol. 1, pp. 1–17 ISBN 9781612097732.
15. Anbar, A.D.; Duan, Y.; Lyons, T.W.; Arnold, G.L.; Kendall, B.; Creaser, R.A.; Kaufman, A.J.; Gordon, G.W.; Scott, C.; Garvin, J.; et al. A Whiff of Oxygen before the Great Oxidation Event? *Science* 2007, 317, 1903–1906, doi:10.1126/science.1140325.
16. Martins, M.C.; Romão, C. v.; Folgosa, F.; Borges, P.T.; Frazão, C.; Teixeira, M. How Superoxide Reductases and Flavodiiron Proteins Combat Oxidative Stress in Anaerobes. *Free Radical Biology and Medicine* 2019, 140, 36–60.
17. Fischer, W.W.; Hemp, J.; Valentine, J.S. How Did Life Survive Earth’s Great Oxygenation? *Current Opinion in Chemical Biology* 2016, 31, 166–178.
18. Stan-Lotter, H.; Fendrihan, S. Halophilic Archaea: Life with Desiccation, Radiation and Oligotrophy over Geological Times. *Life* 2015, 5, 1487–1496, doi:10.3390/life5031487.
19. Korzhenkov, A.A.; Toshchakov, S. v.; Bargiela, R.; Gibbard, H.; Ferrer, M.; Teplyuk, A. v.; Jones, D.L.; Kublanov, I. v.; Golyshin, P.N.; Golyshina, O. v. Archaea
Dominate the Microbial Community in an Ecosystem with Low-to-Moderate Temperature and Extreme Acidity. *Microbiome* 2019, 7, doi:10.1186/s40168-019-0623-8.

20. Blöchl, E.; Rachel, R.; Burggraf, S.; Hafenbradl, D.; Jannasch, H.W.; Stetter, K.O. Pyrolobus Fumarii, Gen. and Sp. Nov., Represents a Novel Group of Archaea, Extending the Upper Temperature Limit for Life to 113°C. *Extremophiles* 1997, doi:10.1007/s007920050010.

21. Pedone, E.; Fiorentino, G.; Bartolucci, S.; Limauro, D. Enzymatic Antioxidant Signatures in Hyperthermophilic Archaea. *Antioxidants* 2020, 9, 1–19.

22. Baker-Austin, C.; Dopson, M. Life in Acid: PH Homeostasis in Acidophiles. *Trends in Microbiology* 2007, 15, 165–171.

23. Dopson, M. Physiological and Phylogenetic Diversity of Acidophilic Bacteria. In *Acidophiles: Life in Extremely Acidic Environments*; Caister Academic Press, 2016; pp. 79–92.

24. Benison, K.C.; O’Neill, W.K.; Blain, D.; Hallsworth, J.E. Water Activities of Acid Brine Lakes Approach the Limit for Life. *Astrobiology* 2021, 21, 729–740, doi:10.1089/ast.2020.2334.

25. Johnson, D.B. Physiology and Ecology of Acidophilic Microorganisms. In *Physiology and Biochemistry of Extremophiles*; ASM Press: Washington, DC, USA, 2014; pp. 255–270.

26. Slonczewski, J.L.; Fujisawa, M.; Dopson, M.; Kruilwich, T.A. Cytoplasmic PH Measurement and Homeostasis in Bacteria and Archaea. *Advances in Microbial Physiology* 2009, 55.

27. Orellana, L.H.; Jerez, C.A. A Genomic Island Provides Acidithiobacillus Ferroxidans ATCC 53993 Additional Copper Resistance: A Possible Competitive Advantage. *Applied Microbiology and Biotechnology* 2011, 92, 761–767, doi:10.1007/s00253-011-3494-x.

28. Orell, A.; Remonsellez, F.; Arancibia, R.; Jerez, C.A. Molecular Characterization of Copper and Cadmium Resistance Determinants in the Biomining Thermoacidophilic Archaeon Sulfolobus Metallicus. *Archaea* 2013, doi:10.1155/2013/289236.

