Article

Evolution of Predicted Acid Resistance Mechanisms in the Extremely Acidophilic *Leptospirillum* Genus

Eva Vergara [1], **Gonzalo Neira** [1], **Carolina González** [1,2], **Diego Cortez** [1], **Mark Dopson** [3] **and David S. Holmes** [1,2,4,*]

[1] Center for Bioinformatics and Genome Biology, Fundación Ciencia & Vida, Santiago 7780272, Chile; evamarilyn.vj@gmail.com (E.V.); gonzalo.neira@ug.uchile.cl (G.N.); carola.mgr@gmail.com (C.G.); diegonahuel8@gmail.com (D.C.)

[2] Centro de Genómica y Bioinformática, Facultad de Ciencias, Universidad Mayor, Santiago 8580745, Chile

[3] Centre for Ecology and Evolution in Microbial Model Systems, Linnaeus University, SE-391 82 Kalmar, Sweden; mark.dopson@lnu.se

[4] Universidad San Sebastian, Santiago 7510156, Chile

[*] Correspondence: dsholmes2000@yahoo.com

Abstract: Organisms that thrive in extremely acidic environments (≤pH 3.5) are of widespread importance in industrial applications, environmental issues, and evolutionary studies. *Leptospirillum* spp. constitute the only extremely acidophilic microbes in the phylogenetically deep-rooted bacterial phylum Nitrospirae. Leptospirilli are Gram-negative, obligatory chemolithoautotrophic, aerobic, ferrous iron oxidizers. This paper predicts genes that Leptospirilli use to survive at low pH and infers their evolutionary trajectory. Phylogenetic and other bioinformatic approaches suggest that these genes can be classified into (i) "first line of defense", involved in the prevention of the entry of protons into the cell, and (ii) neutralization or expulsion of protons that enter the cell. The first line of defense includes potassium transporters, predicted to form an inside positive membrane potential, spermidines, hopanoids, and Slps (starvation-inducible outer membrane proteins). The "second line of defense" includes proton pumps and enzymes that consume protons. Maximum parsimony, clustering methods, and gene alignments are used to infer the evolutionary trajectory that potentially enabled the ancestral *Leptospirillum* to transition from a postulated circum-neutral pH environment to an extremely acidic one. The hypothesized trajectory includes gene gains/loss events driven extensively by horizontal gene transfer, gene duplications, gene mutations, and genomic rearrangements.

Keywords: *Nitrospira*; extreme acidophile; acid mine drainage (AMD); bioleaching; evolution; acid resistance; comparative genomics; horizontal gene transfer (HGT); phylogenetics

1. Introduction

Although microorganisms such as *Escherichia coli*, *Vibrio cholerae*, and *Salmonella* spp. are neutrophiles, some strains survive acid shock during transient passage through low pH conditions such as in the stomach (<pH 3.5). In addition, the neutrophile *Helicobacter pylori* survives and grows in the stomach by creating its own near neutral pH environment via the hydrolysis of urea to produce CO_2 and NH_3 that buffer the acidity of the local environment (reviewed in [1]). Genes and mechanisms that these microorganisms use to survive acid shock and the regulatory networks that control their expression were reviewed in [2–6]. Due to the transitory nature of the acid shock response, these neutrophiles have been dubbed "amateur" acidophiles [5]. In contrast, extreme or "professional" acidophiles are organisms that thrive in extremely low pH environments (≤pH 3.5).

They include Bacteria, Archaea, and Eukarya and are found widely distributed across the Tree of Life (reviewed in [1,7–10]).

Extreme acidophiles maintain a cytoplasmic pH close to neutral while confronted with a $>10^4$ fold proton gradient across the cytoplasmic membrane [1,8,11]. Extreme acidophiles use many of the same mechanisms to cope with acid resistance as the amateur acidophiles, but also have additional mechanisms that allow them to cope with greater proton gradients [1]. Extreme acidophiles have several "first line" of defense mechanisms against low pH that function both within and outside the cytoplasmic membrane to exclude protons from entering the cell. They also have second line of defense mechanisms that consume or expel protons that have penetrated the cytoplasmic membrane. Less is known regarding the acid response mechanisms of extreme acidophiles than their amateur counterparts and especially the evolutionary events that resulted in the ability to grow at very low pH values.

The "first line" of defense of extreme acidophiles includes the development of an inside positive membrane potential (Donnan potential), such that there is a chemiosmotic barrier that inhibits positively charged protons from crossing the membrane and acidifying the cytoplasm. The membrane potential is likely generated by potassium ions (and to a lesser extent, sodium ions) accumulated via the Kch, Kdp, and Trk K^+ channel proteins [12–14]. Kch is a member of the voltage-gated ion channel superfamily mostly studied in eukaryotes, wherein they play a role in membrane potential regulation. The KdpFABC complex is a potassium transporting P-type ATPase with a high affinity to potassium [15], while the products of three unlinked genes (*trk*AEH) are necessary for Trk activity in *E. coli* and similar enterics [16]. Evidence to support the role of potassium ions in forming this inside positive membrane potential includes the observation that removing potassium and/or sodium decreases the ability of *Sulfolobus* spp. [12,17] and *Acidithiobacillus thiooxidans* [18] to cope with acid resistance. All three of these potassium transporters have been previously reported in *Leptospirillum ferriphilum* as part of its pH homeostasis mechanism [13]. The internal positive membrane potential is a key difference between professional acidophiles compared to amateur acidophiles [19]. However, definitive proof of the role of potassium in generating the membrane potential is still lacking.

A second cytoplasmic membrane-associated method to reduce proton penetration into the cell involves the use of spermidine. Spermidine is a positively charged, aliphatic polycation polyamine that together with spermine, putrescine, and cadaverine are found in all three domains of life. These polyamines are involved in a wide range of biological functions including acid stress responses (reviewed in [3,20]). The mechanism by which spermidine protects against acid stress is not completely understood but it is reported that it stimulates RpoS ($\sigma 38$) synthesis in *E. coli* that, in turn, regulates many genes involved in the acid stress response [21–23]. Spermidine has been detected in *Leptospirillum ferrooxidans* [24] and *spe*EH, potentially involved in its biosynthesis, and is also predicted in the *L. ferriphilum* genome [13]. Orthologs of *spe*EH have higher transcript numbers during acid stress in the extreme acidophile *Acidithiobacillus ferrooxidans*, an observation that is consistent with the idea that they may be involved in acid resistance [25].

Hopanoid membrane lipids have also been shown to be involved in the first line acid resistance response [26,27]. Hopanoids have diverse proposed functions including modulating membrane fluidity and permeability. They may also be involved in other stress responses including oxidative stress and high temperature [28–30]. Hopanoid (*hpn*) genes are often found in gene clusters (e.g., [31]) and code for enzymes that either synthesize hopanoids from squalene or modify them, yielding different hopanoid molecules that may have a specific stress related function in various strains [26]. Hopanoids are present in acidophiles [32,33], and their RNA transcripts are detected in acid mine drainage (AMD) [34]. Deletion of hopanoid synthesis genes in neutrophiles resulted in impaired growth at low pH [28,30].

The starvation-inducible outer membrane protein (Slp) has been implicated in acid resistance. For example, in enteropathogenic *E. coli*, *slp* is located in a gene cluster that has been associated with acid resistance [35]. Co-expression with other acid resistance genes of this cluster has been reported [36]. Although the details of the mechanism of action of Slp remain unknown [37], it is proposed that

it prevents the entrance of organic acids such as succinate, lactate, and formate across the outer membrane [38]. Organic acids have been shown to be highly toxic to acidophiles [39].

Second line of defense mechanisms include proton antiporters such as NahP-/NhaA-type Na^+/H^+ exchangers. These proteins constitute a large family of integral membrane proteins with roles related to homeostasis of sodium in high salt environments and regulation of intracellular pH [40]. In addition, voltage gated ClC-type Cl^-/H^+ transporters [41] are essential for the survival of *E. coli* under extreme acid stress [42]. In addition to its proton exporting function, the ClC channel is also proposed to prevent inner-membrane hyperpolarization (inner positive) in *E. coli* at extremely acidic pH as a result of the action of other acid resistance systems such as arginine and glutamate decarboxylation systems [43]. A similar function for ClC in acid resistance has been proposed in the amateuracidophile *Bacillus coagulans* [44].

Another second line of defense mechanism is the use of amino acid decarboxylation systems such as the Gad glutamate decarboxylase system. These systems have two components: (i) an inner membrane amino acid antiporter that imports the amino acid in exchange for its decarboxylated form and (ii) a cytoplasmic decarboxylase that catalyzes the proton-consuming decarboxylation of an amino acid [45]. Glutamate decarboxylase is the major response in *E. coli* under extreme acid conditions [6]. H^+-consuming reactions during acid stress have also been observed in *Acidithiobacillus caldus* [33]. Several amino acid decarboxylases genes are present in *Leptospirillum*, e.g., [13,46], and transcripts coding for glutamate, arginine, and lysine decarboxylation have been detected in an AMD community [34].

In this study, we predict and analyze both first and second line of defense acid resistance systems in the extremely acidophilic genus *Leptospirillum.* We create a global model of pH homeostasis mechanisms that suggests how this genus can survive in hyper-acidic conditions. We use parsimony to infer the evolutionary trajectories of the gene gains/losses and gene mutations that are hypothesized to be involved in acid resistance. The *Leptospirillum* genus was chosen as a model system for this study because it is one of the most extreme bacterial acidophiles known with a growth pH range between 0.7 and 2.2 [47,48]. In addition, fluorescent in situ hybridization and "omic" studies highlight *Leptospirillum* as one of the main active taxa in extremely low pH natural and man-made acid mine drainage environments [49–58] and in commercial copper and gold biorecovery operations [47,59].

2. Methods

2.1. Genomes and Quality Assessment

Fifteen complete and partial genome sequences from the *Leptospirillum* genus were downloaded from the NCBI RefSeq genomic database, National Center for Biotechnology Information (NCBI), and Joint Genome Institute's IMG-M databases (https://img.jgi.doe.gov/) in August 2018. In addition, the *Nitrospira marina* Nb-295 genome (IMG Genome ID 2596583682, https://img.jgi.doe.gov/) was downloaded as an outgroup for the *Leptospirillum* genus. Quality assessment of the 16 Nitrospirae genomes was carried out by CheckM [60], defining >90% base completeness and <10% contamination as high-quality genomes according to a bioinformatics pipeline (Figure 1).
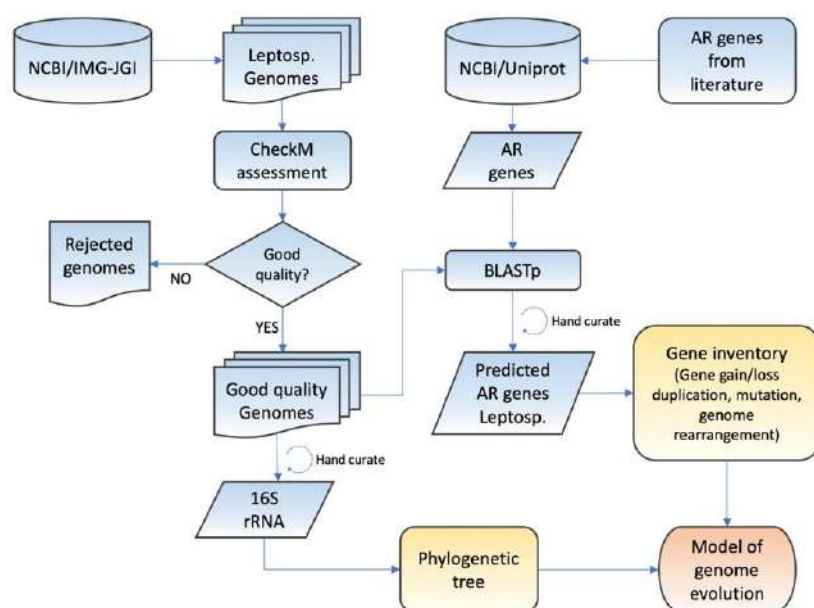
Figure 1. Bioinformatic pipeline outlining the strategy used for developing a model of the evolution of acid resistance (AR) mechanisms in the extremely acidophilic. Leptosp. = *Leptospirillum* genus.

2.2. Phylogenetic Analysis

16S rRNA gene sequences from organisms within the Nitrospirae phylum were identified by a BLASTn comparison against the SILVA [61], RDP [62], and GREENGENES rRNA databases [63] with an E-value threshold of 1E^{-5}. Gene sequences with a minimum length of 1400 nucleotides were selected [64]. Alignments of 16S rRNA sequences were generated with MAFFT with the L-INSI iterative refinements option [65,66], MUSCLE [67], and T-Coffee alignment tools [68]. A maximum-likelihood tree was constructed with IQTREE [69], and the best-suited evolutionary models were selected using the model test tool implemented in IQTREE [70] according to the Bayesian (BIC) and Akaike information criterion (AIC). The robustness of the inferred tree was assessed using the nonparametric bootstrap procedure implemented in IQTREE (1000 replicates of the original datasets) with the ultrafast bootstrap option [71]. The final tree was visualized using Figtree (http://tree.bio.ed.ac.uk/software/figtree/).

2.3. Prediction of Mobile Genetic Elements and Genome Islands

Insertion elements and transposases were predicted and classified using TnpPred [72] and ISsaga [73,74]. Sigi-HMM [75] was used to predict genes obtained by Horizontal Gene Transfer (HGT). IslandViewer 4 [76] was used to predict genomic islands. Genome contexts of predicted acid resistance genes, associated hypothetical genes, and predicted mobile elements were analyzed using STRING [77] and by manual inspection using MAUVE [78] and Artemis [79].

2.4. Identification of Genes Related to Low pH Resistance

Genes reported to be involved in acid resistance were identified through an extensive literature search [4–6,38,80–87]. A search for similar genes in *Leptospirillum* and *Nitrospira marina* Nb-295 genomes was carried out through BLASTp comparison [88] using a minimal *E*-value cutoff of 1e^{-5}. Synteny blocks and conservation of genetic context between *Nitrospiraceae* genomes were determined by MAUVE [78]. Genomic contexts were visualized by Artemis [79]. Conservation of sequences and domains within the *Leptospirillum* genus were analyzed and visualized by the WebLogo [89,90] and AliView [91] alignments tools. Selected *Nitrospiraceae* genes were compared against the UniProt and NCBI databases by BLASTp to identify orthologous proteins in other microorganisms. This collection of sequences was aligned with MAFFT using the L-INSI iterative refinements option [65,66]. IQTREE was used to construct a maximum-likelihood tree with 1000 replicates by the ultrafast bootstrap

option [71] and to identify the best-suited evolutionary model by the Bayesian (BIC) and Akaike information criterion (AIC) [70]. Phylogenetic trees were visualized with Figtree (http://tree.bio.ed.ac.uk/software/figtree/) and iTOL [92].

2.5. Evolutionary Pressures on Acid Resistance Genes

Selective pressure on genes was determined by calculating the ratio of non-synonymous DNA substitutions (D_n) to synonymous DNA substitutions (D_s) in the coding region [93]. Individual genes (DNA and amino acid sequences) from all protein families were extracted using custom Perl scripts. Amino acid alignments were constructed using the MAFFT L-INSI iterative refinements option [65,66] and MUSCLE [67] and used as input for PAL2NAL [94] in conjunction with their nucleotide sequences to obtain the corresponding codon alignments for gene families. D_n/D_s ratios were assigned for all possible pairwise comparisons within a protein family and calculated based on the codon alignments using the SeqinR package of the R project [95]. Mean D_n/D_s ratios were assigned for individual gene families by averaging all pairwise ratios within each family. D_n/D_s ratios of > 1 indicated beneficial mutations, and ratios of <1 indicated purifying selection [96].

2.6. Mapping Evolutionary Events

The inference of branch-site-specific events was made using the 16S rRNA gene tree of *Leptospirillum* genomes with *N. marina* Nb-295 as the outgroup. The presence and absence of genes related to acid resistance and associated genes located within the same genomic context were mapped onto each branch of the phylogenetic tree to model gene gain, loss, and modification events. Inference of evolutionary events was made using maximum parsimony criteria [97,98].

3. Results and Discussion

3.1. Genomic Features of Leptospirillum Genomes

Twelve publicly available genomes of the *Leptospirillum* genus were analyzed together with the genome of the neutrophile *N. marina*. *N. marina* is the closest phylogenetic relative of the Leptospirilli with available genomic data (Table 1). Five of the genomes were complete, while eight (including *N. marina*) were permanent drafts. "*Leptospirillum rubarum*", *Leptospirillum* "5-way CG", *Leptospirillum* "C-75", and "*Leptospirillum ferrodiazotrophum*" were from metagenomic samples. The G + C content of *Leptospirillum* Group I, II, and III genomes were 50.1, 54.0, and 57.5%, respectively. Leptospirilli included both mesophiles and moderate thermophiles and grew between 40- 43 °C. However, *L. ferrooxidans* C2-3, *L. ferriphilum* DSM 14647^T, and *L. ferriphilum* Sp-Cl were mesophiles with growth temperatures between 30°C and 37°C. *Leptospirillum* sp. "UBA BS" (Group IV; NCBI Accession: PRJNA176861 [32]) was not included in this study because the genome did not meet the quality criteria of CheckM [60], exhibiting only 41% completeness with 38% contamination.

Table 1. Properties of organisms and their genomes used in this study.

Genome	Size (Mb)	# Predicted Genes	G + C (%)	pH	Temp (°C)	Status [1]	NCBI Accession [2]	Geographical Origin	Refs
Leptospirillum ferrooxidans C2-3	2.5	2587	50.0	1.8	30	C	NC_017094.1	Mount Oyama, Miyake, Japan	[99]
Leptospirillum ferriphilum DSM 14647	2.6	2687	54.1	1.4–1.8	37	C	NZ_OBMB00000000.1	Enrichment culture, Peru	[13]
Leptospirillum ferriphilum ML-04	2.4	2475	54.6	2.5	40	C	NC_018649.1	Sulfide hot spring, Yunnan, China	[46]
Leptospirillum ferriphilum YSK	2.3	2361	54.1	1.6	40	C	NZ_CP007243.1	Dexing copper mine, JiangXi, China	[100]
Leptospirillum ferriphilum DX	2.3	2381	54.5	1.5	40	D	NZ_MPOJ00000000.1	Dexing copper mine, JiangXi, China	[101]
Leptospirillum ferriphilum ZJ	2.3	2449	54.7	1.5	40	D	NZ_MPOK00000000.1	Zijinshan copper mine, Fujian, China	[101]
Leptospirillum ferriphilum Sp-Cl	2.4	2552	54.4	1.5	37	D	NZ_LGSH00000000.1	Spence mine, Chile	[59]
Leptospirillum sp. "CF-1"	2.7	2731	54.6	1.6–1.7	40	C	NZ_CP012147.1	Iron Mountain, CA, USA	[102]
Leptospirillum sp. "C75"	2.6	2528	54.4	0.7–1.2	40–43	D	GCF_000262365.1	Iron Mountain, CA, USA	[103]
Leptospirillum sp. "5-way CG"	2.7	2633	51.5	0.8	42	D	DS995259.1–DS995275.1	Iron Mountain, CA, USA	[104]
"*Leptospirillum rubarum*"	2.6	2654	54.7	1.1	41	D	GCA_000205145.2	Iron Mountain, CA, USA	[52]
"*Leptospirillum ferrodiazotrophum*"	2.8	2727	57.5	1.1	41	D	GG693851.1–GG693892.1	Iron Mountain, CA, USA	[52]
Nitrospira marina Nb-295	4.6	4276	50	6.4–7.5	30	D	2596583682	Gulf of Maine, USA	‡

[1] Sequence status where C: Complete, D: Draft. [2] NCBI RefSeq accession numbers are provided; if not available, then NCBI RefSeq assembly IDs are provided. ‡ IMG genome ID (https://img.jgi.doe.gov/). pH and temperature for organismal growth come from culture experiments or from in situ environmental measurements.

3.2. Phylogenetic Relatedness between Leptospirillum Species and Other Members of the Nitrospirae Phylum

A phylogenetic tree of the members of the Nitrospirae phylum was developed based on 16S rRNA gene sequences (Figure 2a). The tree was rooted using the validated *Rubrobacter radiotolerans*[T] DSM 5868 as an outgroup. Three species of *Leptospirillum* can be distinguished in the phylogenetic tree: *L. ferrooxidans* (Group I), *L. ferriphilum* (Group II), and "*L. ferrodiazotrophum*" (Group III) with bootstrap support ≥ 60% (Figure 2a). This tree was consistent with published trees of *Leptospirillum* [51,101,105]. The phylogenetic branching points of the members of Group II had insufficient resolution to be separated (Figure 2a). Therefore, the phylogenetic tree was also presented as a cladogram (Figure 2b), showing their predicted branching order with bootstrap values. Based on the measurement of genetic distance, it was shown that *N. marina* Nb-295 was the closest extant relative with a sequenced genome to the *Leptospirillum* genus.
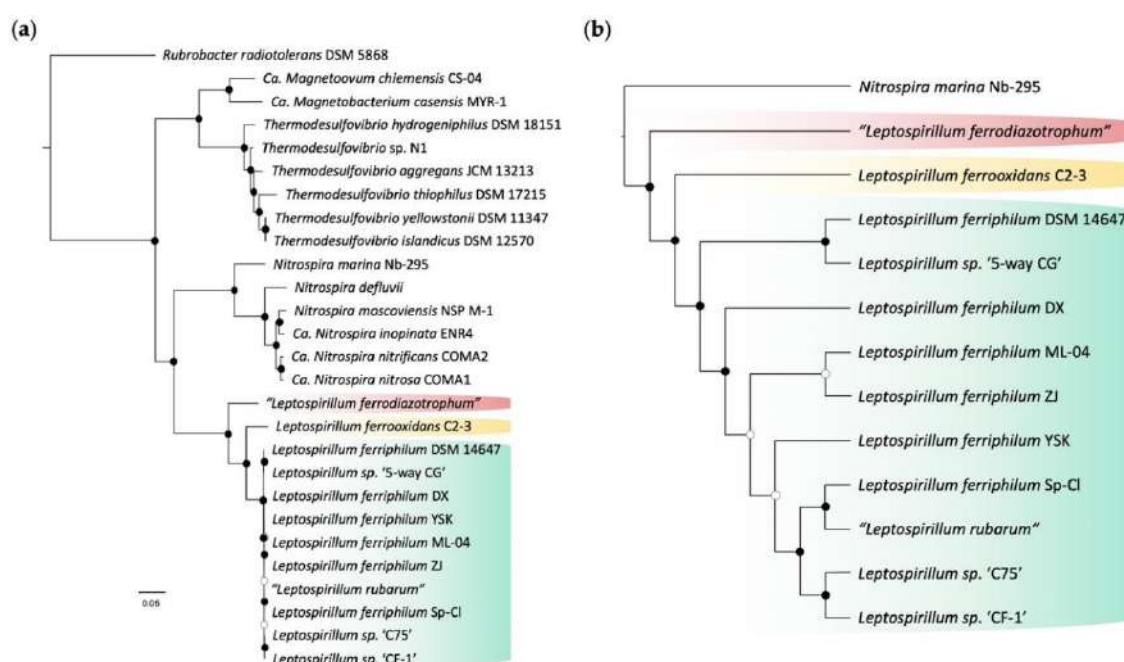

Figure 2. 16S rRNA phylogeny and cladogram of *Leptospirillum* species with sequenced genomes used in this study. Clades belonging to *Leptospirillum* Groups I (yellow), II (green), and III (red) are indicated. (**a**) Phylogeny of the *Leptospirillum* genus inferred from 16S rRNA gene sequences. Filled circles at the nodes indicate bootstrap support ≥ 60% and open circles bootstrap support < 60%. The scale bar indicates the number of nucleotide substitutions per site. (**b**) Cladogram of the *Leptospirillum* genus derived from the 16S rRNA gene phylogenetic tree using *N. marina* as an outgroup. The cladogram shows the predicted branching order of the *Leptospirillum* genus. Bootstrap values are derived from Figure 2a. Color coding is the same as shown in Figure 2a.

3.3. Gene Inventories

Genes with predicted or experimental evidence for functions related with first and second lines of defense to low pH environments were identified in the literature. A list of the genes used, their NCBI accession numbers, and their predicted features are provided in Supplementary Table S1.

3.4. First Line of Defense

3.4.1. Membrane Potential and Potassium Transporters

Kch, potentially encoding a K⁺ channel protein, was identified in all *Leptospirillum* genomes, but not in *N. marina*. One possible explanation for this is that *kch* was incorporated into the *Leptospirillum* lineage by HGT after its separation from the *N. marina* lineage. Alternatively, *kch* was present in the last common ancestor and was lost in the *N. marina* lineage. Examination of the NCBI database using

BLASTp showed that the best hits of the *Leptospirillum* Kch were with proteins of the *Acidithiobacillus* genus together with several other known acidophiles (Supplementary Figure S1). These microorganisms are frequently found in extremely acidic environments populated by *Leptospirillum* (e.g., [34,56,106,107]), strongly suggesting that they have shared *kch* via HGT. In addition, this result was consistent with the contention that *kch* is associated with acid resistance.