29. Dopson, M.; Holmes, D.S. Metal Resistance in Acidophilic Microorganisms and Its Significance for Biotechnologies. *Applied Microbiology and Biotechnology* 2014, 98, 8133–8144.

30. Król, A.; Mizerna, K.; Bożym, M. An Assessment of PH-Dependent Release and Mobility of Heavy Metals from Metallurgical Slag. *Journal of Hazardous Materials* 2020, 384, doi:10.1016/j.jhazmat.2019.121502.

31. Osorio, H.; Martínez, V.; Veloso, F.A.; Pedroso, I.; Valdés, J.; Jedlicki, E.; Holmes, D.S.; Quatrini, R. Iron Homeostasis Strategies in Acidophilic Iron Oxidizers: Studies in Acidithiobacillus and Leptospirillum. *Hydrometallurgy* 2008, 94, 175–179, doi:10.1016/j.hydromet.2008.05.038.

32. Jeon, S.J.; Ishikawa, K. Identification and Characterization of Thioredoxin and Thioredoxin Reductase from Aeropyrum Perinix K1. *European Journal of Biochemistry* 2002, 269, 5423–5430, doi:10.1046/j.1432-1033.2002.03231.x.

33. Amo, T.; Atomi, H.; Imanaka, T. Biochemical Properties and Regulated Gene Expression of the Superoxide Dismutase from the Facultatively Aerobic Hyperthermophile Pyrococcalum Calidifontis. *Journal of Bacteriology* 2003, 185, 6340–6347, doi:10.1128/JB.185.21.6340-6347.2003.

34. Dillard, B.D.; Demick, J.M.; Adams, M.W.W.; Lanzilotta, W.N. A Cryo-Crystallographic Time Course for Peroxide Reduction by Rubrerythrin from
35. Fütterer, O.; Angelov, A.; Liesegang, H.; Gottschalk, G.; Schleper, C.; Schepers, B.; Dock, C.; Antranikian, G.; Liebl, W. Genome Sequence of *Picrophilus Torridus* and Its Implications for Life around PH 0. *Proceedings of the National Academy of Sciences* 2004, 101, 9091–9096, doi:10.1073/pnas.0401356101.

36. Neira, G.; Cortez, D.; Jil, J.; Holmes, D.S. AciDB 1.0: A Database of Acidophilic Organisms, Their Genomic Information and Associated Metadata., doi:10.1093/bioinformatics/btaa638/5875601.

37. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the Quality of Microbial Genomes Recovered from Isolates, Single Cells, and Metagenomes. *Genome Research* 2015, 25, 1043–1055, doi:10.1101/gr.186072.114.

38. Pruitt, K.D.; Tatusova, T.; Maglott, D.R. NCBI Reference Sequences (RefSeq): A Curated Non-Redundant Sequence Database of Genomes, Transcripts and Proteins. *Nucleic Acids Research* 2007, 35, doi:10.1093/nar/gkl842.

39. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* 2014, 30, 2068–2069, doi:10.1093/bioinformatics/btu153.

40. Asnicar, F.; Thomas, A.M.; Beghini, F.; Mengoni, C.; Manara, S.; Manghi, P.; Zhu, Q.; Bolzan, M.; Cumbo, F.; May, U.; et al. Precise Phylogenetic Analysis of Microbial Isolates and Genomes from Metagenomes Using PhyloPhlAn 3.0. *Nature Communications* 2020, 11, doi:10.1038/s41467-020-16366-7.

41. Buchfink, B.; Xie, C.; Huson, D.H. Fast and Sensitive Protein Alignment Using DIAMOND. *Nature Methods* 2014, 12, 59–60.

42. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 2013, 30, 772–780, doi:10.1093/molbev/mst010.

43. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* 2009, 25, 1972–1973, doi:10.1093/bioinformatics/btp348.

44. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution* 2015, 32, 268–274, doi:10.1093/molbev/msu300.