In *Leptospirillum* Group III, *kch* was adjacent to, but divergent from, a gene potentially encoding a phage holin-like protein that is involved in stress response and other functions (reviewed in [108]). In *Leptospirillum* Groups I and II, additional copies of the phage holin were located close to predicted genes encoding DNA uptake competence functions (ComEC). These are thought to be one of the major components involved in HGT (reviewed in [109]). In *Leptospirillum* Groups I and II, *kch* was clustered with two other predicted acid resistance genes, *slp*8 and *gad*A (discussed in Sections 3.4.4 and 3.5.2, respectively). One possibility is that *kch*, *slp*8, and *gad*A entered the *Leptospirillum* genome by HGT, possibly via a phage mediated uptake mechanism.

A *kdp*ABC gene cluster potentially encoding a potassium transporting Kdp P-type ATPase was found in all the *Leptospirillum* genomes (Figure 3). Downstream of the *kdp* cluster, there was a predicted BBP2 porin, a putative *gad*C2a permease, followed by another K⁺ sensing histidine kinase with a response regulator CitB. This cluster may be associated with K⁺ regulation. According to STRING analysis [77], this gene cluster was co-expressed in other species, suggesting that it was an operon. Some Group II genomes contained a predicted transposase (*tnp*3) associated with the insertion of two hypothetical genes. There was also an insertion of two hypothetical genes just upstream of *gad*C2a. The functions of the hypothetical genes remain unknown.
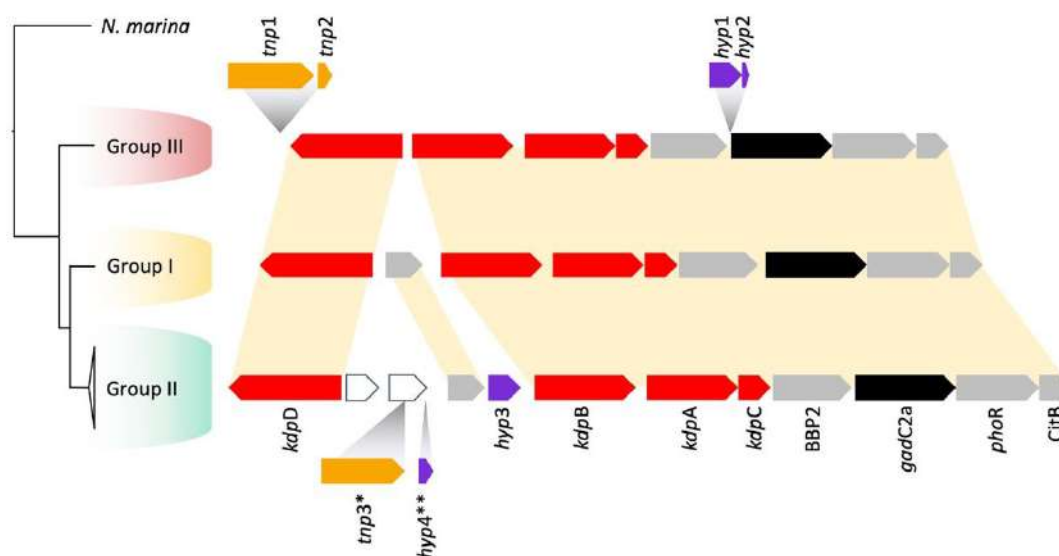

Figure 3. Phylogenetic distribution and genomic contexts of genes predicted for Kdp potassium uptake. *N. marina*, which lacks these genes, is included as an outgroup. Color coding of genes: red = Kdp genes, purple = orphan hypothetical genes, grey = additional genes whose genomic context is conserved, orange = predicted mobile elements and their remnants (* *tnp*3 inserted only in CF-1, ** *hyp*4 inserted only in the YSK strain), and black = *gad*C2A potentially involved proton export (see Section 3.5.2).

*Trk*A was identified in all *Leptospirillum* genomes and in the genome of *N. marina*. Comparative amino acid sequence analysis indicated that *Leptospirillum trk*A was found in a cluster that was separate from other *Nitrospira* (Supplementary Figure S2). An extremely large D_n/D_s ratio was observed (~1) between *trk*A of *Leptospirillum* compared to other *Nitrospira*, suggesting that it could have been vertically inherited from a likely neutrophilic common ancestor of *Leptospirillum* and other *Nitrospira* and then subjected to intense selective pressure to adapt to an acidic environment. Once adapted, it underwent a few additional changes, as shown by extremely low D_n/D_s ratios (~0.05) within the *Leptospirillum* genus (Supplementary Figure S2).

3.4.2. Spermidine Biosynthesis and Associated Genes

The *Leptospirillum* genomes were searched for genes potentially encoding aliphatic polycation polyamines. No genes encoding for spermine or cadaverine biosynthesis were detected in any of the genomes. However, a conserved cluster of four genes potentially encoding spermidine biosynthesis was predicted in all three *Leptospirillum* groups, extending the earlier prediction of spermidine genes in in *L. ferriphilum* [13]. This cluster was not detected in *N. marina* (Figure 4 and Supplementary Figure S3). Three of the genes in the cluster were predicted to encode the biosynthesis of spermidine from S-adenosyl-L-methionine (SAM): *spe*H encoding S-adenosylmethionine decarboxylase, *spe*E encoding spermidine synthase, and an *odc*-like gene predicted to produce putrescine from ornithine (Figure 4a,b). The fourth gene (*hyp*4) encoded a conserved hypothetical protein UPF0182 found in many organisms in the same genomic context, but whose function remains unknown. UPF0182 was predicted to have a signal peptide for protein export and six transmembrane regions and was most likely to be an inner membrane protein.
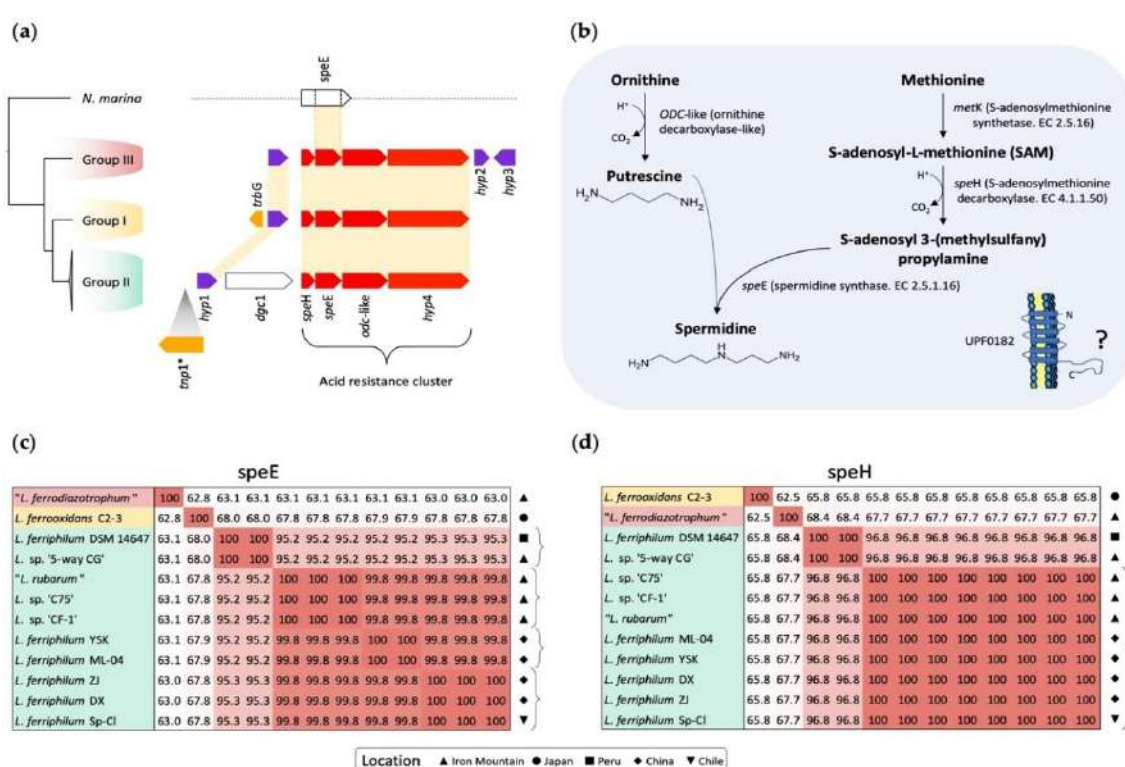

Figure 4. (**a,b**) Predicted spermidine biosynthesis genes and pathways of *Leptospirillum*. (**a**) Phylogenetic distribution and genomic contexts of spermidine biosynthesis genes. Color coding of genes: red = spermidine biosynthesis genes, purple = orphan hypothetical genes, orange = mobile elements and their remnants. * *tnp*1 was observed only in strain "CF-1" of Group II. (**b**) Predicted functions of the spermidine biosynthesis genes. (**c,d**) Heat maps of the nucleotide sequence identity of *spe*E (**c**) and *spe*H (**d**) in *Leptospirillum*. Red coloring indicates 100% nucleotide sequence identity, and numbers indicate the percent nucleotide identity.

An additional upstream gene, termed *dgc*1 (for diguanylate cyclase), was found only in Group II (Figure 4). This gene contained a predicted GGDEF domain and three associated GAF superfamily domains. In other organisms, the GGDEF domain is involved in cyclic diguanosine monophosphate turnover and the production of the secondary messenger C-di-GMP (reviewed in [110]). A functional GGDEF gene was reported in the extremely acidophilic genus *Acidithiobacillus*, although it was associated with the EAL rather than the GAF domains [111,112]. The secondary messenger C-di-GMP was implicated in the regulation of biofilm formation and other functions in many bacteria [113]. Spermidine has also been associated with both the formation and inhibition of biofilms in other bacteria [30,31].

All three *Leptospirillum* groups contained a hypothetical orphan gene (*hyp*1) upstream of the spermidine cluster, and Group III contained, in addition, two predicted hypothetical orphan genes downstream (*hyp*2 and *hyp*3; Figure 4a). Although *hyp*1 remains of unknown function, it has been identified in AMD community proteomes along with the full spermidine "acid resistance cluster" [52,55]. Transcripts for the spermidine genes have been detected in AMD community meta-transcriptomes [34].

Genes potentially involved in HGT and/or genome rearrangement were detected in the neighborhood of the spermidine cluster. These included a predicted P-type conjugative transfer protein TrbG with a signal sequence and lipoprotein signal and a TnpIS5-like sequence (Figure 4a).

Heat maps derived from alignments of DNA sequences of *spe*E and *spe*H in *Leptospirillum* Group II illustrated an important aspect of their evolution (Figure 4c,d). The DNA sequence of *spe*E was 100% conserved between some strains. For example, *spe*E of "*L. rubarum*", *L.* sp. "C75", and *L.* sp. "CF-1" from Iron Mountain, USA, shared 100% nucleotide sequence identity (Figure 4c). Inspection of the position of these strains in the phylogenetic cladogram (Figure 2b) suggested that *spe*E was inherited from their last common ancestor and its sequence subsequently maintained under strong selective pressure within the shared acidic environment of Iron Mountain. The close physical proximity of the strains could also facilitate genetic exchange and homologous recombination, contributing to the maintenance of DNA sequence similarity. The *spe*E of *L. ferriphilum* ZJ, DX, and Sp-Cl formed another cluster with 100% DNA sequence identity different from the Iron Mountain cluster (Figure 4c). Strains ZJ and DX were from China, whereas Sp-Cl was from Chile. In this case, close geographic proximity could not explain the sequence identity, and selective pressure resulting from a similar environment seemed more likely to account for the maintenance of sequence identity. A comparison of the nucleotide sequences of *spe*H showed 100% identity in strains of Group II *Leptospirillum* from Iron Mountain, China, and Chile, with two exceptions (Figure 4d). The exceptions were *L. ferriphilum* DSM 14647 (from Peru) and *L.* sp. '5-way CG' (from Iron Mountain) that formed a second cluster with 100% sequence identity.

It was hypothesized that geographical proximity could potentially explain some of the evolutionary trajectories of *spe*E and *spe*H, perhaps by maintenance of sequence identity via homologous recombination [104]. However, adaptation of vertically inherited genes to similar acidic econiches was a more likely explanation for those strains not geographically juxtaposed (e.g., *L. ferriphilum* Sp-Cl and *L. ferriphilum* DSM 14647).

3.4.3. Hopanoid Biosynthesis

*Hpn*CDE, potentially encoding squalene and a core set of hopanoid biosynthesis genes (*hpn*FGAHROP), were identified in *Leptospirillum* and *N. marina* (Figure 5). All these genes showed considerable syntenic conservation within all *Leptospirillum* Groups and *N. marina*, suggesting that they were inherited from a common ancestor by vertical descent. On the other hand, *hpn*IJ, encoding enzymes that modify hopanoids to bacteriohopanetetrol cyclitol ether, were predicted only in the genomes of *Leptospirillum*. It was hypothesized that *hpn*IJ entered the genome of the ancestral *Leptospirillum* after its divergence from the other *Nitrospira*. Mobile elements (*tnp*1-3) were detected in the neighborhood of the *hpn* gene cluster in *Leptospirillum* Group I (Figure 5b), suggesting that HGT of the cluster into *Leptospirillum* may have occurred. *Hpn*IJ have been shown in *Burkholderia* to be involved in C_{35} extensions of hopanoids including bacteriohopanetetrol (BHT), BHT glucosamine, and BHT cyclitol ether, which are in turn involved in the response to environmental stress conditions including low pH [29].
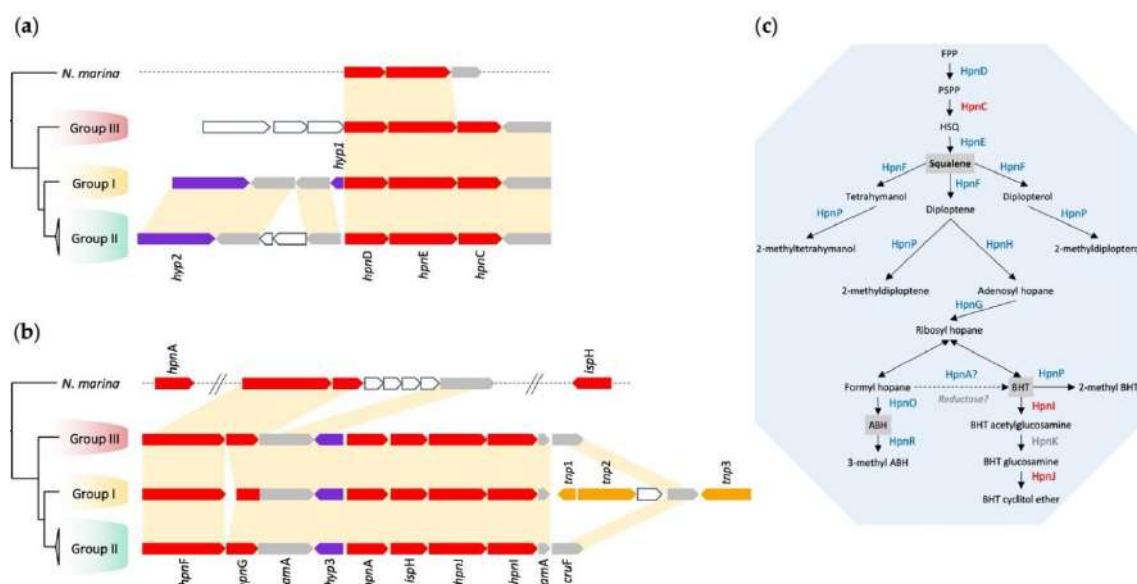
Figure 5. Genomic contexts for genes predicted to participate in (**a**) squalene and (**b**) hopanoid biosynthesis and bacteriohopanetetrol modification. Color coding of genes: red = squalene and hopanoid genes, grey = additional genes whose genomic context is conserved, purple = orphan hypothetical genes, orange = mobile elements and their remnants (*tnp*1-3), and white = other genes. * = *hpn*J1, *hpn*J2, and 3 are located in other genome locations. (**c**) Predicted pathways for squalene and hopanoid biosynthesis in *Leptospirillum* and *N. marina*. Blue lettering = biosynthetic steps encoded in both *Leptospirillum* and *N. marina*, and in red lettering, biosynthetic steps predicted only in *Leptospirillum*. Abbreviations: BHT = bacteriohopanetetrol, ABH = aminobacteriohopanetriol.

3.4.4. Slp Starvation Lipoprotein

Four copies of *slp* were identified in *N. marina* (termed *slp*s1-4) and an additional four copies were discovered in *Leptospirillum* (termed *slp*s5-8). Phylogenetic analysis of their amino acid sequences suggested that all eight copies were distinct and displayed different evolutionary trajectories (Supplementary Figure S4). It is possible that one or more of the *Leptospirillum slp*s were derived by vertical descent from the *slp*s of the inferred ancestor with *N. marina*. However, the long branch lengths derived from the phylogeny made it difficult to pin-point unambiguously the *Nitrospira* ancestral *slp* that gave rise to an ancestral *Leptospirillum slp*.

The evolutionary trajectories of the *slp*s in some clades of *Leptospirillum* could be explained by the similarity of geographic location. For example, *slp*6 and *slp*7 were found in clades belonging mainly to Iron Mountain, USA, and exhibited 100% amino acid sequence identity (Supplementary Figure S5). However, geographical proximity could not explain all trajectories. For example, *slp*5 and *slp*7 of *L. ferriphilum* DSM 14647 from Peru had 100% amino acid sequence identity with the Iron Mountain, USA, clade (Supplementary Figure S5). Furthermore, *slp*8 of *L. ferriphilum* DSM 14647 from Peru had 100% amino acid sequence identity with *slp*8 from a Chinese location (*L. ferriphilum* ML-04) and with *Leptospirillum* sp. "5-way CG" from Iron Mountain, USA. We concluded that the inheritance pattern of the *slp*s could be explained either by geographic proximity or by adaptation to similar acidic econiches similar to that postulated for *spe*E and *spe*H, as described in Section 3.4.2

Several deductions could be inferred from an inspection of the genomic contexts of the four *Leptospirillum slp*s (Figure 6): (i) All four *slp*s in each group displayed a different genomic context and, with the exception of *slp*5, also between species of the same group. The genomic context of *slp*5 was strongly conserved between Groups I and II and slightly conserved with Group III (Figure 6a). (ii) All four *slp*s were located near genes potentially encoding transposase-like functions, tRNAs, and phage-like genes, suggesting that they entered the genomes by HGT or underwent mobile element mediated rearrangement within these genomes. (iii) *slp*7 and *slp*8 were associated with *hpn*R and *kch*, respectively, genes potentially encoding other acid resistance functions (Figure 6b,d). (iv) All four *slp*s were associated with a number (18 in total) of orphan hypothetical genes with no known

function. Given the context of these genes, they may be related to acid resistance or other stress-related functions and can be highlighted for future experimental analysis. Alternatively, they could be unidentified phage remnants.
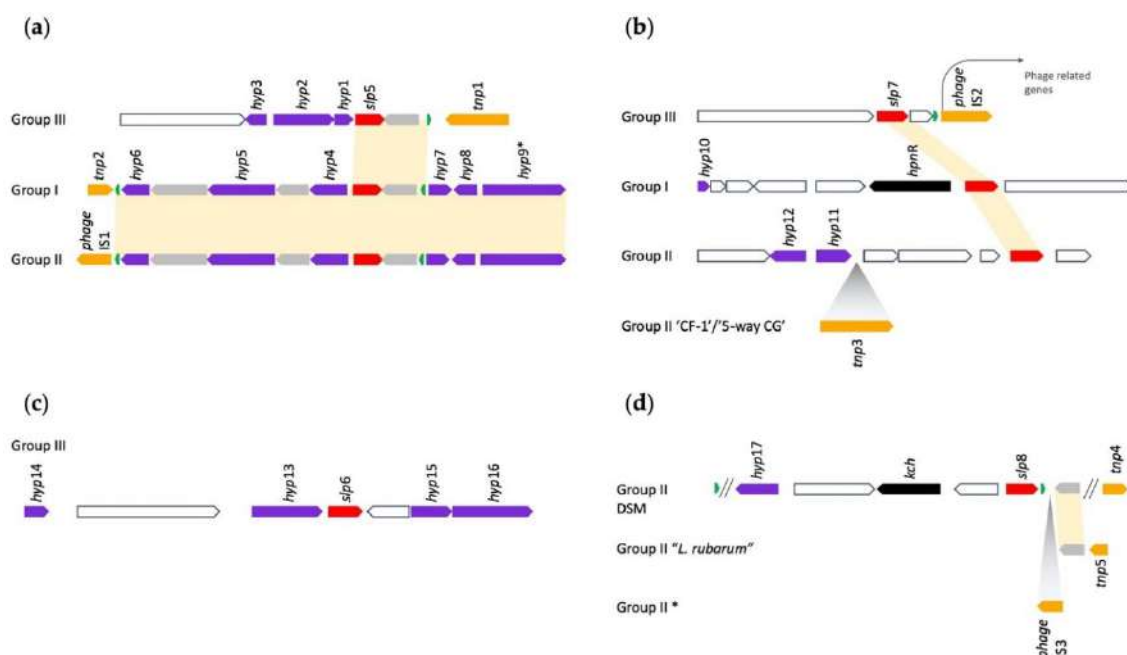

Figure 6. Genetic context of *slp* genes in *Leptospirillum*. (**a**) Genetic context for *slp*5 present in all *Leptospirillum*; (**b**) genetic context *slp*7 present in all *Leptospirillum*; (**c**) genetic context of *slp*6 present only in Group III; and (**d**) genetic context of *slp*8 present only Group II. Color coding of genes: red = *slp*, purple = orphan hypothetical gene, green = tRNA, gray = other conserved genes, orange = mobile elements, *tnp*1 (COG0675), *tnp*2 (PHA02517), *tnp*3 (pfam13701), *tnp*4 (COG3677), *tnp*5 (pfam13751), and *phage*IS, and black = *kch*, potentially involved in acid resistance (see Section 3.4.1). * = present in all strains of Group II except *L ferriphilum* DSM 14467 and "*L. rubarum*".

Although the function(s) of the *Leptospirillum slps* remain(s) unknown, all contain the lipobox motif that is characteristic of *slp*s in other organisms, at the end of a predicted signal peptide (Supplementary Figure S6) characteristic of *slp*s from other organisms [114]. The +2 position after the lipobox was proposed to be the main determinant for protein export such that if it had an Asp amino acid residue, then it was retained at the inner membrane (Supplementary Figure S6). An Asp was identified only in Slp5 from Group III, suggesting all the other proposed Slps might be exported to the outer membrane.

3.5. Second Line of Defense

3.5.1. Proton Antiporters

One gene copy of the putative voltage gated ClC-type chloride/proton antiporter (ClcA*)* was identified in all *Leptospirillum* genomes, but not in *N. marina*. The genomic region around *clc*A was not conserved in any of the *Leptospirillum* Groups. However, each genomic context of *clc*A was associated with mobile elements or their remains. For example, a transposase DDE domain (cl26088) and a DNA recombinase Rci/bacteriophage Hp1-like integrase (cd00796) were in its neighborhood in Groups III and II. Cluster analysis of the ClcA amino acid sequences showed that the evolutionary trajectory of ClcA followed the pattern of the 16S rRNA phylogeny and was most likely inherited by vertical descent within the *Leptospirillum* groups (Supplementary Figure S7). We suggested that the mobile elements associated with *clc*A could have been involved in its chromosomal relocation in each group.