45. Minh, B.Q.; Nguyen, M.A.T.; von Haeseler, A. Ultrafast Approximation for Phylogenetic Bootstrap. *Molecular Biology and Evolution* 2013, 30, 1188–1195, doi:10.1093/molbev/mst024.

46. Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v4: Recent Updates and New Developments. *Nucleic Acids Research* 2019, 47, W256–W259, doi:10.1093/nar/gkz239.

47. Jain, C.; Rodriguez-R, L.M.; Phillippy, A.M.; Konstantinidis, K.T.; Aluru, S. High Throughput ANI Analysis of 90K Prokaryotic Genomes Reveals Clear Species Boundaries. *Nature Communications* 2018, 9, doi:10.1038/s41467-018-07641-9.

48. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The Protein Families Database. *Nucleic Acids Research* 2014, 42.

49. Hunter, S.; Apweiler, R.; Attwood, T.K.; Bairoch, A.; Bateman, A.; Binns, D.; Bork, P.; Das, U.; Daugherty, L.; Duquenne, L.; et al. InterPro: The Integrative Protein Signature Database. *Nucleic Acids Research* 2009, 37, doi:10.1093/nar/gkn785.
Function Classification. *Bioinformatics* **2014**, *30*, 1236–1240, doi:10.1093/bioinformatics/btu031.

51. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics. *Genome Biology* **2019**, 20, doi:10.1186/s13059-019-1832-y.

52. Kelly, S.; Maini, P.K. DendroBLAST: Approximate Phylogenetic Trees in the Absence of Multiple Sequence Alignments. *PLoS ONE* **2013**, 8, doi:10.1371/journal.pone.0058537.

53. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2 - Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **2010**, 5, doi:10.1371/journal.pone.0009490.

54. Yu, N.Y.; Wagner, J.R.; Laird, M.R.; Melli, G.; Rey, S.; Lo, R.; Dao, P.; Sahinalp, S.C.; Ester, M.; Foster, L.J.; et al. PSORTb 3.0: Improved Protein Subcellular Localization Prediction with Refined Localization Subcategories and Predictive Capabilities for All Prokaryotes. *Bioinformatics* **2010**, 26, 1608–1615, doi:10.1093/bioinformatics/btq249.

55. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks. *Nature Biotechnology* **2019**, 37, 420–423, doi:10.1038/s41587-019-0036-z.

56. Maglott, D.; Ostell, J.; Pruitt, K.D.; Tatusova, T. Entrez Gene: Gene-Centered Information at NCBI. *Nucleic Acids Research* **2011**, 39, doi:10.1093/nar/gkq1237.

57. Larsson, A. AliView: A Fast and Lightweight Alignment Viewer and Editor for Large Datasets. *Bioinformatics* **2014**, 30, 3276–3278, doi:10.1093/bioinformatics/btu531.

58. Harrison, K.J.; Crécy-Lagarde, V. de; Zallot, R. Gene Graphics: A Genomic Neighborhood Data Visualization Web Application. *Bioinformatics* **2018**, 34, 1406–1408, doi:10.1093/bioinformatics/btx793.

59. Golysheva, O. v.; Toshchakov, S. v.; Makarova, K.S.; Gavrilyov, S.N.; Korzhukov, A.A.; la Cono, V.; Arcadi, E.; Nechityalno, T.Y.; Ferrer, M.; Kublanov, I. v.; et al. “ARMAN” Archaea Depend on Association with Euryarchaeal Host in Culture and in Situ. *Nature Communications* **2017**, 8, doi:10.1038/s41467-017-00104-7.

60. Murray, A.E.; Freudenstein, J.; Gribaldo, S.; Hatzenpichler, R.; Hugenholtz, P.; Kämpfer, P.; Konstantinidis, K.T.; Lane, C.E.; Papke, R.T.; Parks, D.H.; et al. Roadmap for Naming Uncultivated Archaea and Bacteria. *Nature Microbiology* **2020**, 5, 987–994.