A second mechanism postulated to remove protons from the cytoplasm involved the NhaP sodium/proton antiporter for which two copies, termed *nha*P1 and 2, were identified only in *Leptospirillum* Group II. A cluster analysis of their amino acid sequences suggested they were members of two different families (Supplementary Figure S8). *Nha*P1 was located in a conserved genomic context in all members of Group II. Remnants of a number of transposases and integrases together with tRNA-Arg were detected in the gene neighborhood, suggesting that *nha*P1 was acquired by HGT in an ancestral *Leptospirillum* after its divergence with the last common ancestor. However, the conserved genomic context is consistent with the idea that it was inherited vertically within the different groups after the initial HGT event. A copy of *nha*P2 was found only in *L. ferriphilum* ML-04 (Group II). However, its sequence was interrupted by a transposase insertion (*tnp*1; Figure 7a). It was unlikely that *nha*P2 was functional because the transposase insertion introduced stop codons in its reading frame and split the functional NhaP domain (COG0025) into two parts.
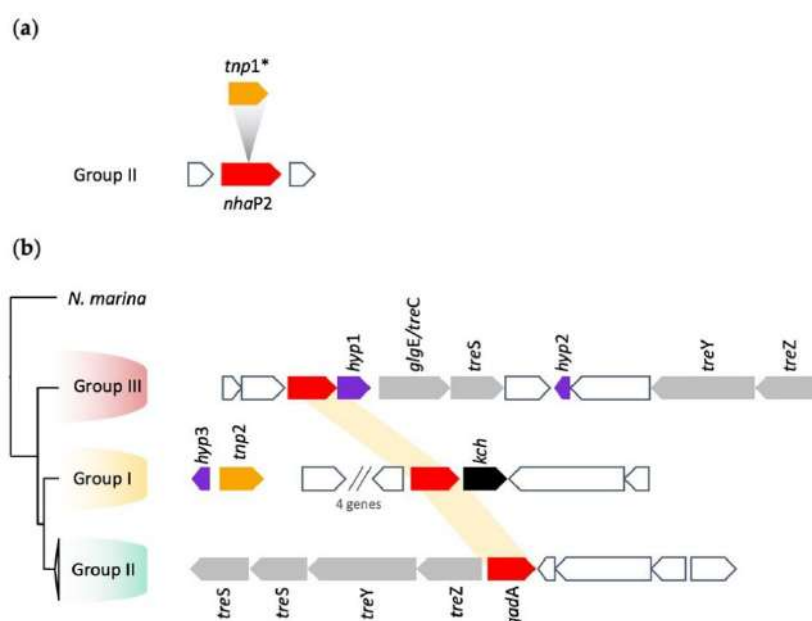

Figure 7. Gene context of (**a**) *nha*P2 showing conservation in all strains of Group II. Color coding of genes: red = *nha*P2, orange mobile element *tnp*1 (transposase DDE_3; pfam13358) found only in strain ML-04. (**b**) Phylogenetic distribution of *gad*A in all *Leptospirillum* Groups. Color coding of genes: red = *gad*A, purple = orphan hypothetical genes, grey = conserved genes predicted to be involved in trehalose biosynthesis, orange = mobile element (*tnp*2 (PHA02517), black = *kch*, potentially involved in acid resistance (see Section 3.4.1), and white = other genes.

3.5.2. Gad Decarboxylase

Four copies of the acid resistance amino acid permease (*gad*C) and one copy of the acid resistance amino acid decarboxylase (*gad*A) were identified in all *Leptospirillum*. In contrast, *N. marina* contained only one putative amino acid permease. The glutamate decarboxylase (*gad*A) did not show synteny in its genome context between the three groups (Figure 7b), and a cluster analysis suggested that the gene was introduced by HGT from Archaea (in separate events, as the top hits for the three groups were different; Supplementary Figure S9). In addition, four predicted *gad*C amino acid permeases encoding genes were identified in a separate genomic location to *gad*A with one of them in the gene context of potassium transporter Kdp system (Figure 3). Cluster analysis suggested one *gad*C copy was similar to the one found in the *N. marina* genome with three additional families present in the *Leptospirillum* genomes (Supplementary Figure S10).

*Gad*A (glutamate decarboxylase) in *Leptospirillum* Groups II and III was associated with a cluster of trehalose biosynthesis genes (Figure 7b). It has been shown that potassium, glutamate, and trehalose form part of a response to osmotic shock and acid stress in *E. coli* (reviewed in [115]), suggesting that a similar response was possible in *Leptospirillum* Groups II and III.

3.6. Model of Leptospirillum Acid Resistance

A model of the *Leptospirillum* acid resistance systems, classified into first and second lines of defense mechanisms, is shown in Figure 8. Transcriptomic and proteomic analyses supported the relationship of the first line of defense genes with low pH adaptation in *Leptospirillum* [11,34,52,55,116]. Evidence also linked the expression of the KdpABC high-affinity potassium transport system and the HpnCDEFGIJHGRP and BamA hopanoid system to acid stress in AMD communities [34,52,55,116]. Environmental expression of genes involved in the second line of defense, such as the glutamate decarboxylase system *gad* and the Na+/H+ antiporter *nha*P, has been detected [34,116,117]. Clearly, additional experiments are required to explore the validity of the model. However, we posit that it provides a platform for helping to circumscribe future experimental directions.
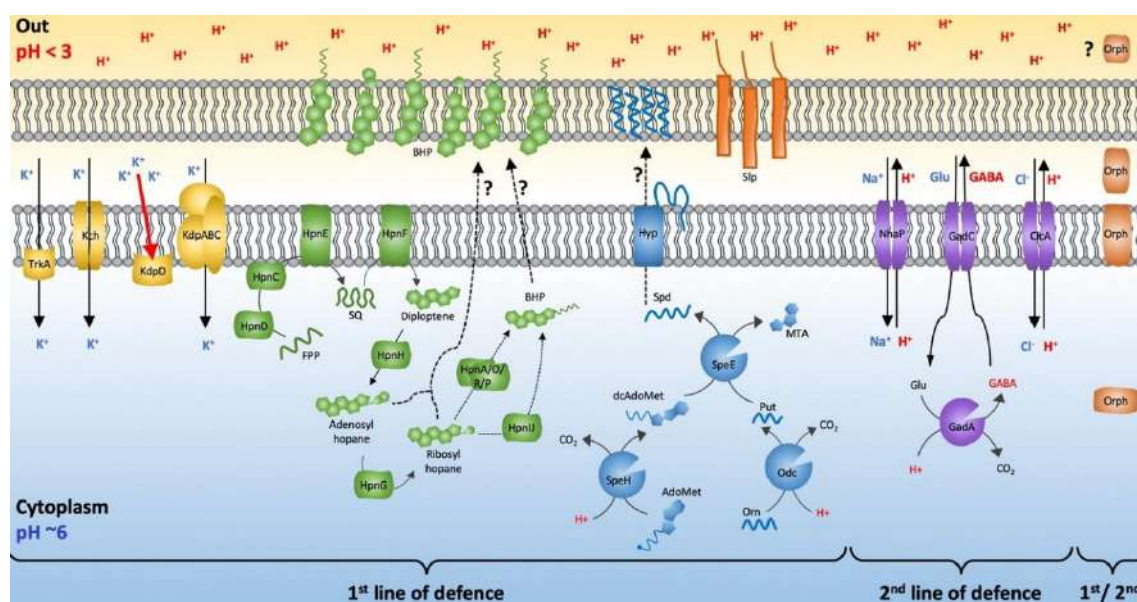

Figure 8. Model of acid resistance in *Leptospirillum*, categorized as either first line of defense involved in preventing the uptake of protons into the cell or second line of defense involved in neutralizing or expelling protons that enter the cell. Multiple hypothetical orphan genes (Orph) cluster with both first and second line of defense genes. The red arrow indicates that expression of KdpABC is activated by external K+ sensed by KdpD. The cytoplasm is conjectured to be about pH 6, as has been found in other extreme acidophiles, although this has not been experimentally determined in *Leptospirillum*.

Three genomic regions of *Leptospirillum* contained potassium transport system *kch* genes in proximity with other acid resistance genes, for example, *gad*C2 (Figure 3), *slp* (Figure 6), and *gad*A (Figure 7). This could allow their coordinated regulation. Each of these systems was associated with multiple mobile elements, suggesting that they could have entered the genome by HGT. Acid resistance genes were also found in close proximity to other stress responsive genes such as those involved in biofilm formation (Figure 4) and trehalose biosynthesis (Figure 7), potentially allowing coordination of genes involved in acid resistance and osmotic stress.

Whereas the inventory of potential mechanisms involved in first and second lines of acid resistance in *Leptospirillum* was quite extensive, they were by no means the only ones used by organisms for acid resistance. For example, in *Leptospirillum*, protons were hypothesized to be exported by the glutamate decarboxylase system, but evolutionary and mechanistically related systems such as ornithine decarboxylases have been implicated in acid stress responses in amateur acidophiles [86]. The extreme acidophile *Ferrovum*, belonging to the Betaproteobacteria class, has been hypothesized to use the Kef-type K+ transport system to help in maintaining a positive inside membrane potential and to utilize urease activity to neutralize its immediate environment [118]. External cellular capsule formation has been speculated to be involved in acid resistance in the Acidithiobacillia class [119]. None of these systems were predicted in *Leptospirillum*. Given that

multiple acid resistance mechanisms were found in different Bacterial classes widely distributed in the Tree of Life, it was most likely that acid resistance evolved independently multiple times, perhaps aided by HGT. A similar conclusion has been made regarding the evolution of acid resistance in Archaea [120].

Genes encoding orphan hypothetical proteins located in genomic contexts associated with both first and second line of defense acid resistance genes may potentially encode unknown acid resistance mechanisms or functions that modify known acid resistance responses. Their bioinformatic prediction highlighted the need for experimental investigation into their functions.

3.7. Phylogenetic Distribution of Acid Resistance Genes and Their Inferred Evolutionary Trajectories

Evolutionary events leading to acid resistance in the *Leptospirillum* genus were inferred using parsimony bioinformatic methods [121] and were mapped onto the branches of the phylogenetic tree of *Leptospirillum* (Figure 9).
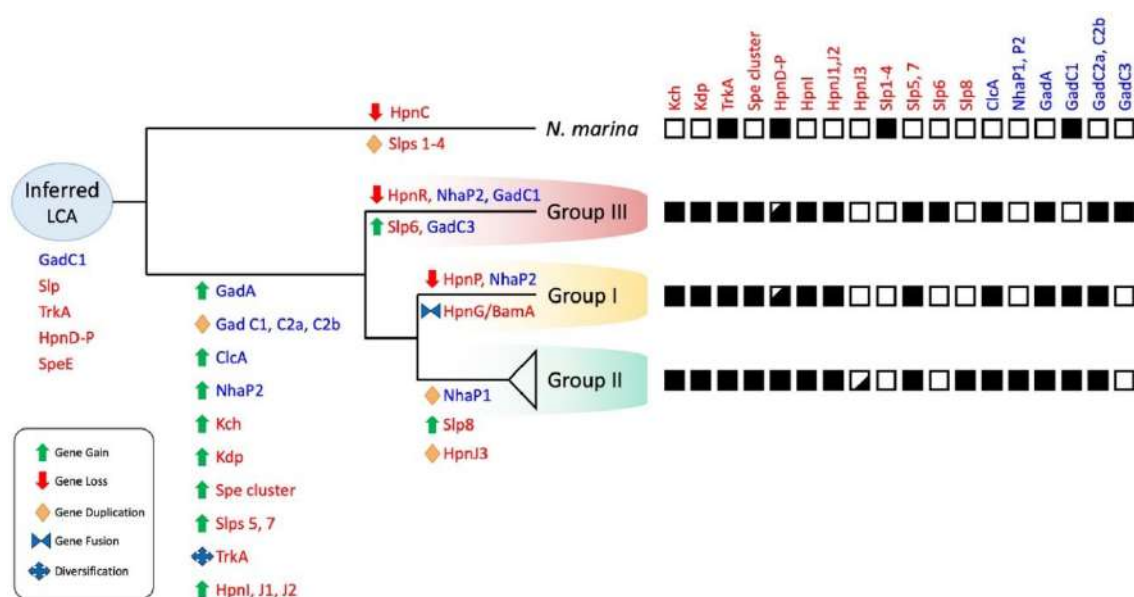

Figure 9. Inferred evolutionary reconstruction of the acid related genes of *Leptospirillum* with the main evolutionary events (gene gain/loss/fusion/duplication). Parsimony was used for the inferences. Names in red = first line of defense mechanisms and blue = second line of defense mechanisms. Black square = presence and white square = absence of mechanism. Black squares with white sections = mechanism is present in some, but not all, strains. LCA = last common ancestor.

Many of the predicted acid resistance genes were hypothesized to have been absent in the inferred last common ancestor of *Leptospirillum* and *N. marina*. These included genes encoding the K$^+$ transporters Kch and Kdp, the four spermidine gene clusters, HpnIJ, ClcA, NhaP, and GadC2 and 3. It was proposed that they entered the *Leptospirillum* ancestral line by HGT via conjugation (e.g., the spermidine four-gene cluster), viruses (e.g., *slp*s7 and 8, *clc*A), and multiple examples involving transposases. HGT has been suggested to be a prevalent mechanism in genome evolution in a wide range of microorganisms [122].