61. Parks, D.H.; Chuvochina, M.; Chaumeil, P.A.; Rinke, C.; Mussig, A.J.; Hugenholtz, P. A Complete Domain-to-Species Taxonomy for Bacteria and Archaea. *Nature Biotechnology* **2020**, 38, 1079–1086, doi:10.1038/s41587-020-0501-8.

62. Sheng, Y.; Abreu, I.A.; Cabelli, D.E.; Maroney, M.J.; Miller, A.F.; Teixeira, M.; Valentine, J.S. Superoxide Dismutases and Superoxide Reductases. *Chemical Reviews* 2014, 114, 3854–3918.

63. dello Russo, A.; Rullo, R.; Nitti, G.; Masullo, M.; Bocchini, V. Iron Superoxide Dismutase from the Archaeon Sulfolobus Solfataricus: Average Hydrophobicity and Amino Acid Weight Are Involved in the Adaptation of Proteins to Extreme Environments. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1997**, 1343, 23–30, doi:10.1016/S0167-4838(97)00105-2.

64. Ferrer, M.; Golysheva, O. v.; Belouqi, A.; Golyshin, P.N.; Timmis, K.N. The Cellular Machinery of Ferroplasma Acidiphilum Is Iron-Protein-Dominated. *Nature* **2007**, 445, 91–94, doi:10.1038/nature05362.

65. Wheaton, G.H.; Mukherjee, A.; Kelly, R.M. Transcriptomes of the Extremely Thermoacidophilic Archaeon Metallosphaera Sedula Exposed to Metal “Shock”
Reveal Generic and Specific Metal Responses. *Applied and Environmental Microbiology* 2016, 82, 4613–4627, doi:10.1128/AEM.01176-16.

66. Jay, Z.Z.; Beam, J.P.; Dlakić, M.; Rusch, D.B.; Kozubal, M.A.; Inskeep, W.P. Marsarchaeota Are an Aerobic Archaeal Lineage Abundant in Geothermal Iron Oxide Microbial Mats. *Nature Microbiology* 2018, 3, 732–740, doi:10.1038/s41564-018-0163-1.

67. Dombrowski, N.; Lee, J.H.; Williams, T.A.; Offre, P.; Spang, A. Genomic Diversity, Lifestyles and Evolutionary Origins of DPANN Archaea. *FEMS Microbiology Letters* 2019, 366, doi:10.1093/femsle/fnz008.

68. Dombrowski, N.; Williams, T.A.; Sun, J.; Woodcroft, B.J.; Lee, J.H.; Minh, B.Q.; Rinke, C.; Spang, A. Undinarchaeota Illuminate DPANN Phylogeny and the Impact of Gene Transfer on Archaeal Evolution. *Nature Communications* 2020, 11, doi:10.1038/s41467-020-17408-w.

69. Lucchetti-Miganeh, C.; Goudenègé, D.; Thybert, D.; Salbert, G.; Barloy-Hubler, F. SORGOdb: Superoxide Reductase Gene Ontology Curated DataBase. *BMC Microbiology* 2011, 11, doi:10.1186/1471-2180-11-105.

70. Nivière, V.; Fontecave, M. Discovery of Superoxide Reductase: An Historical Perspective. *Journal of Biological Inorganic Chemistry* 2004, 9, 119–123.

71. Grunden, A.M.; Jenney, F.E.; Ma, K.; Ji, M.; Weinberg, M.v.; Adams, M.W.W. In Vitro Reconstitution of an NADPH-Dependent Superoxide Reduction Pathway from Pyrococcus Furiosus. *Applied and Environmental Microbiology* 2005, 71, 1522–1530, doi:10.1128/AEM.71.3.1522-1530.2005.

72. Cárdenas, J.P.; Moya, F.; Covarrubias, P.; Shmaryahu, A.; Levicán, G.; Holmes, D.S.; Quatrini, R. Comparative Genomics of the Oxidative Stress Response in Bioleaching Microorganisms. *Hydrometallurgy* 2012, 127–128, 162–167, doi:10.1016/j.hydromet.2012.07.014.