With two exceptions, the donors of the HGT genes were difficult to trace, perhaps because the events occurred so long ago that molecular signals of the donors have been erased with the passage of time. The *Leptospirillum* Kch potassium transporter had several top BLASTp hits with other extreme acidophiles, including the *Acidithiobacillus* genus which shares its low pH environment. *Leptospirillum* is rooted deeper in the Tree of Life than *Acidithiobacillus* and therefore was probably ancestral to it, suggesting that the direction of transfer of Kch was from a *Leptospirillum* ancestor to an *Acidithiobacillus* ancestor. A second example of possible HGT donor identification lied in the comparison of the "*L. ferrodiazotrophum*" and "*L. rubarum*" glutamate decarboxylase acid resistance

system with other members of the AMD community, suggesting that the genes were introduced by HGT into a *Leptospirillum* ancestor from an Archaeal *Ferroplasma* ancestor.

Gene duplication and gene diversification events were identified using a combination of phylogenetic inference based on alignments of families (cluster analysis) [123] and calculations of D_n/D_s [94–96]. Multiple examples of gene duplication events were detected. These included examples of potential vertical descent followed by gene duplication giving rise to paralogs (e.g., HpnJ1-3 in Group II). There were also many cases of gene duplications that were predicted to be xenologs, arising from HGT events. Xenologs were defined as a distinct form of horizontal gene transfer in which a gene was displaced by an ortholog from a different lineage [124], e.g., *slp*s5-8 in *Leptospirillum* replaced *slp*s present in the last common ancestor.

The relative timing of the events leading to the hypothesized transition of the ancestral *Leptospirillum* from a circumneutral or mildly acidic environment to a hyper-acidic one was difficult to assess. Based on what was known about acid stress response mechanisms of moderate ("amateur") acidophiles (pH 3.5–6), it seemed likely that the first transition events involved the development of second line of defense mechanisms including proton expulsion mechanisms such as ClcA, NhaP, and Gad. However, it could not be ruled out that the mechanisms of the first line of defense were also involved in the early transition to very low pH environments. Some of these, such as hopanoids and spermidine, could have been gained initially to provide protection from other stresses such as oxidative stress or high temperature and subsequently adapted for low pH protection.

This paper focused on the potential mechanisms employed by *Leptospirillum* to thrive in extremely low pH environments. One of the major aspects of adaptation that was not investigated was how proteins fold and function in acid conditions. The cytoplasm was hypothesized to be around pH 6 or circumneutral, as was shown for other acidophiles, although this was not experimentally verified for *Leptospirillum*. If this assumption were correct, then only proteins or protein loops outside the periplasmic membrane would be exposed to low pH. Protein adaptations to low pH have been investigated in other acidophiles (e.g., [125,126]), but since no information was available for *Leptospirillum*, the model of its hypothesized transition from a neutral to an acid environment remains incomplete. Other major lacunae in our knowledge of the evolution of acidophilia in *Leptospirillum* were how changes in pH were sensed and transduced into gene regulation and how chaperones could be involved in maintaining protein integrity.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Slonczewski, J.L.; Fujisawa, M.; Dopson, M.; Krulwich, T.A. Cytoplasmic pH measurement and homeostasis in bacteria and archaea. *Adv. Microb. Physiol.* **2009**, *55*, 1–79.

2. Tucker, D.L.; Tucker, N.; Conway, T. Gene expression profiling of the pH response in *Escherichia coli*. *J. Bacteriol.* **2002**, *184*, 6551–6558, doi:10.1128/jb.184.23.6551-6558.2002.

3. Kanjee, U.; Houry, W.A. Mechanisms of acid resistance in *Escherichia coli*. *Annu. Rev. Microbiol.* **2013**, *67*, 65–81, doi:10.1146/annurev-micro-092412-155708.

4. Krulwich, T.A.; Sachs, G.; Padan, E. Molecular aspects of bacterial pH sensing and homeostasis. *Nat. Rev. Microbiol.* **2011**, *9*, 330–343, doi:10.1038/nrmicro2549.

5. Foster, J.W. *Escherichia coli* acid resistance: Tales of an amateur acidophile. *Nat. Rev. Microbiol.* **2004**, *2*, 898–907.

6. Zhao, B.; Houry, W.A. Acid stress response in enteropathogenic Gammaproteobacteria: An aptitude for survival. *Biochem. Cell Biol.* **2010**, *88*, 301–314, doi:10.1139/o09-182.

7. *Acidophiles: Life in Extremely Acidic Environments*; Quatrini, R., Johnson, D.B., Eds.; Caister Academic Press: Norfolk, UK, 2016; doi:10.21775/9781910190333.

8. Baker-Austin, C.; Dopson, M. Life in acid: pH homeostasis in acidophiles. *Trends Microbiol.* **2007**, *15*, 165–171.

9. Hedrich, S.; Schippers, A. Distribution of acidophilic microorganisms in natural and man-made acidic enviroments. In *Acidophiles: Life in Extremely Acidic Environments*; Quatrini, R., Johnson, D.B., Eds.; Caister Academic Press: Norfolk, UK, 2016; pp. 153–175.

10. Quatrini, R.; Johnson, D.B. Microbiomes in extremely acidic environments: Functionalities and interactions that allow survival and growth of prokaryotes at low pH. *Curr. Opin. Microbiol.* **2018**, *43*, 139–147, doi:10.1016/j.mib.2018.01.011.

11. Zammit, C.M.; Watkin, E.L.J. Adaptation to extreme acidity and osmotic stress. In *Acidophiles: Life in Extremely Acidic Environments*; Quatrini, R., Johnson, D.B., Eds.; Caister Academic Press: Norfolk, UK, 2016; pp. 49–62.

12. Buetti-Dinh, A.; Dethlefsen, O.; Friedman, R.; Dopson, M. Transcriptomic analysis reveals how a lack of potassium ions increases *Sulfolobus acidocaldarius* sensitivity to pH changes. *Microbiology (Read. Engl.)* **2016**, *162*, 1422–1434, doi:10.1099/mic.0.000314.

13. Christel, S.; Herold, M.; Bellenberg, S.; El Hajjami, M.; Buetti-Dinh, A.; Pivkin, I.V.; Sand, W.; Wilmes, P.; Poetsch, A.; Dopson, M. Multi-omics reveals the lifestyle of the acidophilic, mineral-oxidizing model species *Leptospirillum ferriphilum*^{T}. *Appl. Environ. Microbiol.* **2018**, *84*, e02091-17, doi:10.1128/aem.02091-17.

14. Cholo, M.C.; van Rensburg, E.J.; Osman, A.G.; Anderson, R. Expression of the genes encoding the Trk and Kdp potassium transport systems of *Mycobacterium tuberculosis* during growth *in vitro*. *Biomed. Res. Int.* **2015**, *2015*, 608682, doi:10.1155/2015/608682.

15. Huang, C.S.; Pedersen, B.P.; Stokes, D.L. Crystal structure of the potassium-importing KdpFABC membrane complex. *Nature* **2017**, *546*, 681–685, doi:10.1038/nature22970.

16. Epstein, W. The roles and regulation of potassium in bacteria. *Prog. Nucleic Acid Res. Mol. Biol.* **2003**, *75*, 293–320, doi:10.1016/s0079-6603(03)75008-9.

17. Schäfer, G. Bioenergetics of the archaebacterium *Sulfolobus*. *Biochim. Et Biophys. Acta* **1996**, *1277*, 163–200.

18. Suzuki, I.; Lee, D.; Mackay, B.; Harahuc, L.; Oh, J.K. Effect of various ions, pH, and osmotic pressure on oxidation of elemental sulfur by *Thiobacillus thiooxidans*. *Appl. Environ. Microbiol.* **1999**, *65*, 5163–5168.

19. Goulbourne, E.J.; Matin, M.; Zychlinsky, E. Mechanism of delta pH maintenance in active and inactive cells of an obligately acidophilic bacterium. *J. Bacteriol.* **1986**, *166*, 59–65.

20. Rhee, H.J.; Kim, E.J.; Lee, J.K. Physiological polyamines: Simple primordial stress molecules. *J. Cell. Mol. Med.* **2007**, *11*, 685–703, doi:10.1111/j.1582-4934.2007.00077.x.

21. Shah, D.H.; Casavant, C.; Hawley, Q.; Addwebi, T.; Call, D.R.; Guard, J. *Salmonella Enteritidis* strains from poultry exhibit differential responses to acid stress, oxidative stress, and survival in the egg albumen. *Foodborne Pathog. Dis.* **2012**, *9*, 258–264, doi:10.1089/fpd.2011.1009.

22. Watson, N.; Dunyak, D.S.; Rosey, E.L.; Slonczewski, J.L.; Olson, E.R. Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. *J. Bacteriol.* **1992**, *174*, 530–540, doi:10.1128/jb.174.2.530-540.1992.

23. Dela Vega, A.L.; Delcour, A.H. Polyamines decrease *Escherichia coli* outer membrane permeability. *J. Bacteriol.* **1996**, *178*, 3715–3721, doi:10.1128/jb.178.13.3715-3721.1996.

24. Hosoya, R.; Yokoyama, Y.; Hamana, K.; Itoh, T. Polyamine analysis within the eubacterial thirteen phyla Acidobacteria, Actinobacteria, Chlorobi, Chloroflexi, Chrysiogenetes, Deferribacteres, Fibrobacteres, Firmicutes, Fusobacteria, Gemmatimonadetes, Nitrospirae, Planctomycetes, and Verrucomicrobia. *Microbiol. Cult. Coll.* **2006**, *22*, 21–33.

25. Li, Q.; Li, N.; Liu, X.; Zhou, Z.; Fang, Y.; Fan, X.; Fu, X.; Liu, Y.D.; Yin, H.Q. Characterization of the acid stress response of *Acidithiobacillus ferrooxidans* ATCC 23270 based on the method of microarray. *J. Biol. Res.-Thessal.* **2012**, *17*, 3–15.

26. Belin, B.J.; Busset, N.; Giraud, E.; Molinaro, A.; Silipo, A.; Newman, D.K. Hopanoid lipids: From membranes to plant–bacteria interactions. *Nat. Rev. Microbiol.* **2018**, *16*, 304, doi:10.1038/nrmicro.2017.173.

27. Sohlenkamp, C.; Geiger, O. Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiol. Rev.* **2016**, *40*, 133–159, doi:10.1093/femsre/fuv008.

28. Welander, P.V.; Hunter, R.C.; Zhang, L.; Sessions, A.L.; Summons, R.E.; Newman, D.K. Hopanoids play a role in membrane integrity and pH homeostasis in *Rhodopseudomonas palustris* TIE-1. *J. Bacteriol.* **2009**, *191*, 6145–6156, doi:10.1128/jb.00460-09.

29. Schmerk, C.L.; Welander, P.V.; Hamad, M.A.; Bain, K.L.; Bernards, M.A.; Summons, R.E.; Valvano, M.A. Elucidation of the *Burkholderia cenocepacia* hopanoid biosynthesis pathway uncovers functions for conserved proteins in hopanoid-producing bacteria. *Environ. Microbiol.* **2015**, *17*, 735–750, doi:10.1111/1462-2920.12509.

30. Kulkarni, G.; Wu, C.H.; Newman, D.K. The general stress response factor EcfG regulates expression of the C-2 hopanoid methylase HpnP in *Rhodopseudomonas palustris* TIE-1. *J. Bacteriol.* **2013**, *195*, 2490–2498, doi:10.1128/jb.00186-13.

31. Pan, J.J.; Solbiati, J.O.; Ramamoorthy, G.; Hillerich, B.S.; Seidel, R.D.; Cronan, J.E.; Almo, S.C.; Poulter, C.D. Biosynthesis of squalene from farnesyl diphosphate in bacteria: Three steps catalyzed by three enzymes. *ACS Cent. Sci.* **2015**, *1*, 77–82, doi:10.1021/acscentsci.5b00115.

32. Jones, D.S.; Albrecht, H.L.; Dawson, K.S.; Schaperdoth, I.; Freeman, K.H.; Pi, Y.; Pearson, A.; Macalady, J.L. Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. *ISME J.* **2012**, *6*, 158–170, doi:10.1038/ismej.2011.75.

33. Mangold, S.; Rao Jonna, V.; Dopson, M. Response of *Acidithiobacillus caldus* toward suboptimal pH conditions. *Extremophiles* **2013**, *17*, 689–696, doi:10.1007/s00792-013-0553-5.

34. Chen, L.X.; Hu, M.; Huang, L.N.; Hua, Z.S.; Kuang, J.L.; Li, S.J.; Shu, W.S. Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME J.* **2015**, *9*, 1579–1592, doi:10.1038/ismej.2014.245.