73. Zamocky, M.; Furtmüller, P.G.; Obinger, C. Evolution of Catalases from Bacteria to Humans. *Antioxidants and Redox Signaling* 2008, 10, 1527–1547.

74. Passardi, F.; Zamocky, M.; Favet, J.; Jakopitsch, C.; Penel, C.; Obinger, C.; Dunand, C. Phylogenetic Distribution of Catalase-Peroxidases: Are There Patches of Order in Chaos? *Gene* 2007, 397, 101–113, doi:10.1016/j.gene.2007.04.016.

75. Saikolappan, S.; Das, K.; Sasindran, S.J.; Jagannath, C.; Dhandayuthapani, S. OsmC Proteins of Mycobacterium Tuberculosis and Mycobacterium Smegmatis Protect against Organic Hydroperoxide Stress. *Tuberculosis* 2011, 91, doi:10.1016/j.tube.2011.10.021.

76. Park, S.C.; Pham, B.P.; van Duyet, L.; Jia, B.; Lee, S.; Yu, R.; Woo Han, S.; Yang, J.K.; Hahn, K.S.; Cheong, G.W. Structural and Functional Characterization of Osmotically Inducible Protein C (OsmC) from Thermococcus Kodakaraensis KOD1. *Biochimica et Biophysica Acta - Proteins and Proteomics* 2008, 1784, 783–788, doi:10.1016/j.bbapap.2008.02.002.

77. Gauss, G.H.; Benas, P.; Wiedenheft, B.; Young, M.; Douglas, T.; Lawrence, C.M. Structure of the DPS-like Protein from Sulfolobus Solfataricus Reveals a Bacterioferritin-like Dimetal Binding Site within a DPS-like Dodecameric Assembly. *Biochemistry* 2006, 45, 10815–10827, doi:10.1021/bi060782u.

78. Wiedenheft, B.; Mosolf, J.; Willits, D.; Yeager, M.; Dryden, K.A.; Young, M.; Douglas, T. An Archaeal Antioxidant: Characterization of a Dps-like Protein from Sulfolobus Solfataricus; 2005;

79. Wiedenbeck, J.; Cohan, F.M. Origins of Bacterial Diversity through Horizontal Genetic Transfer and Adaptation to New Ecological Niches. *FEMS Microbiology Reviews* 2011, 35, 957–976.
80. Weinberg, M. v.; Jenney, F.E.; Cui, X.; Adams, M.W.W. Rubrerythrin from the Hyperthermophilic Archaeon Pyrococcus Furiosus Is a Rubredoxin-Dependent, Iron-Containing Peroxidase. *Journal of Bacteriology* 2004, 186, 7888–7895, doi:10.1128/JB.186.23.7888-7895.2004.

81. Kim, S.A.; Punshon, T.; Lanzirotti, A.; Li, A.; Alonso, J.M.; Ecker, J.R.; Kaplan, J.; Guerinot, M. lio Localization of Iron in Arabidopsis Seed Requires the Vacuolar Membrane Transporter VIT1. *Science* 2006, 314, 1295–1298, doi:10.1126/science.1132563.

82. Cardenas, J.P.; Quatrini, R.; Holmes, D.S. Aerobic Lineage of the Oxidative Stress Response Protein Rubrerythrin Emerged in an Ancient Microaerobic, (Hyper)Thermophilic Environment. *Frontiers in Microbiology* 2016, 7, doi:10.3389/fmicb.2016.01822.

83. Kato, T.; Kumazaki, K.; Wada, M.; Taniguchi, R.; Nakane, T.; Yamashita, K.; Hirata, K.; Ishitani, R.; Ito, K.; Nishizawa, T.; et al. Crystal Structure of Plant Vacuolar Iron Transporter VIT1. *Nature Plants* 2019, 5, 308–315, doi:10.1038/s41477-019-0367-2.