35. Hommais, F.; Krin, E.; Coppee, J.Y.; Lacroix, C.; Yeramian, E.; Danchin, A.; Bertin, P. GadE (YhiE): A novel activator involved in the response to acid environment in *Escherichia coli*. *Microbiology (Read. Engl.)* **2004**, *150*, 61–72, doi:10.1099/mic.0.26659-0.

36. Merrell, D.S.; Camilli, A. Acid tolerance of gastrointestinal pathogens. *Curr. Opin. Microbiol.* **2002**, *5*, 51–55, doi:10.1016/s1369-5274(02)00285-0.

37. Alexander, D.M.; St John, A.C. Characterization of the carbon starvation-inducible and stationary phase-inducible gene *slp* encoding an outer membrane lipoprotein in *Escherichia coli*. *Mol. Microbiol.* **1994**, *11*, 1059–1071, doi:10.1111/j.1365-2958.1994.tb00383.x.

38. Mates, A.K.; Sayed, A.K.; Foster, J.W. Products of the *Escherichia coli* acid fitness island attenuate metabolite stress at extremely low pH and mediate a cell density-dependent acid resistance. *J. Bacteriol.* **2007**, *189*, 2759–2568, doi:10.1128/jb.01490-06.

39. Alexander, B.; Leach, S.; Ingledew, W.J. The relationship between chemiosmotic parameters and sensitivity to anions and organic acids in the acidophile *Thiobacillus ferrooxidans*. *J. Gen. Microbiol.* **1987**, *133*, 1171–1179.

40. Padan, E.; Tzubery, T.; Herz, K.; Kozachkov, L.; Rimon, A.; Galili, L. NhaA of *Escherichia coli*, as a model of a pH-regulated Na^+/H^+ antiporter. *Biochim. Biophys. Acta-Bioenerg.* **2004**, *1658*, 2–13, doi:10.1016/j.bbabio.2004.04.018.

41. Accardi, A.; Walden, M.; Nguitragool, W.; Jayaram, H.; Williams, C.; Miller, C. Separate ion pathways in a Cl^-/H^+ exchanger. *J. Gen. Physiol.* **2005**, *126*, 563–570, doi:10.1085/jgp.200509417.

42. Iyer, R.; Iverson, T.M.; Accardi, A.; Miller, C. A biological role for prokaryotic ClC chloride channels. *Nature* **2002**, *419*, 715–718, doi:10.1038/nature01000.

43. Richard, H.; Foster, J.W. *Escherichia coli* glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *J. Bacteriol.* **2004**, *186*, 6032–6041, doi:10.1128/jb.186.18.6032-6041.2004.

44. McLaggan, D.; Keyhan, M.; Matin, A. Chloride transport pathways and their bioenergetic implications in the obligate acidophile *Bacillus coagulans*. *J. Bacteriol.* **1990**, *172*, 1485–1490.

45. Feehily, C.; Karatzas, K.A. Role of glutamate metabolism in bacterial responses towards acid and other stresses. *J. Appl. Microbiol.* **2013**, *114*, 11–24, doi:10.1111/j.1365-2672.2012.05434.x.

46. Mi, S.; Song, J.; Lin, J.; Che, Y.; Zheng, H.; Lin, J. Complete genome of *Leptospirillum ferriphilum* ML-04 provides insight into its physiology and environmental adaptation. *J. Microbiol.* **2011**, *49*, 890–901.

47. Coram, N.J.; Rawlings, D.E. Molecular relationship between two groups of the genus *Leptospirillum* and the finding that *Leptospirillum ferriphilum* sp. nov. dominates South African commercial biooxidation tanks that operate at 40°C. *Appl. Environ. Microbiol.* **2002**, *68*, 838–845.

48. Hippe, H. *Leptospirillium* gen. nov (ex Markoysan 1972), nom. rev., including *Leptospirillium ferrooxidans* sp nov (ex Markoysan 1972), nom. rev. and *Leptospirilliium thermoferrooxidans* sp nov (Golovacheva et al. 1992). *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 501–503.

49. Bond, P.L.; Smriga, S.P.; Banfield, J.F. Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl. Environ. Microbiol.* **2000**, *66*, 3842–3849.

50. Ram, R.J.; VerBerkmoes, N.C.; Thelen, M.P.; Tyson, G.W.; Baker, B.J.; Blake, R.C.; Shah, M.; Hettich, R.L.; Banfield, J.F. Community proteomics of a natural microbial biofilm. *Science (N. Y. NY)* **2005**, *308*, 1915–1920.

51. Tyson, G.W.; Lo, I.; Baker, B.J.; Allen, E.E.; Hugenholtz, P.; Banfield, J.F. Genome-directed isolation of the key nitrogen fixer *Leptospirillum ferrodiazotrophum* sp. nov. from an acidophilic microbial community. *Appl. Environ. Microbiol.* **2005**, *71*, 6319–6324, doi:10.1128/AEM.71.10.6319-6324.2005.

52. Goltsman, D.S.; Denef, V.J.; Singer, S.W.; VerBerkmoes, N.C.; Lefsrud, M.; Mueller, R.S.; Dick, G.J.; Sun, C.L.; Wheeler, K.E.; Zemla, A.; et al. Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing "*Leptospirillum rubarum*" (Group II) and "*Leptospirillum ferrodiazotrophum*" (Group III) bacteria in acid mine drainage biofilms. *Appl. Environ. Microbiol.* **2009**, *75*, 4599–4615, doi:10.1128/AEM.02943-08.

53. Wilmes, P.; Remis, J.P.; Hwang, M.; Auer, M.; Thelen, M.P.; Banfield, J.F. Natural acidophilic biofilm communities reflect distinct organismal and functional organization. *ISME J.* **2009**, *3*, 266–270, doi:10.1038/ismej.2008.90.

54. Belnap, C.P.; Pan, C.; VerBerkmoes, N.C.; Power, M.E.; Samatova, N.F.; Carver, R.L.; Hettich, R.L.; Banfield, J.F. Cultivation and quantitative proteomic analyses of acidophilic microbial communities. *ISME J.* **2010**, *4*, 520–530, doi:10.1038/ismej.2009.139.

55. Belnap, C.P.; Pan, C.; Denef, V.J.; Samatova, N.F.; Hettich, R.L.; Banfield, J.F. Quantitative proteomic analyses of the response of acidophilic microbial communities to different pH conditions. *ISME J.* **2011**, *5*, 1152–1161, doi:10.1038/ismej.2010.200.

56. Saglam, E.S.; Akcay, M.; Colak, D.N.; Inan Bektas, K.; Belduz, A.O. Generation of acid mine drainage around the Karaerik copper mine (Espiye, Giresun, NE Turkey): Implications from the bacterial population in the Acisu effluent. *Extremophiles* **2016**, *20*, 673–685, doi:10.1007/s00792-016-0857-3.

57. Schrenk, M.O.; Edwards, K.J.; Goodman, R.M.; Hamers, R.J.; Banfield, J.F. Distribution of *Thiobacillus ferrooxidans* and *Leptospirillium ferrooxidans*: Implications for generation of acid mine drainage. *Science (N. Y. NY)* **1998**, *279*, 1519–1522.

58. Baker, B.J.; Banfield, J.F. Microbial communities in acid mine drainage. *FEMS Microbiol. Ecol.* **2003**, *44*, 139–152.

59. Issotta, F.; Galleguillos, P.A.; Moya-Beltrán, A.; Davis-Belmar, C.S.; Rautenbach, G.; Covarrubias, P.C.; Acosta, M.; Ossandon, F.J.; Contador, Y.; Holmes, D.S.; et al. Draft genome sequence of chloride-tolerant *Leptospirillum ferriphilum* Sp-Cl from industrial bioleaching operations in northern Chile. *Stand. Genom. Sci.* **2016**, *11*, 1–7, doi:10.1186/s40793-016-0142-1.

60. 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 Res.* **2015**, *25*, 1043–1055, doi:10.1101/gr.186072.114.

61. Pruesse, E.; Quast, C.; Knittel, K.; Fuchs, B.M.; Ludwig, W.G.; Peplies, J.; Glockner, F.O. SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **2007**, *35*, 7188–7196, doi:Doi 10.1093/Nar/Gkm864.

62. Cole, J.R.; Wang, Q.; Fish, J.A.; Chai, B.; McGarrell, D.M.; Sun, Y.; Brown, C.T.; Porras-Alfaro, A.; Kuske, C.R.; Tiedje, J.M. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **2014**, *42*, D633-642, doi:10.1093/nar/gkt1244.

63. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072, doi:Doi 10.1128/Aem.03006-05.

64. Turenne, C.Y.; Tschetter, L.; Wolfe, J.; Kabani, A. Necessity of quality-controlled 16S rRNA gene sequence databases: Identifying nontuberculous *Mycobacterium* species. *J. Clin. Microbiol.* **2001**, *39*, 3637–3648, doi:10.1128/jcm.39.10.3638-3648.2001.

65. Katoh, K.; Misawa, K.; Kuma, K.i.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066, doi:10.1093/nar/gkf436.

66. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780, doi:10.1093/molbev/mst010.

67. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797, doi:Doi 10.1093/Nar/Gkh340.

68. Di Tommaso, P.; Moretti, S.; Xenarios, I.; Orobitg, M.; Montanyola, A.; Chang, J.M.; Taly, J.F.; Notredame, C. T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* **2011**, *39*, W13–W17, doi:10.1093/nar/gkr245.

69. 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. *Mol. Biol. Evol.* **2015**, *32*, 268–274, doi:10.1093/molbev/msu300.

70. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589, doi:10.1038/nmeth.4285.

71. Hoang, D.T.; Chernomor, O.; von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **2018**, *35*, 518–522, doi:10.1093/molbev/msx281.

72. Riadi, G.; Medina-Moenne, C.; Holmes, D.S. TnpPred: A web service for the robust prediction of prokaryotic transposases. *Comp. Funct. Genom.* **2012**, *2012*, 5, doi:10.1155/2012/678761.

73. Siguier, P.; Perochon, J.; Lestrade, L.; Mahillon, J.; Chandler, M. ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Res.* **2006**, *34*, D32-36, doi:10.1093/nar/gkj014.

74. Varani, A.M.; Siguier, P.; Gourbeyre, E.; Charneau, V.; Chandler, M. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol* **2011**, *12*, R30, doi:10.1186/gb-2011-12-3-r30.

75. Waack, S.; Keller, O.; Asper, R.; Brodag, T.; Damm, C.; Fricke, W.F.; Surovcik, K.; Meinicke, P.; Merkl, R. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. *BMC Bioinform.* **2006**, *7*, 142, doi:10.1186/1471-2105-7-142.

76. Bertelli, C.; Laird, M.R.; Williams, K.P.; Simon Fraser University Research Computing Group; Lau, B.Y.; Hoad, G.; Winsor, G.L.; Brinkman, F.S.L. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* **2017**, *45*, W30–W35, doi:10.1093/nar/gkx343.

77. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613, doi:10.1093/nar/gky1131.

78. Darling, A.E.; Mau, B.; Perna, N.T. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **2010**, *5*, e11147, doi:10.1371/journal.pone.0011147.

79. Carver, T.; Harris, S.R.; Berriman, M.; Parkhill, J.; McQuillan, J.A. Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics (Oxf. Engl.)* **2012**, *28*, 464–469.

80. Casiano-Colon, A.; Marquis, R.E. Role of the arginine deiminase system in protecting oral bacteria and an enzymatic basis for acid tolerance. *Appl. Environ. Microbiol.* **1988**, *54*, 1318–1324.

81. Gajiwala, K.S.; Burley, S.K. HDEA, a periplasmic protein that supports acid resistance in pathogenic enteric bacteria. *J. Mol. Biol.* **2000**, *295*, 605–612, doi:10.1006/jmbi.1999.3347.

82. Jeong, K.C.; Hung, K.F.; Baumler, D.J.; Byrd, J.J.; Kaspar, C.W. Acid stress damage of DNA is prevented by Dps binding in *Escherichia coli* O157:H7. *BMC Microbiol.* **2008**, *8*, 181, doi:10.1186/1471-2180-8-181.

83. Leverrier, P.; Vissers, J.P.; Rouault, A.; Boyaval, P.; Jan, G. Mass spectrometry proteomic analysis of stress adaptation reveals both common and distinct response pathways in *Propionibacterium freudenreichii*. *Arch. Microbiol.* **2004**, *181*, 215–230, doi:10.1007/s00203-003-0646-0.