84. Sorribes-Dauden, R.; Peris, D.; Martínez-Pastor, M.T.; Puig, S. Structure and Function of the Vacuolar Ccc1/VIT1 Family of Iron Transporters and Its Regulation in Fungi. *Computational and Structural Biotechnology Journal* 2020, 18, 3712–3722.

85. Valdes-Stauber, N.; Scherer, S. Isolation and Characterization of Linocin M18, a Bacteriocin Produced by *Brevibacterium Linens*.; 1994; Vol. 60;.

86. Soito, L.; Williamson, C.; Knutson, S.T.; Fetrow, J.S.; Poole, L.B.; Nelson, K.J. PREX: PeroxiRedoxin Classification IndEX, a Database of Subfamily Assignments across the Diverse Peroxiredoxin Family. *Nucleic Acids Research* 2011, 39, doi:10.1093/nar/gkq1060.

87. Limauro, D.; Pedone, E.; Galdi, I.; Bartolucci, S. Peroxiredoxins as Cellular Guardians in Sulfolobus Solfataricus- Characterization of Bcp1, Bcp3 and Bcp4. *FEBS Journal* 2008, 275, 2067–2077, doi:10.1111/j.1742-4658.2008.06361.x.

88. D’Ambrosio, K.; Limauro, D.; Pedone, E.; Galdi, I.; Pedone, C.; Bartolucci, S.; de Simone, G. Insights into the Catalytic Mechanism of the Bcp Family: Functional and Structural Analysis of Bcp1 from Sulfolobus Solfataricus. *Proteins: Structure, Function and Bioinformatics* 2009, 76, 995–1006, doi:10.1002/prot.22408.

89. Sun, C.C.; Dong, W.R.; Shao, T.; Li, J.Y.; Zhao, J.; Nie, L.; Xiang, L.X.; Zhu, G.; Shao, J.Z. Peroxiredoxin 1 (Prx1) Is a Dual-Function Enzyme by Possessing Cys-Independent Catalase-like Activity. *Biochemical Journal* 2017, 474, 1373–1394, doi:10.1042/BCJ20160851.

90. Bang, Y.J.; Oh, M.H.; Choi, S.H. Distinct Characteristics of Two 2-Cys Peroxiredoxins of Vibrio Vulnificus Suggesting Differential Roles in Detoxifying Oxidative Stress. *Journal of Biological Chemistry* 2012, 287, 42516–42524, doi:10.1074/jbc.M112.421214.

91. Groitl, B.; Jakob, U. Thiol-Based Redox Switches. *Biochimica et Biophysica Acta - Proteins and Proteomics* 2014, 1844, 1335–1343.

92. Maupin-Furlow, J.A. Methionine Sulfoxide Reductases of Archaea. *Antioxidants* 2018, 7.

93. Kim, H.S.; Kwak, G.H.; Lee, K.; Jo, C.H.; Hwang, K.Y.; Kim, H.Y. Structural and Biochemical Analysis of a Type II Free Methionine-R- Sulfoxide Reductase from Thermoplasma Acidophilum. *Archives of Biochemistry and Biophysics* 2014, 560, 10–19, doi:10.1016/j.abb.2014.07.009.

94. Juillan-Binard, C.; Picciocchi, A.; Andrieu, J.P.; Dupuy, J.; Petit-Hartlein, I.; Caux-Thang, C.; Vivès, C.; Nivière, V.; Fieschi, F. A Two-Component NADPH Oxidase (NOX)-like System in Bacteria Is Involved in the Electron Transfer Chain to the
Methionine Sulfoxide Reductase MsrP. *Journal of Biological Chemistry* 2017, 292, 2485–2494, doi:10.1074/jbc.M116.752014.

95. Case, A.J. On the Origin of Superoxide Dismutase: An Evolutionary Perspective of Superoxide-Mediated Redox Signaling. *Antioxidants* 2017, 6.