84. Lund, P.; Tramonti, A.; De Biase, D. Coping with low pH: Molecular strategies in neutralophilic bacteria. *FEMS Microbiol. Rev.* **2014**, *38*, 1091–1125, doi:10.1111/1574-6976.12076.

85. Mujacic, M.; Baneyx, F. Regulation of *Escherichia coli hchA*, a stress-inducible gene encoding molecular chaperone Hsp31. *Mol. Microbiol.* **2006**, *60*, 1576–1589, doi:10.1111/j.1365-2958.2006.05207.x.

86. Romano, A.; Ladero, V.; Alvarez, M.A.; Lucas, P.M. Putrescine production via the ornithine decarboxylation pathway improves the acid stress survival of *Lactobacillus brevis* and is part of a horizontally transferred acid resistance locus. *Int. J. Food Microbiol.* **2014**, *175*, 14–19, doi:10.1016/j.ijfoodmicro.2014.01.009.

87. Soksawatmaekhin, W.; Kuraishi, A.; Sakata, K.; Kashiwagi, K.; Igarashi, K. Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. *Mol. Microbiol.* **2004**, *51*, 1401–1412, doi:10.1046/j.1365-2958.2003.03913.x.

88. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.

89. Schneider, T.D.; Stephens, R.M. Sequence logos: A new way to display consensus sequences. *Nucleic Acids Res.* **1990**, *18*, 6097–6100, doi:10.1093/nar/18.20.6097.

90. Crooks, G.E.; Hon, G.; Chandonia, J.M.; Brenner, S.E. WebLogo: A sequence logo generator. *Genome Res.* **2004**, *14*, 1188–1190, doi:10.1101/gr.849004.

91. Larsson, A. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics (Oxf. Engl.)* **2014**, *30*, 3276–3278, doi:10.1093/bioinformatics/btu531.

92. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259, doi:10.1093/nar/gkz239.

93. Jeffares, D.C.; Tomiczek, B.; Sojo, V.; dos Reis, M. A beginners guide to estimating the non-synonymous to synonymous rate ratio of all protein-coding genes in a genome. *Methods Mol. Biol. (Cliftonn. J.)* **2015**, *1201*, 65–90, doi:10.1007/978-1-4939-1438-8_4.

94. Suyama, M.; Torrents, D.; Bork, P. PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **2006**, *34*, W609-W612, doi:10.1093/nar/gkl315.

95. Charif, D.; Lobry, J. SeqinR 1.0-2: A contributed package to the R Project for statistical computing devoted to biological sequences retrieval and analysis. In *Structural Approaches to Sequence Evolution*; Bastolla, U., Porto, M., Roman, H.E., Vendruscolo, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 207–232.

96. Li, W.H.; Wu, C.I.; Luo, C.C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **1985**, *2*, 150–174, doi:10.1093/oxfordjournals.molbev.a040343.

97. Cohen, O.; Ashkenazy, H.; Belinky, F.; Huchon, D.; Pupko, T. GLOOME: Gain loss mapping engine. *Bioinformatics (Oxf. Engl.)* **2010**, *26*, 2914–2915, doi:10.1093/bioinformatics/btq549.

98. Cohen, O.; Pupko, T. Inference of gain and loss events from phyletic patterns using stochastic mapping and maximum parsimony--a simulation study. *Genome Biol. Evol.* **2011**, *3*, 1265–1275, doi:10.1093/gbe/evr101.

99. Fujimura, R.; Sato, Y.; Nishizawa, T.; Oshima, K.; Kim, S.W.; Hattori, M.; Kamijo, T.; Ohta, H. Complete genome sequence of *Leptospirillum ferrooxidans* strain C2-3, isolated from a fresh volcanic ash deposit on the island of Miyake, Japan. *J. Bacteriol.* **2012**, *194*, 4122–4123, doi:10.1128/jb.00696-12.

100. Jiang, H.; Liang, Y.; Yin, H.; Xiao, Y.; Guo, X.; Xu, Y.; Hu, Q.; Liu, H.; Liu, X. Effects of arsenite resistance on the growth and functional gene expression of *Leptospirillum ferriphilum* and *Acidithiobacillus thiooxidans* in pure culture and coculture. *Biomed. Res. Int.* **2015**, *2015*, 203197, doi:10.1155/2015/203197.

101. Zhang, X.; Liu, X.; Liang, Y.; Xiao, Y.; Ma, L.; Guo, X.; Miao, B.; Liu, H.; Peng, D.; Huang, W.; et al. Comparative genomics unravels the functional roles of co-occurring acidophilic bacteria in bioleaching heaps. *Front Microbiol.* **2017**, *8*, 790, doi:10.3389/fmicb.2017.00790.

102. Ferrer, A.; Bunk, B.; Spröer, C.; Biedendieck, R.; Valdés, N.; Jahn, M.; Jahn, D.; Orellana, O.; Levicán, G. Complete genome sequence of the bioleaching bacterium *Leptospirillum* sp. group II strain CF-1. *J. Biotechnol.* **2016**, *222*, 21–22, doi:http://dx.doi.org/10.1016/j.jbiotec.2016.02.008.

103. Denef, V.J.; Banfield, J.F. In situ evolutionary rate measurements show ecological success of recently emerged bacterial hybrids. *Science (N. Y. NY)* **2012**, *336*, 462–466, doi:10.1126/science.1218389.

104. Simmons, S.L.; Dibartolo, G.; Denef, V.J.; Goltsman, D.S.; Thelen, M.P.; Banfield, J.F. Population genomic analysis of strain variation in Leptospirillum group II bacteria involved in acid mine drainage formation. *PLos Biol.* **2008**, *6*, e177, doi:10.1371/journal.pbio.0060177.

105. Goltsman, D.S.; Dasari, M.; Thomas, B.C.; Shah, M.B.; VerBerkmoes, N.C.; Hettich, R.L.; Banfield, J.F. New group in the *Leptospirillum* clade: Cultivation-independent community genomics, proteomics, and transcriptomics of the new species "*Leptospirillum* group IV UBA BS". *Appl. Environ. Microbiol.* **2013**, *79*, 5384–5393, doi:10.1128/aem.00202-13.

106. Zhang, X.; Niu, J.; Liang, Y.; Liu, X.; Yin, H. Metagenome-scale analysis yields insights into the structure and function of microbial communities in a copper bioleaching heap. *BMC Genet.* **2016**, *17*, 21, doi:10.1186/s12863-016-0330-4.

107. Liljeqvist, M.; Ossandon, F.J.; González, C.; Rajan, S.; Stell, A.; Valdes, J.; Holmes, D.S.; Dopson, M. Metagenomic analysis reveals adaptations to a cold-adapted lifestyle in a low-temperature acid mine drainage stream. *FEMS Microb. Ecol.* **2015**, *91*, doi:10.1093/femsec/fiv1011, doi:10.1093/femsec/fiv011.

108. Saier, M.H., Jr.; Reddy, B.L. Holins in bacteria, eukaryotes, and archaea: Multifunctional xenologues with potential biotechnological and biomedical applications. *J. Bacteriol.* **2015**, *197*, 7–17, doi:10.1128/jb.02046-14.

109. Dubnau, D.; Blokesch, M. Mechanisms of DNA uptake by naturally competent bacteria. *Annu. Rev. Genet.* **2019**, *53*, 217–237, doi:10.1146/annurev-genet-112618-043641.

110. Romling, U.; Gomelsky, M.; Galperin, M.Y. C-di-GMP: The dawning of a novel bacterial signalling system. *Mol. Microbiol.* **2005**, *57*, 629–639, doi:10.1111/j.1365-2958.2005.04697.x.

111. Diaz, M.; Castro, M.; Copaja, S.; Guiliani, N. Biofilm formation by the acidophile bacterium *Acidithiobacillus thiooxidans* involves c-di-GMP pathway and Pel exopolysaccharide. *Genes* **2018**, *9*, 10.3390/genes9020113.

112. Moya-Beltran, A.; Rojas-Villalobos, C.; Diaz, M.; Guiliani, N.; Quatrini, R.; Castro, M. Nucleotide second messenger-based signaling in extreme acidophiles of the *Acidithiobacillus* species complex: Partition between the core and variable gene complements. *Front Microbiol.* **2019**, *10*, 381, doi:10.3389/fmicb.2019.00381.

113. Jenal, U.; Reinders, A.; Lori, C. Cyclic di-GMP: Second messenger extraordinaire. *Nat. Rev. Microbiol.* **2017**, *15*, 271–284, doi:10.1038/nrmicro.2016.190.

114. Zuckert, W.R. Secretion of bacterial lipoproteins: Through the cytoplasmic membrane, the periplasm and beyond. *Biochim. Et Biophys. Acta* **2014**, *1843*, 1509–1516, doi:10.1016/j.bbamcr.2014.04.022.

115. Cheng, X.; Guinn, E.J.; Buechel, E.; Wong, R.; Sengupta, R.; Shkel, I.A.; Record, M.T., Jr. Basis of protein stabilization by K glutamate: Unfavorable interactions with carbon, oxygen groups. *Biophys. J.* **2016**, *111*, 1854–1865, doi:10.1016/j.bpj.2016.08.050.

116. Hua, Z.S.; Han, Y.J.; Chen, L.X.; Liu, J.; Hu, M.; Li, S.J.; Kuang, J.L.; Chain, P.S.; Huang, L.N.; Shu, W.S. Ecological roles of dominant and rare prokaryotes in acid mine drainage revealed by metagenomics and metatranscriptomics. *ISME J.* **2015**, *9*, 1280–1294, doi:10.1038/ismej.2014.212.

117. Mosier, A.C.; Li, Z.; Thomas, B.C.; Hettich, R.L.; Pan, C.; Banfield, J.F. Elevated temperature alters proteomic responses of individual organisms within a biofilm community. *ISME J.* **2015**, *9*, 180–194, doi:10.1038/ismej.2014.113.

118. Ullrich, S.R.; González, C.; Poehlein, A.; Tischler, J.S.; Daniel, R.; Schlömann, M.; Holmes, D.S.; Muhling, M. Gene loss and horizontal gene transfer contributed to the genome evolution of the extreme acidophile "*Ferrovum*". *Front Microbiol.* **2016**, *7*, 10.3389/fmicb.2016.00797.

119. Feng, S.; Yang, H.; Wang, W. System-level understanding of the potential acid-tolerance components of *Acidithiobacillus thiooxidans* ZJJN-3 under extreme acid stress. *Extremophiles* **2015**, *19*, 1029–1039, doi:10.1007/s00792-015-0780-z.

120. Colman, D.R.; Poudel, S.; Hamilton, T.L.; Havig, J.R.; Selensky, M.J.; Shock, E.L.; Boyd, E.S. Geobiological feedbacks and the evolution of thermoacidophiles. *ISME J.* **2018**, *12*, 225–236, doi:10.1038/ismej.2017.162.

121. Fitch, W.M. Toward defining the course of evolution: Minimum change for a specific tree topology. *Syst. Zool* **1971**, *20*, 406–416, doi:10.2307/2412116.

122. Mirkin, B.G.; Fenner, T.I.; Galperin, M.Y.; Koonin, E.V. Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. *BMC Evol. Biol.* **2003**, *3*, 2, doi:10.1186/1471-2148-3-2.

123. Vialle, R.A.; Tamuri, A.U.; Goldman, N. Alignment modulates ancestral sequence reconstruction accuracy. *Mol. Biol. Evol.* **2018**, *35*, 1783–1797, doi:10.1093/molbev/msy055.

124. Novichkov, P.S.; Omelchenko, M.V.; Gelfand, M.S.; Mironov, A.A.; Wolf, Y.I.; Koonin, E.V. Genome-wide molecular clock and horizontal gene transfer in bacterial evolution. *J. Bacteriol.* **2004**, *186*, 6575–6585, doi:10.1128/jb.186.19.6575-6585.2004.

125. Guiliani, N.; Jerez, C.A. Molecular cloning, sequencing, and expression of *omp-40*, the gene coding for the major outer membrane protein from the acidophilic bacterium *Thiobacillus ferrooxidans*. *Appl. Environ. Microbiol.* **2000**, *66*, 2318–2324.

126. Chi, A.; Valenzuela, L.; Beard, S.; Mackey, A.J.; Shabanowitz, J.; Hunt, D.F.; Jerez, C.A. Periplasmic proteins of the extremophile *Acidithiobacillus ferrooxidans*: A high throughput proteomics analysis. *Mol Cell Proteom.* **2007**, *6*, 2239–2251, doi:10.1074/mcp.M700042-MCP200.