96. Jenney, F.E.; Verhagen, M.F.J.M.; Cui, X.; Adams, M.W.W. Anaerobic Microbes: Oxygen Detoxification Without Superoxide Dismutase. *Science* 1999, 286, 306–309, doi:10.1126/science.286.5438.306.

97. Koonin, E. v.; Wolf, Y.I. Genomics of Bacteria and Archaea: The Emerging Dynamic View of the Prokaryotic World. *Nucleic Acids Research* 2008, 36, 6688–6719, doi:10.1093/nar/gkn668.

98. Wagner, A.; Whitaker, R.J.; Krause, D.J.; Heilers, J.H.; van Wolferen, M.; van der Does, C.; Albers, S.V. Mechanisms of Gene Flow in Archaea. *Nature Reviews Microbiology* 2017, 15, 492–501.

99. Beard, S.; Ossandon, F.J.; Rawlings, D.E.; Quatrini, R. The Flexible Genome of Acidophilic Prokaryotes. *Current Issues in Molecular Biology* 2020, 40, 231–266, doi:10.21775/cimbb.040.231.

100. Zhang, X.; Liu, X.; Liang, Y.; Guo, X.; Xiao, Y.; Ma, L.; Miao, B.; Liu, H.; Peng, D.; Huang, W.; et al. Adaptive Evolution of Extreme Acidophile Sulfolobus Thermosulfidooxidans Potentially Driven by Horizontal Gene Transfer and Gene Loss. *Applied and Environmental Microbiology* 2017, 83, doi:10.1128/AEM.03098-16.

101. Vergara, E.; Neira, G.; González, C.; Cortez, D.; Dopson, M.; Holmes, D.S. Evolution of Predicted Acid Resistance Mechanisms in the Extremely Acidophilic Leptospirillum Genus. *Genes* 2020, 11, doi:10.3390ogen11040389.

102. Chiancone, E.; Ceci, P. The Multifaceted Capacity of Dps Proteins to Combat Bacterial Stress Conditions: Detoxification of Iron and Hydrogen Peroxide and DNA Binding. *Biochimica et Biophysica Acta - General Subjects* 2010, 1800, 798–805.

103. Yuan, F.; Yin, S.; Xu, Y.; Xiang, L.; Wang, H.; Li, Z.; Fan, K.; Pan, G. The Richness and Diversity of Catalases in Bacteria. *Frontiers in Microbiology* 2021, 12, doi:10.3389/fmicb.2021.645477.

104. Canfield, D.E.; Rosing, M.T.; Bjerrum, C. Early Anaerobic Metabolisms. In Proceedings of the Philosophical Transactions of the Royal Society B: Biological Sciences; Royal Society, October 29 2006; Vol. 361, pp. 1819–1834.

105. Martin, W.F.; Sousa, F.L. Early Microbial Evolution: The Age of Anaerobes. *Cold Spring Harbor Perspectives in Biology* 2016, 8, doi:10.1101/cshperspect.a018127.

106. Ouzounis, C.A.; Kunin, V.; Darzentas, N.; Goldovsky, L. A Minimal Estimate for the Gene Content of the Last Universal Common Ancestor - Exobiology from a Terrestrial Perspective. In Proceedings of the Research in Microbiology; January 2006; Vol. 157, pp. 57–68.

107. Ślesak, I.; Ślesak, H.; Kruk, J. RubisCO Early Oxygenase Activity: A Kinetic and Evolutionary Perspective. *BioEssays* 2017, 1700071, 1700071, doi:10.1002/bies.201700071.

108. Planavsky, N.J.; Asael, D.; Hofmann, A.; Reinhard, C.T.; Lalonde, S. v.; Knudsen, A.; Wang, X.; Ossa Ossa, F.; Pecoits, E.; Smith, A.J.B.; et al. Evidence for Oxygenic Photosynthesis Half a Billion Years before the Great Oxidation Event. *Nature Geoscience* 2014, 7, 283–286, doi:10.1038/ngeo2122.
